Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 11–20

AN ITERATIVE METHOD TO COMPUTE ZEROS OF QUATERNION POLYNOMIALS

ROGÉRIO SERÔDIO, JOSÉ AUGUSTO FERREIRA AND JOSÉ VITÓRIA

Abstract: The aim of this paper is to propose an iterative method to compute the dominant zero of a quaternion polynomial. We prove that the method is convergent in the sense that it generates a sequence of quaternions that converges to the dominant zero of the polynomial. The idea subjacent to the proposed method is the well known method proposed by Sebastião e Silva in "*Sur une méthode d'approximation semblable à celle de Gräffe*", Portugaliae Mathematica, 1941, to compute approximately the zeros of complex polynomials.

Key words: Division algebra, polynomials, zeros of polynomials, quaternion

Mathematics Subject Classification (2000): 12E15, 65H04

1. Introduction

The computation and the characterization of the zeros of quaternion polynomials have been object of research by several authors. Without being exhaustive we mention in [9], [16] and [21] where the topological structure of the zero-sets of quaternion polynomials is discussed. Methods for the computation of zeros of quarternionic polynomials were proposed for instance in [11], [18], [19] being the methods studied in [11], [18] based on Niven's algorithm [13]. It should be point out that the localization of quaternion polynomials were analyzed in [14].

The aim of this paper is to propose an iterative method to compute zeros of an unilateral (left) quaternion polynomial p that allows us to construct a sequence of unilateral (left) quaternion polynomials r_ℓ that converges to p, in a sense that will be specified later, as well as a sequence of quaternions that converges to the dominant zero of p. The motivation for our procedure is the algorithm of Sebastião e Silva [17] which is a method to compute approximately the zeros of complex polynomials, asserting that if a zero is distinct from all other ones, in modulus, the method provides this zero. We point out that this method continues to be object of research by several authors [1, 2, 3, 7, 8, 15, 20].

Received June 6, 2011.

The paper is organized as follows: in Section 2 we present the basic definitions involving quaternions, the norm that will be used and basic results. The main ingredient in the design of our procedure is the quaternionic Vandermonde matrix. Section 3 focused the computation of the an explicit expression of the inverse of the quaternionic Vandermonde matrix. For a quaternion polynomial p with specified requirements, in Section 4 we define a sequence of quaternion polynomials and a sequence of quaternions that converge to a quaternion polynomial, that has the zeros of p except the dominant zero, and to the dominant zero of p, respectively. These convergences are proved and the main result of this paper is established in this section - Theorem 4. Based on this main result we can design an algorithm that can be used to compute an approximation for the dominant zero of a quaternion polynomial.

2. Basic results

Let $\mathbb{H} = \{a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} : a, b, c, d \in \mathbb{R}\}$ be the quaternion field, where $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -1$, $\boldsymbol{ij} = -\boldsymbol{ji} = \boldsymbol{k}$, $\boldsymbol{jk} = -\boldsymbol{kj} = \boldsymbol{i}$, and $\boldsymbol{ki} = -\boldsymbol{ik} = \boldsymbol{j}$. For $\boldsymbol{\lambda} = a+b\boldsymbol{i}+c\boldsymbol{j}+d\boldsymbol{k} \in \mathbb{H}$, the conjugate of $\boldsymbol{\lambda}$ is defined as $\overline{\boldsymbol{\lambda}} = a-b\boldsymbol{i}-c\boldsymbol{j}-d\boldsymbol{k}$. Thus, a, the real part of $\boldsymbol{\lambda}$, denoted by $Re(\boldsymbol{\lambda})$, is given by $a = (\boldsymbol{\lambda} + \overline{\boldsymbol{\lambda}})/2$ and $\boldsymbol{\lambda}\overline{\boldsymbol{\lambda}} = \overline{\boldsymbol{\lambda}}\boldsymbol{\lambda} = a^2 + b^2 + c^2 + d^2 \in \mathbb{R}$. The norm of $\boldsymbol{\lambda}$, denoted by $|\boldsymbol{\lambda}|$, is defined by $\sqrt{\boldsymbol{\lambda}\overline{\boldsymbol{\lambda}}}$.

It can be show the following basic properties.

Proposition 1. *For any* $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \mathbb{H}$, $\overline{\boldsymbol{\lambda}_1 + \boldsymbol{\lambda}_2} = \overline{\boldsymbol{\lambda}_1} + \overline{\boldsymbol{\lambda}_2}$, *and* $\overline{\boldsymbol{\lambda}_1\boldsymbol{\lambda}_2} = \overline{\boldsymbol{\lambda}_2}\ \overline{\boldsymbol{\lambda}_1}$.

An important concept in approximation theory that has an important role in our method is the concept of dominance.

Definition 1. *Let* $\boldsymbol{\lambda}_i \in \mathbb{H}, i = 1, \dots m$. *We say that* $\boldsymbol{\lambda}_1$ *dominates* $\boldsymbol{\gamma}_2$ *if* $|\boldsymbol{\lambda}_1| > |\boldsymbol{\lambda}_2|$. *If* $\boldsymbol{\lambda}_1$ *dominates* $\boldsymbol{\lambda}_i, i = 1 \dots, m$, *then we say that* $\boldsymbol{\lambda}_1$ *is dominant in* $\{\boldsymbol{\lambda}_i, i = 1, \dots, m\}$.

