Algebraic tools for the study of quaternionic behavioral systems

Ricardo Pereira*, Paula Rocha and Paolo Vettori
Department of Mathematics
University of Aveiro
{ricardo, procha, pvettori}@mat.ua.pt

Abstract

In this paper we study behavioral systems whose trajectories are given as solutions of quaternionic difference equations. As happens in the commutative case, it turns out that quaternionic polynomial matrices play an important role in this context. Therefore we pay special attention to such matrices and derive new results concerning their Smith form. Based on these results, we obtain a characterization of system theoretic properties such as controllability and stability of a quaternionic behavior.

1 Introduction

The behavioral approach to dynamical systems, introduced by J. C. Willems [15, 16] in the eighties, considers as the main object of study in a system the set of all the trajectories which are compatible with its laws, known as the system behavior. Whereas the classical approaches start by dividing the trajectories into input, output and/or state space variables, according to some predefined mathematical model (for instance, the input-output or the state space model), the point of view of the behavioral approach is rather innovative. One looks at the set of trajectories without imposing any structure, i.e., without speaking, at an early stage, of inputs and outputs, of causes and effects. This point of view does not only unify the previous approaches,

*Supported by Fundação para a Ciência e a Tecnologia

fitting them within an elegant theory, but it also permits to study a larger class of dynamical systems, including situations where it is not possible or desirable to make any distinction between input and output variables.

During the last two decades the importance of Clifford algebras, and in particular of the quaternion algebra, has been widely recognized. Actually, using this algebra, phenomena occurring in areas such as electromagnetism, quantum physics and robotics may be described by a more compact notation [5, 7].

Systems with quaternionic signals were already investigated in the classic state space approach [4]. Here we aim at laying the foundations of the theory of quaternionic systems in the behavioral approach. Although every quaternionic system can be regarded as a complex or real system of higher dimension with special structure, keeping at the quaternionic level (i.e., viewing it as a system over $\mathbb{H}$) allows higher efficiency in computational terms. Since quaternionic polynomial matrices play an important role in this context, a considerable part of our work is devoted to the study of such matrices and in particular to their Smith form. The obtained results are relevant for the algebraic characterization of system theoretic properties.

The structure of the paper is as follows. In section 2, after presenting the quaternionic skew-field $\mathbb{H}$ and quaternionic matrices, we refer to some examples that show the advantages of using quaternions in the description and solution of well-known physical problems. In section 3 we introduce basic notions of quaternionic behavioral theory and show how to extend usual concepts of commutative linear algebra to the quaternionic algebra. Then, in section 4, we characterize the Smith form of complex adjoint matrices and make its relation to the quaternionic Smith form explicit. Sections 5 and 6 are concerned with the characterization of controllability and stability of quaternionic behaviors.

2 Quaternions and Applications

Let $\mathbb{R}$ denote the field of real numbers. The quaternion skew-field $\mathbb{H}$ is an associative but non-commutative algebra over $\mathbb{R}$ defined as the set

$$\mathbb{H} = \{a + bi + cj + dk \,:\, a, b, c, d \in \mathbb{R}\},$$

where i, j, k are called imaginary units and satisfy

$$i^2 = j^2 = k^2 = ijk = -1.$$

This implies that

$$ij = -ji = k, \qquad jk = -kj = i, \qquad ki = -ik = j.$$

For any $\eta = a + bi + cj + dk \in \mathbb{H}$, we define its *real part* as $\operatorname{Re}\eta = a$, its *imaginary part* as $\operatorname{Im}\eta = bi + cj + dk$, its *conjugate* as $\overline{\eta} = \operatorname{Re}\eta - \operatorname{Im}\eta = a - bi - cj - dk$, and its *norm* as $|\eta| = \sqrt{\overline{\eta}\eta} = \sqrt{a^2 + b^2 + c^2 + d^2}$. Note that $\overline{\eta\nu} = \overline{\nu}\,\overline{\eta}$, $\forall \eta, \nu \in \mathbb{H}$.

Two quaternions η and ν are said to be *similar*, $\eta \sim \nu$, if there exists a nonzero $\alpha \in \mathbb{H}$ such that $\eta = \alpha\nu\alpha^{-1}$. Similarity is an equivalence relation and we denote by $[\nu]$ the equivalence class containing ν. It can be proved [17] that $\eta \sim \nu$ if and only if $\operatorname{Re}\eta = \operatorname{Re}\nu$ and $|\eta| = |\nu|$. Therefore, for instance, all the imaginary units belong to the same equivalence class, i.e., $i \sim j \sim k$. Moreover, for all $\eta \in \mathbb{H}$, $\eta \sim \overline{\eta}$. As a consequence of the characterization of similarity, the following holds, where $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\} \subseteq \mathbb{H}$ is the complex field.

Proposition 2.1. *[17] For all $\eta \in \mathbb{H}$, $[\eta] \cap \mathbb{C} \neq \emptyset$ and $[\eta] \cap \mathbb{R} \neq \emptyset \Leftrightarrow [\eta] = \{\eta\} \Leftrightarrow \eta \in \mathbb{R}$.*

Example 2.2. Consider the quaternion $\eta = 1 - 2i + j + 2k$. The complex $z = 1 + 3i$ and its conjugate $\overline{z} = 1 - 3i$ are similar to η, since $\operatorname{Re} z = \operatorname{Re}\overline{z} = \operatorname{Re}\eta = 1$ and $|z| = |\overline{z}| = |\eta| = \sqrt{10}$.

Given an $m \times n$ matrix with quaternionic entries $A = (a_{st}) \in \mathbb{H}^{m\times n}$, its *conjugate* is defined as $\overline{A} = (\overline{a}_{st})$, its *transpose* as $A^\top = (a_{ts}) \in \mathbb{H}^{n\times m}$, and its *conjugate transpose* as $A^* = \overline{A}^\top \in \mathbb{H}^{n\times m}$.

Since each matrix $A \in \mathbb{H}^{m\times n}$ may be uniquely written as $A = A_1 + A_2 j$, where A_1, $A_2 \in \mathbb{C}^{m\times n}$ we can define an injective $\mathbb{R}$-linear map: $\mathbb{H}^{m\times n} \to \mathbb{C}^{2m\times 2n}$ such that

$$A \mapsto A^c = \begin{bmatrix} A_1 & A_2 \\ -\overline{A_2} & \overline{A_1} \end{bmatrix}. \tag{2.1}$$

The matrix A^c is called the *complex adjoint matrix* of A. In general, any complex matrix with structure (2.1) is said to be a *complex adjoint matrix.*

We may as well define a bijective $\mathbb{R}$-linear map: $\mathbb{H}^{m\times n} \to \mathbb{C}^{2m\times n}$ such that

$$A \mapsto A^{cv} = \begin{bmatrix} A_1 \\ -\overline{A_2} \end{bmatrix}, \tag{2.2}$$

which in particular maps column vectors into column vectors.

It is easy to check that the following properties hold [17]:

$$(I_n)^c = I_{2n}; \quad (A^{-1})^c = (A^c)^{-1}; \quad (AB)^c = A^c B^c; \quad (AB)^{cv} = A^c B^{cv}, \tag{2.3}$$

where I_n is the $n \times n$ identity matrix and A and B are quaternionic matrices of suitable dimensions and, in case, invertible.

Next we present some examples of applications of quaternions which motivated our interest in quaternionic dynamical systems.

Quaternions are a powerful tool in the description of rotations in $\mathbb{R}^3$. Indeed, let $u = \begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}^T$ and $v = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T$ be vectors in $\mathbb{R}^3$. Suppose that the rotation of v by an angle θ about the direction of u yields the vector $\widetilde{v} = \begin{bmatrix} \widetilde{v_1} & \widetilde{v_2} & \widetilde{v_3} \end{bmatrix}^T$.

