ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Research

Cite this article: Xu D, Jahanchahi C, Took CC, Mandic DP. 2015 Enabling quaternion derivatives: the generalized HR calculus. *R. Soc. open sci.* **2**: 150255. http://dx.doi.org/10.1098/rsos.150255

Received: 16 June 2015
Accepted: 28 July 2015

Subject Category:
Engineering

Subject Areas:
electrical engineering/systems theory/mathematical modelling

Keywords:
generalized HR calculus, non-analytic quaternion function, nonlinear quaternion functions, quaternion derivatives, quaternion least mean square

Author for correspondence:
Danilo P. Mandic
e-mail: d.mandic@imperial.ac.uk

THE ROYAL SOCIETY PUBLISHING

Enabling quaternion derivatives: the generalized HR calculus

Dongpo Xu[1], Cyrus Jahanchahi[2], Clive C. Took[3] and Danilo P. Mandic[2]

[1]School of Mathematics and Statistics, Northeast Normal University, Changchun 130024, People's Republic of China
[2]Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK
[3]Department of Computing, University of Surrey, Guildford GU2 7XH, UK

Quaternion derivatives exist only for a very restricted class of analytic (regular) functions; however, in many applications, functions of interest are real-valued and hence not analytic, a typical case being the standard real mean square error objective function. The recent HR calculus is a step forward and provides a way to calculate derivatives and gradients of both analytic and non-analytic functions of quaternion variables; however, the HR calculus can become cumbersome in complex optimization problems due to the lack of rigorous product and chain rules, a consequence of the non-commutativity of quaternion algebra. To address this issue, we introduce the generalized HR (GHR) derivatives which employ quaternion rotations in a general orthogonal system and provide the left- and right-hand versions of the quaternion derivative of general functions. The GHR calculus also solves the long-standing problems of product and chain rules, mean-value theorem and Taylor's theorem in the quaternion field. At the core of the proposed GHR calculus is quaternion rotation, which makes it possible to extend the principle to other functional calculi in non-commutative settings. Examples in statistical learning theory and adaptive signal processing support the analysis.

1. Introduction

Quaternions have become a standard tool in many modern areas, including image processing [1,2], aerospace and satellite tracking [3], modelling of wind profile in renewable energy [4] and in the processing of polarized waves [5,6]. Compared to real vector algebra, quaternion algebra [7] has been shown to both reduce the number of parameters in the modelling and offer advantages in terms of functional simplicity and accuracy [8,9]. The most common optimization approach in applications is based on the gradient of the objective function; one such algorithm

© 2015 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

is the quaternion least mean square (QLMS) [4]. The objective functions in practical applications are typically based on the mean square error (MSE), a real function of quaternion variables, and are thus not analytic according to standard quaternion analysis [10–12]. This is a major obstacle to a more widespread use of quaternions in learning systems.

The existing ways to find the derivative of a real function $f(q)$ with respect to the unknown quaternion variable q are:

— *The pseudo-derivative*, which considers f as a function of the four real components q_a, q_b, q_c and q_d of the quaternion variable q, and then takes componentwise real derivatives with respect to the real variables q_a, q_b, q_c and q_d. In other words, the real-valued function $f : \mathbb{H} \to \mathbb{R}$ is treated as a real differentiable mapping between $\mathbb{R}^4$ and $\mathbb{R}$. This leads to unnecessarily long expressions and is especially cumbersome and tedious in complex optimization problems and when dealing with nonlinear functions.

— *The HR calculus*, which is compact and elegant [13], as it finds formal derivatives of f with respect to the quaternion variables and their involutions in a direct way. This applies to both real functions of quaternion variables and nonlinear functions. This approach is based on the differentials of q, q^i, q^j, q^k, which are independent in the quaternion field, as shown in lemmas A.1 and A.2. The advantage of using HR derivatives is that the computations and analysis are kept in the quaternion domain rather than using quaternion-to-real transformations, and many algorithms can be readily extended from the complex to the quaternion domain.

Although the HR calculus is a significant step forward, the product and chain rules are not defined within the HR calculus, which complicates the calculation of derivatives of, for example, nonlinear quaternion functions. Other functional calculi [10–12,14] in quaternion analysis similarly suffer from this obstacle.

The aim of this work is to revisit the HR calculus [13] and to equip it with the product rule and chain rule in order to solve these long-standing problems in quaternion calculus. Motivated by the complex CR calculus [15–17], we first consider a general orthogonal system which, in conjunction with the HR calculus, introduces the generalized HR (GHR) calculus. The GHR calculus comprises both the left- and right-hand versions of quaternion derivative; these are necessary to consider due to non-commutativity of quaternion product. In particular, we show that for real functions of quaternion variables, such as the standard MSE objective function, the left and right GHR derivatives are identical. An important consequence of this property is that within the GHR calculus, the choice of the left/right GHR derivative is irrelevant for practical applications of quaternion optimization; this is currently a major source of confusion in the quaternion community. Another consequence of the novel product rule is that it not only enables the calculation of the GHR derivatives for general functions of quaternion variables, but also it is generic—if one function within the product is real-valued, this novel product rule degenerates into the traditional product rule, as shown in corollary 4.11. A family of chain rules is also introduced in order to calculate the derivatives of nonlinear functions of quaternion variables, which include complex- and real-valued functions as degenerate quaternion functions. Since at the core of the GHR calculus is the quaternion rotation, this approach can be naturally extended to other functional calculi in non-commutative settings. Finally, we revisit two fundamental theorems in quaternion calculus—the quaternion mean value theorem and quaternion Taylor's theorem—and derive them in a compact and generic form, based on the GHR derivatives. The GHR calculus therefore provides a solution to some long-standing mathematical problems [18] and promises a tool for a more widespread use of quaternions in practical applications. Illustrative examples in statistical signal processing support the analysis.

2. Background on quaternions

2.1. Quaternion algebra

Quaternions are an associative but not commutative algebra over $\mathbb{R}$, defined as

$$\mathbb{H} = \operatorname{span}\{1, i, j, k\} \triangleq \{q_a + iq_b + jq_c + kq_d \mid q_a, q_b, q_c, q_d \in \mathbb{R}\}, \tag{2.1}$$

where $\{1, i, j, k\}$ is a basis of $\mathbb{H}$, and the imaginary units i, j and k satisfy $i^2 = j^2 = k^2 = ijk = -1$, which implies $ij = k = -ji$, $jk = i = -kj$, $ki = j = -ik$. For any quaternion $q = q_a + iq_b + jq_c + kq_d = S_q + V_q$, the scalar (real) part is denoted by $q_a = S_q = \Re(q)$, whereas the vector part $V_q = \Im(q) = iq_b + jq_c + kq_d$ spans the three imaginary parts. For $p, q \in \mathbb{H}$, the quaternion product is given by $pq = S_pS_q - V_p \cdot V_q + S_pV_q +$

$S_qV_p + V_p \times V_q$, where the symbols '$\cdot$' and '$\times$' denote, respectively, the standard inner product and vector product. The presence of vector product makes the quaternion product non-commutative, i.e. in general for $p, q \in \mathbb{H}$, $pq \neq qp$. The conjugate of a quaternion q is defined as $q^* = S_q - V_q$, while the conjugate of a quaternion product satisfies $(pq)^* = q^*p^*$. The modulus of a quaternion is defined as $|q| = \sqrt{qq^*}$, and thus $|pq| = |p||q|$. The inverse of a quaternion $q \neq 0$ is $q^{-1} = q^*/|q|^2$ which yields an important consequence $(pq)^{-1} = q^{-1}p^{-1}$. If $|q| = 1$, we call q a *unit* quaternion. A quaternion q is said to be *pure* if $\Re(q) = 0$. For pure quaternions, $q^* = -q$ and $q^2 = -|q|^2$. Thus, a *pure unit* quaternion is a square root of -1; examples are the imaginary units i, j and k.

Quaternions can also be written in the *polar* form $q = |q|(\cos\theta + \hat{q}\sin\theta)$, where $\hat{q} = Vq/|Vq|$ is a pure unit quaternion and $\theta = \arccos(\Re(q)/|q|) \in \mathbb{R}$ is the angle (or argument) of the quaternion. We shall next introduce the quaternion rotation and involution operators.

Definition 2.1 (quaternion rotation [19]). For any quaternion q, the transformation

$$q^\mu \triangleq \mu q \mu^{-1} \tag{2.2}$$

geometrically describes a three-dimensional rotation of the vector part of q by an angle 2θ about the vector part of μ, where $\mu = |\mu|(\cos\theta + \hat{\mu}\sin\theta)$ is any non-zero quaternion.

Properties of the quaternion rotation (see [6,20] used in this work) are

$$(pq)^\mu = p^\mu q^\mu, \quad pq = q^p p = qp^{(q^*)}, \quad \forall\, p, q \in \mathbb{H} \tag{2.3}$$

and

$$q^{\mu\nu} = (q^\nu)^\mu, \quad q^{\mu *} \triangleq (q^*)^\mu = (q^\mu)^* \triangleq q^{*\mu}, \quad \forall\, \nu, \mu \in \mathbb{H}. \tag{2.4}$$

Note that the representation in (2.1) can be generalized to a general orthogonal basis $\{1, i^\mu, j^\mu, k^\mu\}$, where the following properties hold [19]:

$$i^\mu i^\mu = j^\mu j^\mu = k^\mu k^\mu = i^\mu j^\mu k^\mu = -1. \tag{2.5}$$

Definition 2.2 (quaternion involution [21]). The involution of a quaternion q around a pure unit quaternion η is given by

$$q^\eta = \eta q \eta^{-1} = \eta q \eta^* = -\eta q \eta$$

and represents a rotation of q about η by π.

Of particular interest to this work are quaternion involutions around the imaginary units i, j, k, given by [21]

$$\left.\begin{aligned} q = q_a + iq_b + jq_c + kq_d, \quad q^i = -iqi = q_a + iq_b - jq_c - kq_d \\ q^j = -jqj = q_a - iq_b + jq_c - kq_d, \quad q^k = -kqk = q_a - iq_b - jq_c + kq_d, \end{aligned}\right\} \tag{2.6}$$

which allows us to express the four real-valued components of a quaternion q as [13,21]

$$q_a = \frac{1}{4}(q + q^i + q^j + q^k), \quad q_b = \frac{1}{4i}(q + q^i - q^j - q^k), \tag{2.7}$$

$$q_c = \frac{1}{4j}(q - q^i + q^j - q^k), \quad q_d = \frac{1}{4k}(q - q^i - q^j + q^k). \tag{2.8}$$

This is analogous to the complex case, where $x = \frac{1}{2}(z + z^*)$ and $y = -(i/2)(z - z^*)$ for any $z = x + iy \in \mathbb{C}$ [22]. Note that the quaternion conjugation operator $(\cdot)^*$ is also an involution and can be written in terms of q, q^i, q^j and q^k as

$$q^* = \tfrac{1}{2}(-q + q^i + q^j + q^k). \tag{2.9}$$

2.2. Analytic functions in $\mathbb{H}$

To arrive at the notion of analytic (regular, monogenic) function in $\mathbb{H}$, recall that due to the non-commutativity of quaternion product, there are two ways to express the quotient in the definition of quaternion derivative, as shown below.

Proposition 2.3 (Sudbery [10]). *Let $D \subseteq \mathbb{H}$ be a simply connected domain of definition of the function $f : D \to \mathbb{H}$. If for any $q \in D$*

$$\lim_{h\to 0}[(f(q+h) - f(q))h^{-1}] \tag{2.10}$$

exists in $\mathbb{H}$, then necessarily $f(q) = \omega q + \lambda$ for some $\omega, \lambda \in \mathbb{H}$. If for any $q \in D$

$$\lim_{h\to 0}[h^{-1}(f(q+h) - f(q))] \tag{2.11}$$

exists in $\mathbb{H}$, then necessarily $f(q) = q\nu + \lambda$ for some $\nu, \lambda \in \mathbb{H}$.

