Journal of Complexity 29 (2013) 302–322

Contents lists available at SciVerse ScienceDirect

Journal of Complexity

journal homepage: www.elsevier.com/locate/jco

Algorithms for quaternion polynomial root-finding

Bahman Kalantari

Department of Computer Science, Rutgers University, Piscataway, NJ, USA

ARTICLE INFO

Article history:
Received 14 December 2011
Accepted 26 February 2013
Available online 30 April 2013

Keywords:
Quaternions
Polynomial roots
Fundamental theorem of algebra
Newton's method
Recurrence relation
Polynomiography

ABSTRACT

In 1941 Niven pioneered root-finding for a quaternion polynomial $P(x)$, proving the fundamental theorem of algebra (FTA) and proposing an algorithm, practical if the norm and trace of a solution are known. We present novel results on theory, algorithms and applications of quaternion root-finding. Firstly, we give a new proof of the FTA resulting in explicit formulas for both exact and approximate quaternion roots of $P(x)$ in terms of exact and approximate complex roots of the real polynomial $F(x) = P(x)\overline{P}(x)$, where $\overline{P}(x)$ is the conjugate polynomial. In particular, if $|F(c)| \leq \epsilon$, then for a computable quaternion conjugate q of c, $|P(q)| \leq \sqrt{\epsilon}$. Consequences of these include relevance of root-finding methods for complex polynomials, computation of bounds on zeros, and algebraic solution of special quaternion equations. Secondly, working directly in the quaternion space, we develop Newton and Halley methods and analyze their local behavior. Surprisingly, even for a quadratic quaternion polynomial Newton's method may not converge locally. Finally, we derive an analogue of the Bernoulli method in the quaternion space for computing the dominant root in certain cases. This requires the development of an independent theory for the solution of quaternion homogeneous linear recurrence relations. These results also lay a foundation for quaternion polynomiography.

1. Introduction

The set of quaternions is an extension of the complex field $\mathbb{C}$, first described by Sir William Rowan Hamilton in 1843. They are defined as

$$\mathbb{H} = \{q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mid a, b, c, d \in \mathbb{R}\}, \quad \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1. \tag{1}$$

E-mail address: kalantari@cs.rutgers.edu.

0885-064X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jco.2013.03.001

The equations in (1) imply $\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \mathbf{ki} = -\mathbf{ik} = \mathbf{j}$. The quaternions form a division ring and can be identified with $\mathbb{R}^4 : a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \Longleftrightarrow (a, b, c, d)$. While there is a vast literature for polynomial root-finding over the real and complex fields (see e.g. Jenkins and Traub [18], Traub [47], Pan [40], McNamee [33], Kalantari [27]), the published results for quaternions are rather sparse. We will review the existing literature. In this article we are interested in algorithms for the computation of roots of a quaternion polynomial

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad a_i \in \mathbb{H},\ a_n \neq 0. \tag{2}$$

These are sometimes referred as *unilateral* quaternion polynomials. Our previous studies of complex polynomial root-finding with iteration functions, including their visualization, called *polynomiography* (see e.g. [27]), make it natural to ask to what extent can the existing results be generalized to quaternion polynomials. As a result of this study we will need to consider new topics over the quaternions, extending many properties of complex polynomials and developing new theories. We will extend some existing results on complex polynomial root-finding to quaternions polynomial root-finding. In particular, we will investigate the applicability of specific members of a fundamental family of iterations, called the *Basic Family*, extensively studied over the complex field (see [21–30]). Given a complex polynomial $p(z)$, specific members of the family are $B_2(z)$ (Newton) and $B_3(z)$ (Halley)

$$B_2(z) \equiv z - \frac{p(z)}{p'(z)}, \qquad B_3(z) \equiv z - \frac{p(z)}{p'(z)} - \frac{p''(z)p^2(z)}{p'(z)(2p'^2(z) - p''(z)p(z))}. \tag{3}$$

For a history of these two methods, see Traub [47], Kalantari [27]. The Basic Family was studied by Schröder [44], and extensively in [27] where in particular several different equivalent formulations are shown, as well as many other properties including their use in polynomiography. In this article we will develop Newton and Halley methods over the quaternions. Surprisingly, even for a quadratic quaternion polynomial Newton's method may not converge locally.

Over the quaternions there is a very limited study of iteration functions, perhaps primarily due to the difficulties caused by noncommutativity of multiplication. Janovska and Opfer [17] study Newton's method for a very special case of quaternionic roots. Bedding and Briggs [4] study iteration functions over the quaternions, but the focus is essentially on the Mandelbrot set. The Mandelbrot set and the quadratic iterations that define it and their properties do not capture the properties of iteration functions for polynomial root-finding over the complex domain. Analogously, quadratic iterations over the quaternions do not capture polynomial-root-finding iterations over this division ring. There is need for much more research, both from the theoretical point of view, as well as from the practical point of view. In particular, despite some previous work on the visualization of quadratic quaternion equations and their Julia sets (see e.g. Norton [38]), polynomiography over the quaternions is undoubtedly interesting. Not only it could result in a better understanding of iteration functions such as Newton and Halley methods, but it could lead to a rich set of 2D and 3D images.

The article is organized as follows. In Section 2, we describe basic properties of quaternions. In Section 3, we define quaternion polynomials and their evaluations. In Section 4, we review basics of quaternion polynomials and the relevant literature. In Section 5, we give a constructive proof of the fundamental theorem of algebra for quaternion polynomials and consider its applications. In Section 6, we give a method for approximating zeros of a quaternion polynomial based on the approximate complex zeros of a real polynomial. In Section 7, we show how to derive bounds on the norm of zeros of a quaternion polynomial. In Section 8, we describe a decomposition theorem for a quaternion polynomial. In Section 9, we describe a Taylor's theorem for quaternion polynomials and develop Newton and Halley methods and analyze their local behavior. We also describe a simple method for constructing a quaternion polynomial with prescribed zeros. In Section 10, we define homogeneous linear recurrence relation over the quaternions and develop a theory for its solution, and describe its connections to polynomial root-finding. We conclude the article by describing future work which in particular gives rise to various quaternion polynomiography.

2. Preliminary quaternion properties

In this section, we introduce some basic properties of quaternions. Several relevant references are Lam [31], Niven [36], and Zhang [48]. A quaternion and its *conjugate* are defined as

$$q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d, \qquad \overline{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d. \tag{4}$$

The number a is called the *real part* of q, written as Re q. The *trace* and *norm* of q are, respectively

$$t(q) = q + \overline{q} = 2\,\mathrm{Re}\,q, \qquad \nu(q) = |q| = \sqrt{a^2 + b^2 + c^2 + d^2}. \tag{5}$$

The *inverse* of a nonzero quaternion q is the unique quaternion denoted by q^{-1} such that $qq^{-1} = q^{-1}q = 1$. It follows that $q^{-1} = \overline{q}/|q|^2$. It is easy to verify that given quaternions q_1, q_2, we have

$$\overline{q_1 + q_2} = \overline{q_1} + \overline{q_2}, \qquad \overline{q_1 q_2} = \overline{q_2}\,\overline{q_1}. \tag{6}$$

Conjugation is linear over the reals but does not define a homomorphism on $\mathbb{H}$. From (6) we get

$$|q_1 q_2| = \sqrt{q_1 q_2 \overline{q_1 q_2}} = \sqrt{q_1 q_2 \overline{q_2}\,\overline{q_1}} = |q_2|\sqrt{q_1 \overline{q_1}} = |q_1|\,|q_2|. \tag{7}$$

The division of a quaternion q_1 by $q_2 \neq 0$ must be specified either as $q_1 q_2^{-1}$ or as $q_2^{-1} q_1$. If q_1, q_2 are nonzero then $(q_1 q_2)^{-1} = q_2^{-1} q_1^{-1}$.

Two quaternions q and q' are said to be *congruent* or *equivalent*, written $q \sim q'$, if for some quaternion $w \neq 0$ we have, $q' = wqw^{-1}$. The *congruence class* of $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, denoted by $[q]$, is the set

$$[q] = \{q' \in \mathbb{H} \mid q' \sim q\}. \tag{8}$$

Since the norm of the product of quaternions is the product of the norms, we have $|q'| = |wqw^{-1}| = |q|$. It is straightforward to verify that q and q' have the same real parts. Thus,

$$[q] \subset \{q' \in \mathbb{H} \mid \mathrm{Re}\,q' = \mathrm{Re}\,q,\ |q'| = |q|\}. \tag{9}$$

On the other hand, any quaternion is congruent to a complex number with the same real part and norm (see e.g. Zhang [48]). Specifically, $q \sim a + \mathbf{i}\sqrt{b^2 + c^2 + d^2}$. Thus we have

$$[q] = \{a + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k} \mid x_2^2 + x_3^2 + x_4^2 = b^2 + c^2 + d^2\}. \tag{10}$$

From (10) it follows that $[q]$ is a singleton element if and only if q is a real number. If q is not real, its congruent class is the three-dimensional sphere in the coordinate space of x_2, x_3, x_4, centered at the point $(a, 0, 0, 0)$, having radius equal to $\sqrt{b^2 + c^2 + d^2}$.

For a given $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H} - \mathbb{R}$ the *characteristic polynomial* and *characteristic equation* are

$$P_q(x) = x^2 - t(q)x + \nu^2(q), \qquad P_q(x) = 0. \tag{11}$$

From (10) it follows that for any $q' \in [q]$ we have

$$P_{q'}(x) = x^2 - t(q')x + \nu^2(q') = (x - q')(x - \overline{q'}) = P_q(x). \tag{12}$$

The *discriminant* of $P_q(x)$ is

$$\Delta = t^2(q) - 4\nu^2(q) = -4(b^2 + c^2 + d^2) < 0. \tag{13}$$

Denoting the set of zeros of $P_q(x)$ by $\mathbf{Z}_q$, it is easy to show that for $q \in \mathbb{H} - \mathbb{R}$, $\mathbf{Z}_q = [q]$. In particular, the quadratic polynomial $x^2 + 1$ is the characteristic polynomial of $\pm\mathbf{i}$, $\pm\mathbf{j}$, $\pm\mathbf{k}$, and its zeros constitute the unit sphere in the space of x_2, x_3, x_4, centered at the origin $(0, 0, 0, 0)$:

$$\{x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k} \in \mathbb{H} \mid x_2^2 + x_3^2 + x_4^2 = 1\}. \tag{14}$$

Conversely, given reals $t \geq 0$ and ν satisfying $t^2 - 4\nu^2 < 0$, the polynomial $G(x) = x^2 - tx + \nu^2$ has roots $\theta = \frac{1}{2}(t + \mathbf{i}\sqrt{t^2 - 4\nu^2})$ and its conjugate $\overline{\theta}$. The set of roots of $G(x)$ in $\mathbb{H}$ is $[\theta]$. For $q \in \mathbb{H}$, let

$$C(q) = \{q' \in \mathbb{H} \mid qq' = q'q\}, \tag{15}$$

i.e. the set of quaternion that commute with q. It is not difficult to show that for $q \in \mathbb{H} - \mathbb{R}$ we have, $C(q) = \mathbb{R}[q]$, the set of all polynomials in q with real coefficients. Since $q \in \mathbb{H} - \mathbb{R}$ is a root of its characteristic polynomial, $q^2 = \alpha q + \beta$, for some real numbers α, β. It follows that for any nonzero integer t, q^t can be written as a linear polynomial in q and hence

$$C(q) = \mathbb{R}[q] = \{\alpha + \beta q \mid \alpha, \beta \in \mathbb{R}\}, \tag{16}$$

a field isomorphic to $\mathbb{C}$.