Identifying v and u with the quaternions $v_1 i + v_2 j + v_3 k$ and $u_1 i + u_2 j + u_3 k$, respectively, and letting q be the quaternion

$$q = \cos\frac{\theta}{2} + \frac{u}{|u|}\sin\frac{\theta}{2}, \tag{2.4}$$

we have that $qv\overline{q} = \widetilde{v_1} i + \widetilde{v_2} j + \widetilde{v_3} k$ (see [8]). Actually, for any $q, v \in \mathbb{H}$, the quaternion $\widetilde{v} = qv\overline{q}$ has zero real part whenever so has v. When q is unitary, i.e., $|q| = 1$, the action of the operator $L_q(v) = qv\overline{q}$ consists in a rotation of v by an angle θ about the direction specified by the vector u, where u and θ are given, up to periodicity, by formula (2.4).

Note that the operator L_q is defined for any $q \in \mathbb{H}$ and acts as a rotation composed with a norm variation. So, to obtain only a rotation, when q is not unitary, the operator $\widetilde{L}_q : v \mapsto qvq^{-1}$ is used, which is defined for any $q \neq 0$. Indeed, it is easy to check that $\widetilde{L}_q = L_{\frac{q}{|q|}}$.

It is not uncommon to find situations where the rotation of a rigid body is dependent on time, and this dynamics is advantageously written in terms of quaternionic differential or difference equations (see [3, 12]).

Using quaternionic notation it is also possible to find an elegant solution of the differential equation which describes the orbits of the planets, i.e., to solve the "Kepler problem" (see [13]). Quaternions, compared to vectors in $\mathbb{R}^3$, have an extra degree of freedom which may be exploited to simplify the equations. Indeed, by choosing conveniently the free parameter, the problem is reduced to the resolution of the simple quaternionic differential equation $\ddot{q} + q = 0$. The solution of this equation is $q = e^{it}\alpha + e^{-it}\beta$, where α and β are suitable constant quaternions.

Furthermore, many are the physical theories, from electromagnetism to relativity, that can be formulated naturally using quaternions. As for quantum mechanics, in particular, the introduction of quaternionic potentials gives rise to a new theory whose predictions still have to be confirmed (see [1, 2] and the references therein).

3 Quaternionic Behavioral Systems

In this section we study behaviors that can be described as solution sets of quaternionic matrix difference equations, i.e., those which are the kernel of some suitable matrix difference operator. Such equations arise either directly or from the digital implementation of problems described by differential equations. We first provide the necessary results about quaternionic polynomials and polynomial matrices. Then, we show that as in the real and the in complex case, two matrices represent the same behavior if and only if each one is a left multiple of the other.

Definition 3.1. [10, Def. 1.3.1] A *dynamical system* Σ is defined as a triple

$$\Sigma = (\mathbb{T}, \mathbb{W}, \mathcal{B}),$$

with $\mathbb{T}$ a set called the *time axis*, $\mathbb{W}$ a set called the *signal space*, and $\mathcal{B}$ a subset of $\mathbb{W}^{\mathbb{T}}$ called the *behavior*, where $\mathbb{W}^{\mathbb{T}} = \{f :\ \mathbb{T} \to \mathbb{W}\}$.

In this paper $\mathbb{T} = \mathbb{Z}$ and $\mathbb{W} = \mathbb{H}^r$, for some $r \in \mathbb{N}$. This class of systems is called *discrete-time quaternionic systems.*

We will assume that the system behavior $\mathcal{B}$ can be described by means of matrix difference equations, i.e., the trajectories w in $\mathcal{B}$ are the solutions of an equation of the form

$$R_M w(t+M) + R_{M+1} w(t+M+1) + \cdots + R_N w(t+N) = 0,\ \forall\, t \in \mathbb{Z}, \quad (3.1)$$

where $R_p \in \mathbb{H}^{g\times r}$, $p = M, \ldots, N$, $M \leq N$, $M, N \in \mathbb{Z}$.

If we define the *shift operator* by $(\sigma^\tau w)(t) = w(t+\tau)$, $t, \tau \in \mathbb{Z}$, the left-hand side of equation (3.1) can be written in the more compact form

$$R(\sigma, \sigma^{-1}) w(t) = \sum_{l=M}^{N} R_l \sigma^l w(t) = \sum_{l=M}^{N} R_l w(t+l). \quad (3.2)$$

This notation also reveals that we may describe $\mathcal{B}$ as the kernel of the difference operator $R(\sigma, \sigma^{-1})$ acting on $(\mathbb{H}^r)^{\mathbb{Z}}$, i.e.,

$$\mathcal{B} = \ker R(\sigma, \sigma^{-1}) := \left\{ w \in (\mathbb{H}^r)^{\mathbb{Z}} : R(\sigma, \sigma^{-1}) w = 0 \right\}. \quad (3.3)$$

Note that behaviors which can be written as the kernel of some difference operator are *linear on the right*, i.e., for any $w_1, w_2 \in \mathcal{B}$ and $\alpha_1, \alpha_2 \in \mathbb{H}$, $w_1\alpha_1 + w_2\alpha_2 \in \mathcal{B}$, and *shift-invariant*, i.e., if $w \in \mathcal{B}$ then $\sigma^\tau w \in \mathcal{B}$, $\forall \tau \in \mathbb{Z}$, or, equivalently, $\sigma^\tau \mathcal{B} = \mathcal{B}$.

The form of the operator $R(\sigma, \sigma^{-1})$ in (3.2) suggests, as it is usual within the behavioral approach, to consider the polynomial matrix (in s and s^{-1})

$$R(s, s^{-1}) = \sum_{l=M}^{N} R_l s^l, \tag{3.4}$$

which is called a *kernel representation* of the behavior (3.3), and try to relate its algebraic properties to dynamical properties of $\mathcal{B}$.

However, unlike the real or complex case, there is not a unique way to define quaternionic polynomials. The one we will choose is determined by the following consideration. The variable of the polynomial represents the shift operator σ, which clearly commutes with any quaternionic value. Thus, also the variable s has to commute with the coefficients. This leads to the following definition.

Definition 3.2. A *quaternionic Laurent-polynomial* (or *L-polynomial*) $p(s, s^{-1})$ is defined by

$$p(s, s^{-1}) = \sum_{l=M}^{N} p_l s^l,\ p_l \in \mathbb{H},\ \ M, N \in \mathbb{Z}, M \leq N,$$

where s (and s^{-1}) commute with the coefficients.

If $p_N \neq 0 \neq p_M$, the *degree* of $p(s, s^{-1})$ is $\deg p = N - M$. If $M \geq 0$, $p(s, s^{-1})$ is simply said to be a *quaternionic polynomial* and we denote it by $p(s)$. In this case, the degree of the polynomial $p(s)$ is N.

The set of quaternionic L-polynomials and quaternionic polynomials are denoted by $\mathbb{H}[s, s^{-1}]$ and by $\mathbb{H}[s]$, respectively. As usual, $\mathbb{H}^{m\times n}[s, s^{-1}]$ and $\mathbb{H}^{m\times n}[s]$ are the sets of $m \times n$ matrices with entries in $\mathbb{H}[s, s^{-1}]$ and $\mathbb{H}[s]$, respectively.

Quaternionic (L-) polynomials endowed with the usual operations are non-commutative rings. Note that, since s commutes with the coefficients, as we said before, $(\alpha s^n)(\beta s^m) = \alpha\beta s^{n+m}$, $\alpha, \beta \in \mathbb{H}$. With this definition, moreover, it is easily proved that the (L-) polynomial corresponding to the composition of operators $p(\sigma, \sigma^{-1}) \circ q(\sigma, \sigma^{-1})$ is equal to the product $p(s, s^{-1})q(s, s^{-1})$.

To simplify the notation, in the sequel we may omit the indeterminates s and s^{-1}. We will also indicate the product of polynomials $p(s, s^{-1})$ and $q(s, s^{-1})$ as $pq(s, s^{-1})$.