Proposition 2.3 is discussed in [10,23] and indicates that the traditional definitions of derivative in (2.10) and (2.11) are too restrictive and apply only to linear functions of quaternions. One attempt to relax this constraint is due to Fueter [24,25], whose analyticity condition is termed the Cauchy–Riemann–Fueter (CRF) equation, and is given by [10,11]

$$\left.\begin{aligned} \text{Left CRF}: &\quad \frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k = 0 \\ \text{Right CRF}: &\quad \frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d} = 0. \end{aligned}\right\} \tag{2.12}$$

The limitations of the CRF condition were pointed out by Gentili & Struppa in [14,26], who showed that general polynomial functions (even the identity $f(q) = q$) satisfy neither the left CRF nor the right CRF. To further relax the analyticity condition, a *local* analyticity condition (LAC) was proposed in [12], by using the polar form of a quaternion, to give

$$\left.\begin{aligned} \text{Left LAC}: &\quad \frac{\partial f}{\partial q_a} + \left(q_b \frac{\partial f}{\partial q_a} + q_c \frac{\partial f}{\partial q_c} + q_d \frac{\partial f}{\partial q_d}\right) \frac{V_q}{|V_q|^2} = 0 \\ \text{Right LAC}: &\quad \frac{\partial f}{\partial q_a} + \frac{V_q}{|V_q|^2}\left(q_b \frac{\partial f}{\partial q_a} + q_c \frac{\partial f}{\partial q_c} + q_d \frac{\partial f}{\partial q_d}\right) = 0, \end{aligned}\right\} \tag{2.13}$$

where $q = q_a + V_q$ and $V_q = iq_b + jq_c + kq_d$. The theory of local analyticity is now well developed, and we refer the reader to [14,26,27] for the slice regular functions and to [28] for applications. More recent work in this area includes [29,30], and references therein. The advantage of the LAC is that both the polynomial functions of q and some elementary functions satisfy either the left LAC or the right LAC. However, in general, the products and compositions of two LAC functions f and g no longer meet the local analytic condition. For example, if $f(q) = q$ and $g(q) = \omega q$, $\omega \in \mathbb{H}$, then f and g satisfy the left LAC, but the product $fg = q\omega q$ does not satisfy the left LAC. The same applies for the right LAC, when the function g is written as $g(q) = q\omega$.

Quaternion derivative is defined only for analytic functions; however, in optimization it is often required that the objective function to be minimized or maximized is real-valued. A typical example is the mean square type objective function given by

$$J(\mathbf{w}) = \|f(\mathbf{w})\|^2. \tag{2.14}$$

Note that according to the definition of analytic (regular) function given in [10–12,14,26,27], the function J is not analytic. In order to take its derivatives (but not limited to only such real quadratic functions), we can, however, use the HR calculus [13] which extends the complex CR calculus [15,17,31] to the quaternion field. This generalization is not trivial, and the rules of the CR and HR calculus are different; for more detail see §3.

Remark 2.4. It is important to note that the left (right) terminology introduced in (2.12), (2.13) and in the sequel differs from that in [10–12,14,27]. We adopt the use of the left (right) terminology based on the position of $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$, rather than on the positions of imaginary units i, j, k. Although this is only a notational difference, we later show that the left derivatives (using this convenience) in definitions 3.2 and 4.1 are in this way equipped with a left constant rule (4.3), that is, the *left* constant comes out from the *left* derivative of product, and the left derivatives stand on the left side of the quaternion differential in (A 12) and (A 16). This also allows for a consistent, intuitive and physically meaningful use of terminology.

3. The HR calculus

Optimization problems involving quaternions arise in a number of applications in control theory, signal processing, robotics and biomechanics. Solutions often require a first- or second-order approximation of the objective function; however, real functions of quaternion variables are essentially non-analytic. The recently proposed HR calculus [13] solves these issues through the use of quaternion involutions. The HR derivatives (the derivation of HR calculus is given in appendix A) are introduced below.

Definition 3.1 (real-differentiability [10]). A function $f(q) = f_a(q_a, q_b, q_c, q_d) + if_b(q_a, q_b, q_c, q_d) + jf_c(q_a, q_b, q_c, q_d) + kf_d(q_a, q_b, q_c, q_d)$ is called real differentiable when $f_a(q_a, q_b, q_c, q_d)$, $f_b(q_a, q_b, q_c, q_d)$, $f_c(q_a, q_b, q_c, q_d)$ and $f_d(q_a, q_b, q_c, q_d)$ are differentiable as functions of real variables q_a, q_b, q_c and q_d.

Definition 3.2 (the HR derivatives [13]). If $f : \mathbb{H} \to \mathbb{H}$ is real-differentiable, then the formal left HR derivatives of the function f with respect to $\{q, q^i, q^j, q^k\}$ and $\{q^*, q^{i*}, q^{j*}, q^{k*}\}$ are defined as

$$\begin{bmatrix} \frac{\partial f}{\partial q}, & \frac{\partial f}{\partial q^*} \\ \frac{\partial f}{\partial q^i}, & \frac{\partial f}{\partial q^{i*}} \\ \frac{\partial f}{\partial q^j}, & \frac{\partial f}{\partial q^{j*}} \\ \frac{\partial f}{\partial q^k}, & \frac{\partial f}{\partial q^{k*}} \end{bmatrix}^{\mathrm{T}} = \frac{1}{4} \begin{bmatrix} \frac{\partial f}{\partial q_a}, & \frac{\partial f}{\partial q_a} \\ \frac{\partial f}{\partial q_b}, & -\frac{\partial f}{\partial q_b} \\ \frac{\partial f}{\partial q_c}, & -\frac{\partial f}{\partial q_c} \\ \frac{\partial f}{\partial q_d}, & -\frac{\partial f}{\partial q_d} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -i & -i & i & i \\ -j & j & -j & j \\ -k & k & k & -k \end{bmatrix} \tag{3.1}$$

and the formal right HR derivatives of the function f are defined as

$$\begin{bmatrix} \frac{\partial_r f}{\partial q}, & \frac{\partial_r f}{\partial q^*} \\ \frac{\partial_r f}{\partial q^i}, & \frac{\partial_r f}{\partial q^{i*}} \\ \frac{\partial_r f}{\partial q^j}, & \frac{\partial_r f}{\partial q^{j*}} \\ \frac{\partial_r f}{\partial q^k}, & \frac{\partial_r f}{\partial q^{k*}} \end{bmatrix} = \frac{1}{4} \begin{bmatrix} 1 & -i & -j & -k \\ 1 & -i & j & k \\ 1 & i & -j & k \\ 1 & i & j & -k \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial q_a}, & \frac{\partial f}{\partial q_a} \\ \frac{\partial f}{\partial q_b}, & -\frac{\partial f}{\partial q_b} \\ \frac{\partial f}{\partial q_c}, & -\frac{\partial f}{\partial q_c} \\ \frac{\partial f}{\partial q_d}, & -\frac{\partial f}{\partial q_d} \end{bmatrix}, \tag{3.2}$$

where $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$ are the partial derivatives of f with respect to q_a, q_b, q_c and q_d.

Remark 3.3. It is important to note that the right HR derivatives exist if and only if the left HR derivatives also exist. The only difference between the left HR derivatives and the right HR derivatives is in the position of the partial derivative $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$. Within the left HR derivatives, $\partial f/\partial q_a, \partial f/\partial q_b, \partial f/\partial q_c$ and $\partial f/\partial q_d$ stand on the left side and imaginary units i, j, k on the right side (cf. (3.1)). It is exactly the opposite case for right HR derivatives. Note that the terms $\partial f/\partial q_a, \partial f/\partial q_b, \partial f/\partial q_c$ and $\partial f/\partial q_d$ cannot swap positions with the imaginary units i, j, k because of the non-commutative nature of quaternion product.

3.1. The validity of the traditional product rule

A straightforward use of the HR derivatives may become too tedious for complicated functions, for example, for the power function $f(q) = q^n$. This is because the HR calculus does not satisfy the traditional product rule which would simplify the calculation. Indeed, for two quaternion functions $f(q)$ and $g(q)$, in general we have $\partial(fg)/\partial q \neq f(\partial g/\partial q) + \partial f/\partial q g$. We shall illustrate this difficulty through the following two examples.

Example 3.4. Find the HR derivative of the function $f : \mathbb{H} \to \mathbb{H}$ given by

$$f(q) = q^2 = q_a^2 - (q_b^2 + q_c^2 + q_d^2) + 2q_a(iq_b + jq_c + kq_d), \tag{3.3}$$

where $q = q_a + iq_b + jq_c + kq_d$, $q_a, q_b, q_c, q_d \in \mathbb{R}$.

Solution: By definition 3.2, the left HR derivative of q^2 becomes

$$\begin{aligned}\frac{\partial(q^2)}{\partial q} &= \frac{1}{4}\left(\frac{\partial q^2}{\partial q_a} - \frac{\partial q^2}{\partial q_b}i - \frac{\partial q^2}{\partial q_c}j - \frac{\partial q^2}{\partial q_d}k\right)\\ &= \frac{1}{4}(2q_a + 2(iq_b + jq_c + kq_d) - (-2q_b + 2q_a i)i - (-2q_c + 2q_a j)j - (-2q_d + 2q_a k)k)\\ &= \frac{1}{4}(8q_a + 4q_b i + 4q_c j + 4q_d k) = q + \Re(q). \end{aligned} \tag{3.4}$$

Alternatively, $q(\partial q/\partial q) + (\partial q/\partial q)q = 2q$. This shows the traditional product rule is not valid.

Example 3.5. Find the HR derivative of the function $f : \mathbb{H} \to \mathbb{H}$ given by

$$f(q) = |q|^2 = qq^*. \tag{3.5}$$

Solution: By definition 3.2, the HR derivative of $|q|^2$ is

$$\frac{\partial(|q|^2)}{\partial q} = \frac{1}{4}\left(\frac{\partial |q|^2}{\partial q_a} - \frac{\partial |q|^2}{\partial q_b}i - \frac{\partial |q|^2}{\partial q_c}j - \frac{\partial |q|^2}{\partial q_d}k\right) = \frac{1}{4}(2q_a - 2q_b i - 2q_c j - 2q_d k) = \frac{1}{2}q^*, \tag{3.6}$$

while $q(\partial q^*/\partial q) + (\partial q/\partial q)q^* = -q/2 + q^*$, and thus the traditional product rule is not valid.

Remark 3.6. Examples 3.4 and 3.5 show that the left HR derivatives do not admit the traditional product rule. Similarly, the traditional product rule is not applicable for the right HR derivative in definition 3.2.

4. The generalization of HR calculus: generalized HR derivatives

We now introduce the novel GHR derivatives which comprise both the product and chain rules. This is archived by replacing the basis $\{1, i, j, k\}$ in definition 3.2 with a general orthogonal basis $\{1, i^\mu, j^\mu, k^\mu\}$, see also (2.5). The derivation of the GHR calculus is similar to that of the HR calculus in appendix A and is omitted for space considerations.