On the one hand, in contrast with complex polynomials, a quaternion polynomial can have infinitely many roots. On the other hand, a quaternion polynomial could have fewer roots than its degree without multiplicity. For example, the only root of the quadratic polynomial $x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k}$ is $\mathbf{j}$, having multiplicity of one. We will justify this later.

We conclude this section by pointing out that the quaternions can be written as a direct sum, $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}\mathbf{j}$, so that an algebraic system such as Maple with built-in complex arithmetic can represent quaternions as linear polynomials. It suffices to write a special routine for multiplication, e.g.

$$(\alpha + \beta\mathbf{j})(\gamma + \delta\mathbf{j}) = (\alpha\gamma - \beta\overline{\delta}) + (\alpha\delta + \beta\overline{\gamma})\mathbf{j}. \tag{17}$$

3. Quaternion polynomials and their evaluation

We will consider general polynomials over the quaternions. Let $\mathbb{H}[x]$ denote the polynomial ring over $\mathbb{H}$, i.e. the set of all polynomials

$$P(x) = a_n x^n + \cdots + a_1 x + a_0, \quad a_i \in \mathbb{H},\ a_n \neq 0,\ n \geq 1. \tag{18}$$

The *quaternion conjugate* of $P(x)$, or just *conjugate*, is

$$\overline{P}(x) = \overline{a}_n x^n + \cdots + \overline{a}_1 x + \overline{a}_0. \tag{19}$$

The addition and multiplication of quaternion polynomials is defined as in the commutative case where the variable x is assumed to commute with the quaternion coefficients. Such quaternion polynomials are sometimes referred as *unilateral* quaternionic polynomials. For properties of such polynomials over general division rings see Lam [31]. For a given $q \in \mathbb{H}$, $P(q)$ denotes the evaluation of P at q, i.e.

$$P(q) = a_n q^n + \cdots + a_1 q + a_0. \tag{20}$$

Unlike complex polynomials, the evaluation map $q \mapsto P(q)$ over the quaternions does not define a ring homomorphism from $\mathbb{H}[x]$ to $\mathbb{H}$. This in a sense is a source of difficulty in polynomial root-finding over the quaternions. If $P(q) = 0$ we say q is a *zero* (or a *root*) of $P(x)$. The set of all zeros of $P(x)$ will be denoted by $\mathbf{Z}_P$. While the variable x in $P(x)$ is assumed to commute with the coefficients, in its evaluation at a particular $q \in \mathbb{H}$ the coefficients do not necessarily commute with q.

A quaternion q is a root of $P(x)$ if and only if $P(x) = Q(x)(x - q)$ for some $Q(x) = \sum_{i=0}^{n-1} q_i x^i$. To prove this, suppose $P(x) = Q(x)(x - q)$. Then

$$P(x) = \sum_{i=0}^{n-1} q_i x^i (x - q) = \sum_{i=0}^{n-1} q_i x^{i+1} - \sum_{i=0}^{n-1} q_i q x^i. \tag{21}$$

Then $P(q) = 0$. Conversely, if $P(q) = 0$ by the Euclidean algorithm $P(x) = Q(x)(x - q) + r$. But we must have $r = 0$. However, if $P(q') = 0$ for some $q' \neq q$, it does not necessarily follow that $Q(q') = 0$. As an example if $P(x) = (x - \mathbf{i})(x - \mathbf{j}) = x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k}$, then $\mathbf{j}$ is a root but $\mathbf{i}$ is not a root. The example shows that if $(x - q)$ is a factor of $P(x)$ on the right it may not be a factor on the left (see example right after Theorem 5.3).

4. Basics of quaternion polynomials and literature review

Niven [36] pioneered solving polynomial equations over the quaternions by proving the fundamental theorem of algebra (FTA): a nonconstant quaternion polynomial must have a root in $\mathbb{H}$. In fact Eilenberg and Niven [9] gave a topological proof of the fundamental theorem of algebra over the quaternions for the general case where x and the constants do not necessarily commute and the polynomial consists of a highest term of the form $b_1xb_2x\cdots xb_n$, plus the sum of lower order terms of the same type. In particular, a unilateral polynomial $P(x) \in \mathbb{H}[x]$ of positive degree must have at least a root in $\mathbb{H}$.

In this section, we will review some basic properties of quaternion polynomials as well as some existing algorithms for computing their roots. In what follows, we first describe what has become known as Niven's algorithm [36]. Given a quaternion polynomial $P(x)$ and a nonzero quaternion q with trace t and norm ν, Niven showed that by division we have

$$P(x) = Q(x)(x^2 - tx + \nu^2) + f(a_i, t, \nu^2)x + g(a_i, t, \nu^2), \tag{22}$$

where f and g are polynomials in t, ν^2 and the coefficients a_i, and Q is some quaternion polynomial. If q is a zero of $P(x)$, from the Factor Theorem (Theorem 4.1) it follows that by substituting into (22) we get

$$P(q) = fq + g = 0. \tag{23}$$

Clearly, $f = 0$ if and only if $g = 0$. Thus if $f = 0$, then $(x^2 - tx + \nu^2)$ is a factor of $P(x)$ and this in fact implies $[q] \subset \mathbf{Z}_P$. If $f \neq 0$, solving for q gives a root of $P(x)$

$$q = -f^{-1}g. \tag{24}$$

As an example consider the polynomial $P(x) = (x - \mathbf{i})(x - \mathbf{j}) = x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k}$ defined earlier. Then $\mathbf{j}$ is a root of this polynomial and its characteristic polynomial is $x^2 + 1$. Dividing $P(x)$ by $x^2 + 1$ gives

$$x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k} = (x^2 + 1) + (\mathbf{i} + \mathbf{j})x + (\mathbf{k} - 1). \tag{25}$$

This shows $\mathbf{j}$ is a root as we get

$$x = (\mathbf{i} + \mathbf{j})^{-1}(\mathbf{k} - 1) = -\frac{1}{2}(\mathbf{i} + \mathbf{j})(\mathbf{k} - 1) = -\frac{1}{2}(-\mathbf{j} - \mathbf{i} + \mathbf{i} - \mathbf{j}) = \mathbf{j}. \tag{26}$$

In fact $\mathbf{j}$ is the only solution to $P(x) = 0$. This we justify after describing the Factor Theorem, Theorem 4.1.

In general, if the trace and norm of a root of $P(x)$ are known, dividing $P(x)$ by the characteristic polynomial of the root, either the characteristic polynomial is a factor of $P(x)$, or we determine a linear equation which gives a root. In the latter case $f \neq 0$, and

$$N(t, \nu) = \nu^2\bar{f}f - \bar{g}g = 0, \qquad T(t, \nu) = t\bar{f}f + \bar{f}g + \bar{g}f = 0. \tag{27}$$

Clearly, one cannot easily determine the trace and norm of a root of $P(x)$. However, as Niven showed, conversely if we consider the equations in (27) as simultaneous equations in the real pair of variables (t, ν), then there is a real pair of solutions. Furthermore, any such real pair of solutions gives the trace and norm of a zero q of $P(x)$. Thus Niven's algorithm consists of solving (27) and once the trace and norm are known, to proceed as described earlier.

Despite the elegance of Niven's algorithm, this approach is computationally impractical as it requires a simultaneous solution of two equations of degree $2n-1$. Equivalently, such system through the use of resolvent can be reduced to a single real polynomial of degree $O(n^2)$, and could be inefficient to solve when n is large. The division in Niven's algorithm however is the basis for several algorithms in quaternion root-finding literature. Furthermore, it can be used to obtain some interesting results about quaternion polynomials. For instance, one can use Niven's division to conclude the following result proved in [46], and in more generality for division rings in Bray and Whaples [6] and Gordon and Motzkin [10].

Proposition 4.1. *If q and a distinct conjugate q' are both roots of $P(x)$, then so is any element of the conjugacy class $[q]$. In particular, $P(x) = Q(x)P_q(x)$ for some quaternion polynomial $Q(x)$.*

Proof. Suppose $P_q(x)$ is not a factor of $P(x)$. Then applying Niven's division it follows that both q and q' are solutions of the corresponding equation $fx+g = 0$, with $f, g \neq 0$. But there is a unique solution, a contradiction. □

A key result on polynomials over a division ring which can be proved in a straightforward fashion is the following (see Lam [31]).

Theorem 4.1 (*Factor Theorem*)**.**

(i) *Given $P(x) \in \mathbb{H}[x]$, $q \in \mathbb{H}$ is a root if and only if there exists $Q(x) \in \mathbb{H}[x]$ such that*

$$P(x) = Q(x)(x - q). \tag{28}$$

(ii) *Suppose $P(x) = Q(x)H(x)$ for some polynomials $Q(x), H(x) \in \mathbb{H}[x]$. Given $q \in \mathbb{H}$, suppose $a = H(q) \neq 0$. Then*

$$P(q) = Q(aqa^{-1})H(q). \tag{29}$$

If q is a root of $P(x)$ and $a = H(q) \neq 0$, then aqa^{-1} is a root of $Q(x)$. □

In the next section, we will use the Factor Theorem (Theorem 4.1) to give a novel proof of the FTA and then in subsequent sections we describe several practical applications of the theorem. We will now proceed to give some definitions for the classification of zeros of a quaternion polynomial.

Definition 4.1. Let $P(x)$ be a quaternion polynomial with q a root.

(i) We say q is a *root of multiplicity k* if $P(x) = Q(x)(x-q)^k$, where $Q(q) \neq 0$. In particular, if $k = 1$, q is a simple root.
(ii) We say q is an isolated root if there exists a neighborhood of q that contains no other root of $P(x)$.
(iii) We say q is a *spherical root* if $[q]$ is contained in Z_P (the set of zeros of $P(x)$). We call $[q]$ a *sphere of zeros* for $P(x)$.
(iv) We say q is a *distinguished spherical root* if q is a spherical zero of $P(x)$, but q has no distinct conjugate as a root with the same multiplicity.
(v) We say $[q]$ is a *sphere of zeros of order k* if each $q' \in [q]$ is a root of multiplicity k.

An isolated root may have multiplicity larger than one. As an example of distinguished spherical root, consider $\mathbf{i}$ as a root of $P(x) = (x - \mathbf{i})(x^2 + 1)$. The following proposition justifies the definition of distinguished spherical roots and those of multiplicity higher than one. It is a consequence of Proposition 4.1 and the Factor Theorem (Theorem 4.1).