The notions of conjugacy and of similarity for quaternionic (L-) polynomials are naturally defined as follows. The *conjugate* of $p(s) = p_N s^N + p_{N-1} s^{N-1} + \cdots + p_M s^M \in \mathbb{H}[s, s^{-1}]$ is $\overline{p}(s) = \overline{p}_N s^N + \overline{p}_{N-1} s^{N-1} + \cdots + \overline{p}_M s^M$. As regards the similarity of L-polynomials, we say that $p(s, s^{-1}) \sim q(s, s^{-1})$ if there exists a nonzero $\alpha \in \mathbb{H}$ such that $p(s, s^{-1}) = \alpha q(s, s^{-1})\alpha^{-1}$. Clearly, this is an equivalence relation. We denote by $[q(s, s^{-1})]$ the equivalence class containing $q(s, s^{-1})$.

Properties related to conjugation of quaternions extend to polynomials as we show in the following propositions.

Proposition 3.3. *Let* $p_1, \ldots, p_n \in \mathbb{H}[s, s^{-1}]$. *Then* $\overline{p_1 p_2 \cdots p_n} = \overline{p}_n \cdots \overline{p}_2 \overline{p}_1$.

Proof. The proof is trivial since $\overline{\eta\nu} = \overline{\nu}\,\overline{\eta}$, $\forall \eta, \nu \in \mathbb{H}$ and s commutes with the coefficients. □

Proposition 3.4. *Let* $p, q \in \mathbb{H}[s, s^{-1}]$. *Then*

(*i*) $p\overline{p} = \overline{p}p \in \mathbb{R}[s, s^{-1}]$.

(*ii*) *If* $pq \in \mathbb{R}[s, s^{-1}]$, *then* $pq = qp$.

Proof. (*i*) By Proposition 3.3 we have that $\overline{\overline{p}p} = \overline{p}\,\overline{\overline{p}} = \overline{p}p$, i.e., $\overline{p}p \in \mathbb{R}[s, s^{-1}]$. Therefore it commutes with p, i.e., $p\overline{p}p = \overline{p}pp$. Hence $p\overline{p} = \overline{p}p$.

(*ii*) Real polynomials commute with any polynomial. Thus, by (*i*),

$$\overline{q}pq = pq\overline{q} = q\overline{q}p = \overline{q}qp,$$

and so, $pq = qp$. □

The definition of the complex adjoint matrix of $R(s, s^{-1}) \in \mathbb{H}^{m \times n}[s, s^{-1}]$, $R^c(s, s^{-1})$, is analogous to the constant case. Similarly, we extend the map (2.2) to sequences and define for any behavior $\mathcal{B}$ the complex behavior $\mathcal{B}^{\mathbb{C}} = \{w^{cv} : w \in \mathcal{B}\}$, where $w^{cv}(t) = (w(t))^{cv}$. $\mathcal{B}^{\mathbb{C}}$ is called the *complex form* of $\mathcal{B}$ and, as the following proposition shows, admits a kernel representation which can be derived from any kernel representation of $\mathcal{B}$.

Proposition 3.5. *Let* $R(s, s^{-1}) \in \mathbb{H}^{m \times n}[s, s^{-1}]$. *Then* $(\ker R(\sigma, \sigma^{-1}))^{\mathbb{C}} = \ker R^c(\sigma, \sigma^{-1})$.

Proof. Let $v \in (\ker R)^{\mathbb{C}}$. Then, by definition there exists $w \in \ker R$ such that $v = w^{cv}$. Since $Rw = 0$ then $R^c v = R^c w^{cv} = 0$. Hence $v \in \ker R^c$. Conversely, let $v \in \ker R^c$. This uniquely determines w such that $v = w^{cv}$. Then $(Rw)^{cv} = R^c w^{cv} = R^c v = 0$, which implies that $Rw = 0$ and also $v = w^{cv} \in (\ker R)^{\mathbb{C}}$. □

Proposition 3.5 shows that the analysis of $\mathcal{B}$ is equivalent to the analysis of its complex form $\mathcal{B}^{\mathbb{C}}$. In the same way, quaternionic (L-) polynomial matrices share many algebraic properties with their complex adjoint matrices, as we show in the following statements, where *unimodular* matrices are defined analogously to the commutative case and *full row rank (frr)* matrices are (L-) polynomial matrices R such that for any (L-) polynomial row vector X, $XR = 0$ implies $X = 0$.

Lemma 3.6. *A quaternionic (L-) polynomial matrix R is frr if and only if R^c is frr. More generally, for every quaternionic (L-) polynomial matrix R,* $\operatorname{rank} R = n$ *if and only if* $\operatorname{rank} R^c = 2n$.

Proof. Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$. First we prove that R is frr if and only if R^c is frr.

"If" part. Suppose that R is not frr. Then there exists a nonzero row vector $X \in \mathbb{H}^{1\times g}[s, s^{-1}]$ such that $XR = 0$, hence $X^c R^c = 0$, with $X^c \neq 0$, i.e., R^c is not frr.

"Only if" part. Suppose that R^c is not frr. Then there exists a nonzero complex polynomial row vector $Y = \begin{bmatrix} Y_1 & Y_2 \end{bmatrix}$, with $Y_1, Y_2 \in \mathbb{C}^{1\times g}[s, s^{-1}]$ such that $YR^c = 0$. Define $X \in \mathbb{H}^{1\times g}[s, s^{-1}]$ as $X = Y_1 + Y_2 j$. It is easy to verify that $XR = 0$. Since $X \neq 0$, R is not frr.

The general case follows using the fact that for any $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ there exists a unimodular matrix $U \in \mathbb{H}^{g\times g}[s, s^{-1}]$ such that

$$UR = \begin{bmatrix} \widetilde{R} \\ 0 \end{bmatrix}$$

with $\widetilde{R} \in \mathbb{H}^{\tilde{g}\times r}[s, s^{-1}]$ frr [10, Thm. 2.5.23]. □

Proposition 3.7. *Given two quaternionic (L-) polynomial matrices A and B, if the equation*

$$A^c = MB^c \tag{3.5}$$

holds with a complex (L-) polynomial matrix M, then there exists a quaternionic (L-) polynomial matrix T such that $A = TB$. Moreover, if B is frr then $M = T^c$.

Proof. Let $A = A_1 + A_2 j,\ B = B_1 + B_2 j,$ and $M = \left[\begin{smallmatrix} T_1 & T_2 \\ T_3 & T_4 \end{smallmatrix}\right]$. Then,

$$A^c = MB^c \Leftrightarrow \left\{ \begin{array}{rcl} A_1 & = & T_1 B_1 - T_2 \overline{B}_2 \\ A_2 & = & T_1 B_2 + T_2 \overline{B}_1 \\ -\overline{A}_2 & = & T_3 B_1 - T_4 \overline{B}_2 \\ \overline{A}_1 & = & T_3 B_2 + T_4 \overline{B}_1 \end{array} \right. . \tag{3.6}$$

Let $T = T_1 + T_2 j$. From (3.6) it follows that $A^c = T^c B^c$ and so $A = TB$. Now suppose that B is frr. By Lemma 3.6 we have that B^c is also a frr matrix and, since $(M - T^c)B^c = 0$, we obtain that $M = T^c$. □

Corollary 3.8. *Let $U \in \mathbb{H}^{r\times r}[s, s^{-1}]$. Then U is unimodular if and only if $U^c \in \mathbb{C}^{2r\times 2r}[s, s^{-1}]$ is unimodular.*

Proof. "Only if" part. Let $U \in \mathbb{H}^{r\times r}[s, s^{-1}]$ be unimodular. Then there exists $V \in \mathbb{H}^{r\times r}[s, s^{-1}]$ such that $VU = I_r$, which is equivalent to $V^c U^c = (VU)^c = I_r^c = I_{2r}$, i.e., U^c is unimodular.

"If" part. If U^c is unimodular, there exists $W \in \mathbb{C}^{2r\times 2r}[s, s^{-1}]$ such that $I_{2r} = WU^c$. From Proposition 3.7 we conclude that there exists V such that $V^c = W$ and $VU = I_r$, hence U is unimodular. □

In the sequel we investigate a fundamental equivalence relation for kernel representations.