Definition 4.1 (the generalized HR derivatives). If $f : \mathbb{H} \to \mathbb{H}$ is real-differentiable, then the left GHR derivatives of the function f with respect to q^μ and $q^{\mu *}$ ($\mu \neq 0, \mu \in \mathbb{H}$) are defined as

$$\left.\begin{aligned}\frac{\partial f}{\partial q^\mu} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}i^\mu - \frac{\partial f}{\partial q_c}j^\mu - \frac{\partial f}{\partial q_d}k^\mu\right)\\ \text{and}\qquad \frac{\partial f}{\partial q^{\mu *}} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b}i^\mu + \frac{\partial f}{\partial q_c}j^\mu + \frac{\partial f}{\partial q_d}k^\mu\right),\end{aligned}\right\} \tag{4.1}$$

while the right GHR derivatives are defined as

$$\left.\begin{aligned}\frac{\partial_r f}{\partial q^\mu} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - i^\mu\frac{\partial f}{\partial q_b} - j^\mu\frac{\partial f}{\partial q_c} - k^\mu\frac{\partial f}{\partial q_d}\right)\\ \text{and}\qquad \frac{\partial_r f}{\partial q^{\mu *}} &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + i^\mu\frac{\partial f}{\partial q_b} + j^\mu\frac{\partial f}{\partial q_c} + k^\mu\frac{\partial f}{\partial q_d}\right),\end{aligned}\right\} \tag{4.2}$$

where $q = q_a + iq_b + jq_c + kq_d, q_a, q_b, q_c, q_d \in \mathbb{R}$, $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$ are the partial derivatives of f with respect to q_a, q_b, q_c and q_d, while the set $\{1, i^\mu, j^\mu, k^\mu\}$ is an orthogonal basis of $\mathbb{H}$.

Remark 4.2. The GHR derivatives are more concise and physically more intuitive than the HR derivatives, which are a special case of the GHR derivatives for $\mu = \{1, i, j, k\}$ (definitions 3.2 and 4.1). The concept of the GHR derivative can also be applied to other orthogonal systems, such as $\{1, \eta, \eta', \eta''\}$ in [32].

Proposition 4.3 (constant rule). *Let $f : \mathbb{H} \to \mathbb{H}$ be real-differentiable. Then, the following holds:*

$$\frac{\partial(\nu f)}{\partial q^\mu} = \nu\frac{\partial f}{\partial q^\mu}, \quad \frac{\partial(f\nu)}{\partial q^\mu} = \frac{\partial f}{\partial q^{\nu\mu}}\nu \tag{4.3}$$

and

$$\frac{\partial_r(f\nu)}{\partial q^\mu} = \frac{\partial_r f}{\partial q^\mu}\nu, \quad \frac{\partial_r(\nu f)}{\partial q^\mu} = \nu\frac{\partial_r f}{\partial q^{\nu^*\mu}}, \tag{4.4}$$

where $\mu, \nu \in \mathbb{H}$ are non-zero quaternion constants.

Proof. By the definition of the left GHR derivative in (4.1), we have

$$\begin{aligned}\frac{\partial(\nu f)}{\partial q^\mu} &= \frac{1}{4}\left(\frac{\partial(\nu f)}{\partial q_a} - \frac{\partial(\nu f)}{\partial q_b} i^\mu - \frac{\partial(\nu f)}{\partial q_c} j^\mu - \frac{\partial(\nu f)}{\partial q_d} k^\mu\right) \\ &= \frac{1}{4}\left(\nu\frac{\partial f}{\partial q_a} - \nu\frac{\partial f}{\partial q_b} i^\mu - \nu\frac{\partial f}{\partial q_c} j^\mu - \nu\frac{\partial f}{\partial q_d} k^\mu\right) = \nu\frac{\partial f}{\partial q^\mu}. \end{aligned} \tag{4.5}$$

Using the equality $(q^\mu)^\nu = q^{\nu\mu}$ in (2.4), the second equality of (4.3) is proved as

$$\begin{aligned}\frac{\partial(f\nu)}{\partial q^\mu} &= \frac{1}{4}\left(\frac{\partial(\nu f)}{\partial q_a} - \frac{\partial(\nu f)}{\partial q_b} i^\mu - \frac{\partial(\nu f)}{\partial q_c} j^\mu - \frac{\partial(\nu f)}{\partial q_d} k^\mu\right) \\ &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a}\nu - \frac{\partial f}{\partial q_b}\nu i^\mu - \frac{\partial f}{\partial q_c}\nu j^\mu - \nu\frac{\partial f}{\partial q_d}\nu k^\mu\right) \\ &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i^{\nu\mu} - \frac{\partial f}{\partial q_c} j^{\nu\mu} - \nu\frac{\partial f}{\partial q_d} k^{\nu\mu}\right)\nu = \frac{\partial f}{\partial q^{\nu\mu}}\nu. \end{aligned} \tag{4.6}$$

Hence, (4.3) immediately follows, and (4.4) can be proved in a similar way. ■

Remark 4.4. It is important to note that if a function f is premultiplied by a constant η in the first equality of (4.3), then the left GHR derivative of the product is equal to the left GHR derivative of f premultiplied by the constant, but not for postmultiplication. In other words, the left constant ν can come out from the derivative of the product; for this reason we refer to (4.1) as the left GHR derivative. The equalities in (4.4) complement those in (4.3). Thus, we refer to the derivatives in (4.2) as the right GHR derivatives, denoted by ∂_r in order to distinguish them from the left GHR derivatives in (4.1).

Proposition 4.5 (conjugate rule). *Let $f : \mathbb{H} \to \mathbb{H}$ be real-differentiable. Then, the following holds:*

$$\frac{\partial f}{\partial q^\mu} = \left(\frac{\partial_r f^*}{\partial q^{\mu *}}\right)^* \quad \textit{and} \quad \frac{\partial f}{\partial q^{\mu *}} = \left(\frac{\partial_r f^*}{\partial q^\mu}\right)^*. \tag{4.7}$$

Proof. By the definition of the right GHR derivative in (4.2), we have

$$\left(\frac{\partial_r f^*}{\partial q^{\mu *}}\right)^* = \frac{1}{4}\left(\frac{\partial f^*}{\partial q_a} + i^\mu\frac{\partial f^*}{\partial q_b} + j^\mu\frac{\partial f^*}{\partial q_c} + k^\mu\frac{\partial f^*}{\partial q_d}\right)^*. \tag{4.8}$$

Using the equalities $(pq)^* = q^*p^*$ and $(q^*)^\mu = (q^\mu)^*$ in (2.4), the above equality becomes

$$\left(\frac{\partial_r f^*}{\partial q^{\mu *}}\right)^* = \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i^\mu - \frac{\partial f}{\partial q_c} j^\mu - \frac{\partial f}{\partial q_d} k^\mu\right) = \frac{\partial f}{\partial q^\mu}. \tag{4.9}$$

Hence, the first part of (4.7) follows, and the second part can be proved in a similar way. ■

Proposition 4.6 (rotation rule). *Let $f : \mathbb{H} \to \mathbb{H}$ be real-differentiable. Then the following holds:*

$$\left(\frac{\partial f}{\partial q^\mu}\right)^\nu = \frac{\partial f^\nu}{\partial q^{\nu\mu}} \quad \textit{and} \quad \left(\frac{\partial f}{\partial q^{\mu *}}\right)^\nu = \frac{\partial f^\nu}{\partial q^{\nu\mu *}}. \tag{4.10}$$

Proof. Using the equalities in (2.4) and the left GHR derivative in (4.1), we have

$$\begin{aligned}\left(\frac{\partial f}{\partial q^\mu}\right)^\nu &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i^\mu - \frac{\partial f}{\partial q_c} j^\mu - \frac{\partial f}{\partial q_d} k^\mu\right)^\nu \\ &= \frac{1}{4}\left(\frac{\partial f^\nu}{\partial q_a} - \frac{\partial f^\nu}{\partial q_b} i^{\nu\mu} - \frac{\partial f^\nu}{\partial q_c} j^{\nu\mu} - \frac{\partial f^\nu}{\partial q_d} k^{\nu\mu}\right) = \frac{\partial f^\nu}{\partial q^{\nu\mu}}. \end{aligned} \tag{4.11}$$

Hence, the first part of (4.10) follows, and the second part can be proved in a similar way. ■

Corollary 4.7. *Let $f : \mathbb{H} \to \mathbb{H}$ be real-differentiable. Then, the following holds:*

$$\left(\frac{\partial f}{\partial q^\eta}\right)^\eta = \frac{\partial f^\eta}{\partial q} \quad \textit{and} \quad \left(\frac{\partial f}{\partial q^{\eta *}}\right)^\eta = \frac{\partial f^\eta}{\partial q^*}, \quad \forall\, \eta \in \{1, i, j, k\}. \tag{4.12}$$

Proof. Since $\partial f/\partial q^{\eta\eta} = \partial f/\partial q$ for any $\eta \in \{1, i, j, k\}$, then the proof follows directly from proposition 4.6. ■

Proposition 4.8. *Let $f:\mathbb{H}\to\mathbb{R}$ be real-differentiable. Then, the following holds:*

$$\frac{\partial f}{\partial q^{\mu}}=\left(\frac{\partial f}{\partial q^{\mu *}}\right)^{*}=\frac{\partial_{r} f}{\partial q^{\mu}} \quad \text{and} \quad \frac{\partial f}{\partial q^{\mu *}}=\left(\frac{\partial f}{\partial q^{\mu}}\right)^{*}=\frac{\partial_{r} f}{\partial q^{\mu *}}. \tag{4.13}$$

Proof. Since f is real-valued, its partial derivatives $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$ are real numbers, which yields $(\partial f/\partial \xi)^{*}=\partial f/\partial \xi$, where $\xi\in\{q_a,q_b,q_c,q_d\}$. Using the equality $(q^{\mu})^{*}=(q^{*})^{\mu}$ in (2.4) and the left GHR derivative in (4.1), we have

$$\begin{aligned}\left(\frac{\partial f}{\partial q^{\mu *}}\right)^{*}&=\frac{1}{4}\left(\frac{\partial f}{\partial q_a}+\frac{\partial f}{\partial q_b}i^{\mu}+\frac{\partial f}{\partial q_c}j^{\mu}+\frac{\partial f}{\partial q_d}k^{\mu}\right)^{*}\\&=\frac{1}{4}\left(\frac{\partial f}{\partial q_a}-i^{\mu}\frac{\partial f}{\partial q_b}-j^{\mu}\frac{\partial f}{\partial q_c}-k^{\mu}\frac{\partial f}{\partial q_d}\right)=\frac{\partial_{r} f}{\partial q^{\mu}}\\&=\frac{1}{4}\left(\frac{\partial f}{\partial q_a}-\frac{\partial f}{\partial q_b}i^{\mu}-\frac{\partial f}{\partial q_c}j^{\mu}-\frac{\partial f}{\partial q_d}k^{\mu}\right)=\frac{\partial f}{\partial q^{\mu}}.\end{aligned} \tag{4.14}$$

Hence, the first part of (4.13) follows, and the second part can be proved in a similar way. ■

Remark 4.9. From the identity (4.13), observe that the left GHR derivative is equal to the right GHR derivative if the function f is real-valued. This result is instrumental for practical applications of quaternion optimization, where the objective function (or cost function) is frequently real-valued. By virtue of the GHR calculus, the choice of the left/right GHR derivative therefore becomes obsolete as shown in (4.13). In the sequel, without loss in generality we shall mainly focus on the left GHR derivatives.

4.1. The novel product rule

We now introduce a novel product rule into quaternion analysis and show that the traditional product rule is a special case of the proposed product rule in corollary 4.11.