It is easy to prove the next proposition:

Proposition 2. *If* $\boldsymbol{\lambda}_1 \in \mathbb{H}$ *dominates* $\boldsymbol{\lambda}_2 \in \mathbb{H}$, *then* $\lim\limits_{\ell\to\infty} \boldsymbol{\lambda}_1^{-\ell}\boldsymbol{\lambda}_2^{\ell} = 0$.

We introduce in $\mathbb{H}$ the relation $\sim$ defined by

$$\text{given } \boldsymbol{\lambda}, \boldsymbol{\lambda}' \in \mathbb{H}, \boldsymbol{\lambda} \sim \boldsymbol{\lambda}' \text{ if there exist } \sigma \in \mathbb{H} \text{ such that } \boldsymbol{\lambda}' = \sigma\boldsymbol{\lambda}\sigma^{-1}. \tag{1}$$

If $\boldsymbol{\lambda} \sim \boldsymbol{\lambda}'$ then we say that $\boldsymbol{\lambda}$ is similar to $\boldsymbol{\lambda}'$. The relation $\sim$ defined by (1) is an equivalence relation on $\mathbb{H}$. The conjugacy class of $\boldsymbol{\lambda} \in \mathbb{H}$ is defined by $\{x \in \mathbb{H} : x \sim \boldsymbol{\lambda}\}$is it is denoted by $[\boldsymbol{\lambda}]$.

[21] N. Topuridze, "*On roots of quaternion polynomials*", Journal of Mathematical Sciences, Springer, 160:843–855, 2009.
[22] F. Zhang, "*Quaternions and matrices of quaternions*", Linear Algebra Appl., 251:21–57, 1997.

Rogério Serôdio
Center of Mathematics-Department of Mathematics, University of Beira Interior, Rua Marquês d' Ávila e Bolama, 6201-001 Covilhã, Portugal
E-mail address: rserodio@ubi.pt

José Augusto Ferreira
CMUC-Department of Mathematics, University of Coimbra, Apartado 3008, 3001-454 Coimbra, Portugal
E-mail address: ferreira@mat.uc.pt
URL: http://www.mat.uc.pt/~ferreira

José Vitória
Department of Mathematics, University of Coimbra, Apartado 3008, 3001-454 Coimbra, Portugal
E-mail address: jvitoria@mat.uc.pt

Let $\mathbb{H}[x]$ denote the ring of the unilateral left polynomials $p(x) = \boldsymbol{\alpha}_n x^n + \boldsymbol{\alpha}_{n-1}x^{n-l} + \cdots + \boldsymbol{\alpha}_1 x + \boldsymbol{\alpha}_0$ over $\mathbb{H}$, where $n \in \mathbb{N}_0$, where $\boldsymbol{\alpha}_i \in \mathbb{H}$, for $i = 1, \ldots, n$, and $\boldsymbol{\alpha}_n \neq 0$.The addition and the multiplication of polynomials are defined in the same way as in the commutative case, where the variable x is assumed to commute with the quaternion coefficients. By pq we represent the polynomial product of $p, q \in \mathbb{H}[x]$. For every $\boldsymbol{\lambda} \in \mathbb{H}$, the value of p at $\boldsymbol{\lambda}$ is defined by $p(\boldsymbol{\lambda}) = \boldsymbol{\alpha}_n\boldsymbol{\lambda}^n + \boldsymbol{\alpha}_{n-1}\boldsymbol{\lambda}^{n-1} + \ldots + \boldsymbol{\alpha}_1\boldsymbol{\lambda} + \boldsymbol{\alpha}_0$. If $\boldsymbol{\lambda} \in \mathbb{H}$ is such that $p(\boldsymbol{\lambda}) = 0$ then $\boldsymbol{\lambda}$ is said to be a zero of p. The set of all zeros of p is denoted by $Zero(p)$.

We remark that as $\boldsymbol{\alpha}_n \neq 0$, then $p(\boldsymbol{\lambda}) = 0$ if and only if $\boldsymbol{\alpha}_n^{-1}p(\boldsymbol{\lambda}) = 0$. Thus, for the sake of simplicity, we can always assume that the polynomial is monic, i.e., $\boldsymbol{\alpha}_n = 1$.

Some results about quaternion polynomials are presented here.

If $p, q, r \in \mathbb{H}[X]$ are such that $p = qr$ then we say that r divides at left p or r is a left divisor of p. In what follows we will use the term "divides" and "divisor" to refer "divides at left" and "left divisor" respectively.

Theorem 1 ([10]). *$\boldsymbol{\lambda} \in \mathbb{H}, \boldsymbol{\lambda} \in Zero(p)$ if, and only if, $x - \boldsymbol{\lambda}$ is a divisor of p.*

Theorem 2 ([10]). *Let $p, q, r \in \mathbb{H}[X]$ such that $p = qr$ and $\boldsymbol{\lambda} \in \mathbb{H}$. If $\boldsymbol{\gamma} = r(\boldsymbol{\lambda}) \neq 0$, then*
$$p(\boldsymbol{\lambda}) = q(\boldsymbol{\gamma}\boldsymbol{\lambda}\boldsymbol{\gamma}^{-1})r(\boldsymbol{\lambda}).$$

Theorem 3 (Wedderburn's Theorem [10]). *All non-constant quaternion polynomial can be factorized into a product of linear factors.*

Corollary 3.1. *All non-constant quaternion polynomial of degree m can be factorized into a product of m linear factors.*