Proposition 4.2. *Each sphere of zeros of a quaternion polynomial has at most one spherical root of higher multiplicity than its distinct conjugates.* □

The following structural theorem on the zeros of a quaternion polynomial, stated in Pogorui and Shapiro [41], follows easily as a consequence of the Factor Theorem (Theorem 4.1) and Proposition 4.1. Later we will prove a more precise decomposition theorem, Theorem 8.1.

Theorem 4.2. *Let $P(x)$ be a quaternion polynomial of degree n. Then the number of isolated zeros, plus twice the number of spheres of zeros, counted with multiplicity, is at most n.* □

We now use the Factor Theorem (Theorem 4.1) to argue that the quadratic equation $P(x) = x^2 - (\mathbf{i}+\mathbf{j})x + \mathbf{k} = (x - \mathbf{i})(x - \mathbf{j})$ has only one simple root, namely $\mathbf{j}$. From the Factor Theorem any other root of $P(x)$ must be a conjugate of $\mathbf{i}$. But since this root also has the same characteristic polynomial as $\mathbf{j}$, Niven's division leads to the same solution.

In the remainder of this section, we review some known results on quaternion polynomial root-finding algorithms. Unlike the complex case, over the quaternions even solving a quadratic equation is nontrivial. Huang and So [16] give complete formulas for a quadratic equation $x^2 + bx + c = 0$, breaking it into several cases. Applications of this special case of quaternion equation are described in Heidrich [11] and Huang and So [15]. Prior to the formulas of Huang and So [16], others have studied solutions of quadratic quaternion equations (see e.g. Niven [36] and Zhang and Mu [49]).

Indeed our results, to be proved in the next section, suggest a new and simple approach for computing the solutions of a quadratic polynomial, simplifying Huang and So's [16] formulas. This simplification is in the sense that solving a quaternion quadratic equation is reduced to solving a real quartic equation for which closed formulas are well-known.

The quadratic equations where x and the coefficients may not commute are yet more complicated. Porter [42] considers this case but when one solution is known. A method for solving more general nonunilateral quadratic equations is described in Jia et al. [19].

For general unilateral quaternion polynomials, there are computational approaches such as those proposed by Serodio et al. [45], who consider the companion matrix of the quaternion polynomial and convert the root-finding problem into an eigenvalue problem and then use Niven's algorithm. De Leo et al. [8] also consider the computation of the companion matrix and the eigenvector connection. Janovska and Opfer [17] consider solving the equation $x^n - a$, i.e. quaternionic roots by Newton's method and obtain computational results. More specifically, they study the convergence properties of a formal Newton's method and cases where it may or may not converge locally. Since a derivative in the sense of complex analysis does not exist, they also make use of the Jacobian matrix and the Gateaux derivative. Bedding and Briggs [4] have examined the iterations of quaternion functions and the issues in the extension of the notion of regularity of functions (i.e. applicability of rules of differentiation) from complex domain to quaternion domain. In particular, they examine the Mandelbrot set [32] in the quaternion space and consider extensions of the notion of stability of cycles needed in the analysis of this set. This is no longer determined by the derivative, rather by the eigenvalues of the Jacobian map.

In this article, we develop Newton and Halley methods for general quaternion polynomials and study their local behavior of the iterations near a root. We will compare our results to the previously known results. These previous results include the structure of zeros studied by Pogorui and Shapiro [41], bound on zeros studied by Opfer [39], and properties of Vandermonde matrix over the quaternions studied by Opfer [39] and Renmin et al. [43]. The general interest in solving roots of quaternion polynomials is due to its connection to quantum mechanical problems (see Adler [1]). In this article our interest lies in the study of theoretical aspects, however we will also investigate applications. In particular, the results will be useful in the visualization of quaternion polynomial root-finding through polynomiography, to be considered in forthcoming articles.

5. A constructive proof of quaternion FTA and its applications

This section presents some of our main results. Here we develop a new proof of the fundamental theorem of algebra for quaternion polynomials. Indeed using the Factor Theorem (Theorem 4.1) and the ordinary FTA for complex polynomials, not only do we prove the existence of a root of a quaternion polynomial, but directly connect its roots, and hence their computation, to algorithms for complex polynomial root-finding. In this sense our proof is more direct, constructive, and algorithmically more useful than the previously known results.

More precisely, given a quaternion polynomial $P(x)$ we define its conjugate $\overline{P}(x)$ and show that the function $F(x) = P(x)\overline{P}(x)$ is a polynomial with real coefficients. We will then establish a precise relationship between the exact and approximate roots of $P(x)$ and $F(x)$. These are accomplished in Theorems 5.1 and 6.2. This in particular results in a simple and straightforward proof of the FTA. It is interesting that in his pioneering article, Niven [36] proved the fundamental theorem of algebra by considering the function $\overline{P}(x)P(x)$, which is the conjugate of $F(x)$. But since $F(x)$ has real coefficients this conjugate coincides with $F(x)$. However Niven seems to have followed a different route to prove the FTA, an approach that would make it more difficult to compute a root. Indeed others have used the function $\overline{P}(x)P(x)$ in different contexts, including its use in computing the roots of $P(x)$, e.g. Serodio and Siu [46] and Pogorui and Shapiro [41]. Theorem 5.1, while resulting in a simple proof of the FTA, also gives explicit information on the connections between the zeros of $P(x)$ and $F(x)$, and we will even make use of this connection further in the next sections. Theorem 5.1 also motivates our approximation result, Theorem 6.2, to be proved in the next section. Theorem 6.2 gives the needed accuracy in approximation of $F(x)$ via a complex polynomial root-finding algorithm, in order to have $P(x)$ within a prescribed accuracy. The following is straightforward.

Proposition 5.1. *Let $G(x)$ be a quaternion polynomial and $H(x)$ a real polynomial (i.e. having real coefficients). Then they commute, i.e*

$$G(x)H(x) = H(x)G(x). \tag{30}$$

More strongly, the center of $\mathbb{H}[x]$ is $\mathbb{R}[x]$. □

The following theorem gives an algebraic reduction of FTA for quaternion polynomials to the usual FTA. It is much more direct and simpler than Niven's proof outlined earlier.

Theorem 5.1 (*FTA for Quaternion Polynomials*)**.** *Consider $P(x) = \sum_{i=0}^{n} a_i x^i$ and its conjugate $\overline{P}(x) = \sum_{i=0}^{n} \overline{a}_i x^i$. Define*

$$F(x) = P(x)\overline{P}(x). \tag{31}$$

Then $F(x)$ has real coefficients. Let $\theta \in \mathbb{C}$ be a root of $F(x)$. Then $P(x)$ has a root in $[\theta]$. Furthermore, if $q \in \mathbb{H}$ is a root of $P(x)$, then there exists a complex number $\theta \in [q]$ which is a root of $F(x)$. More specifically,

(i) *Suppose $\theta \in \mathbb{C}$ is a root of $F(x)$. If θ is not a root of $\overline{P}(x)$, then*

$$\overline{P}(\theta)\theta\overline{P}(\theta)^{-1} \tag{32}$$

is a root of $P(x)$.

(ii) *Suppose $\theta \in \mathbb{C}$ is a root of $F(x)$. If θ is a root of $\overline{P}(x)$, and $\overline{\theta}$ is not a root of $\overline{P}(x)$, then*

$$\overline{P}(\overline{\theta})\overline{\theta}\,\overline{P}(\overline{\theta})^{-1} \tag{33}$$

is a root of $P(x)$.

(iii) *Suppose $\theta \in \mathbb{C}$ is a root of $F(x)$. If θ and $\overline{\theta}$ are both roots of $\overline{P}(x)$, then both θ and $\overline{\theta}$ are also roots of $P(x)$. In particular, if $\theta \in \mathbb{R}$, $P(\theta) = 0$.*

(iv) *Suppose $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$ is a root of $P(x)$. Then $\theta = a + \mathbf{i}\sqrt{b^2 + c^2 + d^2}$ is a root of $F(x)$.*

Proof. We first prove that $F(x)$ has real coefficients. We have

$$F(x) = P(x)\overline{P}(x) = \left(\sum_{i=0}^{n} a_i x^i\right)\left(\sum_{j=0}^{n} \overline{a}_j x^j\right) = \sum_{k=0}^{2n} \sum_{i+j=k} a_i \overline{a}_j x^k. \tag{34}$$

The coefficient of x^k thus satisfies

$$\sum_{i+j=k} a_i \overline{a}_j = \sum_{i+j=k} a_j \overline{a}_i = \sum_{i+j=k} \overline{a_i \overline{a}_j}. \tag{35}$$

Hence $F(x)$ has real coefficients.

Proof of (i): this follows from the direct application of the Factor Theorem (Theorem 4.1) to $F(x)$.

Proof of (ii): since $F(x)$ has real coefficients, the complex conjugate of a root is also a root. Hence $\overline{\theta}$ is also a root of $F(x)$ and we can apply part (i).

Proof of (iii): it can be shown that if $P(x) = U(x)V(x)$ for some polynomials $U(x)$ and $V(x)$, then $\overline{P}(x) = \overline{V}(x)\overline{U}(x)$. If $\theta \in \mathbb{C} - \mathbb{R}$, using the fact that both θ and $\overline{\theta}$ are roots of $\overline{P}(x)$ and Proposition 4.1, we conclude

$$\overline{P}(x) = G(x)P_\theta(x) = P_\theta(x)G(x). \tag{36}$$

Thus we have

$$P(x) = \overline{G}(x)P_\theta(x). \tag{37}$$

But this implies that θ and $\overline{\theta}$ are roots of $P(x)$. If $\theta \in \mathbb{R}$, then $\overline{P}(\theta) = \overline{P(\theta)}$. Thus $\overline{P}(\theta) = 0$ if and only if $P(\theta) = 0$.

Proof of (iv): if q is a real number from the Factor Theorem (Theorem 4.1), q is also a root of $F(x)$. Suppose q is not a real number. From the Factor Theorem, $P(x) = Q(x)(x-q)$. Thus $\overline{P}(x) = (x-\overline{q})\overline{Q}(x)$. Using this, Proposition 5.1, and that q and θ have the same characteristic polynomial gives

$$F(x) = P(x)\overline{P}(x) = Q(x)(x-q)(x-\overline{q})\overline{Q}(x) = Q(x)\overline{Q}(x)P_q(x) = Q(x)\overline{Q}(x)P_\theta(x). \tag{38}$$

Hence θ is a root of $F(x)$. □

Theorem 5.1 is a powerful result and can be used to construct not only the roots of a quaternion polynomial from those of a real polynomial, but their approximation, as well as several other applications we will demonstrate later in this article. First, we consider an example of its application.

Example 5.1 (*Computing n-th Roots of a Quaternion*)**.** Consider $P(x) = x^n - q$, $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. We consider the solutions of $F(x) = P(x)\overline{P}(x) = x^{2n} - (q+\overline{q})x^n + q\overline{q}$. If $q \in \mathbb{R}$, then $F(x) = (x^n - q)^2$. Its roots are the complex n-th roots of q. However, since $P(x)$ and $\overline{P}(x)$ are identical, for $n \geq 3$ the complex roots of $F(x)$ come in conjugate pairs and part (iii) of Theorem 5.1 implies that the roots of $P(x)$ also come in conjugate pairs. Thus, there are infinitely many roots to $P(x)$. In fact in this case one can obtain an exact factorization of $P(x)$: it suffices to consider only two special cases, when $q = \pm 1$, since other values of q can only change the norm of the roots. If $n \geq 3$ is odd, the roots of $x^n - 1$ consist of a single real root, namely 1, as well as $(n-1)/2$ spheres of zeros. If $n \geq 4$ is even, the roots of $x^n - 1$ consist of $n/2$ spheres of zeros.