Theorem 4.10 (product rule of left GHR). *If the functions $f,g:\mathbb{H}\to\mathbb{H}$ are real-differentiable, then so too is their product fg, so that*

$$\frac{\partial (fg)}{\partial q^{\mu}}=f\frac{\partial g}{\partial q^{\mu}}+\frac{\partial f}{\partial q^{g\mu}}g \quad \text{and} \quad \frac{\partial (fg)}{\partial q^{\mu *}}=f\frac{\partial g}{\partial q^{\mu *}}+\frac{\partial f}{\partial q^{g\mu *}}g, \tag{4.15}$$

where $\partial f/\partial q^{g\mu}$ and $\partial f/\partial q^{g\mu *}$ can be obtained by replacing μ with $g\mu$ in definition 4.1.

Proof. The proof is given in appendix B. ■

Corollary 4.11. *If the functions $f:\mathbb{H}\to\mathbb{H}$ and $g:\mathbb{H}\to\mathbb{R}$ are real-differentiable, then their product fg satisfies the traditional product rule*

$$\frac{\partial (fg)}{\partial q^{\mu}}=f\frac{\partial g}{\partial q^{\mu}}+\frac{\partial f}{\partial q^{\mu}}g \quad \text{and} \quad \frac{\partial (fg)}{\partial q^{\mu *}}=f\frac{\partial g}{\partial q^{\mu *}}+\frac{\partial f}{\partial q^{\mu *}}g, \tag{4.16}$$

where $\partial f/\partial q^{\mu}$ and $\partial f/\partial q^{\mu *}$ are the left GHR derivatives in definition 4.1.

Proof. Since $q^{g\mu}=q^{\mu}$ and $q^{g\mu *}=q^{\mu *}$ for a real function $g:\mathbb{H}\to\mathbb{R}$, the corollary follows. ■

We now present some GHR derivatives of nonlinear quaternion functions enabled by the GHR calculus. These are very useful in applications, such as in nonlinear adaptive filters and quaternion-valued neural networks.

Example 4.12 (split-quaternion function). Find the GHR derivative of a split-quaternion function $f:\mathbb{H}\to\mathbb{H}$ given by

$$f(q)=\varphi(q_a)+j\varphi(q_b)+j\varphi(q_c)+k\varphi(q_d), \tag{4.17}$$

where $q=q_a+iq_b+jq_c+kq_d\in\mathbb{R}$ and $\varphi:\mathbb{R}\to\mathbb{R}$ is a real-valued differentiable function.

Solution: By the definition of the left GHR derivatives in (4.1), it follows that

$$\begin{aligned}\frac{\partial f(q)}{\partial q} &= \frac{1}{4}\left(\frac{\partial f(q)}{\partial q_a} - \frac{\partial f(q)}{\partial q_a}i - \frac{\partial f(q)}{\partial q_a}j - \frac{\partial f(q)}{\partial q_a}k\right)\\ &= \frac{1}{4}(\varphi'(q_a) + \varphi'(q_b) + \varphi'(q_c) + \varphi'(q_d))\end{aligned} \tag{4.18}$$

and

$$\begin{aligned}\frac{\partial f(q)}{\partial q^*} &= \frac{1}{4}\left(\frac{\partial f(q)}{\partial q_a} + \frac{\partial f(q)}{\partial q_a}i + \frac{\partial f(q)}{\partial q_a}j + \frac{\partial f(q)}{\partial q_a}k\right)\\ &= \frac{1}{4}(\varphi'(q_a) - \varphi'(q_b) - \varphi'(q_c) - \varphi'(q_d)).\end{aligned} \tag{4.19}$$

This shows that the GHR derivatives of the split-quaternion function are real-valued.

Example 4.13 (power function). Find the GHR derivative of the power function $f : \mathbb{H} \to \mathbb{H}$ given by

$$f(q) = q^n, \tag{4.20}$$

where n is any positive integer.

Solution: Using the product rule in theorem 4.10, it follows that

$$\frac{\partial q^n}{\partial q^\mu}\mu = \frac{\partial(qq^{n-1})}{\partial q^\mu}\mu = q\frac{\partial q^{n-1}}{\partial q^\mu}\mu + \frac{\partial q}{\partial q^{q^{n-1}\mu}}q^{n-1}\mu = q\frac{\partial q^{n-1}}{\partial q^\mu}\mu + \Re(q^{n-1}\mu), \tag{4.21}$$

where the term $(\partial q/\partial q^\mu)\mu$, given in table 1, was used in the last equality. Note that the above expression is recurrent in $(\partial q^n/\partial q^\mu)\mu$. Upon expanding this expression and using the initial condition $(\partial q/\partial q^\mu)\mu = \Re(\mu)$, this yields

$$\frac{\partial q^n}{\partial q^\mu}\mu = \sum_{m=1}^{n} q^{n-m}\Re(q^{m-1}\mu). \tag{4.22}$$

In a similar manner, we have

$$\frac{\partial q^n}{\partial q^{\mu*}}\mu = \frac{\partial(qq^{n-1})}{\partial q^{\mu*}}\mu = q\frac{\partial q^{n-1}}{\partial q^{\mu*}}\mu + \frac{\partial q}{\partial q^{q^{n-1}\mu*}}q^{n-1}\mu = q\frac{\partial q^{n-1}}{\partial q^{\mu*}}\mu - \frac{1}{2}(q^{n-1}\mu)^*, \tag{4.23}$$

which is equivalent to

$$\frac{\partial q^n}{\partial q^{\mu*}}\mu = -\frac{1}{2}\sum_{m=1}^{n} q^{n-m}(q^{m-1}\mu)^*. \tag{4.24}$$

Example 4.14 (exponential function). Find the GHR derivative of the function $f : \mathbb{H} \to \mathbb{H}$ given by

$$\exp(q) \triangleq \sum_{n=0}^{+\infty}\frac{q^n}{n!}. \tag{4.25}$$

Solution: From (4.22), it follows that

$$\frac{\partial \exp(q)}{\partial q^\mu}\mu = \sum_{n=0}^{+\infty}\frac{1}{n!}\sum_{m=1}^{n} q^{n-m}\Re(q^{m-1}\mu). \tag{4.26}$$

In a similar manner, we have

$$\frac{\partial \exp(q)}{\partial q^{\mu*}}\mu = -\frac{1}{2}\sum_{n=0}^{+\infty}\frac{1}{n!}\sum_{m=1}^{n} q^{n-m}(q^{m-1}\mu)^*. \tag{4.27}$$

Remark 4.15. The exponential function is the most important elementary function, as both trigonometric functions and hyperbolic functions can be expressed in terms of the exponential function. The elementary function in example 4.14 is a power series, and it does not change the direction of the vector part of quaternion. Therefore, such elementary functions can swap positions with a quaternion q, i.e. $f(q)q = qf(q)$, giving an important property, $f^*(q) = f(q^*)$, which can be used in practical applications, such as quaternion neural networks [33] and quaternion nonlinear adaptive filters [28]. It is important to note that if a quaternion variable q degenerates into a real variable x in the definitions of elementary functions in this subsection, then the GHR derivatives simplify into the standard real derivatives, e.g. the GHR derivative of the power function in (4.22) will become nx^{n-1}. Therefore, the GHR derivatives are a generalized form of the real derivatives and the real derivatives are a special case of the GHR derivatives.

Table 1. Summary of the rules for the left GHR derivatives.

$f(q)$	$\frac{\partial f}{\partial q^\mu}\mu$	$\frac{\partial f}{\partial q^{\mu *}}\mu$	note
q	$\Re(\mu)$	$-\frac{1}{2}\mu^*$	—
ωq	$\omega\Re(\mu)$	$-\frac{1}{2}\omega\mu^*$	$\forall\,\omega\in\mathbb{H}$
$q\nu$	$\Re(\nu\mu)$	$-\frac{1}{2}(\nu\mu)^*$	$\forall\,\nu\in\mathbb{H}$
$\omega q\nu+\lambda$	$\omega\Re(\nu\mu)$	$-\frac{1}{2}\omega(\nu\mu)^*$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
q^*	$-\frac{1}{2}\mu^*$	$\Re(\mu)$	—
ωq^*	$-\frac{1}{2}\omega\mu^*$	$\omega\Re(\mu)$	$\forall\,\omega\in\mathbb{H}$
$q^*\nu$	$-\frac{1}{2}(\nu\mu)^*$	$\Re(\nu\mu)$	$\forall\,\nu\in\mathbb{H}$
$\omega q^*\nu+\lambda$	$-\frac{1}{2}\omega(\nu\mu)^*$	$\omega\Re(\nu\mu)$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
q^{-1}	$-q^{-1}\Re(q^{-1}\mu)$	$\frac{1}{2}q^{-1}\mu^*(q^*)^{-1}$	—
$(q^*)^{-1}$	$\frac{1}{2}(q^*)^{-1}\mu^*q^{-1}$	$-(q^*)^{-1}\Re((q^*)^{-1}\mu)$	—
$(\omega q\nu+\lambda)^{-1}$	$-f\omega\Re(\nu f\mu)$	$\frac{1}{2}f\omega(\nu f\mu)^*$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
$(\omega q^*\nu+\lambda)^{-1}$	$\frac{1}{2}f\omega(\nu f\mu)^*$	$-f\omega\Re(\nu f\mu)$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
q^2	$q\Re(\mu)+\Re(q\mu)$	$-\frac{1}{2}q\mu^*-\frac{1}{2}(q\mu)^*$	—
$(q^*)^2$	$-\frac{1}{2}q^*\mu^*-\frac{1}{2}(q^*\mu)^*$	$q^*\Re(\mu)+\Re(q^*\mu)$	—
$(\omega q\nu+\lambda)^2$	$g\omega\Re(\nu\mu)+\omega\Re(\nu g\mu)$	$-\frac{1}{2}g\omega(\nu\mu)^*-\frac{1}{2}\omega(\nu g\mu)^*$	$g=\omega q\nu+\lambda$
$(\omega q^*\nu+\lambda)^2$	$-\frac{1}{2}g\omega(\nu\mu)^*-\frac{1}{2}\omega(\nu g\mu)^*$	$g\omega\Re(\nu\mu)+\omega\Re(\nu g\mu)$	$g=\omega q^*\nu+\lambda$
$\Re(q)$	$\frac{1}{4}\mu$	$\frac{1}{4}\mu$	—
$\Re(\omega q\nu+\lambda)$	$\frac{1}{4}\mu\nu\omega$	$\frac{1}{4}\mu\omega^*\nu^*$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
$\Re(\omega q^*\nu+\lambda)$	$\frac{1}{4}\mu\omega^*\nu^*$	$\frac{1}{4}\mu\nu\omega$	$\forall\,\omega,\nu,\lambda\in\mathbb{H}$
$\lvert q\rvert$	$\frac{1}{4\lvert q\rvert}\mu q^*$	$\frac{1}{4\lvert q\rvert}\mu q$	—
$\lvert q\rvert^2$	$\frac{1}{2}\mu q^*$	$\frac{1}{2}\mu q$	—
$\lvert\omega q\nu+\lambda\rvert$	$\frac{1}{4\lvert g\rvert}\mu\nu g^*\omega$	$\frac{1}{4\lvert g\rvert}\mu\omega^*g\nu^*$	$g=\omega q\nu+\lambda$
$\lvert\omega q^*\nu+\lambda\rvert$	$\frac{1}{4\lvert g\rvert}\mu\omega^*g\nu^*$	$\frac{1}{4\lvert g\rvert}\mu\nu g^*\omega$	$g=\omega q^*\nu+\lambda$
$\lvert\omega q\nu+\lambda\rvert^2$	$\frac{1}{2}\mu\nu g^*\omega$	$\frac{1}{2}\mu\omega^*g\nu^*$	$g=\omega q\nu+\lambda$
$\lvert\omega q^*\nu+\lambda\rvert^2$	$\frac{1}{2}\mu\omega^*g\nu^*$	$\frac{1}{2}\mu\nu g^*\omega$	$g=\omega q^*\nu+\lambda$

(*Continued*.)