We define in $\mathbb{H}[X]$ a norm that will be used to introduce in $\mathbb{H}[X]$ the concept of distance. Let $p(x) = \boldsymbol{\alpha}_n x^n + \boldsymbol{\alpha}_{n-1}x^{n-l} + \cdots + \boldsymbol{\alpha}_1 x + \boldsymbol{\alpha}_0 \in \mathbb{H}[X]$. By $\|p\|_\infty$ we represent the following quantity
$$\|p\|_\infty = \max\{|\boldsymbol{\alpha}_i|, i = 0, \ldots, n\}. \tag{2}$$

It can be easily shown that $\|.\|_\infty : \mathbb{H}[X] \to \mathbb{R}_0^+$ is in fact a norm and we write $\lim\limits_{\ell\to\infty} p_\ell(x) = p(x)$ when $\lim\limits_{\ell\to\infty} \|p_\ell - p\|_\infty = 0$.

3. The quaternionic Vandermonde matrix

The computation of the sequence that converges to the dominant zero of $p \in \mathbb{H}[X]$ is obtained solving a linear system in $\mathbb{H}$ with the Vandermonde matrix. The explicit expression of such sequence requires the computation of the solution of the linear system. This expression is obtained constructing the inverse of the Vandermonde matrix. By $\mathcal{M}_m(\mathbb{H})$ we represent the set of all quaternionic matrices of order m.

Definition 2. *Given the quaternions $\boldsymbol{\lambda}_i \in \mathbb{H}, i = 1, \ldots, m$, the matrix $V_m = [\boldsymbol{\lambda}_j^{m-i}] \in \mathcal{M}_m(\mathbb{H})$ is called quaternionic Vandermonde matrix associated with $\boldsymbol{\lambda}_i, i = 1, \ldots, m$.*

From now on, we will consider only sets of distinct quaternions belonging at most two to each conjugacy class. Also, we will use the notation $\prod_{i=1}^{m} q_i$ to denote the product $q_m \ldots q_2.q_1$.

Let p_m be a monic quaternion polynomial such that $\{\boldsymbol{\lambda}_i, i = 1, \ldots, m\} \subseteq Zero(p_m)$. By Corollary 3.1 we can factorize p_m into the product of m linear factors $p_m(x) = \prod_{i=1}^{m}(x - \boldsymbol{\mu}_i)$, where $\boldsymbol{\mu}_i, i = 1, \ldots, m$, belong to the conjugacy classes of $\boldsymbol{\lambda}_i, i = 1, \ldots, m$, respectively. By Theorems 1 and 2 we can write

$$\boldsymbol{\mu}_1 = \boldsymbol{\lambda}_1,\ \boldsymbol{\mu}_i = p_{i-1}(\boldsymbol{\lambda}_i)\boldsymbol{\lambda}_i p_{i-1}(\boldsymbol{\lambda}_i)^{-1}, i = 2, \ldots, m,$$

where $p_i(x) = \prod_{j=1}^{i}(x - \boldsymbol{\mu}_j)$.

We now introduce several notations that will lead to a simplified presentation. Let $\mathcal{G}^{(k)}$ represent the set of all applications $g : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, m\}$ such that $g(\ell) < g(\ell + 1), \ell = 1, \ldots, k - 1$. Given $\boldsymbol{\lambda}_i \in \mathbb{H}, i = 1, \ldots, m$, we define $\boldsymbol{\sigma}_k(m)$ by

$$\boldsymbol{\sigma}_k(m) = \sum_{g \in \mathcal{G}^{(k)}} \prod_{\ell=1}^{k} \boldsymbol{\mu}_{g(\ell)}.$$

When no doubts subsists, we will represent $\boldsymbol{\sigma}_k(m)$ simply by $\boldsymbol{\sigma}_k$.

Using the above notation we can establish a representation of $p_m(x) = \prod_{i=1}^{m}(x - \boldsymbol{\mu}_i) \in \mathbb{H}[X]$.

of this paper it is established a set of conditions that guarantee that the two mentioned sequences converge.

It is worth noting that this method can be recursively applied if the conditions of convergence are fulfilled to the each deflated monic quaternion polynomial.