If $q \in \mathbb{H} - \mathbb{R}$, then $F(x) = P_q(x^n)$. To compute the roots of $F(x)$, first we compute the complex roots of $P_q(x)$. These are the complex numbers $\theta = a + \mathbf{i}\sqrt{b^2+c^2+d^2}$ and $\overline{\theta} = a - \mathbf{i}\sqrt{b^2+c^2+d^2}$. Thus the roots of $F(x)$ are the complex n-th roots of θ and the complex n-th roots of $\overline{\theta}$. The complex n-th roots of θ are $\eta\omega^i$, $i = 0, 1, \ldots, n-1$, where η is any of the n-th roots of θ and $1, w, \ldots, w^{n-1}$ are the n-th roots of unity.

Quaternion n-th roots have been studied before, e.g. by Brand [5] and Cho [7], and Niven [37]. However, our approach demonstrates the automatic formulas for these roots based on the complex roots of $F(x)$. For a quadratic quaternion equation, we can also make use of the theorem to give exact formulas as well as computational algorithms. We shall demonstrate these and their polynomiography in a separate article.

From Theorem 5.1, the fundamental theorem of algebra for complex polynomials, and the Factor Theorem (Theorem 4.1), we have

Theorem 5.2. *Given a monic polynomial $P(x) \in \mathbb{H}[x]$ of degree n, for some set of quaternions $q_i \in \mathbb{H}$, $i = 1, \ldots, n$, we have*

$$P(x) = (x-q_n)\cdots(x-q_2)(x-q_1). \quad \square \tag{39}$$

From Theorem 5.2, the definition of $F(x)$, and repeated application of Proposition 4.1, we can state a precise decomposition for $\overline{P}(x)$ and $F(x)$:

Theorem 5.3. *Given a monic polynomial $P(x) \in \mathbb{H}[x]$ of degree n, for some set of quaternions $q_i \in \mathbb{H}$, $i = 1, \ldots, n$ we have*

$$P(x) = (x-q_n)\cdots(x-q_2)(x-q_1), \tag{40}$$

$$\overline{P}(x) = (x-\overline{q}_1)\cdots(x-\overline{q}_{n-1})(x-\overline{q}_n), \tag{41}$$

$$F(x) = P(x)\overline{P}(x) = P_1(x)P_2(x)\cdots P_n(x), \tag{42}$$

where for $i = 1, \ldots, n$, $P_i(x)$ is the characteristic polynomial of q_i, if q_i is not real, and $(x-q_i)^2$ otherwise. □

From Theorem 5.3 and the Factor Theorem (Theorem 4.1), it follows that q_1 is a zero of $P(x)$. But it does not necessarily follow that q_2 is also a root of $P(x)$. As an example, we have seen that while $\mathbf{j}$

is a root of $(x - \mathbf{i})(x - \mathbf{j}) = x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k}$, $\mathbf{i}$ is not a root. We may also construct an example of a quadratic $(x - q_2)(x - q_1)$ where q_1 and q_2 are from different conjugacy classes and q_2 is not a root. We may ask: given the factorization of $P(x)$ into linear terms what are its roots? As we have seen the factorization into linear terms is not unique, e.g. $(x^2 + 1)$, or the case of characteristic polynomial of a quaternion $\mathbb{H} - \mathbb{R}$. We have the following theorem.

Theorem 5.4 (*Serodio and Siu [46]*). *Suppose $P(x) = (x - q_n) \cdots (x - q_2)(x - q_1)$, $q_i \in \mathbb{H}$. Then for each $i = 1, \ldots, n$, $P(x)$ has a zero in $[q_i]$.* □

In fact using Theorem 5.1 we may state a precise version of Theorem 5.4.

Theorem 5.5. *Suppose*

$$P(x) = (x - q_n) \cdots (x - q_2)(x - q_1), \quad q_i = a_i + b_i\mathbf{i} + c_i\mathbf{j} + d_i\mathbf{k} \in \mathbb{H}. \tag{43}$$

Then if for each $i = 1, \ldots, n$ we set $\theta_i = a_i + \mathbf{i}\sqrt{b_i^2 + c_i^2 + d_i^2}$, then either $\overline{P}(\theta_i)\theta_i\overline{P}(\theta_i)^{-1}$ is a root of $P(x)$, or $\overline{P}(\overline{\theta}_i)\overline{\theta}_i\overline{P}(\overline{\theta}_i)^{-1}$ is a root of $P(x)$. □

6. Approximate zeros of a quaternion polynomial

In this section, we connect the approximation of $F(x) = P(x)\overline{P}(x)$ to those of $P(x)$ and $\overline{P}(x)$. Here our goal is to approximate $|P(x)|$ to prescribed accuracy $\epsilon \in (0, 1)$, based on approximation of $|F(x)|$.

Proposition 6.1. *Let $P(x)$ be a quaternion polynomial of degree n. Given $q \in \mathbb{H} - \mathbb{R}$ and its characteristic polynomial $P_q(x) = (x^2 - tx + \nu^2)$, suppose*

$$P(x) = Q(x)P_q(x) + \alpha x + \beta. \tag{44}$$

Then

$$\overline{P}(x) = \overline{Q}(x)P_q(x) + \overline{\alpha}x + \overline{\beta}. \tag{45}$$

In particular, if $\theta \in \mathbb{C}$ is a root of the characteristic polynomial then

$$P(\theta) = \alpha\theta + \beta, \qquad P(\overline{\theta}) = \alpha\overline{\theta} + \beta, \qquad \overline{P}(\theta) = \overline{\alpha}\theta + \overline{\beta}, \qquad \overline{P}(\overline{\theta}) = \overline{\alpha}\,\overline{\theta} + \overline{\beta}. \tag{46}$$

Proof. The first formula in $P(x)$ is a consequence of Niven's division. The second for $\overline{P}(x)$ is a consequence of conjugation of the first and the fact that the conjugate of $Q(x)P_q(x)$ is $P_q(x)\overline{Q}(x) = \overline{Q}(x)P_q(x)$. □

Proposition 6.2. *Given arbitrary quaternions $\alpha = \alpha_1 + \mathbf{i}\alpha_2 + \mathbf{j}\alpha_3 + \mathbf{k}\alpha_4$, $\beta = \beta_1 + \mathbf{i}\beta_2 + \mathbf{j}\beta_3 + \mathbf{k}\beta_4$, we have*

$$\overline{\alpha}\beta + \overline{\beta}\alpha = \alpha\overline{\beta} + \beta\overline{\alpha} = 2\sum_{i=1}^{4} \alpha_i\beta_i. \tag{47}$$

Proof. The correctness of the first equation follows from conjugation. The second equality is easy to verify. □

Theorem 6.1. *Let $P(x)$ be a quaternion polynomial. For any q we have,*

$$|P(q)|^2 + |P(\overline{q})|^2 = |\overline{P}(q)|^2 + |\overline{P}(\overline{q})|^2. \tag{48}$$

Proof. If q is real, this is trivial. Thus we assume q is not real. Let α, β be as in Proposition 6.1. It is easy to verify that the left-hand-side of the equation in the proposition coincides with

$$2|q|^2\alpha^2 + 2\,\mathrm{Re}(q)(\alpha\overline{\beta} + \beta\overline{\alpha}) + 2|\beta|^2, \tag{49}$$

while the right-hand-side coincides with

$$2|q|^2\alpha^2 + 2\,\mathrm{Re}(q)(\overline{\alpha}\beta + \overline{\beta}\alpha) + 2|\beta|^2. \tag{50}$$

Thus by Proposition 6.2 they are equal. □

Theorem 6.2. *Let $\epsilon \in (0, 1)$. Suppose $\theta \in \mathbb{C}$ satisfies $|F(\theta)| \le \epsilon^2$. Let*

$$q = \overline{P}(\theta)\theta\overline{P}(\theta)^{-1}, \qquad q' = \overline{P}(\overline{\theta})\overline{\theta}\,\overline{P}(\overline{\theta})^{-1}. \tag{51}$$

(i) *If $|\overline{P}(\theta)| > \epsilon$, then $|P(q)| < \epsilon$.*
(ii) *If $|\overline{P}(\overline{\theta})| > \epsilon$, then $|P(q')| < \epsilon$.*
(iii) *If $|\overline{P}(\theta)| < \epsilon$, and $|\overline{P}(\overline{\theta})| < \epsilon$, then $|P(\theta)| < \sqrt{2}\epsilon$ and $|P(\overline{\theta})| < \sqrt{2}\epsilon$.*

Proof. Proof of (i) follows trivially from the fact that $F(\theta) = P(q)\overline{P}(\theta)$, hence $|F(\theta)| = |P(q)|\,|\overline{P}(\theta)|$.

To prove (ii), we observe that since $F(x)$ has real coefficients, $|F(\theta)| = |F(\overline{\theta})|$. Thus applying the Factor Theorem (Theorem 4.1) we have $F(\overline{\theta}) = P(q')\overline{P}(\overline{\theta})$.

To prove (iii), we use Theorem 6.1 to conclude

$$|P(\theta)|^2 + |P(\overline{\theta})|^2 = |\overline{P}(\theta)|^2 + |\overline{P}(\overline{\theta})|^2. \tag{52}$$

It thus follows that

$$|P(\theta)|^2 \le 2\epsilon^2, \qquad |P(\overline{\theta})|^2 \le 2\epsilon^2. \quad \square \tag{53}$$

Remark 6.1. From Theorem 6.2, in order to compute a quaternion q so that $|P(q)| \le \epsilon$, it requires computing a complex number c so that $|F(c)| \le \epsilon^2$. Suppose that we have computed such q and that ξ is the closets root of $P(x)$ to ξ. Then from (29) in the Factor Theorem (Theorem 4.1) if $a = (q - \xi) \neq 0$, $P(q) = Q(aqa^{-1})(q - \xi)$. Thus, $|q - \xi| \le \epsilon/|Q(aqa^{-1})|$. If ξ is an isolated root of $P(x)$ and the roots of $P(x)$ are not too near, then we would expect $|Q(aqa^{-1})|$ not to be too small. In practice, once we are reasonably close to a root of a complex polynomial, a few iterations of Newton's method will suffice to bring it to a highly accurate approximation. The connection between approximation of $F(x)$ and $P(x)$ suggests the application of known polynomial root-finding complexities and algorithms, e.g. see Pan [40], Neff and Reif [35], Jenkins and Traub [18], Pan [40], McNamee [33] and Kalantari [27]. We will however also consider direct methods for the approximation of zeros of $P(x)$.