Table 1. (*Continued*.)

$f(q)$	$\frac{\partial f}{\partial q^{\mu}}\mu$	$\frac{\partial f}{\partial q^{\mu *}}\mu$	note
$\frac{q}{\lvert q\rvert}$	$\frac{1}{\lvert q\rvert}\Re(\mu)-\frac{1}{4\lvert q\rvert^3}q\mu q^*$	$-\frac{1}{2\lvert q\rvert}\mu^*-\frac{1}{4\lvert q\rvert^3}q\mu q$	—
$\frac{q^*}{\lvert q\rvert}$	$-\frac{1}{2\lvert q\rvert}\mu^*-\frac{1}{4\lvert q\rvert^3}q^*\mu q^*$	$\frac{1}{\lvert q\rvert}\Re(\mu)-\frac{1}{4\lvert q\rvert^3}q^*\mu q$	—
$\frac{\omega q\nu+\lambda}{\lvert\omega q\nu+\lambda\rvert}$	$\frac{\omega}{2\lvert g\rvert}\Re(\nu\mu)+\frac{g}{4\lvert g\rvert^3}\nu^*(\omega^* g\mu)^*$	$-\frac{\omega}{4\lvert g\rvert}(\nu\mu)^*-\frac{g}{2\lvert g\rvert^3}\nu^*\Re(\omega^* g\mu)$	$g=\omega q\nu+\lambda$
$\frac{\omega q^*\nu+\lambda}{\lvert\omega q^*\nu+\lambda\rvert}$	$-\frac{\omega}{2\lvert g\rvert}(\nu\mu)^*-\frac{f}{\lvert g\rvert}\frac{\partial\lvert g\rvert}{\partial q^{\mu}}\mu$	$\frac{\omega}{\lvert g\rvert}\Re(\nu\mu)-\frac{f}{\lvert g\rvert}\frac{\partial\lvert g\rvert}{\partial q^{\mu *}}\mu$	$g=\omega q^*\nu+\lambda$

4.2. The chain rule

Another advantage of the GHR derivatives is that they admit the chain rule, which is formulated in the following theorem.

Theorem 4.16 (chain rule of left GHR). *Let $S\subseteq\mathbb{H}$ and let $g:S\to\mathbb{H}$ be real-differentiable at an interior point q of the set S. Let $T\subseteq\mathbb{H}$ be such that $g(q)\in T$ for all $q\in S$. Assume that $f:T\to\mathbb{H}$ is real-differentiable at an interior point $g(q)\in T$, then the composite function $f(g(q))$ satisfies the following chain rules:*

$$\left.\begin{aligned}\frac{\partial f(g(q))}{\partial q^{\mu}}&=\frac{\partial f}{\partial g^{\nu}}\frac{\partial g^{\nu}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q^{\mu}}\\ \frac{\partial f(g(q))}{\partial q^{\mu *}}&=\frac{\partial f}{\partial g^{\nu}}\frac{\partial g^{\nu}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q^{\mu *}}\end{aligned}\right\}\tag{4.28}$$

and

$$\left.\begin{aligned}\frac{\partial f(g(q))}{\partial q^{\mu}}&=\frac{\partial f}{\partial g^{\nu *}}\frac{\partial g^{\nu *}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu i *}}\frac{\partial g^{\nu i *}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu j *}}\frac{\partial g^{\nu j *}}{\partial q^{\mu}}+\frac{\partial f}{\partial g^{\nu k *}}\frac{\partial g^{\nu k *}}{\partial q^{\mu}}\\ \frac{\partial f(g(q))}{\partial q^{\mu *}}&=\frac{\partial f}{\partial g^{\nu *}}\frac{\partial g^{\nu *}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu i *}}\frac{\partial g^{\nu i *}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu j *}}\frac{\partial g^{\nu j *}}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^{\nu k *}}\frac{\partial g^{\nu k *}}{\partial q^{\mu *}},\end{aligned}\right\}\tag{4.29}$$

where $\mu,\nu\in\mathbb{H}$, $\mu\nu\neq 0$, and $\partial f/\partial g^{\nu}=\partial f(g)/\partial g^{\nu}$ and $\partial f/\partial g^{\nu *}=\partial f(g)/\partial g^{\nu *}$ are the left GHR derivatives.

Proof. The proof of theorem 4.16 is given in appendix C. ■

Theorem 4.16 is also valid for complex-valued and real-valued composite functions of quaternion variables, as stated in the following two corollaries, the proofs of which are similar to that of theorem 4.16, and are thus omitted.

Corollary 4.17. *Let $S\subseteq\mathbb{H}$ and let $g:S\to\mathbb{C}$ be real-differentiable at an interior point q of the set S. Let $T\subseteq\mathbb{C}$ be such that $g(q)\in T$ for all $q\in S$. Assume that $f:T\to\mathbb{C}$ is real-differentiable at an interior point $g(q)\in T$, then the left GHR derivatives of the composite function $f(g(q))$ are as follows*:

$$\frac{\partial f(g(q))}{\partial q^{\mu}}=\frac{\partial f}{\partial g}\frac{\partial g}{\partial q^{\mu}}+\frac{\partial f}{\partial g^*}\frac{\partial g^*}{\partial q^{\mu}}\quad\text{and}\quad\frac{\partial f(g(q))}{\partial q^{\mu *}}=\frac{\partial f}{\partial g}\frac{\partial g}{\partial q^{\mu *}}+\frac{\partial f}{\partial g^*}\frac{\partial g^*}{\partial q^{\mu *}},\tag{4.30}$$

where $\mu\in\mathbb{H}$, $\mu\neq 0$, and $\partial f/\partial g=\partial f(g)/\partial g$ and $\partial f/\partial g^*=\partial f(g)/\partial g^*$ are the complex CR derivatives within the CR calculus.

Corollary 4.18. *Let $S\subseteq\mathbb{H}$ and let $g:S\to\mathbb{R}$ be real-differentiable at an interior point q of the set S. Let $T\subseteq\mathbb{R}$ be such that $g(q)\in T$ for all $q\in S$. Assume that $f:T\to\mathbb{R}$ is real-differentiable at an interior point $f(q)\in T$, then the left GHR derivatives of the composite function $f(g(q))$ are as follows*:

$$\frac{\partial f(g(q))}{\partial q^{\mu}}=f'(g)\frac{\partial g}{\partial q^{\mu}}\quad\text{and}\quad\frac{\partial f(g(q))}{\partial q^{\mu *}}=f'(g)\frac{\partial g}{\partial q^{\mu *}},\tag{4.31}$$

where $\mu\in\mathbb{H}$, $\mu\neq 0$ and $f'(g)$ is the real derivative of a real-valued function.

Theorem 4.19 (chain rule of right GHR). *Let $S\subseteq\mathbb{H}$ and let $g:S\to\mathbb{H}$ be real-differentiable at an interior point q of the set S. Let $T\subseteq\mathbb{H}$ be such that $g(q)\in T$ for all $q\in S$. Assume that $f:T\to\mathbb{H}$ is real-differentiable at an*

inner point $g(q) \in T$, then the right GHR derivatives of the composite function $f(g(q))$ are as follows:

$$\left.\begin{aligned}\frac{\partial_r f(g(q))}{\partial q^\mu} &= \frac{\partial_r g^\nu}{\partial q^\mu}\frac{\partial_r f}{\partial g^\nu} + \frac{\partial_r g^{\nu i}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu i}} + \frac{\partial_r g^{\nu j}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu j}} + \frac{\partial_r g^{\nu k}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu k}}\\ \frac{\partial_r f(g(q))}{\partial q^{\mu *}} &= \frac{\partial_r g^\nu}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^\nu} + \frac{\partial_r g^{\nu i}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu i}} + \frac{\partial_r g^{\nu j}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu j}} + \frac{\partial_r g^{\nu k}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu k}}\end{aligned}\right\} \tag{4.32}$$

and

$$\left.\begin{aligned}\frac{\partial_r f(g(q))}{\partial q^\mu} &= \frac{\partial_r g^{\nu *}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu *}} + \frac{\partial_r g^{\nu i *}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu i *}} + \frac{\partial_r g^{\nu j *}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu j *}} + \frac{\partial_r g^{\nu k *}}{\partial q^\mu}\frac{\partial_r f}{\partial g^{\nu k *}}\\ \frac{\partial_r f(g(q))}{\partial q^{\mu *}} &= \frac{\partial_r g^{\nu *}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu *}} + \frac{\partial_r g^{\nu i *}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu i *}} + \frac{\partial_r g^{\nu j *}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu j *}} + \frac{\partial_r g^{\nu k *}}{\partial q^{\mu *}}\frac{\partial_r f}{\partial g^{\nu k *}},\end{aligned}\right\} \tag{4.33}$$

where $\mu, \nu \in \mathbb{H}$, $\mu\nu \neq 0$, and $\partial_r f/\partial g^\nu = \partial_r f(g)/\partial g^\nu$ and $\partial_r f/\partial g^{\nu *} = \partial_r f(g)/\partial g^{\nu *}$ are the right GHR derivatives.

Proof. The proof of theorem 4.19 is similar to that of theorem 4.16 and is thus omitted. ■

4.3. Mean value theorem

The mean value theorem is one of the most important tools in calculus, and we next introduce its compact version for general quaternion-valued functions of quaternion variables.

Theorem 4.20 (mean value theorem of left kind). *Consider a continuous function $f : S \subseteq \mathbb{H} \to \mathbb{H}$ for which the left GHR derivatives exist and are continuous in the set S. Then, for any $q_0, q_1 \in S$ for which the segment joining them also lies in S, we have*

$$f(q_1) - f(q_0) = \sum_\mu \int_0^1 \frac{\partial f(q_0 + t\lambda)}{\partial q^\mu}\lambda^\mu \,\mathrm{d}t = \sum_\mu \int_0^1 \frac{\partial f(q_0 + t\lambda)}{\partial q^{\mu *}}\lambda^{\mu *}\,\mathrm{d}t, \tag{4.34}$$

where $\mu \in \{1, i, j, k\}$, $\lambda = q_1 - q_0$, and $\partial f(q_0 + t\lambda)/\partial q^\mu = \partial f(q)/\partial q^\mu|_{q=q_0+t\lambda}$ is the left GHR derivative as in definition 4.1.