References

[1] D.A. Bini, G. Latouche, B. Meini, "*Solving matrix polynomial equations arising in queuing problems*", Linear Algebra Appl. 340:225–244, 2002.
[2] J-P. Cardinal, "*On two iterative methods for approximating the roots of a polynomial*", The mathematics of numerical analysis (Park City, UT, 1995), 165–188, Lectures in Appl. Math., 32, Amer. Math. Soc., Providence, RI, 1996.
[3] S.P. Chung, "*Generalization and acceleration of an algorithm of Sebastião e Silva and its duals*", Numer. Math., 25, No. 4, 365–377, 1975/76.
[4] A. Damiano, G. Gentil and D. Struppa, "*Computations in the ring of quaternionic polynomials*", J. Symbolic Comput., 45: 38–45, 2010.
[5] J. E. Dennis, Jr., J. F. Traub and R. P. Weber, "*Algorithms for solvents of matrix polynomials*", SIAM J. Numer. Anal., 15(3):523–533, 1978.
[6] J. E. Dennis, Jr., J. F. Traub and R. P. Weber, "*On the matrix polynomial, lambda-matrix and block eigenvalue problems*", Computer Science Department Technical Report, Cornell University, Ithaca, New York, and Carnigie-Mellon University, Pittsburgh, Pennsylvania, 1971.
[7] A.S. Householder, "*Generalization of an algorithm of Sebastião e Silva*", Numer. Math. 16:375–382, 1971.
[8] A.S. Householder, "*Multigradients and the zeros of transcendental functions*", Linear Algebra and Appl., 4:175–182, 1971.
[9] D. Janovská and G. Opfer, "*A note on the computation of all zeros of simple quaternionic polynomials*", SIAM J. Numer. Anal. 48: 244–256, 2010.
[10] T.Y. Lam **A First Course in Noncommutative Rings**, Springer, Berlin,, 1991.
[11] S. De Leo, G. Ducati and V. Leonardi, "*Zeros of unilateral quaternionic polinomials*", Eletron. J. Linear Algebra, 15:297–313, 2006.
[12] G. M. Menanno and N. Le Bihan, "*Quaternion polynomial matrix diagonalization for the separation of polarized convolutive mixture*", Journal Signal Processing, 90:2219–2231, 2010.
[13] I. Niven, "*Equations in quaternions*", Amer. Math. Monthly, 48:645–661, 1941.
[14] G. Opfer, "*Polynomials and Vandermonde matrices over the field of quaternions*", Electron. Trans. Numer. Anal., 36:9–16, 2009.
[15] V.Y. Pan, "*The amended DSeSC power method for polynomial root-finding*", Comput. Math. Appl., 49, No. 9-10, 1515–1524, 2005.
[16] A. Pogorui and M. Shapiro, "*On the structure of the set of zeros of quaternionic polynomials*", Complex Var. Elliptic Equ., 49:379–389, 2004.
[17] J. Sebastião e Silva, "*Sur une méthode d'approximation semblable à celle de Gräffe*", Port. Math., 2:271–279, 1941.
[18] R. Serôdio, E. Pereira and J. Vitória, "*Computing the zeros of quaternion polynomials*", Comput. Math. Appl., 42:1229–1237, 2001.
[19] R. Serôdio and Lok-Shun Siu, "*Zeros of quaternion polynomials*", Appl. Math. Lett., 14:237–239, 2001.
[20] G.W. Stewart, "*On the convergence of Sebastião e Silva's method for finding a zero of a polynomial*", SIAM Rev., 12, 458–460, 1970.

we obtain $\ell = 41$. *The polynomial* $\widetilde{r}_{41}(x) = x^3 + \boldsymbol{\alpha}_{2,41}x^2 + \boldsymbol{\alpha}_{1,41}x + \boldsymbol{\alpha}_{0,41}$, *with coefficients*

$$\begin{aligned}
\boldsymbol{\alpha}_{2,41} &= -4.406 \times 10^{-27} - 0.19408957238586511121824229261\ \boldsymbol{i} - \\
&\quad -0.11926915103890469073904450534 93\ \boldsymbol{j} + \\
&\quad +0.07462758520946825435086535132 78\ \boldsymbol{k}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\alpha}_{1,41} &= 1.97126741551366726124475727911 + \\
&\quad +1.26837969435472207491683942100\ \boldsymbol{i} - \\
&\quad -2.11984765945138118883478763260\ \boldsymbol{j} + \\
&\quad +0.56717928939883334136817239768\ \boldsymbol{k}
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{\alpha}_{0,41} &= -1.04898037892301017210625273053 + \\
&\quad +2.58718603866364556717928939570\ \boldsymbol{i} + \\
&\quad +2.51487248710408330521139660902\ \boldsymbol{j} + \\
&\quad +1.97015860772308730656124958591\ \boldsymbol{k}
\end{aligned}$$

satisfies

$$\|\tilde{r}_{41} - p_4^{(1)}\|_\infty = 2.16 \times 10^{-26} < 10^{-6}.$$

Furthermore

$$\begin{aligned}
\boldsymbol{\lambda}_{1,41} &= -1.99999999999999999999999999986 - 3.00000000000000000000000001316\ \boldsymbol{i} + \\
&\quad +6.99999999999999999999999997647\ \boldsymbol{j} + 3.00000000000000000000000002871\ \boldsymbol{k},
\end{aligned}$$

is such that

$$|\boldsymbol{\lambda}_{1,41} - \boldsymbol{\lambda}_1| = 3.94 \times 10^{-26} < 10^{-6}.$$

We remark that in this case the stop criteria (19) and (22) enable us to obtain the approximations $\boldsymbol{\lambda}_{1,41}$ *and* $\tilde{r}_{41}$ *with the prescribed accuracy.*

Conclusions

The characterization and the computation of the zeros of quaternion polynomials are problems that have been object of research as it can be seen in the references of this paper. However, until now we have no algorithms that enable us to generate a sequence of polynomials and a sequence of quaternions that converge to a specified polynomial and to the dominant zero of p_m, respectively, depending on the initial conditions. In this paper it is proposed an iterative method that satisfies the previous properties. In the main result

Lemma 1. *The polynomial* $p_m(x) = \prod_{i=1}^{m}(x - \boldsymbol{\mu}_i) \in \mathbb{H}[X]$ *admits the representation*

$$p_m(x) = x^m + \sum_{i=1}^{m}(-1)^i \boldsymbol{\sigma}_i(m) x^{m-i}.$$

Proof:

For $m = 1$ we have $p_1(x) = x - \boldsymbol{\mu}_1$, where $\boldsymbol{\sigma}_1(1) = \boldsymbol{\mu}_1$. Thus we conclude that Lemma 1 holds for $m = 1$.