7. Bound on zeros of quaternion polynomials

Computing a priori bound on the zeros of complex polynomials is an important problem with many theoretical and practical applications, for example see McNamee [33]. Kalantari [25] describes an infinite family of bounds on zeros of a complex polynomial based on the use of Basic Family of iteration functions. These are efficiently computable, see Jin [20], and have found to be practical, see McNamee [33] for a computational result. The following result implies the application of these bounds to a quaternion polynomial.

Theorem 7.1. *Let $P(x)$ be a quaternion polynomial, $F(x) = P(x)\overline{P}(x)$. For each root of $P(x)$ there is a root of $F(x)$ having the same norm. Conversely, for each root of $F(x)$ there is a root of $P(x)$ having the same norm.*

Proof. We only need to apply Theorem 5.1 and the fact that conjugate quaternions have the same norm. □

We remark here that direct bounds on zeros of $P(x)$ are also possible. For instance, Opfer [39] obtains one such bound. But the effectiveness of these bounds is not investigated. Indeed, by Theorem 7.1, in order to compute effective bounds on the norm of zeros of $P(x)$ we must also determine such bounds on the zeros of $F(x)$. To develop effective bounds, we only need $O(n^2)$ preprocessing time to compute the coefficients of $F(x)$.

8. Decomposition of a quaternion polynomial

We have already stated a characterization for zeros of a quaternion polynomial and bound on the number of its zeros (Theorem 4.2). In this section we give a more precise characterization than that of Theorem 4.2, in terms of a decomposition theorem we prove. First, we give a definition.

Definition 8.1. We say a monic quaternion polynomial $P(x)$ of degree $n_1 \geq 0$ is a type I polynomial if either $P(x) = 1$ or

$$P(x) = (x - r_1)^{m_1} \cdots (x - r_s)^{m_s}, \tag{54}$$

where $m_i \geq 1$, $r_i \in \mathbb{H}$, satisfying $[r_i] \neq [r_j]$, $i \neq j$, and $\sum_{i=1}^{s} m_i = n_1$.

We say a monic quaternion polynomial $P(x)$ of degree $n_2 \geq 0$ is a type II polynomial if either $P(x) = 1$ or there exist real numbers $t_i \geq 0$ and ν_i, $i = 1, \ldots, l$, satisfying $t_i^2 - 4\nu_i^2 < 0$ such that

$$P(x) = (x^2 - t_1 x + \nu_1^2)^{\mu_1} \cdots (x^2 - t_l x + \nu_l^2)^{\mu_l}, \tag{55}$$

where $\mu_i \geq 1$, and $\sum_{i=1}^{l} \mu_i = n_2$.

The following is a characterization of quaternion polynomial factorization.

Theorem 8.1. *Let $P(x)$ be a monic quaternion polynomial of degree $n \geq 1$. Then*

$$P(x) = P_0(x)P_1(x)P_2(x), \tag{56}$$

where $P_1(x)$ is a type I polynomial of degree $n_1 \geq 0$, $P_2(x)$ is a type II polynomial of degree $n_2 \geq 0$, $n_1 + n_2 \geq 1$, and $P_0(x)$ is a polynomial of degree $n_0 \geq 0$ where either $P_0(x) = 1$ or $n_0 \geq 1$, in which case $P_0(x)$ has no root in common with $P(x)$ but is such that for each root q of $P_0(x)$ there is a conjugate $q' \neq q$, where q' is a root of $P_1(x)$ or a root of $P_2(x)$. Furthermore, $n_0 + n_1 + 2n_2 = n$, and $P_1(x)$ and $P_2(x)$ commute, i.e. for all quaternions q

$$P_1(q)P_2(q) = P_2(q)P_1(q). \tag{57}$$

In particular, $P(q) = 0$ if and only if either $P_1(q) = 0$, or $P_2(q) = 0$.

Proof. The proof essentially follows from Proposition 4.1, the Factor Theorem (Theorem 4.1), and the FTA for quaternion polynomials. Pick a root q of $P(x)$. If it has a conjugate that is also a root, then $P_q(x)$ can be factored out from $P(x)$. We consider the quotient and repeat this process until no such spherical root is found. This results in $P_2(x)$, with the possibility that $P_2(x) = 1$. Thus, $P(x) = Q(x)P_2(x)$ for some polynomial $Q(x)$. The polynomial $Q(x)$ and $P_2(x)$ commute. This follows from the fact that the coefficients of $P_2(x)$ are real, thus for any quaternion q, we have $P_2(q)qP_2(q)^{-1} = q$. Thus by the Factor Theorem (Theorem 4.1), $Q(q)P_2(q) = P_2(q)Q(q)$. Now $Q(x)$ can have no spherical roots. Thus factoring its roots we get $Q(x) = P_0(x)P_1(x)$, where $P_1(x)$ and $P_0(x)$ are as described. □

Definition 8.2. We say a quaternion polynomial $P(x) = P_0(x)P_1(x)P_2(x)$ of degree $n \geq 1$ is root-deficient if $P_0(x)$ is nontrivial.

We thus conclude the following theorem on the structure of the zeros of a quaternion polynomial.

Theorem 8.2. *Let $P(x)$ be a quaternion polynomial of degree $n \geq 1$. Then the roots of $P(x)$ consist of a set of $n_1 \geq 0$ points in $\mathbb{H}$ (counted with multiplicity), and $n_2 \geq 0$ spheres of zeros centered at points of*

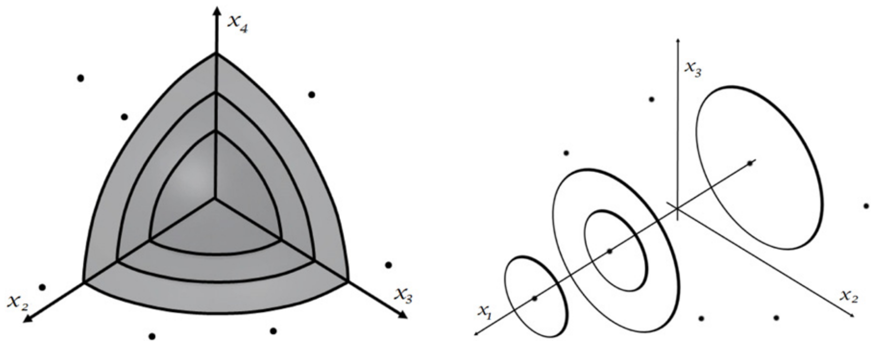

Fig. 1. Three-dimensional cross-sectional view of zeros of a quaternion polynomial.

the form $(a, 0, 0, 0)$ *(counted with multiplicity), satisfying*

$$1 \leq n_1 + 2n_2 \leq n. \tag{58}$$

If $P(x)$ *is not root-deficient, then* $n = n_1 + 2n_2$. *If* $P(x)$ *has only simple roots, then the roots consist of* $n_1 \geq 0$ *distinct points conjugate to no other roots, and* $n_2 \geq 0$ *distinct spheres of zeros.* □

Remark 8.1. To visualize the roots of a quaternion polynomial $P(x)$ we should keep in mind that each root is either isolated or spherical but it could have multiplicity higher than one. Isolated roots of $P(x)$ can be points located anywhere in $\mathbb{R}^4$, but the spherical roots are all centered on the x_1-axis in $\mathbb{R}^4$. These could have different x_1-coordinates, or some could have the same x_1-coordinate, in which case they will be concentric spheres.

Three dimensional view of these zeros can be done in the following way. Suppose for a particular value $x_1 = a$ there are spherical zeros centered at $(a, 0, 0, 0)$ and isolated zeros of the type (a, b, c, d) (i.e. with the same x_1 coordinate value), then viewing the three dimensional space of x_2, x_3, x_4, we would see a set of zeros as concentric spheres centered at the origin, and a set of isolated points at arbitrary locations (see Fig. 1, left image).

Suppose for a particular value in a coordinate other that x_1, say $x_4 = d$, we view the three dimensional space in the variables x_1, x_2, x_3, then we could see a set of zeros as single or concentric circles centered at different x_1 values, i.e. circles of the form $\{(x_1, x_2, x_3) \mid x_1 = a,\ x_2^2 + x_3^2 = r\}$, as well as isolated zeros at arbitrary locations (see Fig. 1, right image). In particular, there is a finite description of the zeros of a quaternion polynomial.

9. Taylor's theorem and iterative methods over quaternions

Our aim in this section is to the develop Newton's method $B_2(x)$, and Halley's methods $B_3(x)$ (see (3)) for quaternion polynomial root-finding and analyze their local behavior. To this end, we first need to define formal derivatives of a polynomial. We will consider a given quaternion polynomial of degree n, $P(x) = \sum_{i=0}^{n} a_i x^i$.

Definition 9.1. For any $k = 0, \ldots, n$, the k-th formal derivative of $P(x)$ is

$$P^{(k)}(x) \equiv \sum_{j=k}^{n} \frac{j!}{(k-j)!} a_j x^{j-k}. \tag{59}$$

Next, we state a pointwise version of Taylor's theorem for quaternions.

Theorem 9.1. *Let ξ be a root of $P(x)$ and q any quaternion. Define*

$$E(\xi, q) \equiv \sum_{k=0}^{n} \frac{P^{(k)}(q)}{k!}(\xi - q)^k. \tag{60}$$

If q commutes with ξ, $E(\xi, q) = 0$. In particular, $E(\xi, q)$ is independent of the coefficients of $P(x)$.

Proof. For complex polynomials an algebraic-combinatorial proof of the theorem which makes use of the binomial theorem is given in [27]. The same proof could apply to the case of quaternions. □

Remark 9.1. The quantity $E(\xi, q)$ may turn out to be nonzero. Nevertheless we can argue that, since $E(\xi, \xi) = 0$, given that $|q - \xi|$ is small, so is $|E(\xi, q)|$. As an example consider $P(x) = x^2 + bx + c$. Then $P'(x) = 2x + b$, $P''(x) = 2$. Thus if ξ is a root of $P(x)$, for an arbitrary q (after simplification) we have,

$$E(\xi, q) = P(q) + P'(q)(\xi - q) + \frac{P''(q)}{2}(\xi - q)^2 = q\xi - \xi q. \tag{61}$$

Thus if $q = \xi + \epsilon$, then $|E(\xi, q) \leq 2|\epsilon|\,|\xi|$. As stated in Theorem 9.1, $E(\xi, q)$ is independent of the coefficient b and c. This fact is useful in defining Newton and Halley methods and in the analysis of their local behavior.

9.1. Newton's method for quaternion polynomials

Using Theorem 9.1, we derive Newton's iteration and a corresponding expansion formula. More specifically we have,

Theorem 9.2. *Suppose ξ is a simple root of $P(x)$. Given q with $P'(q) \neq 0$,*

$$B_2(q) \equiv q - P'(q)^{-1}P(q) = \xi + \sum_{k=2}^{n} P'(q)^{-1}\frac{P^{(k)}(q)}{k!}(\xi - q)^k - P'(q)^{-1}E(\xi, q). \quad \square \tag{62}$$

Definition 9.2. Given a seed q_0, let the Newton's fixed point iteration be

$$q_k = B_2(q_{k-1}) = q_{k-1} - P'(q_{k-1})^{-1}P(q_{k-1}), \quad k \geq 1. \tag{63}$$

We remark that the iteration is well-defined almost everywhere on $\mathbb{H}$ since the set of zeros of $P'(x)$ is a set of measure zero in $\mathbb{H}$. For a complex polynomial, $B_2(q)$ is Newton's iterate. Newton's method for complex polynomials is the first member of the Basic Family, see [27].