Definition 3.9. Let $R_l \in \mathbb{H}^{g_l\times r}[s, s^{-1}]$, $l = 1, 2$. Then R_1 and R_2 are said to be *equivalent representations* if $\ker R_1(\sigma, \sigma^{-1}) = \ker R_2(\sigma, \sigma^{-1})$.

Example 3.10. Consider the following quaternionic polynomial matrices

$$R_1 = \begin{bmatrix} s & -i \\ 0 & s-k \end{bmatrix}, \quad R_2 = \begin{bmatrix} s+k & 0 \\ j & 1 \end{bmatrix}. \tag{3.7}$$

These are equivalent representations of the same behavior which, as it is easy to check, is

$$\ker R_1 = \ker R_2 = \left\{ \begin{bmatrix} j \\ 1 \end{bmatrix} k^t q,\ q \in \mathbb{H} \right\}.$$

A straightforward calculation shows that $R_1 = UR_2$, where

$$U = \begin{bmatrix} 1 & -i \\ -j & s-k \end{bmatrix}$$

is an unimodular L-polynomial matrix.

We will show that, as in the real and in the complex case, two representations are equivalent if and only if each one is a left multiple of the other, as in the previous example. This main result is a consequence of the following more general statement.

Theorem 3.11. *Let R_1 and R_2 be two quaternionic (L-) polynomial matrices. Then* $\ker R_1 \subseteq \ker R_2$ *if and only if there exists a quaternionic (L-) polynomial matrix X such that $XR_1 = R_2$.*

Proof. "If" part. Let $w \in \ker R_1$. Then, $R_2 w = XR_1 w = 0$, and therefore $\ker R_1 \subseteq \ker R_2$.

"Only if" part. We want to prove that there exists a matrix X such that $XR_1 = R_2$. By Proposition 3.5,

$$\ker R_1 \subseteq \ker R_2 \Leftrightarrow \ker R_1^c \subseteq \ker R_2^c.$$

As stated in [14, Section 4], there exists a complex matrix Y such that

$$YR_1^c = R_2^c.$$

But, from Proposition 3.7, there also exists a quaternionic matrix X such that

$$XR_1 = R_2,$$

thus proving the theorem. □

Corollary 3.12. *Two quaternionic representations $R_1(s, s^{-1})$, $R_2(s, s^{-1})$ are equivalent if and only if there exist $X_1(s, s^{-1})$ and $X_2(s, s^{-1})$ such that $R_1 = X_1 R_2$ and $R_2 = X_2 R_1$. Moreover, if both matrices are frr then $X_1 = X_2^{-1}$, i.e., X_1 and X_2 are unimodular matrices.*

Proof. The first part of the corollary is a trivial consequence of Theorem 3.11.

Suppose now that R_1 and R_2 are frr. Since $R_1 = X_1 R_2$ and $R_2 = X_2 R_1$, then $R_1 = X_1 X_2 R_1$ and $R_2 = X_2 X_1 R_2$. Hence we have that $X_1 X_2 = X_2 X_1 = I$, i.e., $X_1 = X_2^{-1}$ and X_1 and X_2 are unimodular. □

Remark 3.13. Since s^l is an invertible element in $\mathbb{H}[s, s^{-1}]$, it follows that

$$\ker R(\sigma, \sigma^{-1}) = \ker \sigma^l R(\sigma, \sigma^{-1}).$$

As a consequence, it is always possible to choose a polynomial representation for any behavior $\mathcal{B}$. Indeed, if $\mathcal{B}$ has a representation $R(s, s^{-1})$, then, for an adequate integer $M \geq 0$, $s^M R(s, s^{-1}) \in \mathbb{H}^{g \times r}[s]$ is still a representation of $\mathcal{B}$. Therefore, for the sake of simplicity, we shall choose polynomial kernel representations, although always regarding them as L-polynomial representations.

As in the commutative case, the quaternionic Smith form plays an important role in the study of quaternionic behavioral systems, in particular in the characterization of controllability and stability. Thus, we dedicate the following section to a detailed analysis of this canonical form.

4 Quaternionic Smith Form

The main result of this section is the characterization of the Smith form of complex adjoint matrices and its relation to the quaternionic Smith form. We assume that the reader is already familiar with the Smith form for complex (L-) polynomial matrices.

We start by giving some results about quaternionic polynomials. Results and definitions are trivially generalized to L-polynomials.

A polynomial $d(s)$ is a *left divisor* of $p(s)$, i.e., $d(s)\,|_l\, p(s)$, or $p(s)$ is a *right multiple* of $d(s)$, if there exists a polynomial $q(s)$ such that $p(s) = d(s)q(s)$. The definition of *right divisor* (*left multiple*) is analogous and we will use the notation $d(s)\,|_r\, p(s)$ to indicate that $d(s)$ divides $p(s)$ on the right. A polynomial $d(s)$ is a *divisor* of a polynomial $p(s)$, which we will denote by $d(s)\,|\,p(s)$, and $p(s)$ is a *multiple* of $d(s)$, if $d(s)\,|_l\, p(s)$ and $d(s)\,|_r\, p(s)$.

It can be proved [6] that $\mathbb{H}[s]$ is a principal ideal domain and therefore also left and right division algorithms can be defined.

We say that $d(s)$ is a *total divisor* of $p(s)$ if $[d(s)]\,|\,[p(s)]$, i.e., if for any $d'(s) \sim d(s)$ and $p'(s) \sim p(s)$, $d'(s)\,|\,p'(s)$. The *greatest real factor* of the polynomial p, $\mathrm{grf}(p) \in \mathbb{R}[s]$, which is defined as the monic real factor of p with maximal degree, is always a total divisor of p.

We show that the definition of total divisor is equivalent to similar but simpler conditions.

Lemma 4.1. *Let $p, q \in \mathbb{H}[s]$. Then*

$$[p]\,|\,[q] \quad \Leftrightarrow \quad p\,|\,[q] \quad \Leftrightarrow \quad [p]\,|\,q.$$

Proof. Obviously the total divisor condition is sufficient. We prove that it is also necessary.

Suppose that $p\,|\,[q]$ and let $p' \sim p$ and $q' \sim q$. We shall prove that $p'\,|\,q'$.

By definition we know that there exists $\alpha \in \mathbb{H}$ such that $p' = \alpha p \alpha^{-1}$ and, by hypothesis, there exists $d \in \mathbb{H}[s]$ such that $\alpha^{-1} q' \alpha = pd$. Therefore, if we put $d' = \alpha d \alpha^{-1}$, we get that

$$q' = \alpha p \alpha^{-1} \alpha d \alpha^{-1} = p'd'.$$

For right divisions the proof is similar, thus $[p] \,|\, [q]$.

Analogously it is proved that also $[p] \,|\, q$ is a sufficient condition. $\square$

Unlike the commutative case, evaluation of polynomials is not a ring homomorphism, i.e., we may have $pq(\lambda) \neq p(\lambda)q(\lambda)$, for some $\lambda \in \mathbb{H}$. Consequently, if we define the *zeros* of $p(s)$ as the values $\lambda \in \mathbb{H}$ such that $p(\lambda) = 0$, the relation between factors and zeros of $p(s)$ is not as simple as for real or complex polynomials.

Actually, if $d(s)$ is a right divisor of $p(s)$, then its zeros are also zeros of $p(s)$ [9, Proposition 16.2]. However, if $d(s)$ is a left divisor of $p(s)$, the zeros of $d(s)$ are not necessarily zeros of $p(s)$. Indeed, let $d(s) = s - i$ and $p(s) = d(s)j = js - k$. Then

$$d(i) = 0 \quad \text{but} \quad p(i) = ji - k = -2k \neq 0.$$

The following lemma collects some basic results about zeros of quaternionic polynomials. First we define the *minimal polynomial of the equivalence class* $[\lambda]$ as the real polynomial

$$\Psi_{[\lambda]}(s) = (s - \lambda)(s - \overline{\lambda}) = s^2 - 2(\operatorname{Re}\lambda)s + |\lambda|^2.$$

Lemma 4.2. *Let $p \in \mathbb{H}[s]$. Then*

1. $\Psi_{[\lambda]} = \Psi_{[\lambda']}$ *if and only if* $\lambda \sim \lambda'$ *(i.e., the definition of* $\Psi_{[\lambda]}$ *is well-posed).*
2. *If* $p(\lambda) = p(\nu) = 0$ *with* $\lambda \neq \nu$, $\lambda \sim \nu$, *then* $\Psi_{[\lambda]} \,|\, p$.
3. *If* $p(\lambda) = 0$ *then* $\Psi_{[\lambda]} \,|\, \overline{p}p$.
4. *If* $\Psi_{[\lambda]} \,|\, \overline{p}p$ *then there exists* $\lambda' \in [\lambda]$ *such that* $p(\lambda') = 0$.