Suppose that this result is valid for $m - 1$. Thus, for m, we have successively

$$\begin{aligned}
p_m(x) &= \prod_{i=1}^{m}(x - \boldsymbol{\mu}_i) \\
&= (x - \boldsymbol{\mu}_m)\left(x^{m-1} + \sum_{i=1}^{m-1}(-1)^i \boldsymbol{\sigma}_i(m-1)x^{m-1-i}\right) \\
&= x^m + \sum_{i=1}^{m-1}(-1)^i \big(\boldsymbol{\sigma}_i(m-1) \\
&\qquad + \boldsymbol{\mu}_m \boldsymbol{\sigma}_{i-1}(m-1)\big)x^{m-i} + (-1)^m \boldsymbol{\mu}_m \boldsymbol{\sigma}_{m-1}(m-1),
\end{aligned}$$

where $\boldsymbol{\sigma}_0(m-1) = 1$.
As

$$\boldsymbol{\sigma}_i(m-1) + \boldsymbol{\mu}_m \boldsymbol{\sigma}_{i-1}(m-1) = \boldsymbol{\sigma}_i(m), \quad \text{and} \quad \boldsymbol{\mu}_m \boldsymbol{\sigma}_{m-1}(m-1) = \boldsymbol{\sigma}_m(m),$$

we conclude the proof. □

In order to clarify some notation that we use in what follows, we represent by $\mathcal{F}_i^{(k)}$ the set of all applications $f_i : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, i-1, i+1, \ldots, m\}$ such that $f_i(\ell) < f_i(\ell + 1), \ell = 1, \ldots, k-1$. Given $\boldsymbol{\lambda}_i \in \mathbb{H}, i =, 1 \ldots, m$, let $\boldsymbol{\sigma}_k^{(i)}$ be defined by

$$\boldsymbol{\sigma}_k^{(i)} = \sum_{f_i \in \mathcal{F}_i^{(k)}} \prod_{\ell=1}^{k} \boldsymbol{\lambda}_{f_i(\ell)}.$$

If $[\boldsymbol{\lambda}_i] \neq [\boldsymbol{\lambda}_j], i \neq j$, we denote by $p_m^{(i)}$ the i-deflated monic quaternion polynomial of degree $m - 1$, such that

$$\{\boldsymbol{\lambda}_j, j = 1, \ldots, m, j \neq i\} \subseteq Zero(p_m^{(i)}).$$

Proposition 3. *Let $\boldsymbol{\lambda}_i \in \mathbb{H}, i = 1, \ldots, m$, such that $\boldsymbol{\lambda}_i \neq \boldsymbol{\lambda}_j, i \neq j$, and $[\boldsymbol{\lambda}_i] \cap [\boldsymbol{\lambda}_j] \cap [\boldsymbol{\lambda}_k] = \emptyset$ for any distinct triplet i, j, k. Let V_m be the quaternionic Vandermonde matrix associated with $\boldsymbol{\lambda}_i \in \mathbb{H}, i = 1, \ldots, m$. Then V_m^{-1} exists and $V_m^{-1} = [(-1)^j p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1} \boldsymbol{\sigma}_j^{(i)}]_{i=1,\ldots,m, j=0,\ldots,m-1}$, with $\boldsymbol{\sigma}_0^{(i)} = 1, i = 1, \ldots, m$.*

Proof:

Since $p_m^{(i)}(\boldsymbol{\lambda}_i) \neq 0$, $i = 1, \ldots, m$, the matrix V_m^{-1} exists. We will show now that $V_m^{-1} V_m = I$, where $I \in \mathcal{M}_m(\mathbb{H})$ denotes de identity matrix. We have

$$(V_m^{-1} V_m)_{ij} = p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1} \sum_{\ell=0}^{m-1} (-1)^\ell \boldsymbol{\sigma}_\ell^{(i)} \boldsymbol{\lambda}_j^{m-1-\ell},$$

and from Lemma 1 we obtain

$$(V_m^{-1} V_m)_{ij} = p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1} p_m^{(i)}(\boldsymbol{\lambda}_j).$$

As $p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1} p_m^{(i)}(\boldsymbol{\lambda}_j)^{-1} = \delta_{ij}$, where δ_{ij} denotes de Kronecker symbol, we conclude that $V_m^{-1} V_m = I$. Considering now Proposition 4.1 of [22] we conclude the proof. □

4. An iterative algorithm to compute the dominant zero

In this section we present a method that allows us to compute approximately the dominant zero of a quaternion polynomial. The method is based on Sebastião e Silva methods for complex polynomials p_m published in [17] which is based on a sequence of polynomials r_ℓ defined by

$$x^\ell \; = \; q_\ell(x) p_m(x) + r_\ell(x). \tag{3}$$

that converges to a polynomial with the zeros of the initial polynomial except the dominant zero. We remark that Dennis et al. in [5, 6] extended this method for matrix polynomials.