To understand the behavior of Newton's method for quaternion polynomials, first we consider an example. Suppose that $P(x) = x^2 - (\mathbf{i} + \mathbf{j})x + \mathbf{k}$. It was shown earlier that $\mathbf{j}$ is the only solution of $P(x) = 0$. If we set $q_0 = \mathbf{j} + \epsilon$, then it can be seen that for the next iterate we have, $|q_1 - \mathbf{j}| \approx \epsilon^2 = |q_0 - \mathbf{j}|^2$.

Now suppose that $q_0 = \mathbf{j} + \epsilon_1 + \epsilon_2\mathbf{i} + \epsilon_3\mathbf{k}$. Then $|q_0 - \mathbf{j}| = \sqrt{\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2}$. We will approximate $|q_1 - \mathbf{j}|$. From the expansion formula (62) we have, $|q_1 - \mathbf{j}| \approx |P'(q_0)^{-1}(q_0\mathbf{j} - \mathbf{j}q_0)|$. We have, $q_0\mathbf{j} - \mathbf{j}q_0 = 2\epsilon_2\mathbf{k} - 2\epsilon_3\mathbf{i}$, and $P'(q_0) = 2\mathbf{j} + 2\epsilon_1 + 2\epsilon_2\mathbf{i} + 2\epsilon_3\mathbf{k} - \mathbf{i} - \mathbf{j} = \mathbf{j} + 2\epsilon_1 + (2\epsilon_2 - 1)\mathbf{i} + 2\epsilon_3\mathbf{k}$. Thus,

$$|P'(q_0)^{-1}(q_0\mathbf{j} - \mathbf{j}q_0)| = 2\sqrt{\epsilon_2^2 + \epsilon_3^2}/\sqrt{1 + \epsilon_1^2 + (1 - 2\epsilon_2)^2 + \epsilon_3^2}. \tag{64}$$

We can select $\epsilon_1, \epsilon_2, \epsilon_3$ so that $|q_1 - \mathbf{j}| < |q_0 - \mathbf{j}|$. However, we can also select these so that $|q_1 - \mathbf{j}| > |q_0 - \mathbf{j}|$. Hence in every neighborhood of the root $\mathbf{j}$ there is a repulsive direction, i.e. a direction where the Newton iterate gets farther away from the root than the current iterate. This is contrary to the case of a complex polynomial where each root is at least an attractive fixed point of

Newton's iteration function. This implies the behavior of Newton or other iterations is even more chaotic than the complex case. We now give a definition to formalize these observations.

Definition 9.3. Suppose that ξ is a root of the quaternion polynomial $P(x)$. We say a quaternion d with $|d| = 1$ is an attractive (repulsive) Newton direction at ξ if there exists $\alpha > 0$ such that

$$|B_2(\xi + \alpha d) - \xi| < |(\xi + \alpha d) - \xi| = \alpha \quad (|B_2(\xi + \alpha d) - \xi| > \alpha). \tag{65}$$

From the Taylor expansion (62) we may easily conclude:

Proposition 9.1. *Suppose that ξ is a simple root of the polynomial $P(x)$ and d a quaternion with $|d| = 1$. Let*

$$L = \lim_{\epsilon \to 0} \frac{|E(\xi, \xi + \epsilon d)|}{\epsilon |P'(\xi)|} = \lim_{\epsilon \to 0} \frac{|B_2(\xi + \epsilon d)|}{\epsilon}. \tag{66}$$

If $L < 1$, then d is an attractive Newton direction at ξ and if $L > 1$, d is a repulsive Newton direction at ξ. When $L < 1$ we call L the contraction factor along d at ξ. □

Remark 9.1. The quantity L is the magnitude of the directional derivative of $B_2(x)$ along the direction d. It is easy to argue that the set of all attractive directions at ξ is an open subset of the unit ball in the quaternion space, centered at ξ. We also note that this set is nonempty since we can choose q_0 to belong to $C(\xi) = \{q \in \mathbb{H} \mid \xi q = q\xi\}$. Several results can now be stated, where we omit the proofs since they follow in a straightforward fashion.

Theorem 9.3. *Suppose ξ is a root of a quaternion polynomial $P(x)$. Suppose for any $q' \in C(\xi) = \{q \in \mathbb{H} \mid \xi q = q\xi\}$, both $P(q')$ and $P'(q')$ belong to $C(\xi)$. There exists a neighborhood of ξ in $C(\xi)$ such that for any q_0 in this neighborhood all Newton's iterations are well-defined and remain in $C(\xi)$. Furthermore, if ξ is a simple root, the fixed point iteration converges to ξ having a quadratic rate. In particular, if $P(x)$ has real coefficients there is a neighborhood of ξ in $C(\xi)$ such that for any q_0 in this neighborhood, the fixed point iterations converge to ξ.* □

9.2. Halley's method for quaternion polynomials

We consider developing Halley's method for quaternion polynomials (see (3)). This requires a more tedious derivation than Newton's. We will omit this derivation for brevity and state the result.

Theorem 9.4. *For a q assume $P(q) \neq 0$, $P'(q) \neq 0$, $P'(q) \neq 1$. Then,*

$$\begin{aligned} B_3(q) &\equiv q - {P'}^{-1}P - {P'}^{-1}\frac{P''}{2}\Delta^{-1}P{P'}^{-1}P \\ &= \sum_{k=3}^{n+1} {P'}^{-1}\left[\frac{P^{(k)}}{k!} - \frac{P''}{2}\Delta^{-1}\left(\frac{P^{(k-1)}}{(k-1)!} - P{P'}^{-1}\frac{P^{(k)}}{k!}\right)\right](\xi - q)^k + E_3, \end{aligned} \tag{67}$$

where,

$$\Delta = \left(P' - P{P'}^{-1}\frac{P''}{2}\right) \tag{68}$$

$$E_3 = {P'}^{-1}\frac{P''}{2}\left(\Delta^{-1}P{P'}^{-1}E(\xi, q) - \Delta^{-1}E(\xi, q)(\xi - q)\right) - {P'}^{-1}E(\xi, q). \quad \square \tag{69}$$

Definition 9.4. Given a seed q_0, let the *Halley's fixed point iteration* be

$$q_k = B_3(q_{k-1}), \quad k \geq 1. \tag{70}$$

With respect to $B_3(x)$, analogous to Definition 9.1, attractive and repulsive directions can be defined. Furthermore, an analogous Proposition 9.1 can be stated with respect to the directional derivatives of $B_3(x)$ at ξ. Also, analogous Theorem 9.3 can be stated with respect to $B_3(x)$, where under suitable assumptions quadratic convergence can be replaced with cubic convergence. We avoid stating a separate theorem for these. One can also derive $B_m(x)$, with $m \geq 4$ but with successively more complicated derivation.

9.3. Constructing a quaternion polynomial with prescribed zeros

Here we will describe a procedure to construct quaternion polynomials with prescribed zeros. This will be useful both in terms of test problems for the methods described here, as well as for their visualization via polynomiography which will be carried out in subsequent work. Unlike the case of complex polynomials where it is straightforward to construct a polynomial with prescribed zeros, the quaternion case takes special care.

Proposition 9.2. *Let $r_1, \ldots, r_k$ be a set of quaternions satisfying $[r_i] \neq [r_j]$ if $i \neq j$. Let $w_1 = r_1$, and set $P_1(x) = (x - w_1)$. Suppose we have computed $w_2, \ldots, w_{k-1}$ such that the polynomial $P_{k-1}(x) = (x - w_{k-1}) \cdots (x - w_1)$ has roots $r_1, \ldots, r_{k-1}$. Set $w_k = P_{k-1}(r_k) r_k P_{k-1}(r_k)^{-1} = (r_k - w_{k-1}) \cdots (r_k - w_1) r_k (r_k - w_1)^{-1} \cdots (r_k - w_{k-1})^{-1}$. Then the polynomial $P_k(x) = (x - w_k) P_{k-1}(x)$ has roots $r_1, \ldots, r_{k-1}, r_k$.*

Proof. Since r_i's are from distinct conjugacy classes, $P_{k-1}(r_k) \neq 0$. Then by the Factor Theorem (Theorem 4.1) w_k is a root of $P_k(x)$. □

Even in case the conjugacy classes are not disjoint a polynomial with desired roots can be constructed. For instance, if $[r_i] = [r_j]$ for some $i \neq j$, the characteristic polynomial of r_i (equivalently r_j) is necessarily a factor of the polynomial. Stronger versions of interpolation are given in [3].

10. Quaternion homogeneous linear recurrence relation

Consider a homogeneous linear recurrence relation (HLRR) of degree n:

$$a_m = c_1 a_{m-1} + c_2 a_{m-2} + \cdots + c_n a_{m-n}, \tag{71}$$

where $c_1, \ldots, c_n$ are given constants in $\mathbb{H}$, and $c_n \neq 0$.

In this section we extend known results on HLRR over the complex numbers to the quaternions. For results on HLLR over the real or complex numbers, see Henrici [12], Hildebrand [13], Householder [14], and Kalantari [23,27].

We shall refer to the following as the *Basic Initial Conditions*:

$$a_0 = 1, \qquad a_{-1} = a_{-2} = \cdots = a_{-n+1} = 0. \tag{72}$$

We shall refer to the solution $\{a_m\}_{m=1}^{\infty}$ of (71) and (72) as the *Fundamental Solution* of HLRR. The significance of this particular set of initial conditions and the Fundamental Solution will become evident later.

The *characteristic polynomial* and *characteristic equation* of HLRR are

$$Q(x) = x^n - (c_1 x^{n-1} + c_2 x^{n-2} + \cdots + c_{n-1} x + c_n), \qquad Q(x) = 0. \tag{73}$$

For example, consider the HLRR $a_m = a_{m-1} + a_{m-2}$. Its Fundamental Solution is the Fibonacci numbers $F_m = F_{m-1} + F_{m-2}$, $F_0 = 1$, $F_{-1} = 0$. However, with $F_0 = 1$, $F_{-1} = 2$ we get the Lucas numbers.

Definition 10.1. Let the negative reciprocal of $Q(x)$ to be

$$P(x) = Q(x^{-1}) x^n = c_n x^n + \cdots + c_1 x - 1. \tag{74}$$

Proposition 10.1. $Q(\eta) = 0$ *if and only if* $P(\eta^{-1}) = 0$. □

We will show that using HLRR we can compute the dominant root of a quaternion polynomial, assuming some regularity condition on the roots. In a sense this gives a generalization of the Bernoulli method for computing the extreme roots of a complex polynomial.