Proof. 1. Simply note that, by its definition, $\Psi_{[\lambda]}$ depends only on $\operatorname{Re}\lambda$ and on $|\lambda|^2$ that uniquely characterize the class $[\lambda]$.

2. See [9, Lemma 16.17].

3. As we said before, $\overline{p}p(\lambda) = 0$ since $p(s)$ is a right factor. If $\lambda \in \mathbb{R}$ then by Proposition 2.1 the result follows. If this is not the case, since the polynomial is real by Proposition 3.4, also $\overline{p}p(\overline{\lambda}) = 0$. By point 2., the proof is concluded.

4. If $p(\lambda) = 0$ the statement is proved. Otherwise, if $p(\lambda) \neq 0$, by [9, Proposition 16.3] there exists $\lambda'' \in [\lambda]$ such that $0 = \overline{p}p(\lambda) = \overline{p}(\lambda'')p(\lambda)$. This implies that $\overline{p}(\lambda'') = 0$. By [9, Theorem 16.4] there exists $\lambda' \in [\lambda''] = [\lambda]$ such that $p(\lambda') = 0$. □

Note that in [6] a different definition of total divisor is given, which is precisely the third condition of the following theorem where we state the equivalence of this and other useful conditions. For that purpose it is necessary to first introduce the notion of two-sided ideal.

Definition 4.3. If $\mathcal{I}$ is a subring of $\mathbb{H}[s]$ and $\mathbb{H}[s]\,\mathcal{I} \subseteq \mathcal{I}$ $(\mathcal{I}\,\mathbb{H}[s] \subseteq \mathcal{I})$ then $\mathcal{I}$ is called a *left (right) ideal* of $\mathbb{H}[s]$. If $\mathcal{I}$ is both a left and a right ideal, then $\mathcal{I}$ is said to be a *two-sided ideal.*

Theorem 4.4. *Let $p, q \in \mathbb{H}[s]$. Then the following conditions are equivalent:*

(*i*) $[p] \,|\, [q]$;

(*ii*) $\mathbb{H}[s]q\mathbb{H}[s] \subseteq p\mathbb{H}[s] \cap \mathbb{H}[s]p$;

(*iii*) $\mathbb{H}[s]q \subseteq \mathcal{I} \subseteq \mathbb{H}[s]p$ *for some two-sided ideal* $\mathcal{I}$;

(*iv*) $q = abp$ *with* $bp \in \mathbb{R}[s]$ *and* $a, b \in \mathbb{H}[s]$.

Proof. We will show that the implications $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (i)$ hold true.

$(i) \Rightarrow (ii)$ We first prove that

$$\mathbb{H}q\mathbb{H} \subseteq p\mathbb{H}[s] \cap \mathbb{H}[s]p. \tag{4.1}$$

Indeed, by the hypothesis and by Lemma 4.1, $p \,|\, [q]$ and so, for any nonzero $\alpha \in \mathbb{H}$, there exists $d \in \mathbb{H}[s]$ such that $\alpha q \alpha^{-1} = pd$. Therefore, for any $\beta \in \mathbb{H}$,

$$\alpha q \beta = \alpha q \alpha^{-1} \alpha \beta = p d \alpha \beta \in p\mathbb{H}[s].$$

Analogously, we can prove that $\alpha q \beta \in \mathbb{H}[s]p$. Thus we only need to prove that (4.1) implies condition (*ii*).

Actually, for any $a = \sum \alpha_n s^n, b = \sum \beta_m s^m \in \mathbb{H}[s]$, equation (4.1) implies that there exist polynomials $l_{nm}, r_{nm} \in \mathbb{H}[s]$ such that

$$\alpha_n q \beta_m = l_{nm} p = p r_{nm}.$$

Therefore, recalling that the variable commutes with the coefficients,

$$aqb = \sum \alpha_n q \beta_m s^{n+m} = \sum l_{nm} s^{n+m} p = p \sum r_{nm} s^{n+m},$$

showing that (ii) holds.

$(ii) \Rightarrow (iii)$ The condition is satisfied with $\mathcal{I}$ being the smallest ideal containing $\mathbb{H}[s]q\mathbb{H}[s]$, which can be shown to be contained in $p\mathbb{H}[s] \cap \mathbb{H}[s]p$.

$(iii) \Rightarrow (iv)$ We first show that the monic left and right generators of any two-sided ideal $\mathcal{I}$ are the same.

Suppose that $\mathcal{I} = \mathbb{H}[s]g = g'\mathbb{H}[s]$ with g and g' monic. Then $g = g'h'$ and $g' = hg$ for some $h, h' \in \mathbb{H}[s]$. Thus, $g = hgh'$ which, by examining the degree of the polynomials, implies that h and h' are constant. Since g and g' are monic, $h = h' = 1$ and therefore $g = g'$.

Now we show that $g \in \mathbb{R}[s]$. If $r = \mathrm{grf}(g)$, then $g = rd$ for some $d \in \mathbb{H}[s]$. Suppose by contradiction that $g \notin \mathbb{R}[s]$, i.e., d can be factorized as $d = d'(s - \alpha)$, for some $\alpha \in \mathbb{H}\backslash\mathbb{R}$. Note that $d(\beta) \neq 0$ for any $\beta \in [\alpha]$ such that $\beta \neq \alpha$. Actually, by Lemma 4.2.2, this would imply that the minimal polynomial of $[\alpha]$, $\Psi_{[\alpha]}$, divides d, which is impossible by the definition of r.

Consider now the polynomial $g(s - \alpha') \in \mathcal{I}$. Since $\mathcal{I}$ is a two-sided ideal, there must exist $x \in \mathbb{H}[s]$ such that $xg = g(s - \alpha')$. Therefore,

$$rxd = xrd = xg = g(s - \alpha') = rd(s - \alpha') \;\Rightarrow\; xd = d(s - \alpha').$$

Choose now $\alpha' \in [\alpha]$ such that $\alpha' \neq \alpha$ and $\alpha' \neq \overline{\alpha}$. Since $xd(\alpha) = 0$, the contradiction is achieved if we prove that α cannot be a zero of $d(s - \alpha')$. Indeed, if this were the case, by Lemma 4.2.2, $\Psi_{[\alpha']} \,|\, d(s-\alpha')$ and consequently $(s - \overline{\alpha'}) \,|\, d$ and thus $d(\overline{\alpha'}) = 0$. But, as we said before, d cannot have zeros different from α within its equivalence class.

As $\mathcal{I} \subseteq \mathbb{H}[s]p$, we have that $g = bp \in \mathbb{R}[s]$, for some $b \in \mathbb{H}[s]$. Finally, since $q \in \mathcal{I}$, there exists $a \in \mathbb{H}[s]$ such that $q = ag = abp$.

$(iv) \Rightarrow (i)$ By Lemma 4.1 we just need to prove that $[p] \,|\, q$. By Proposition 3.4 (ii), $bp = pb \in \mathbb{R}[s]$ and thus for any nonzero $\eta \in \mathbb{H}$

$$q = abp = apb = pba = \eta\eta^{-1}pba = \eta pb\eta^{-1}a = \eta p\eta^{-1}\eta b\eta^{-1}a.$$

This means that $[p] \,|_l\, q$. Similarly we can prove that $[p] \,|_r\, q$. □

Due to the equivalences stated in Theorem 4.4, the following result is a consequence of [6, Theorem 3.16].