Let $p_m \in \mathbb{H}[X]$ be given, $p_m(x) = \sum_{j=0}^{m} \boldsymbol{\beta}_j x^j$ with $\boldsymbol{\beta}_m = 1^\dagger$. We introduce quaternion polynomial sequences (r_ℓ) and (q_ℓ) defined by

$$x^\ell \; = q_\ell(x) p_m(x) + r_\ell(x), \tag{4}$$

†If the polynomial p_m is not monic ($\boldsymbol{\beta}_m = 1$), we can multiply it by $\boldsymbol{\beta}_m^{-1}$ and turn it monic.

where ϵ is a prescribed tolerance. As before, the stop criterium (21) should be complemented with

$$\|\tilde{r}_\ell - \tilde{r}_{\ell-1}\|_\infty < \epsilon, \tag{22}$$

and

$$\frac{\|\tilde{r}_\ell - \tilde{r}_{\ell-1}\|_\infty}{\|\tilde{r}_{\ell-1}\|_\infty} < \epsilon. \tag{23}$$

In order to compute an approximation to $p_m^{(1)}$ with a certain accuracy we apply the described algorithm with (18), (19) and (20) replaced by (21), (22) and (23).

Under the assumption of Theorem 4, the sequences constructed by the previous algorithm converge to the dominant zero of p_m and to $p_m^{(1)}$, respectively. The following example illustrates the application of the introduced algorithm.

Example 4.1. *Consider the following quaternion polynomial*

$$p_4(x) = x^4 + (2 + 3\boldsymbol{i} - 7\boldsymbol{j} - 3\boldsymbol{k})x^3 + (2 - 2\boldsymbol{j} - \boldsymbol{k})x^2 + (-14 + \boldsymbol{i} - 21\boldsymbol{j} - \boldsymbol{k})x$$
$$+ 13 - 4\boldsymbol{i} - 2\boldsymbol{j} + 33\boldsymbol{k}$$

whose zeros belong to the conjugacy classes

$$[-1 + \sqrt{2}\ \boldsymbol{i}], \qquad [1 + \boldsymbol{i}], \qquad [\sqrt{3}\ \boldsymbol{i}], \qquad [-2 + \sqrt{67}\ \boldsymbol{i}].$$

The dominant zero is $\boldsymbol{\lambda}_1 = -2 - 3\boldsymbol{i} + 7\boldsymbol{j} + 3\boldsymbol{k}$. Thus, we have a dominant class and four distinct zeros. Applying Theorem 4, $(\widetilde{r}_\ell(x))$ converges to $p_4^{(1)}(x) = x^3 + \boldsymbol{\gamma}_2 x^2 + \boldsymbol{\gamma}_1 x + \boldsymbol{\gamma}_0$, with

$$\boldsymbol{\gamma}_2 \; = \; -\frac{4026\boldsymbol{i} + 2474\boldsymbol{j} - 1548\boldsymbol{k}}{20743}$$

$$\boldsymbol{\gamma}_1 \; = \; \frac{40890 + 26310\boldsymbol{i} - 43972\boldsymbol{j} + 11765\boldsymbol{k}}{20743}$$

$$\boldsymbol{\gamma}_0 \; = \; -\frac{21759 - 53666\boldsymbol{i} - 52166\boldsymbol{j} - 40867\boldsymbol{k}}{20743}.$$

Using the stop criteria

$$\|\tilde{r}_\ell(x) - \tilde{r}_{\ell-1}\|_\infty < 10^{-6},$$

and

$$|\boldsymbol{\lambda}_{1,\ell} - \boldsymbol{\lambda}_{1,\ell-1}| < 10^{-6},$$

Since $\boldsymbol{\lambda}_1$ is dominate, by Proposition 2 we conclude that

$$|\boldsymbol{\alpha}_{m-1,\ell+1}\left(\boldsymbol{\alpha}_{m-1,\ell}\right)^{-1}-\boldsymbol{\lambda}_1|\longrightarrow 0$$

which leads to (10). $\square$

The sequences (r_ℓ), (q_ℓ) can be computed using the Euclidean division algorithm for quaternion polynomials. However we can use the recursive relation (7). If $p_m \in \mathbb{H}[X]$ satisfies the assumptions of Theorem 4, then this result guarantees that the sequence $(\boldsymbol{\alpha}_{m-1,\ell})$ of the coefficients of the remainder r_ℓ of the division of x^ℓ by p_m determines the approximation $\boldsymbol{\lambda}_{1,\ell}=\boldsymbol{\alpha}_{m-1,\ell+1}\left(\boldsymbol{\alpha}_{m-1,\ell}\right)^{-1}$ for the dominant zero of p_m.

The natural stop criterium that we can use is

$$|p_m(\boldsymbol{\lambda}_{1,\ell})|<\epsilon, \tag{18}$$

where ϵ is a given tolerance. However this stop criterium can be verified being $|\boldsymbol{\lambda}_{1,\ell}-\boldsymbol{\lambda}_1| >> \epsilon$. Another stop criterium that we can consider consist in the computation of an approximation $\boldsymbol{\lambda}_{1,\ell}$ for $\boldsymbol{\lambda}_1$ such that

$$|\boldsymbol{\lambda}_{1,\ell}-\boldsymbol{\lambda}_{1,\ell-1}|<\epsilon. \tag{19}$$

But it is also possible that the approximation $\boldsymbol{\lambda}_{1,\ell}$ satisfies (19) being $|\boldsymbol{\lambda}_{1,\ell}-\boldsymbol{\lambda}_1| >> \epsilon$. This fact motivates the replacement of the stop criteria (18) and (19) by

$$|\boldsymbol{\lambda}_{1,\ell}-\boldsymbol{\lambda}_{1,\ell-1}||\boldsymbol{\lambda}_{1,\ell-1}^{-1}|<\epsilon. \tag{20}$$

This means that the relative error of $\boldsymbol{\lambda}_{1,\ell}$ (with respect to $\boldsymbol{\lambda}_{1,\ell-1}$) is less than the prescribed tolerance ϵ.