Proof. Denote $F(t) = f(g(t))$, where $g(t) = q_0 + t\lambda$ and $0 \le t \le 1$. Then $F(t)$ is continuous on $[0, 1]$ and has derivatives in $(0, 1)$. By theorem 4.16, the derivative of $F(t)$ is

$$\begin{aligned}F'(t) &= \frac{\partial f(g)}{\partial g}\lambda + \frac{\partial f(g)}{\partial g^i}\lambda^i + \frac{\partial f(g)}{\partial g^j}\lambda^j + \frac{\partial f(g)}{\partial g^k}\lambda^k\\ &= \left.\frac{\partial f(q)}{\partial q}\right|_{q=q_0+t\lambda}\lambda + \left.\frac{\partial f(q)}{\partial q^i}\right|_{q=q_0+t\lambda}\lambda^i + \left.\frac{\partial f(q)}{\partial q^j}\right|_{q=q_0+t\lambda}\lambda^j + \left.\frac{\partial f(q)}{\partial q^k}\right|_{q=q_0+t\lambda}\lambda^k\\ &= \frac{\partial f(q_0 + t\lambda)}{\partial q}\lambda + \frac{\partial f(q_0 + t\lambda)}{\partial q^i}\lambda^i + \frac{\partial f(q_0 + t\lambda)}{\partial q^j}\lambda^j + \frac{\partial f(q_0 + t\lambda)}{\partial q^k}\lambda^k.\end{aligned} \tag{4.35}$$

Upon substituting (4.35) into $F(1) - F(0) = \int_0^1 F'(t)\,\mathrm{d}t$, with $F(0) = f(q_0)$ and $F(1) = f(q_1)$, the theorem follows. The second equality can be proved in a similar manner. ■

Corollary 4.21. *Consider a continuous function $f : S \subseteq \mathbb{H} \to \mathbb{R}$ for which the left GHR derivatives exist and are continuous in the set S. Then, for any $q_0, q_1 \in S$ for which the segment joining them also lies in S, we have*

$$f(q_1) - f(q_0) = 4\int_0^1 \Re\left(\frac{\partial f(q_0 + t\lambda)}{\partial q}\lambda\right)\mathrm{d}t = 4\int_0^1 \Re\left(\frac{\partial f(q_0 + t\lambda)}{\partial q^*}\lambda^*\right)\mathrm{d}t, \tag{4.36}$$

where $\lambda = q_1 - q_0$ and $\partial f(q_0 + t\lambda)/\partial q^\mu = \partial f(q)/\partial q^\mu|_{q=q_0+t\lambda}$ is the left GHR derivative as in definition 4.1.

Proof. Function f is real-valued, and therefore $\partial f/\partial q = \partial f^\eta/\partial q$, where $\eta \in \{1, i, j, k\}$. From (2.3) and (4.10), we now have

$$\frac{\partial f}{\partial q^\eta}\lambda^\eta = \frac{\partial f^\eta}{\partial q^\eta}\lambda^\eta = \left(\frac{\partial f}{\partial q}\lambda\right)^\eta. \tag{4.37}$$

The corollary then follows from (2.7) and theorem 4.20, while the second equality can be derived using $\Re(pq) = \Re(p^*q^*)$. ■

For λ which is sufficiently small in the modulus, the right-hand side of (4.34) can be approximated as

$$f(q_1) - f(q_0) \approx \sum_{\mu\in\{1,i,j,k\}} \frac{\partial f(q_0)}{\partial q^\mu}\lambda^\mu. \tag{4.38}$$

If the left GHR derivatives of f are Lipschitz continuous in the vicinity of q and q_1, with the Lipschitz constant L, we can estimate the error in this approximation as

$$\begin{aligned}&\left|f(q_1) - f(q_0) - \sum_{\mu\in\{1,i,j,k\}} \frac{\partial f(q_0)}{\partial q^\mu}\lambda^\mu\right| \\ &\quad\le \sum_{\mu\in\{1,i,j,k\}} \left|\int_0^1 \left(\frac{\partial f(q_0+t\lambda)}{\partial q^\mu} - \frac{\partial f(q_0)}{\partial q^\mu}\right)\lambda^\mu \,\mathrm{d}t\right| \le 4\int_0^1 Lt|\lambda|^2\,\mathrm{d}t = 2L|\lambda|^2.\end{aligned} \tag{4.39}$$

4.4. Taylor's theorem

We can now introduce a novel, rigorous version of Taylor's theorem for quaternion-valued functions of quaternion variables, as a generalization of the standard univariate Taylor's theorem.

Lemma 4.22 (Apostol [34]). *Consider a $(k+1)$-times continuously differentiable function $f : D \subseteq \mathbb{R} \to \mathbb{R}$. If $x \in D$, then*

$$f(x_0+h) = f(x_0) + f'(x_0)h + f''(x_0)\frac{h^2}{2} + \cdots + f^{(k)}(x_0)\frac{h^k}{k!} + R_k, \tag{4.40}$$

where the remainder R_k is given by

$$R_k = \int_{x_0}^{x_0+h} \frac{f^{(k+1)}(t)}{k!}(x_0+h-t)^k\,\mathrm{d}t. \tag{4.41}$$

Theorem 4.23 (Taylor's theorem of left kind). *Consider a third-order continuous real-differentiable function $f : S \subseteq \mathbb{H} \to \mathbb{H}$. If $q_0, q_0+\lambda \in S$ such that the segment joining them also lies in S, then*

$$\begin{aligned}f(q_0+\lambda) - f(q_0) &= \sum_\mu \frac{\partial f(q_0)}{\partial q^\mu}\lambda^\mu + \frac{1}{2}\sum_{\mu,\nu}\frac{\partial^2 f(q_0)}{\partial q^\nu \partial q^\mu}\lambda^\nu\lambda^\mu + O(\lambda^3) \\ &= \sum_\mu \frac{\partial f(q_0)}{\partial q^{\mu*}}\lambda^{\mu*} + \frac{1}{2}\sum_{\mu,\nu}\frac{\partial^2 f(q_0)}{\partial q^{\nu*}\partial q^{\mu*}}\lambda^{\nu*}\lambda^{\mu*} + O(\lambda^3) \quad \text{as } \lambda \to 0,\end{aligned} \tag{4.42}$$

where $\mu, \nu \in \{1, i, j, k\}$, and $\partial^2 f/\partial q^\nu \partial q^\mu$ and $\partial^2 f/\partial q^{\nu*}\partial q^{\mu*}$ are the second-order left GHR derivatives.

Proof. Define an auxiliary function $g(t) = f(q_0+t\lambda)$ with $0 \le t \le 1$. Using the chain rule in theorem 4.16, we obtain

$$\left.\begin{aligned}g'(t) &= \sum_\mu \frac{\partial f(q_0+t\lambda)}{\partial q^\mu}\lambda^\mu \\ \text{and}\qquad g''(t) &= \sum_{\mu,\nu}\frac{\partial^2 f(q_0+t\lambda)}{\partial q^\nu \partial q^\mu}\lambda^\nu\lambda^\mu, \quad g'''(t) = \sum_{\mu,\nu,\eta}\frac{\partial^3 f(q_0+t\lambda)}{\partial q^\eta \partial q^\nu \partial q^\mu}\lambda^\eta\lambda^\nu\lambda^\mu,\end{aligned}\right\} \tag{4.43}$$

where $\mu, \nu, \eta \in \{1, i, j, k\}$. The second-order Taylor polynomial in lemma 4.22 then gives

$$g(1) = g(0) + g'(0) + \tfrac{1}{2}g''(0) + R_2 \tag{4.44}$$

which is equivalent to

$$f(q_0+\lambda) = f(q_0) + \sum_\mu \frac{\partial f(q_0)}{\partial q^\mu}\lambda^\mu + \frac{1}{2}\sum_{\mu,\nu}\frac{\partial^2 f(q_0)}{\partial q^\nu \partial q^\mu}\lambda^\nu\lambda^\mu + R_2, \tag{4.45}$$

where

$$R_2 = \int_0^1 \frac{(1-t)^2}{2}g'''(t)\,\mathrm{d}t = \int_0^1 \frac{(1-t)^2}{2}\sum_{\mu,\nu,\eta}\frac{\partial^3 f(q_0+t\lambda)}{\partial q^\eta \partial q^\nu \partial q^\mu}\lambda^\eta\lambda^\nu\lambda^\mu\,\mathrm{d}t. \tag{4.46}$$

This integral contains three factors of λ, while the remaining factors are bounded. Therefore, R_2 is of the order $|\lambda|^3$, making the fraction $|R_2|/|\lambda|^3$ bounded as $\lambda \to 0$. Hence, the first equality of the theorem follows, and the second equality can be proved in the same way. ∎

Corollary 4.24. *Consider a third-order continuous real-differentiable function $f : S \subseteq \mathbb{H} \to \mathbb{R}$. If $q_0, q_0 + \lambda \in S$ such that the segment joining them also lies in S, then*

$$\begin{aligned} f(q_0+\lambda) - f(q_0) &= 4\Re\left(\frac{\partial f(q_0)}{\partial q}\lambda\right) + 2\sum_{\nu}\Re\left(\frac{\partial^2 f(q_0)}{\partial q^{\nu}\partial q}\lambda^{\nu}\lambda\right) + O(|\lambda|^3) \\ &= 4\Re\left(\frac{\partial f(q_0)}{\partial q^*}\lambda^*\right) + 2\sum_{\nu}\Re\left(\frac{\partial^2 f(q_0)}{\partial q^{\nu *}\partial q^*}\lambda^{\nu *}\lambda^*\right) + O(|\lambda|^3) \quad \text{as } \lambda \to 0, \end{aligned} \tag{4.47}$$

where $\nu \in \{1, i, j, k\}$, and $\partial^2 f/\partial q^{\nu}\partial q$ and $\partial^2 f/\partial q^{\nu *}\partial q^*$ are the second-order left GHR derivatives.

Proof. Using the result for $\partial\Re(q)/\partial q$ in table 1 and the chain rule in theorem 4.16, this corollary is proved similarly to the proof of corollary 4.21. ■

Theorem 4.25 (Taylor's theorem of centre kind). *Consider a third-order continuous real-differentiable function $f : S \subseteq \mathbb{H} \to \mathbb{R}$. If $q_0, q_0 + \lambda \in S$ such that the segment joining them also lies in S, then*

$$\begin{aligned} f(q_0+\lambda) - f(q_0) &= \sum_{\mu}\frac{\partial f(q_0)}{\partial q^{\mu}}\lambda^{\mu} + \frac{1}{2}\sum_{\mu,\nu}\lambda^{\mu *}\frac{\partial^2 f(q_0)}{\partial q^{\nu}\partial q^{\mu *}}\lambda^{\nu} + O(\lambda^3) \\ &= \sum_{\mu}\frac{\partial f(q_0)}{\partial q^{\mu *}}\lambda^{\mu *} + \frac{1}{2}\sum_{\mu,\nu}\lambda^{\mu}\frac{\partial^2 f(q_0)}{\partial q^{\nu *}\partial q^{\mu}}\lambda^{\nu *} + O(\lambda^3) \quad \text{as } \lambda \to 0, \end{aligned} \tag{4.48}$$

where $\mu, \nu \in \{1, i, j, k\}$, and $\partial^2 f/\partial q^{\nu}\partial q^{\mu *}$ and $\partial^2 f/\partial q^{\nu}\partial q^{\mu *}$ are the second-order left GHR derivatives.

Proof. Define an auxiliary function $g(t) = f(q_0 + t\lambda)$ with $0 \le t \le 1$. Using the chain rule in theorem 4.19 and using the property in (4.13), we obtain

$$g'(t) = \sum_{\mu}\lambda^{\mu *}\frac{\partial_r f(q_0 + t\lambda)}{\partial q^{\mu *}} = \sum_{\mu}\lambda^{\mu *}\frac{\partial f(q_0 + t\lambda)}{\partial q^{\mu *}}. \tag{4.49}$$

Using the constant rule (4.3), we arrive at

$$g''(t) = \sum_{\mu}\lambda^{\mu *}\left(\frac{\partial f(q_0 + t\lambda)}{\partial q^{\mu *}}\right)' = \sum_{\mu,\nu}\lambda^{\mu *}\frac{\partial^2 f(q_0 + t\lambda)}{\partial q^{\nu}\partial q^{\mu *}}\lambda^{\nu}, \tag{4.50}$$

where $\mu, \nu \in \{1, i, j, k\}$. The rest of the proof is almost the same as that in theorem 4.23 and is thus omitted. ■

Remark 4.26. The Taylor expansion in theorem 4.23 is concisely expressed using the GHR derivatives. This is different from the Taylor expansion given by Schwartz [18], which decomposes a quaternion q into two mutually perpendicular quaternions in a local coordinate system. In contrast, our approach treats the quaternion q as an augmented quaternion based on quaternion involutions [21]. Schwartz has also stated that his Taylor expansion would cause trouble when the function has terms $q\omega q$, where ω is a general quaternion. Note that there are no such restrictions in theorem 4.23, which only requires the real-differentiability condition in definition 3.1, that is, the functions $f(q)$ should admit the partial derivatives with respect to the four real components q_a, q_b, q_c and q_d.