10.1. Characterization of solution of HLRR

In this section, we will characterize the solutions of HLRR under the assumption of simplicity of roots of the characteristic equation and further assumption that it is not root-deficient (see Definition 8.2). This is accomplished in the next theorem, a well-known result for the case of HLRR over the real or complex numbers. According to the theorem, the solution of HLRR can be represented in terms of the roots of the characteristic polynomial. However, noncommutativity in quaternion multiplication requires a more careful analysis than that of the case of complex numbers.

Theorem 10.1. *Consider HLRR (71) with (72) and its characteristic polynomial (73). Suppose that $Q(x)$ is not root-deficient and its roots are simple. Let $\eta_1, \ldots, \eta_n$ be the set of distinct roots of $Q(x)$, pairwise distinct and such that no three of them belong to the same conjugacy class. Then there exists a unique set of nonzero constants $\alpha_i \in \mathbb{H}$, $i = 1, \ldots, n$ such that for all $m \geq -n+1$*

$$a_m = \eta_1^m \alpha_1 + \cdots + \eta_n^m \alpha_n. \tag{75}$$

Proof. First, we justify that for any set of constants α_i, $a_m = \sum_{i=1}^n \eta_i^m \alpha_i$ satisfies (71). We have

$$a_m - \sum_{j=1}^n c_j a_{m-j} = \sum_{i=1}^n \eta_i^m \alpha_i - \sum_{j=1}^n c_j \sum_{i=1}^n \eta_i^{m-j} \alpha_i = \sum_{i=1}^n \eta_i^{m-n} Q(\eta_i)\alpha_i = 0. \tag{76}$$

Next, we need to show that there exists specific set of α_i's so that a_m satisfies the given Basic Initial Conditions. Substituting the Basic Initial Conditions (72) into (75), we get the following system of linear equations:

$$\begin{pmatrix} \eta_1^{-n+1} & \eta_2^{-n+1} & \cdots & \eta_n^{-n+1} \\ \eta_1^{-n+2} & \eta_2^{-n+2} & \cdots & \eta_n^{-n+2} \\ \vdots & \vdots & \cdots & \vdots \\ \eta_1^{-1} & \eta_2^{-1} & \cdots & \eta_n^{-1} \\ 1 & 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{n-1} \\ \alpha_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{77}$$

The coefficient matrix in (77) is a Vandermonde matrix we denote by $V(\eta_1^{-1}, \ldots, \eta_n^{-1})$. To prove the existence of a unique solution, we invoke a result due to Renmin et al. [43]: necessary and sufficient condition that a linear system with a Vandermonde matrix $V(\beta_1, \ldots, \beta_m)$ of coefficients have a unique solution is that the double determinant (see Zhang [48] for the definition of double determinant) of the Vandermonde matrix is nonzero. This double determinant is nonzero if and only if β_i's are pairwise distinct and no three of them are from the same conjugacy set.

To define double determinant, following Zhang [48], for a given $n \times n$ matrix $A = (a_{ij})$ with quaternion entries, its *conjugate* is defined as $\overline{A} = (\overline{a_{ij}})$, its *transpose* is $A^T = (a_{ji})$, and its *conjugate transpose* is $A^* = (\overline{A})^T$.

The double determinant of A, denoted by $|A|_d$ is $\det(A^*A)$. For results on double determinant and its equivalence to the notion of q-determinant of A, the determinant of the $2n \times 2n$ complex adjoint matrix of A, see Zhang [48, Theorem 8.2].

In our case, the set $\{\beta_i = \eta_i^{-1}, i = 1, \ldots, n\}$ precisely satisfies this above mentioned necessary and sufficient condition. Hence the proof of existence of unique solutions in α_i.

Now suppose that some α_i is zero. Without loss of generality we may assume $\alpha_1 = 0$. Substituting $\alpha_1 = 0$ in (77), it is easy to see that the first $n-1$ equations gives a linear system whose solution is $\alpha_1 = \cdots = \alpha_n = 0$, a contradiction. Hence $\alpha_i \neq 0$ for all $i = 1, \ldots, n$. □

Proposition 10.2. *Consider the HLRR under the assumption of Theorem 10.1. Assume that $Q(x)$ has a unique dominant root η_r and let $\theta = \eta_r^{-1}$. Then*

$$\lim_{m\to\infty} a_{m-1}^{-1} a_m = \alpha_r^{-1}\eta_r\alpha_r, \qquad \lim_{m\to\infty} a_m^{-1} a_{m-1} = \alpha_r^{-1}\theta_r\alpha_r. \tag{78}$$

Proof. We have

$$a_m = \sum_{i=1}^{n} \eta_i^m \alpha_i = \eta_r^m \left(\sum_{i=1}^{n} \eta_r^{-m} \eta_i^m \alpha_i \right) = \eta_r^m \left(\alpha_r + \sum_{i=1, i \neq r}^{n} \eta_r^{-m} \eta_i^m \alpha_i \right). \tag{79}$$

Let $u_m = \alpha_r + \sum_{i=1,i\neq r}^{n} \eta_r^{-m} \eta_i^m \alpha_i$. Since $\alpha_r \neq 0$ and $|\eta_r| > |\eta_i|$ for $i \neq r$, $\lim_{m\to\infty} u_m = \alpha_r$. From the definition of u_m, and since inverse of products is the product of inverses in reverse, we have $a_{m-1}^{-1} = u_{m-1}^{-1} \eta_r^{-m+1}$. Thus

$$a_{m-1}^{-1} a_m = u_{m-1}^{-1} \eta_r u_m. \tag{80}$$

From (80), and since u_m converges to α_r the first limit in (78) and the second limit follows by inversion. □

The following result shows yet another application of Niven's division.

Proposition 10.3. *Given* $\lim_{m\to\infty} a_{m-1}^{-1} a_m = q$, *we can determine* η_r *so that* $q = \alpha_r^{-1} \eta_r \alpha_r$.

Proof. If η_r is real, the limit already converges to η_r. If η_r is not real, since it is the unique dominant root of $Q(x)$, it cannot be a spherical root. Thus we can perform Niven's division to get

$$Q(x) = R(x) P_{\eta_r}(x) + fx + g. \tag{81}$$

Then η_r must satisfy $\eta_r = -f^{-1} g$. □

10.2. Solution of general HLRR

In this section we will consider HLRR where the characteristic polynomial $Q(x)$ may have multiple roots, or be root-deficient. However, for simplicity we will assume the roots of $Q(x)$ to have different norm.

Theorem 10.2. *Consider HLRR (71) with (72) and its characteristic polynomial (73). Suppose that the roots of* $Q(x)$ *have distinct norm, but not necessarily simple. Let* $R = \{\eta_1, \ldots, \eta_t\}$ *be the set of roots of* $Q(x)$ *with multiplicities* n_i, $i = 1, \ldots, t$. *Then for all* $m \geq 0$, *we have*

$$a_m = \sum_{i=1}^{t} \eta_i^m \alpha_i(m) \tag{82}$$

where $\alpha_i(x)$ *is either identically zero or a polynomial of degree at most* $n_i - 1$,

$$\alpha_i(x) = \alpha_{i,0} + \alpha_{i,1} x + \cdots + \alpha_{i,n_i-1} x^{n_i-1}. \tag{83}$$

When $\alpha_i(x)$ *is not identically zero, we denote its leading coefficient with* α_i^*.

10.3. Proof of Theorem 10.2

As in the case of simple roots, we first need to justify that any set of coefficients defining a_m, as defined above, satisfies (71). This can be done as in the special case, so we skip the proof. Next, we need to show that there exists specific set of α_i's so that a_m satisfies the given Basic Initial Conditions (72). To do so, we first need some auxiliary results.

Lemma 10.1. *Suppose* $\eta_1, \ldots, \eta_t$ *are nonzero quaternions satisfying* $|\eta_i| \neq |\eta_j|$. *Let* $\gamma_{ij} = |\eta_i|/|\eta_j|$. *Let* N *be a natural number. There exists a natural number* k_0 *such that if* $m_0 = 2^{k_0}$, *for all* $1 \leq i, j \leq t$ *and for any two nonnegative integers* $n_1, n_2 \leq N$ *we have*

$$m_0^{n_1} |\eta_i|^{m_0} \neq m_0^{n_2} |\eta_j|^{m_0}. \tag{84}$$

Proof. Let $\gamma_M = \min\{\gamma_{ij} \mid \gamma_{ij} > 1\}$. Let k_0 be so that $m_0 = 2^{k_0}$ satisfies $N(\ln m_0/m_0) < \ln \gamma_M$. We claim k_0 satisfies (84). If $n_1 = n_2$, clearly (84) holds. Assume that $n_1 \neq n_2$. If (84) is not satisfied, there exist i, j, and $0 \leq n_1, n_2 \leq N$, say $n_1 > n_2$, thus $|\eta_i| < |\eta_j|$, such that $m_0^{(n_1-n_2)/m_0} = |\eta_j|/|\eta_i|$. Taking the logarithm of both sides in this, the corresponding left-hand-side is less than $\ln \gamma_M$, while the corresponding right-hand-side is greater than or equal to $\ln \gamma_M$, a contradiction. □

Corollary 10.1. *Let $\eta_1, \ldots, \eta_t$ be a set of nonzero distinct quaternions satisfying $|\eta_i| \neq |\eta_j|$. Given a set of natural numbers $N_1, \ldots, N_t$, let $N = N_1 + \cdots + N_t$. Let k_0 be as in the previous lemma. Let $\delta = 2^{\frac{k_0}{2^{k_0}}}$. Then the elements of the following set have distinct norm:*

$$R = \{\eta_i, \delta\eta_i, \delta^2\eta_i, \ldots, \delta^{N_i}\eta_i\}_{i=1}^{t}. \quad \square \tag{85}$$

We now complete the proof of Theorem 10.2. Analogous to the case of HLRR over the real or complex number numbers, in order to prove the particular representation of the a_m, it suffices to prove the validity over a subset of the indices, which would allow the computation of polynomial coefficients. From Corollary 10.1 and our choice of k_0, it follows that for an appropriate set of indices m_i we can obtain a system of linear equations that can be written as $V\alpha = a$, where V is the Vandermonde matrix corresponding to the set R in (85), and where a is the vector corresponding to $a(m_i)$, and α is the vector of unknown polynomial coefficients. By our construction, this system has a unique solution. □

Theorem 10.3. *Consider the homogeneous linear recurrence relation (71) with the Basic Initial Conditions (72). Assume that the roots of $Q(x)$ satisfy the condition of Theorem 10.2. Let r be the index satisfying $|\eta_r| = \max\{|\eta_i| \,|\alpha_i^* \neq 0,\ i = 1, \ldots, t\}$,*

(i) *r is well-defined.*
(ii) *If all the roots of $Q(x)$ are simple, then α_i^* is nonzero for all i so that η_r is the dominant root of $Q(x)$.*
(iii) *a_m can equal zero for at most a finite number of indices.*
(iv) *$\lim_{m\to\infty} a_{m-1}^{-1} a_m = (\alpha_r^*)^{-1}\eta_r\alpha_r^*$.*

Proof. We may write $a_m = \eta_r^m u_m$, where

$$u_m = \alpha_r(m) + v_m, \qquad v_m = \sum_{i=1, i\neq r}^{n} \eta_r^{-m}\eta_i^m\alpha_i(m). \tag{86}$$

Since $a_0 = 1$, we conclude that not all α_i^*s are zero. Hence r is well-defined, proving (i). The proof of (ii) is already given in Theorem 10.1.