The use of the three stop criteria (18), (19) and (20) leads to more accurate results than the use of only one criterion.

Algorithm I: Given $p_m \in \mathbb{H}[X]$, let $r_0(x)=1$ and ϵ be a given tolerance. For $\ell=1,2,\ldots$, compute $r_{\ell+1}$ using (7) where $\boldsymbol{\alpha}_{m-1,\ell}$ is the coefficient of the highest term of r_ℓ. If (18), (19) and (20) hold then stop, else continue.

Theorem 4 also guarantees that the sequence $(\tilde{r}_\ell)$ converges to the polynomial $p_m^{(1)}$. Then we can compute $(\tilde{r}_\ell)$ such that

$$\|\tilde{r}_\ell\|_\infty<\epsilon, \tag{21}$$

where

$$r_\ell(x)=\sum_{j=0}^{m-1}\boldsymbol{\alpha}_{j,\ell}x^j. \tag{5}$$

and $r_0(x)=1$.

The sequences (q_ℓ) can be computed using the Euclidean division algorithm for quaternion polynomials. That is, for each ℓ, the computation of r_ℓ and q_ℓ can be made using the Euclidean division algorithm performed as in the commutative case but adapted to the right multiplication (see Proposition 4 in [4]). In this case (q_ℓ) satisfies the following assumptions: $degree(q_{\ell+1}) = degree(q_\ell)+1$,

$$q_{\ell+1}(x)=q_\ell(x)x+\boldsymbol{\alpha}_{m-1,\ell}. \tag{6}$$

Under such assumptions we establish that (r_ℓ) defined by (4) satisfies a more convenient recursion relation.

Proposition 4. *For $p_m \in \mathbb{H}[X]$, let (r_ℓ), (q_ℓ) be two quartenionic polynomials sequences defined by (4), (5) and (6). Then*

$$r_{\ell+1}(x)=r_\ell(x)x-\boldsymbol{\alpha}_{m-1,\ell}\,p_m(x). \tag{7}$$

Proof:

From (4) we obtain

$$x^{\ell+1} \;=\; q_\ell(x)p_m(x)x+r_\ell(x)x. \tag{8}$$

Considering (4) for ℓ and for $\ell+1$ we establish

$$r_{\ell+1}(x)=r_\ell(x)+(q_\ell-q_{\ell+1}(x))p_m(x).$$

Finally, using the assumption $q_{\ell+1}(x)=q_\ell(x)+\boldsymbol{\alpha}_{m-1,\ell}$ we conclude the proof. $\square$

In the next result we establish that if $p_m \in \mathbb{H}[X]$ satisfies convenient assumptions, then the sequence (r_ℓ) defined by (4), (5) and (6) converges to a polynomial with specified properties and it allows us to define a sequence in $\mathbb{H}$ that converges for the dominant zero of p_m.

Theorem 4. *Let $p_m \in \mathbb{H}[X]$ be a monic quaternion polynomial of degree m such that*

(a) $\{\boldsymbol{\lambda}_i, i=1,\ldots,m\}\subseteq Zero(p_m)$, where $\boldsymbol{\lambda}_i\neq\boldsymbol{\lambda}_j, i\neq j$, and $[\boldsymbol{\lambda}_i]\cap[\boldsymbol{\lambda}_j]\cap[\boldsymbol{\lambda}_k]=\emptyset$ for any distinct triplet i,j,k,

(b) $\boldsymbol{\lambda}_1$ is a dominant zero.

Let (r_ℓ) and (q_ℓ) be two quaternion polynomials sequences that satisfies (4), (5) and (6). Then

$$\lim_{\ell\to\infty} \widetilde{r}_\ell(x) = p_m^{(1)}(x), \tag{9}$$

where $\widetilde{r}_\ell$ is the monic quaternion polynomial associated with r_ℓ, and

$$\lim_{\ell\to\infty} \boldsymbol{\alpha}_{m-1,\ell+1} \left(\boldsymbol{\alpha}_{m-1,\ell}\right)^{-1} = \boldsymbol{\lambda}_1. \tag{10}$$

Proof:

We start by considering the sequences (r_ℓ) and (q_ℓ) that satisfy (4), (5) and (6). For all $x = \boldsymbol{\gamma} \neq \boldsymbol{\lambda}_i$, $i = 1, \ldots, m$ we have

$$\boldsymbol{\gamma}^\ell = q_\ell(\boldsymbol{\beta}\boldsymbol{\gamma}\boldsymbol{\beta}^{-1}) p_m(\boldsymbol{\gamma}) + r_\ell(\boldsymbol{\gamma}), \tag{11}$$

where $\boldsymbol{\beta} = p_m(\boldsymbol{\gamma})$. But, for $\boldsymbol{\gamma} = \boldsymbol{\lambda}_i$, $i = 1, \ldots, m$, applying Theorem 2 to (11) we obtain

$$r_\ell(\boldsymbol{\lambda}_i) = \boldsymbol{\lambda}_i^\ell,\ i = 1, \ldots, m,$$

that is, by (5),

$$\sum_{j=1}^{m} (-1)^{j-1} \boldsymbol{\alpha}_{m-j,\ell} \boldsymbol{\lambda}_i^{m-j} = \boldsymbol{\lambda}_i^\ell,\ i = 1, \ldots, m. \tag{12}$$