5. Applications of the generalized HR calculus

We now illustrate the utility of the GHR calculus in optimization, statistics, signal processing, machine learning and other application areas.

5.1. Derivation of the widely linear quaternion least mean square algorithm

The widely linear QLMS (WL-QLMS) algorithm is based on the quaternion widely linear model $y(n) = \mathbf{w}^{\mathrm{T}}(n)\mathbf{p}(n)$ which deals with the generality of quaternion signals (both proper and improper) [8,22,35], where $\mathbf{p} = (\mathbf{x}^{\mathrm{T}}(n), \mathbf{x}^{i T}(n), \mathbf{x}^{j T}(n), \mathbf{x}^{k T}(n))^{\mathrm{T}}$ is the augmented input vector and $\mathbf{w} = (\mathbf{h}^{\mathrm{T}}(n), \mathbf{g}^{\mathrm{T}}(n), \mathbf{u}^{\mathrm{T}}(n), \mathbf{v}^{\mathrm{T}}(n))^{\mathrm{T}}$ is the associated weight (parameter) vector. The cost function to be minimized

is a real-valued function of quaternion variables

$$J(n) = |e(n)|^2, \tag{5.1}$$

where $e(n) = d(n) - y(n)$ is the error between the desired signal $d(n)$ and the filter output $y(n)$. The weight update of WL-QLMS is then given by

$$\Delta \mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) = -\alpha \nabla_{\mathbf{w}^*} J(n) = -\alpha \left(\frac{\partial J(n)}{\partial \mathbf{w}^*} \right)^{\mathrm{T}}, \tag{5.2}$$

where α is the step size, $(\cdot)^{\mathrm{T}}$ denotes the transpose and the gradient $\nabla_{\mathbf{w}^*} J(n)$ defines the direction of the maximum rate of change of J [13,36]. By the product rule within the GHR calculus, given in theorem 4.10, we have

$$\frac{\partial J}{\partial \mathbf{w}^*} = e^* \frac{\partial e}{\partial \mathbf{w}^*} + \frac{\partial e^*}{\partial \mathbf{w}^{e*}} e, \tag{5.3}$$

where the time index 'n' is omitted. Note the convention that $\partial f / \partial \mathbf{w}$ is a row vector whose nth element is $\partial f / \partial w_n$. The above GHR derivatives are calculated as

$$\left.\begin{aligned} &\frac{\partial e}{\partial \mathbf{w}^*} = \frac{\partial (d - \mathbf{w}^{\mathrm{T}}\mathbf{p})}{\partial \mathbf{w}^*} = -\frac{\partial (\mathbf{w}^{\mathrm{T}}\mathbf{p})}{\partial \mathbf{w}^*} = \frac{1}{2}\mathbf{p}^H \\ \text{and} \quad &\frac{\partial e^*}{\partial \mathbf{w}^{e*}} e = \frac{\partial \left(d^* - \mathbf{p}^H \mathbf{w}^*\right)}{\partial \mathbf{w}^{e*}} e = -\frac{\partial \left(\mathbf{p}^H \mathbf{w}^*\right)}{\partial \mathbf{w}^{e*}} e = -\mathbf{p}^H \Re(e), \end{aligned}\right\} \tag{5.4}$$

where the terms $\partial(qv)/\partial q^*$ and $\partial(\omega q^*)/\partial q^{\mu *}$ are given in table 1 and are used in the last equalities in the expressions above. Substituting (5.4) into (5.3) yields

$$\frac{\partial J}{\partial \mathbf{w}^*} = \frac{1}{2} e^* \mathbf{p}^H - \mathbf{p}^H \Re(e) = \left(\frac{1}{2} e^* - \Re(e) \right) \mathbf{p}^H = -\frac{1}{2} e \mathbf{p}^H. \tag{5.5}$$

Finally, the update of the adaptive weight vector of WL-QLMS becomes

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha e(n) \mathbf{p}^*(n), \tag{5.6}$$

where the constant $\frac{1}{2}$ in (5.5) is absorbed into α.

Remark 5.1. There are many variations of WL-QLMS algorithms, such as the WL-QLMS algorithms based on variants $\{\mathbf{x}^*, \mathbf{x}^{i*}, \mathbf{x}^{j*}, \mathbf{x}^{k*}\}$, $\{\mathbf{x}^{\mu}, \mathbf{x}^{\mu i}, \mathbf{x}^{\mu j}, \mathbf{x}^{\mu k}\}$ and $\{\mathbf{x}^*, \mathbf{x}^{\mu i *}, \mathbf{x}^{\mu j *}, \mathbf{x}^{\mu k *}\}$. Note that if we start from $y(n) = \mathbf{w}^H(n)\mathbf{p}(n)$, the final update rule would become $\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha \mathbf{p}(n) e^*(n)$.

5.2. Derivation of quaternion nonlinear adaptive filtering algorithms

Tools of the GHR calculus allow us to concisely derive quaternion nonlinear adaptive filtering algorithms, a basis for fast-growing area of quaternion learning system. The same real-valued quadratic cost function as in real LMS and complex LMS is used, that is

$$J(n) = |e(n)|^2 = e_a^2(n) + e_b^2(n) + e_c^2(n) + e_d^2(n), \tag{5.7}$$

where $e(n) = e_a(n) + i e_b(n) + j e_c(n) + k e_d(n)$, $e(n) = d(n) - \Phi(s(n))$, $s(n) = \mathbf{w}^{\mathrm{T}}(n)\mathbf{x}(n)$, and Φ is the quaternion nonlinearity. The weight update is given by

$$\Delta \mathbf{w}(n) = \mathbf{w}(n+1) - \mathbf{w}(n) = -\alpha \nabla_{\mathbf{w}^*} J(n) = -\alpha \left(\frac{\partial J(n)}{\partial \mathbf{w}^*} \right)^{\mathrm{T}}, \tag{5.8}$$

where $\alpha > 0$ is the real step size, $(\cdot)^{\mathrm{T}}$ denotes the transpose, and the gradient $\nabla_{\mathbf{w}^*} J(n)$ defines the direction of the maximum rate of change of J [13,36]. By using the chain rule in theorem 4.16, the above gradient can be calculated as

$$\frac{\partial J}{\partial \mathbf{w}^*} = \sum_{\mu \in \{1,i,j,k\}} \frac{\partial J}{\partial s^{\mu *}} \frac{\partial s^{\mu *}}{\partial \mathbf{w}^*}, \tag{5.9}$$

where time index 'n' is omitted for convenience. Note the convention that $\partial f / \partial \mathbf{w}$ is a row vector whose nth element is $\partial f / \partial w_n$. Using the term $\partial(\omega q^*)/\partial q^*$ in table 1, we have

$$\frac{\partial s^*}{\partial \mathbf{w}^*} = \frac{\partial (\mathbf{x}^H \mathbf{w}^*)}{\partial \mathbf{w}^*} = \mathbf{x}^H. \tag{5.10}$$

Upon applying the second equality in (4.12) and using the term $\partial(\omega q^*)/\partial q^{\mu *}\mu$ in table 1, this yields

$$\frac{\partial s^{\mu *}}{\partial \mathbf{w}^*} = \left(\frac{\partial s^*}{\partial \mathbf{w}^{\mu *}}\right)^{\mu} = -\mu \frac{\partial \left(\mathbf{x}^H \mathbf{w}^*\right)}{\partial \mathbf{w}^{\mu *}} \mu = -\mu \mathbf{x}^H \Re(\mu) = 0, \quad \forall\, \mu \in \{i, j, k\}. \tag{5.11}$$

Using the chain rule in corollary 4.18, we have

$$\begin{aligned} \frac{\partial J}{\partial s^*} &= 2e_a \frac{\partial e_a}{\partial s^*} + 2e_b \frac{\partial e_b}{\partial s^*} + 2e_c \frac{\partial e_c}{\partial s^*} + 2e_d \frac{\partial e_d}{\partial s^*} \\ &= -2e_a \frac{\partial \Phi_a(s)}{\partial s^*} - 2e_b \frac{\partial \Phi_b(s)}{\partial s^*} - 2e_c \frac{\partial \Phi_c(s)}{\partial s^*} - 2e_d \frac{\partial \Phi_d(s)}{\partial s^*}. \end{aligned} \tag{5.12}$$

Upon substituting (5.10)–(5.12) into (5.9), we arrive at

$$\frac{\partial J}{\partial \mathbf{w}^*} = -2\left(e_a \frac{\partial \Phi_a(s)}{\partial s^*} + e_b \frac{\partial \Phi_b(s)}{\partial s^*} + e_c \frac{\partial \Phi_c(s)}{\partial s^*} + e_d \frac{\partial \Phi_d(s)}{\partial s^*}\right) \mathbf{x}^H. \tag{5.13}$$

Finally, the weight update for this quaternion nonlinear adaptive filtering algorithm becomes

$$\Delta \mathbf{w}(n) = \alpha \left(e_a(n) \frac{\partial \Phi_a(s(n))}{\partial s^*(n)} + e_b(n) \frac{\partial \Phi_b(s(n))}{\partial s^*(n)} + e_c(n) \frac{\partial \Phi_c(s(n))}{\partial s^*(n)} + e_d(n) \frac{\partial \Phi_d(s(n))}{\partial s^*(n)}\right) \mathbf{x}^*(n), \tag{5.14}$$

where the constant 2 in (5.13) is absorbed into the step size α.

For illustration, consider an example where Φ is a nonlinear split-quaternion function $\Phi(s) = \varphi(s_a) + i\varphi(s_b) + j\varphi(s_c) + \varphi(s_d)$ and $\varphi : \mathbb{R} \to \mathbb{R}$ is a real-valued differentiable function. Then,

$$\frac{\partial \Phi_a(s)}{\partial s^*} = \frac{\partial \varphi(s_a)}{\partial s^*} = \frac{1}{4}\varphi'(s_a). \tag{5.15}$$

In a similar manner, we have

$$\frac{\partial \Phi_b(s)}{\partial s^*} = \frac{i}{4}\varphi'(s_b), \quad \frac{\partial \Phi_c(s)}{\partial s^*} = \frac{j}{4}\varphi'(s_c) \quad \text{and} \quad \frac{\partial \Phi_d(s)}{\partial s^*} = \frac{k}{4}\varphi'(s_d). \tag{5.16}$$

The weight update for such a quaternion nonlinear adaptive filtering algorithm becomes

$$\Delta \mathbf{w}(n) = \alpha(e_a(n)\varphi'(s_a(n)) + ie_b(n)\varphi'(s_b(n)) + je_c(n)\varphi'(s_c(n)) + ke_d(n)\varphi'(s_d(n)))\mathbf{x}^*(n), \tag{5.17}$$

where the constant $\frac{1}{4}$ in (5.15) and (5.16) is absorbed into the step size α.

Another example is when Φ is a quaternion linear function $\Phi(s) = s$, that is $\varphi(x) = x$ in (5.17). Then, the update of the adaptive weight vector within the QLMS algorithm becomes

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \alpha e(n)\mathbf{x}^*(n) \tag{5.18}$$

illustrating the generic nature of the GHR calculus.