By $\operatorname{diag}(a_1, \ldots, a_n)$ we mean a (not necessarily square) matrix with suitable dimensions whose first elements on the main diagonal are $a_1, \ldots, a_n$ and all the other entries are zero.

Theorem 4.5. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$. Then there exist L-polynomial unimodular matrices U and V such that*

$$URV = \Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \in \mathbb{H}^{g\times r}[s],$$

where n is the rank of R, γ_l, $l = 1, \ldots, n$, are monic polynomials with nonzero independent term and $[\gamma_l] \mid [\gamma_{l+1}]$, $l = 1, \ldots, n-1$.

The matrix Γ introduced in Theorem 4.5 is called a *quaternionic Smith form* of R.

If $R \in \mathbb{H}^{g\times r}[s]$ we can not guarantee that the polynomials $\gamma_l(s)$ have nonzero independent term.

Note that, unless it is real, the quaternionic Smith form is not unique. The source of nonuniqueness is characterized in [6]. In the end of this section we will show a necessary condition for two matrices to be quaternionic Smith forms of the same quaternionic matrix.

The following example shows that a quaternionic Smith form of a complex matrix does not coincide with its complex Smith form.

Example 4.6. Let $R = \begin{bmatrix} s+i & 0 \\ 0 & s+i \end{bmatrix}$.

In the complex case this polynomial matrix is a Smith form. However, $s+i$ does not divide its own equivalence class. Indeed, $s-i = j(s+i)j^{-1} \sim s+i$ but $s-i = (s+i) - 2i$. Therefore, R is not a quaternionic Smith form. A simple calculation shows that

$$\Gamma = URV = \begin{bmatrix} 1 & 0 \\ 0 & s^2+1 \end{bmatrix},$$

where

$$U = \begin{bmatrix} \frac{k}{2} & \frac{i}{2} \\ js+k & s-i \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} -j & -\frac{k}{2}(s+i) \\ -1 & \frac{i}{2}(s-i) \end{bmatrix}$$

are unimodular polynomial matrices, is the quaternionic Smith form of R.

Before stating the main theorem about quaternionic and complex Smith forms, we give an auxiliary result. The definition of equivalent matrices is analogous to the commutative case.

Proposition 4.7. *For all monic $q \in \mathbb{H}[s]$ there exists $p \in \mathbb{C}[s]$ such that q^c and p^c are equivalent. Furthermore, for all monic $p \in \mathbb{C}[s]$, the complex Smith form of p^c is* $\mathrm{diag}(r, rc\bar{c})$*, where $r = \mathrm{grf}(p)$ and c is such that $p = rc$.*

Remark 4.8. In other words, this result states that for any monic $q \in \mathbb{H}[s]$ there exists $c \in \mathbb{C}[s]$ such that, if $r = \mathrm{grf}(q)$ and $p = rc$, the following matrices are equivalent:

$$q^c, \quad \begin{bmatrix} p & 0 \\ 0 & \overline{p} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} r & 0 \\ 0 & rc\overline{c} \end{bmatrix}.$$

Proof. First we show that for any monic $p \in \mathbb{C}[s]$, the complex Smith form of p^c is $\mathrm{diag}(r, rc\bar{c})$, where $r = \mathrm{grf}(p)$ and c is such that $p = rc$. By hypothesis, the polynomials c and $\bar{c}$ are coprime and therefore there exist $x, y \in \mathbb{C}[s]$ such that $xc + y\bar{c} = 1$. The complex polynomial matrices

$$U = \begin{bmatrix} c & -y \\ \bar{c} & x \end{bmatrix} \quad \text{and} \quad V = \begin{bmatrix} 1 - \bar{c}y & \bar{c}y \\ -1 & 1 \end{bmatrix}$$

are unimodular and $p^c = U\,\mathrm{diag}(r, rc\bar{c})V$. So $\mathrm{diag}(r, rc\bar{c})$ is the complex Smith form of p^c.

As for the first fact, let now $r = \mathrm{grf}(q)$ and d be such that $q = rd = r(d_1 + d_2 j)$, for some $d_1, d_2 \in \mathbb{C}[s]$. By definition of r, $gcd(d_1, d_2, \overline{d}_1, \overline{d}_2) = 1$. Therefore, the complex Smith form of q^c is $\Gamma = \mathrm{diag}(r, x)$, where

$$rx = \det(q^c) = r^2(d_1\overline{d}_1 + d_2\overline{d}_2). \tag{4.2}$$

By direct calculation, $d\overline{d} = d_1\overline{d_1} + d_2\overline{d_2} \in \mathbb{R}[s]$ and hence this polynomial has no real zeros, because these should be common to d_1 and d_2, which is impossible by the definition of r. Therefore, there exists $c \in \mathbb{C}[s]$ such that $d\overline{d} = c\bar{c}$, i.e., $q\bar{q} = p\bar{p}$ with $p = rc \in \mathbb{C}[s]$. This shows, by (4.2), that $x = rc\bar{c}$, which proves that q^c and $\mathrm{diag}(r, rc\bar{c})$ are equivalent. By the first part of the proof, the result follows. □

Remark 4.9. Note that Proposition 4.7 implies that for any monic $q \in \mathbb{H}[s]$ there always exists a $p \in \mathbb{C}[s]$ such that $p\bar{p} = q\bar{q}$ and $\mathrm{grf}(p) = \mathrm{grf}(q)$. Moreover, if $p \in \mathbb{R}[s]$, then also $q \in \mathbb{R}[s]$ and $q = p$ (cfr. Proposition 2.1).

The following theorem characterizes the complex Smith form of polynomial complex adjoint matrices and gives its relation with their quaternionic Smith forms. The result is trivially generalized to (L-) polynomial matrices.

Theorem 4.10. *1. A polynomial matrix*

$$\Delta = \operatorname{diag}(\delta_1, \delta_1', \ldots, \delta_n, \delta_n') \in \mathbb{C}^{2g\times 2r}[s],$$

is the complex Smith form of the complex adjoint matrix R^c, for some $R \in \mathbb{H}^{g\times r}[s]$, if and only if Δ is a real matrix, $\delta_1|\delta_1'|\cdots|\delta_n|\delta_n'$ and δ_l, δ_l' are monic polynomials with exactly the same real zeros, $l = 1, \ldots, n$.

2. If $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_m) \in \mathbb{H}^{g\times r}[s]$ is a quaternionic Smith form of R, then $m = n$, $\delta_l = \operatorname{grf}(\gamma_l)$ and

$$\gamma_l\overline{\gamma}_l = \delta_l\delta_l'.$$

Proof. 1. "If" part. It follows from the hypothesis that there exist complex polynomials c_l, with no real zeros, such that $\delta_l' = \delta_l c_l \overline{c}_l$. Therefore, since $\delta_l = \operatorname{grf}(\delta_l c_l)$, $\operatorname{diag}(\delta_l, \delta_l') = \operatorname{diag}(\delta_l, \delta_l c_l \overline{c}_l)$ is equivalent to $\operatorname{diag}(\delta_l c_l, \overline{\delta_l c_l})$ by Proposition 4.7. Hence, Δ is equivalent to

$$\operatorname{diag}(\delta_1 c_1, \overline{\delta_1 c_1}, \ldots, \delta_n c_n, \overline{\delta_n c_n}),$$

which, in turn, is equivalent to the complex adjoint matrix of

$$R = \operatorname{diag}(\delta_1 c_1, \ldots, \delta_n c_n) \in \mathbb{H}^{g\times r}[s].$$

"Only if" part. Let Δ be the complex Smith form of R^c. Suppose that $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_n) \in \mathbb{H}^{g\times r}[s]$ is a quaternionic Smith form of R. By Lemma 3.6 it is clear that $m = n$. Let $\gamma_l = r_l d_l$, where $r_l = \operatorname{grf}(\gamma_l)$. By Proposition 4.7, there exists $c_l \in \mathbb{C}[s]$ with no real zeros such that γ_l^c is equivalent to $\operatorname{diag}(r_l, r_l c_l \overline{c}_l)$ and consequently, Γ^c is equivalent to

$$\Delta' = \operatorname{diag}(r_1, r_1 c_1 \overline{c}_1, \ldots, r_n, r_n c_n \overline{c}_n). \tag{4.3}$$

Next we show that Δ' is the complex Smith form of R^c, and hence $\Delta = \Delta'$. Since Δ' is equivalent to R^c, we only need to show that it satisfies the required division properties. Obviously, $r_l \,|\, r_l c_l \overline{c}_l,\ l = 1, \ldots, n$.