As system (12) is equivalent to

$$[(-1)^{i-1}\boldsymbol{\alpha}_{m-i,\ell}]^t V_m = [\boldsymbol{\lambda}_i^\ell]^t \tag{13}$$

where V_m denotes the quaternionic Vandermonde matrix, by Proposition 3 we establish

$$\boldsymbol{\alpha}_{i,\ell} = \sum_{j=1}^{m} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \boldsymbol{\sigma}_{m-1-i}^{(j)},\ i = 0, \ldots, m-1. \tag{14}$$

The coefficients of the monic polynomial $\widetilde{r}_\ell$, $\widetilde{r}_\ell(x) = \boldsymbol{\alpha}_{m-1,\ell}^{-1} r_\ell(x)$, are then easily obtained

$$\boldsymbol{\gamma}_{i,\ell} = \left(\sum_{j=1}^{m} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1}\right)^{-1} \sum_{j=1}^{m} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \boldsymbol{\sigma}_{m-1-i}^{(j)}, \tag{15}$$

for $i = 0, \ldots, m-2$.

We remark that, for $i = 0, \ldots, m-2$, $\boldsymbol{\gamma}_{i,\ell}$ admits the representation

$$\boldsymbol{\gamma}_{i,\ell} = \left(\boldsymbol{\lambda}_1^\ell \sum_{i=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_i^\ell p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1}\right)^{-1} \left(\boldsymbol{\lambda}_1^\ell \sum_{i=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_i^\ell p_m^{(i)}(\boldsymbol{\lambda}_i)^{-1} \boldsymbol{\sigma}_{m-1-i}^{(i)}\right)$$

and then Proposition 1 allows us to conclude that

$$\boldsymbol{\gamma}_{i,\ell} = \left(\sum_{i=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1}\right)^{-1} \sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \boldsymbol{\sigma}_{m-1-i}^{(j)}. \tag{16}$$

Finally, using the dominance of $\boldsymbol{\lambda}_1$ in $\{\boldsymbol{\lambda}_i, i = 1, \ldots, m\}$, we get by Proposition 2,

$$\boldsymbol{\gamma}_{i,\ell} \longrightarrow \left(p_m^{(1)}(\boldsymbol{\lambda}_1)^{-1}\right)^{-1} p_m^{(1)}(\boldsymbol{\lambda}_1)^{-1} \boldsymbol{\sigma}_{m-1-i}^{(1)},$$

that is

$$\boldsymbol{\gamma}_{i,\ell} \longrightarrow \boldsymbol{\sigma}_{m-1-i}^{(1)},\ i = 0, \ldots, m-2, \tag{17}$$

which implies (9).

In what follows we prove (10). Using (15) for ℓ and $\ell+1$ and Proposition 1 we easily establish

$$\begin{aligned}
&\boldsymbol{\alpha}_{m-1,\ell+1} \left(\boldsymbol{\alpha}_{m-1,\ell}\right)^{-1} \\
&= \boldsymbol{\lambda}_1^{\ell+1} \left(\sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-(\ell+1)} \boldsymbol{\lambda}_j^{\ell+1} p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1}\right) \left(\sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1}\right)^{-1} \boldsymbol{\lambda}_1^{-\ell}.
\end{aligned}$$

Thus we have successively

$$\begin{aligned}
&\left|\boldsymbol{\alpha}_{m-1,\ell+1} \left(\boldsymbol{\alpha}_{m-1,\ell}\right)^{-1} - \boldsymbol{\lambda}_1\right| = \\
&= |\boldsymbol{\lambda}_1|^{\ell+1} \left| \left(\sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-(\ell+1)} \boldsymbol{\lambda}_j^{\ell+1} p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1}\right) \left(\sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(i)}(\boldsymbol{\lambda}_j)^{-1}\right)^{-1} - 1 \right| |\boldsymbol{\lambda}_1|^{-\ell} \\
&= |\boldsymbol{\lambda}_1| \left| \sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-(\ell+1)} \boldsymbol{\lambda}_j^{\ell+1} p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} - \sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)} \boldsymbol{\lambda}_j^{-1} \right| \left| \sum_{j=1}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \right|^{-1} \\
&= |\boldsymbol{\lambda}_1| \left| \sum_{j=2}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \left(\boldsymbol{\lambda}_1^{-1}\boldsymbol{\lambda}_j - 1\right) \right| \left| p_m^{(1)}(\boldsymbol{\lambda}_1)^{-1} + \sum_{j=2}^{m} \boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \right|^{-1} \\
&\leq |\boldsymbol{\lambda}_1| \sum_{j=2}^{m} |\boldsymbol{\lambda}_1^{-\ell} \boldsymbol{\lambda}_j^\ell| \left| p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \left(\boldsymbol{\lambda}_1^{-1}\boldsymbol{\lambda}_j - 1\right) \right| \left| p_m^{(1)}(\boldsymbol{\lambda}_1)^{-1} + \sum_{j=2}^{m} (\boldsymbol{\lambda}_1^{-1}\boldsymbol{\lambda}_j)^\ell p_m^{(j)}(\boldsymbol{\lambda}_j)^{-1} \right|^{-1}.
\end{aligned}$$