From the assumption that $|\eta_i| \neq |\eta_j|$ and the definition of η_r, for any $i \neq r$ such that $\alpha_i^* \neq 0$, we have $|\eta_i|/|\eta_r| < 1$, and since for any such index $\alpha_i(m)$ is a polynomial in m, it follows that

$$\lim_{m\to\infty} \eta_r^{-m}\eta_i^m\alpha_i(m) = 0. \tag{87}$$

Suppose $\{a_m\}$ has an infinite sequence of zeros. If all but one of the $\alpha_i(x)$s are identically zero, then from (71) we have, $a_m = \eta_l^m\alpha_l(m)$, for some l and $\alpha_l(m) \not\equiv 0$. Thus, $a_m \neq 0$ for all $m \geq 0$. Hence there must exist at least two different indices for which $\alpha_i(x)$ is not identically zero. For any pair of quaternions $a, b \in \mathbb{H}$, we have $|a \pm b| \geq |a| - |b|$. We may write

$$|\eta_r^{-m_k} a_{m_k}| = |u_{m_k}| \geq |\alpha_r(m_k)| - |v_{m_k}|. \tag{88}$$

Taking the limit in (88) and using (87), that v_{m_k} converges to zero, we get

$$0 \geq \lim_{k\to\infty} |\alpha_r(m_k)|. \tag{89}$$

But the limit in (89) is either a nonzero number, or infinity. This is a contradiction. Thus a_m can be zero at most for a finite set of indices m.

Since $\alpha_i(x)$ is a polynomial, if it is not identically zero then we have

$$\lim_{m\to\infty} \alpha_i^{-1}(m-1)\alpha_i(m) = 1. \tag{90}$$

Thus $a_{m-1}^{-1}a_m = u_{m-1}^{-1}\eta_r u_m$, proving (iii). □

Remark 10.1. The proof of Theorem 10.3 is somewhat analogous to our analysis for complex polynomials in [27]. While the theorem proves that the quotient of the terms converges to a quaternion conjugate of a root of the characteristic polynomial, rather than a quaternion congruent to the dominant root itself, in fact for complex polynomials it converges to the dominant root. Finally, we note that once we have a congruent of a root, we can apply Niven's division to find the actual root.

Concluding remarks. The extension of polynomial root-finding from the complex field to the quaternions is far from straightforward. In the course of our study of quaternion root-finding, not only have we become aware of the wealth of known results, but were led to the discovery of new findings and the wealth of directions of research that lie ahead. Hopefully the findings here will make a worthy contribution to the theory and practice of quaternion polynomial root-finding. In fact, our study of quaternion root-finding has already inspired new approaches in polynomiography of a complex polynomial itself. These and several new research areas are the subject of forthcoming articles. As shown here the local behavior of Newton and Halley methods for quaternion polynomials is quite different from the case of complex polynomials. Even a quaternion quadratic polynomial could exhibit chaotic behavior. We will carry our experimentation of these iterations in forthcoming articles and complement these with their polynomiography. More generally, while the dynamics of the iterations of rational functions over the complex field is well studied (see e.g. Beardon [2] and Milnor [34]), iterations of rational functions over the quaternion are far more complicated. We shall examine their polynomiography as well.

Acknowledgments

I would like to thank Helaman Ferguson for his encouragement to investigate polynomial root-finding and polynomiography over the quaternions. I would also like to thank two anonymous Referees for their extremely valuable feedback. In particular, I am indebted to a Referee for a very detailed list of constructive comments, insights and corrections. The Referee pointed out a mistake on an earlier analysis of Newton's method and motivated the analysis of the local behavior as applied to a quadratic polynomial, carried out in Section 9. His insightful suggestions have substantially improved the presentation of some parts of the article. The Referee also suggested to include algebraic facts that make possible the implementation of quaternion root-finding algorithms on the existing systems such as Maple that already support complex arithmetic. I would also like to thank the Editor, Professor Jin-Yi Cai for his helpful comments and considerations.

References

[1] S. Adler, Quaternionic Quantum Mechanics and Quantum Fields, Oxford University Press, New York, 1995.
[2] A.F. Beardon, Iteration of Rational Functions: Complex Analytic Dynamical Systems, Springer-Verlag, New York, 1991.
[3] B. Beck, Sur les équations polynomiales dans les quaternions, Enseign. Math. 25 (1979) 193–201.
[4] S. Bedding, K. Briggs, Iteration of quaternion functions, Amer. Math. Monthly 103 (1996) 654–664.
[5] L. Brand, The roots of a quaternion, Amer. Math. Monthly 48 (1942) 519–520.
[6] U. Bray, G. Whaples, Polynomials with coefficients from a division ring, Canad. J. Math. 35 (1983) 509–515.
[7] E. Cho, De Moivre's formula for quaternions, Appl. Math. Lett. 11 (1998) 33–35.
[8] S. De Leo, G. Ducati, V. Leonardi, Zeros of unilateral quaternionic polynomials, Electron. J. Linear Algebra 15 (2006) 297–313.
[9] S. Eilenberg, I. Niven, The "Fundamental Theorem of Algebra" for quaternions, Bull. Amer. Math. Soc. 50 (1944) 246–248.
[10] B. Gordon, T.S. Motzkin, On the zeros of polynomials over division rings, Trans. Amer. Math. Soc. 116 (1965) 218–226.
[11] R. Heidrich, G. Jank, On the iteration of quaternionic Möbius transformations, Complex Var. 29 (1996) 313–318.
[12] P. Henrici, Applied and Computational Complex Analysis, Volume I, Wiley, New York, 1974.
[13] F.B. Hildebrand, Introduction to Numerical Analysis, second ed., McGraw-Hill, New York, 1974.
[14] A.S. Householder, The Numerical Treatment of a Single Nonlinear Equation, McGraw-Hill, New York, 1970.
[15] L. Huang, W. So, On left eigenvalues of quaternionic equations, Linear Algebra Appl. 323 (2001) 105–116.
[16] L. Huang, W. So, Quadratic formulas for quaternions, Appl. Math. Lett. 15 (2002) 533–540.

[17] D. Janovska, G. Opfer, Computing quaternionic roots by Newton's method, Electron. Trans. Numer. Anal. 26 (2007) 82–102.
[18] M.A. Jenkins, J.F. Traub, A three-stage variable-shift iteration for polynomial zeros and its relation to generalized Rayleigh iteration, Numer. Math. 14 (1970) 252–263.
[19] Z. Jia, X. Cheng, M. Zhao, A new method for roots of monic quaternionic quadratic polynomial, Comput. Math. Appl. 58 (2007) 1852–1858.
[20] Y. Jin, On efficient computation and asymptotic sharpness of Kalantari's bounds for zeros of polynomials, Math. Comp. 75 (2006) 1905–1912.
[21] B. Kalantari, Generalization of Taylor's theorem and Newton's method via a new family of determinantal interpolation formulas and its applications, J. Comput. Appl. Math. 126 (2000) 287–318.
[22] B. Kalantari, A new visual art medium: polynomiography, ACM SIGGRAPH Comput. Graph. Q. 38 (2004) 21–23.
[23] B. Kalantari, On homogeneous linear recurrence relations and approximation of zeros of complex polynomials, in: Unusual Applications in Number Theory, in: DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 64, 2004, pp. 125–143.
[24] B. Kalantari, Polynomiography and applications in art, education, and science, Comput. Graph. 28 (2004) 417–430.
[25] B. Kalantari, An infinite family of bounds on zeros of analytic functions and relationship to Smale's bound, Math. Comp. 74 (2005) 841–852.
[26] B. Kalantari, Polynomiography: from the fundamental theorem of algebra to art, Leonardo 38 (2005) 233–238.
[27] B. Kalantari, Polynomial Root-Finding and Polynomiography, World Scientific, New Jersey, 2008.
[28] B. Kalantari, Voronoi diagrams and polynomial root-finding, in: Proceedings of the Sixth Annual International Symposium on Voronoi Diagrams in Science and Engineering, IEEE, Copenhagen, Denmark, 2009, pp. 31–40.
[29] B. Kalantari, Polynomial root-finding methods whose basins of attraction approximate Voronoi diagram, Discrete Comput. Geom. 46 (2011) 187–203.
[30] B. Kalantari, I. Kalantari, R. Zaare-Nahandi, A basic family of iteration functions for polynomial root finding and its characterizations, J. Comput. Appl. Math. 80 (1997) 209–226.
[31] T.Y. Lam, A First Course in Noncommutative Rings, Springer-Verlag, New York, 1991.
[32] B.B. Mandelbrot, The Fractal Geometry of Nature, Freeman, San Francisco, 1982.
[33] J.M. McNamee, Numerical Methods for Roots of Polynomials: Part I, Volume 14, Elsevier, Amsterdam, 2007.
[34] J. Milnor, Dynamics in One Complex Variable: Introductory Lectures, third ed., vol. 160, Princeton University Press, New Jersey, 2006.
[35] C.A. Neff, J.H. Reif, An efficient algorithm for the complex roots problem, J. Complexity 12 (1996) 81–115.
[36] I. Niven, Equations in quaternions, Amer. Math. Monthly 48 (1941) 654–661.
[37] I. Niven, The roots of a quaternion, Amer. Math. Monthly 49 (1942) 386–388.
[38] A. Norton, Julia sets in the quaternions, Comput. Graph. 13 (1989) 267–278.
[39] G. Opfer, Polynomials and Vandermonde matrices over the field of quaternions, Electron. Trans. Numer. Anal. 36 (2009) 9–16.
[40] V.Y. Pan, Solving a polynomial equation: some history and recent progress, SIAM Rev. 39 (1997) 187–220.
[41] A. Pogorui, M. Shapiro, On the structure of the set of zeros of quaternion polynomials, Complex Var. 49 (2004) 379–389.
[42] M. Porter, Quaternionic linear and quadratic equations, J. Natur. Geom. 11 (1997) 101–106.
[43] H. Renmin, Z. Xuqiang, W. Linagtao, The double determinant of Vandermonde's type over quaternion field, Appl. Math. Mech. 20 (1999) 1046–1053.
[44] E. Schröder, On infinitely many algorithms for solving equations, Math. Ann. 2 (1870) 317–365 (in German), English translation by G.W. Stewart, TR-92-121, Institute for Advanced Computer Studies, University of Maryland, College Park, MD, 1992.
[45] R. Serodio, E. Pereira, J. Vitoria, Computing the zeros of quaternion polynomials, Comput. Math. Appl. 42 (2001) 1229–1237.
[46] R. Serodio, L. Siu, Zeros of quaternion polynomials, Appl. Math. Lett. 14 (2001) 237–239.
[47] J.F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, NJ, 1964.
[48] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra Appl. 251 (1997) 21–57.
[49] F. Zhang, D.L. Mu, Quadratic equations over noncommutative rings, J. Math. Res. Exposition 14 (1994) 260–264.