We will prove that $r_l c_l \overline{c}_l \,|\, r_{l+1}$. By Theorem 4.4 we know that

$$\gamma_{l+1} = ab\gamma_l,\ b\gamma_l \in \mathbb{R}[s],\ a, b \in \mathbb{H}[s]. \tag{4.4}$$

The fact that $\gamma_l = r_l d_l$ divides $b\gamma_l \in \mathbb{R}[s]$ implies that also the least real multiple of γ_l, i.e., $r_l d_l \overline{d}_l$, is a factor of $b\gamma_l$, and hence, by (4.4), a factor of γ_{l+1}. Note that $a \,|\, b \Rightarrow \operatorname{grf}(a) \,|\, \operatorname{grf}(b)$ and therefore we have that $r_l d_l \overline{d}_l \,|\, \operatorname{grf}(\gamma_{l+1}) = r_{l+1}$. However, by the reasoning in the proof of Proposition 4.7, we know that

$$r_l^2 d_l \overline{d}_l = \gamma_l \overline{\gamma}_l = r_l^2 c_l \overline{c}_l, \tag{4.5}$$

and thus $r_l c_l \bar{c}_l = r_l d_l \bar{d}_l \,|\, r_{l+1}$. Therefore, $\Delta = \Delta'$, i.e., $\delta_l = r_l$ and $\delta'_l = r_l c_l \bar{c}_l$, $l = 1, \ldots, n$, and consequently $\delta_1 | \delta'_1 | \cdots | \delta_n | \delta'_n$. It is obvious that Δ is a real matrix. Moreover, since the polynomials c_l have no real zeros, we have that δ_l and δ'_l do have the same real zeros.

2. In the previous point we have seen that $m = n$, and $\delta_l = r_l = \mathrm{grf}(\gamma_l)$. Finally, note that by equation (4.5), $\delta_l \delta'_l = \gamma_l \overline{\gamma}_l$. □

Remark 4.11. Since the complex Smith form is unique, it follows from Theorem 4.10 that if

$$\Gamma = \mathrm{diag}(\gamma_1, \ldots, \gamma_m) \quad \text{and} \quad \Gamma' = \mathrm{diag}(\gamma'_1, \ldots, \gamma'_m)$$

are quaternionic Smith forms of a quaternionic matrix R, then $\gamma_l \overline{\gamma}_l = \gamma'_l \overline{\gamma'}_l$, $l = 1, \ldots, m$.

However, the reciprocal fact is not true. For instance, let $r = s^2 + 1 \in \mathbb{H}[s]$, $\gamma = r$ and $\gamma' = (s + i)(s + j)$. It is easily checked that $\gamma\overline{\gamma} = (s^2 + 1)^2 = \gamma'\overline{\gamma'}$. Obviously, γ is a quaternionic Smith form of r, but, as γ' is not equivalent to γ, γ' is not a quaternionic Smith form of r.

5 Controllability

In this section we recall the concept of controllability, which plays a fundamental role in systems theory. Roughly speaking, we call a behavior controllable if it is possible to switch from one trajectory to any other trajectory within the behavior in finite time.

Definition 5.1. [10, Def. 5.2.2] A behavior $\mathcal{B}$ of a time-invariant dynamical system is called *controllable* if for any two trajectories $w_1, w_2 \in \mathcal{B}$, and any time instant t_1, there exists $t_2 > t_1$ and a trajectory $w \in \mathcal{B}$ such that

$$w(t) = \begin{cases} w_1(t), & t \leq t_1; \\ w_2(t), & t \geq t_2. \end{cases} \tag{5.1}$$

When property (5.1) holds we say that w_1 and w_2 are *concatenable* in $\mathcal{B}$. Therefore $\mathcal{B}$ is controllable if all its trajectories are concatenable in $\mathcal{B}$.

Lemma 5.2. *Let $R \in \mathbb{H}^{g \times r}[s, s^{-1}]$ and $\mathcal{B} = \ker R$. Then $\mathcal{B}$ is controllable if and only if $\mathcal{B}^{\mathbb{C}}$ is controllable.*

Proof. This result follows immediately from the definition of controllability and from the isomorphism between $\mathcal{B}$ and $\mathcal{B}^{\mathbb{C}}$. □

The following theorem gives characterizations of controllability system. We recall that a matrix is left-prime if it admits only unimodular left factors.

Theorem 5.3. *Let $R \in \mathbb{H}^{g\times r}[s, s^{-1}]$ be frr and $\mathcal{B} = \ker R$. Then the following conditions are equivalent:*

- *(i) $\mathcal{B}$ is controllable;*
- *(ii) R is left prime;*
- *(iii) the quaternionic Smith form of R is $\begin{bmatrix} I & 0 \end{bmatrix}$;*
- *(iv) there exists an image representation, i.e., $\exists M \in \mathbb{H}^{r\times m}[s, s^{-1}]$ such that $\mathcal{B} = \operatorname{Im} M$.*

Proof. The proof of the equivalence between (i), (iii) and (iv) is analogous to the one given for [15, Proposition 4.3]. The equivalence between (ii) and (iii) is proved similarly to the commutative case, see for instance [11, 4.1.19]. $\square$

In the commutative case, Theorem 5.3 holds as well. Moreover, $\mathcal{B} = \ker R$ is controllable if and only if $\operatorname{rank} R(\lambda, \lambda^{-1})$ is constant for all $0 \neq \lambda \in \mathbb{C}$. However, in the quaternionic case, this is not true. In the following example it is shown that if U is unimodular, then $U(\lambda, \lambda^{-1})$ is not necessarily invertible for all $0 \neq \lambda \in \mathbb{H}$. Clearly, in this case $\ker U(\sigma, \sigma^{-1}) = \{0\}$ is controllable, but $\operatorname{rank} U(\lambda, \lambda^{-1})$ is not constant for all $\lambda \in \mathbb{H}\backslash\{0\}$.

Example 5.4. Let

$$U(s, s^{-1}) = \begin{bmatrix} -is + k & js \\ -i & j \end{bmatrix} \quad \text{and} \quad V(s, s^{-1}) = \begin{bmatrix} -k & ks \\ 1 & -s - j \end{bmatrix}.$$

Since $UV = I$, U and V are unimodular matrices. However, $U(\lambda, \lambda^{-1})$ is not invertible when $\lambda = \frac{1}{2}j$. Indeed,

$$U\left(\tfrac{1}{2}j, \left(\tfrac{1}{2}j\right)^{-1}\right) \begin{bmatrix} 1 \\ k \end{bmatrix} = \begin{bmatrix} \frac{1}{2}k & -\frac{1}{2} \\ -i & j \end{bmatrix} \begin{bmatrix} 1 \\ k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Remark 5.5. Lemma 5.2 says that it is possible to check whether a quaternionic behavior is controllable by analysing its complex form. However, in general, this method, besides increasing the size of the matrices involved and hence the computational complexity, may also transform the problem into a less intuitive one. For instance, let $R = \begin{bmatrix} s + j & i(s + j) \end{bmatrix}$ and $\mathcal{B} = \ker R$. It is possible to conclude immediately that $\mathcal{B}$ is not controllable since $\begin{bmatrix} s + j & 0 \end{bmatrix}$ is obviously a quaternionic Smith form of R. On the other hand, looking at the corresponding complex adjoint matrix $R^c = \left[\begin{smallmatrix} s & is & 1 & i \\ -1 & i & s & -is \end{smallmatrix}\right]$ it is not so evident whether $\mathcal{B}^{\mathbb{C}}$ is controllable or not.

6 Stability

In this section we extend the characterization of stability to quaternionic behaviors. First we give the concept of behavioral stability.

Definition 6.1. [10, Def. 7.2.1] A dynamical system with behavior $\mathcal{B}$ is called *(asymptotically) stable* if for every trajectory $w \in \mathcal{B}$, $\lim_{t \to +\infty} w(t) = 0$.

Note that $\mathcal{B}$ is stable if and only if $\mathcal{B}^{\mathbb{C}}$ is stable. The characterization of stability for a complex behavior $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ is given by the next result, which is the discrete version of [10, Thm. 7.2.2].

Theorem 6.2. *Let $\mathcal{B} \subseteq (\mathbb{C}^r)^{\mathbb{Z}}$ be a behavior given as $\mathcal{B} = \ker R(\sigma, \sigma^{-1})$, with $R \in \mathbb{C}^{g \times r}[s, s^{-1}]$. Then $\mathcal{B}$ is stable if and only if* $\operatorname{rank} R(\lambda, \lambda^{-1}) = r$, $\forall \lambda \in \mathbb{C}$ *such that* $|\lambda| \geq 1$.

For quaternionic behaviors the following result holds.

Theorem 6.3. *Let $\mathcal{B} \subseteq (\mathbb{H}^r)^{\mathbb{Z}}$ be a behavior given as $\mathcal{B} = \ker R(\sigma, \sigma^{-1})$, with $R \in \mathbb{H}^{g \times r}[s, s^{-1}]$ and let $\Gamma = \operatorname{diag}(\gamma_1, \ldots, \gamma_r)$ be a quaternionic Smith form of R. Then*

$$\mathcal{B} \text{ is stable } \Leftrightarrow \gamma_r(\lambda) = 0 \Rightarrow |\lambda| < 1,\ \lambda \in \mathbb{H}.$$

Proof. As we mentioned, $\mathcal{B}$ is stable if and only if $\mathcal{B}^{\mathbb{C}}$ is stable. Let $\Delta = \operatorname{diag}(\delta_1, \ldots, \delta_{2r}) \in \mathbb{R}^{2g \times 2r}[s]$ be the complex Smith form of R^c. Since R^c and Δ are equivalent, by Theorem 6.2 we have that $\mathcal{B}^{\mathbb{C}}$ is stable if and only if $\delta_{2r}(\mu) = 0$ implies $|\mu| < 1,\ \mu \in \mathbb{C}$. Moreover, this is equivalent to $\delta_{2r}(\lambda') = 0 \Rightarrow |\lambda'| < 1,\ \lambda' \in \mathbb{H}$. Indeed, let $\lambda' \in \mathbb{H}$ be such that $\delta_{2r}(\lambda') = 0$ and let $\mu \in \mathbb{C}$, $\mu \in [\lambda']$. Since $\delta_{2r} \in \mathbb{R}[s]$, we have that also $\delta_{2r}(\mu) = 0$. Thus $|\mu| < 1$. But if $\mu \in [\lambda']$, $|\lambda'| = |\mu|$ and hence we conclude that $|\lambda'| < 1$. The reciprocal implication is obvious. Therefore we just need to show that

$$\delta_{2r}(\lambda') = 0 \Rightarrow |\lambda'| < 1,\ \lambda' \in \mathbb{H} \ \Leftrightarrow\ \gamma_r(\lambda) = 0 \Rightarrow |\lambda| < 1,\ \lambda \in \mathbb{H}.$$

Recall that by Theorem 4.10 he have

$$\gamma_r \overline{\gamma}_r = \delta_{2r} \delta_{2r-1}. \tag{6.1}$$

"$\Rightarrow$" Let $\lambda \in \mathbb{H}$ be such that $\gamma_r(\lambda) = 0$. By Lemma 4.2.3 we have that $\gamma_r \overline{\gamma}_r(\lambda) = 0$ which by (6.1) implies $\delta_{2r}\delta_{2r-1}(\lambda) = 0$. As $\delta_r \in \mathbb{R}[s]$, then $\delta_{2r}\delta_{2r-1}(\lambda) = \delta_{2r}(\lambda)\delta_{2r-1}(\lambda)$. Therefore $\delta_{2r}(\lambda) = 0 \vee \delta_{2r-1}(\lambda) = 0$, which, since $\delta_{2r-1} \mid \delta_{2r}$, is equivalent to have $\delta_{2r}(\lambda) = 0$, and by hypothesis $|\lambda| < 1$.

"$\Leftarrow$" Let $\lambda' \in \mathbb{H}$ be such that $\delta_{2r}(\lambda') = 0$. This implies that $\delta_{2r}\delta_{2r-1}(\lambda') = 0$ and by (6.1) we have that $\gamma_r\overline{\gamma}_r(\lambda') = 0$. By Lemma 4.2.4 there exists $\mu \in [\lambda']$ such that $\gamma_r(\mu) = 0$, and since $|\lambda'| = |\mu|$, by hypothesis $|\lambda'| < 1$. □

7 Conclusion

In this paper usual concepts of commutative linear algebra were extended to quaternionic algebra in the context of the study of quaternionic behavioral systems. In particular, a relation was obtained between the quaternionic Smith form of a quaternionic matrix and the complex Smith form of its complex adjoint matrix. These results were used to characterize system theoretic properties such as controllability and stability of quaternionic systems in the behavioral approach.

Acknowledgement

The present research was partially supported by *Unidade de Investigação Matemática e Aplicações*, University of Aveiro, Portugal, through *Programa Operacional "Ciência, Tecnologia, Inovação"* (POCTI) of the *Fundação para a Ciência e Tecnologia* (FCT), co-financed by the European Community fund FEDER.

References

[1] S. P. Brumby and G. C. Joshi. Experimental status of quaternionic quantum mechanics. *Chaos Solitons Fractals*, 7(5):747–752, 1996.

[2] S. De Leo and G. Ducati. Quaternionic differential operators. *J. Math. Phys.*, 42(5):2236–2265, 2001.

[3] A. El-Gohary and E. R. Elazab. Exponential control of a rotational motion of a rigid body using quaternions. *Appl. Math. Comput.*, 137(2-3):195–207, 2003.

[4] M. Hazewinkel, J. Lewis, and C. Martin. Symmetric systems with semisimple structure algebra: the quaternionic case. *Systems Control Lett.*, 3(3):151–154, 1983.

[5] D. Hestenes. *New foundations for classical mechanics*. Kluwer Academic Publishers Group, Dordrecht, The Netherlands, 1986.

[6] N. Jacobson. *The Theory of Rings*. American Mathematical Society, New York, 1943.

[7] B. Jancewicz. *Multivectors and Clifford algebra in electrodynamics*. World Scientific Publishing Co. Inc., Teaneck, NJ, 1988.

[8] J. B. Kuipers. *Quaternions and rotation sequences*. Princeton University Press, Princeton, NJ, 1999.

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

[10] J. W. Polderman and J. C. Willems. *Introduction to Mathematical Systems Theory: A Behavioral Approach*, volume 26 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 1997.

[11] M. Vidyasagar. *Control system synthesis*. MIT Press, Cambridge, MA, 1985.

[12] P. B. Visscher and X. Feng. Quaternion-based algorithm for micromagnetics. *Phys. Rev. B*, 65(104412), 2002.

[13] J. Vrbik. A novel solution to Kepler's problem. *European J. Phys.*, 24(6):575–583, 2003.

[14] J. C. Willems. From time series to linear system — part I. finite dimensional linear time invariant systems. *Automatica—J. IFAC*, 22(5):561–580, 1986.

[15] J. C. Willems. Models for dynamics. In U. Kirchgraber and H. O. Walther, editors, *Dynamics Reported*, volume 2, pages 171–269. John Wiley & Sons Ltd., Chichester, 1989.

[16] J. C. Willems. Paradigms and puzzles in the theory of dynamical systems. *IEEE Trans. Automat. Control*, 36(3):259–294, Mar. 1991.

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