Remark 5.2. From (5.17) and (5.18), we note that quaternion linear and nonlinear adaptive filtering algorithms have been developed in a unified form. This also shows that the GHR calculus gives us much more freedom in the design, as the nonlinear function Φ is not required to satisfy the odd-symmetry condition [28,37]. We can derive many other algorithms, such as the augmented quaternion nonlinear gradient descent algorithm [38], in the same way. In the interest of space, we leave this to the interested reader.

Remark 5.3. The QLMS algorithm (5.18) is different from the QLMS in [4,13], due to the use of different product rule. The traditional product rule was used in [4,13] to simplify the derivation; however, our examples in §3a illustrate that the traditional product rule is not applicable to the HR derivatives. On the other hand, the chain rules within the GHR calculus result in the QLMS in (5.18), which has the same generic form as that of the complex LMS [39]. For the performance comparison and steady-state analysis of the existing QLMS algorithms, we refer the reader to [40] for more details.

6. Conclusion

A novel and rigorous framework for the efficient computation of quaternion derivatives, referred to as the GHR calculus, has been established. The GHR methodology has been shown to greatly relax the existence conditions for the derivatives of general nonlinear functions of quaternion variables, and to simplify the calculation of quaternion derivatives through its novel product and chain rules. We have shown that, unlike the existing quaternion derivatives, the GHR calculus is general and can be used for both analytic and non-analytic functions of quaternion variables. The core of the GHR calculus is the use of quaternion rotations in order to overcome the non-commutativity of quaternion product, and the use

of quaternion involutions to obtain an elegant quaternion basis. Through the analysis and examples, the proposed framework has been shown to allow for real- and complex-valued optimization algorithms to be extended to the quaternion field in a generic, compact and intuitive way. Application case studies in statistical signal processing and learning systems demonstrate the effectiveness of the proposed GHR framework.

Authors' contributions. D.X. designed the study, carried out the derivations and analysis and drafted the manuscript; C.J. helped with the extensions of HR calculus and the analysis; C.C.T. helped with the application case studies and proofread the manuscript; D.P.M. conceived and coordinated the study, was involved in the design and analysis, and helped in drafting the manuscript. All authors gave final approval for publication.

Competing interests. The authors have no competing interests.

Funding. D.X. was supported by the National Natural Science Foundation of China (no. 61301202), and by the Research Fund for the Doctoral Program of Higher Education of China (no. 20122304120028). D.P.M. was supported by EPSRC grant no. EP/H026266/1.

Acknowledgements. The authors wish to thank the anonymous reviewers for their insightful comments and suggestions on how to improve this work.

Appendices

Several results and proofs of statements from the main body are next detailed.

Appendix A. The derivation of the HR calculus

The following lemmas are useful to distinguish between the GHR derivatives and the standard quaternion differential.

Lemma A.1. *Let $f_n : \mathbb{H} \to \mathbb{H}\,(n = 1, 2, 3, 4)$ be any arbitrary quaternion-valued functions and $\mu \in \mathbb{H}, \mu \neq 0$. If for the left case*

$$f_1 dq^{\mu} + f_2 dq^{\mu i} + f_3 dq^{\mu j} + f_4 dq^{\mu k} = 0 \tag{A 1}$$

or for the right case

$$dq^{\mu} f_1 + dq^{\mu i} f_2 + dq^{\mu j} f_3 + dq^{\mu k} f_4 = 0, \tag{A 2}$$

then $f_n = 0$ for $n \in \{1, 2, 3, 4\}$.

Proof. The left case. By applying the rotation transformation on both sides of (2.6), it follows that

$$\left.\begin{aligned} q^{\mu} &= q_a + i^{\mu} q_b + j^{\mu} q_c + k^{\mu} q_d, \quad q^{\mu i} = q_a + i^{\mu} q_b - j^{\mu} q_c - k^{\mu} q_d \\ q^{\mu j} &= q_a - i^{\mu} q_b + j^{\mu} q_c - k^{\mu} q_d, \quad q^{\mu k} = q_a - i^{\mu} q_b - j^{\mu} q_c + k^{\mu} q_d. \end{aligned}\right\} \tag{A 3}$$

Upon applying the differential operator to the above expressions and substituting $dq^{\mu}, dq^{\mu i}, dq^{\mu j}$ and $dq^{\mu k}$ into (A 1), we have

$$\begin{aligned} &f_1(dq_a + i^{\mu} dq_b + j^{\mu} dq_c + k^{\mu} dq_d) + f_2(dq_a + i^{\mu} dq_b - j^{\mu} dq_c - k^{\mu} dq_d) \\ &\quad + f_3(dq_a - i^{\mu} dq_b + j^{\mu} dq_c - k^{\mu} dq_d) + f_4(dq_a - i^{\mu} dq_b - j^{\mu} dq_c + k^{\mu} dq_d) = 0. \end{aligned} \tag{A 4}$$

This is equivalent to

$$\begin{aligned} &(f_1 + f_2 + f_3 + f_4) dq_a + (f_1 + f_2 - f_3 - f_4) i^{\mu} dq_b \\ &\quad + (f_1 - f_2 + f_3 - f_4) j^{\mu} dq_c + (f_1 - f_2 - f_3 + f_4) k^{\mu} dq_d = 0. \end{aligned} \tag{A 5}$$

Since the differentials dq_a, dq_b, dq_c and dq_d are independent, this yields

$$\left.\begin{aligned} f_1 + f_2 + f_3 + f_4 = 0, \quad f_1 + f_2 - f_3 - f_4 = 0, \\ f_1 - f_2 + f_3 - f_4 = 0, \quad f_1 - f_2 - f_3 + f_4 = 0. \end{aligned}\right\} \tag{A 6}$$

Hence, it follows that $f_1 = f_2 = f_3 = f_4 = 0$. The right case can be proved in a similar way. ■

Lemma A.2. *Let $f_n : \mathbb{H} \to \mathbb{H}\,(n = 1, 2, 3, 4)$ be any arbitrary quaternion-valued function and $\mu \in \mathbb{H}, \mu \neq 0$. If for the left case*

$$f_1 dq^{\mu *} + f_2 dq^{\mu i *} + f_3 dq^{\mu j *} + f_4 dq^{\mu k *} = 0$$

or for the right case

$$dq^{\mu *} f_1 + dq^{\mu i *} f_2 + dq^{\mu j *} f_3 + dq^{\mu k *} f_4 = 0,$$

then $f_n = 0$ for $n \in \{1, 2, 3, 4\}$.

Proof. The proof of lemma A.2 follows from that of lemma A.1 and is omitted. ■

We now provide the derivation of the HR calculus. For any quaternion-valued function $f(q) \in \mathbb{H}$, we can state (since the fields $\mathbb{H}$ and $\mathbb{R}^4$ are isomorphic) that

$$f(q) = f_a(q_a, q_b, q_c, q_d) + if_b(q_a, q_b, q_c, q_d) + jf_c(q_a, q_b, q_c, q_d) + kf_d(q_a, q_b, q_c, q_d), \tag{A 7}$$

where $f_a(\cdot)$, $f_b(\cdot)$, $f_c(\cdot)$, $f_d(\cdot) \in \mathbb{R}$. Then, the function f can be equally seen as a function of the four independent real variables q_a, q_b, q_c and q_d, and the differential of f can be expressed as [10]

$$[\text{Left}]: \quad df = \frac{\partial f}{\partial q_a} dq_a + \frac{\partial f}{\partial q_b} dq_b + \frac{\partial f}{\partial q_c} dq_c + \frac{\partial f}{\partial q_d} dq_d \tag{A 8}$$

and

$$[\text{Right}]: \quad df = dq_a \frac{\partial f}{\partial q_a} + dq_b \frac{\partial f}{\partial q_b} + dq_c \frac{\partial f}{\partial q_c} + dq_d \frac{\partial f}{\partial q_d}, \tag{A 9}$$

where $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$ are the partial derivatives of f with respect to q_a, q_b, q_c and q_d, respectively. Note that the two equations are identical since dq_a, dq_b, dq_c and dq_d are real quantities. As a result, both equations are equally valid as a starting point for the derivation of the HR calculus.

The left case (A 8). There are two ways to link the real and quaternion differentials; these are based on (2.7) and its conjugate which correspond to the HR derivatives and conjugate HR derivatives.

A.1. The left HR derivatives

From (2.7), the differentials of the components of a quaternion can be formulated as

$$\left.\begin{aligned} dq_a &= \frac{1}{4}(dq + dq^i + dq^j + dq^k), \quad dq_b = -\frac{i}{4}(dq + dq^i - dq^j - dq^k) \\ dq_c &= -\frac{j}{4}(dq - dq^i + dq^j - dq^k), \quad dq_d = -\frac{k}{4}(dq - dq^i - dq^j + dq^k). \end{aligned}\right\} \tag{A 10}$$

By inserting (A 10) into (A 8), the differential of the function f becomes

$$\begin{aligned} df &= \frac{1}{4}\frac{\partial f}{\partial q_a}(dq + dq^i + dq^j + dq^k) - \frac{1}{4}\frac{\partial f}{\partial q_b} i(dq + dq^i - dq^j - dq^k) \\ &\quad - \frac{1}{4}\frac{\partial f}{\partial q_c} j(dq - dq^i + dq^j - dq^k) - \frac{1}{4}\frac{\partial f}{\partial q_d} k(dq - dq^i - dq^j + dq^k) \\ &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right) dq + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right) dq^i \\ &\quad + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right) dq^j + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right) dq^k. \end{aligned} \tag{A 11}$$

Now, we can define the formal left derivatives $\partial f/\partial q, \partial f/\partial q^i, \partial f/\partial q^j$ and $\partial f/\partial q^k$, so that

$$df = \frac{\partial f}{\partial q} dq + \frac{\partial f}{\partial q^i} dq^i + \frac{\partial f}{\partial q^j} dq^j + \frac{\partial f}{\partial q^k} dq^k \tag{A 12}$$

holds. Comparing (A 12) with (A 11) and applying lemma A.1, yields the left HR derivatives

$$\begin{bmatrix} \dfrac{\partial f}{\partial q} \\ \dfrac{\partial f}{\partial q^i} \\ \dfrac{\partial f}{\partial q^j} \\ \dfrac{\partial f}{\partial q^k} \end{bmatrix}^{\mathrm{T}} = \frac{1}{4} \begin{bmatrix} \dfrac{\partial f}{\partial q_a} \\ \dfrac{\partial f}{\partial q_b} \\ \dfrac{\partial f}{\partial q_c} \\ \dfrac{\partial f}{\partial q_d} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ -i & -i & i & i \\ -j & j & -j & j \\ -k & k & k & -k \end{bmatrix}. \tag{A 13}$$

A.2. The left conjugate HR derivatives

Upon applying the conjugate operator to both sides of (2.7), the differentials of the components of a quaternion can be formulated as

$$\left.\begin{aligned} dq_a &= \frac{1}{4}(dq^* + dq^{i*} + dq^{j*} + dq^{k*}), \quad dq_b = \frac{i}{4}(dq^* + dq^{i*} - dq^{j*} - dq^{k*}),\\ dq_c &= \frac{j}{4}(dq^* - dq^{i*} + dq^{j*} - dq^{k*}), \quad dq_d = \frac{k}{4}(dq^* - dq^{i*} - dq^{j*} + dq^{k*}). \end{aligned}\right\} \tag{A 14}$$

By inserting (A 14) into (A 8), the differential of f can be written as

$$\begin{aligned} df &= \frac{1}{4}\frac{\partial f}{\partial q_a}(dq^* + dq^{i*} + dq^{j*} + dq^{k*}) + \frac{1}{4}\frac{\partial f}{\partial q_b} i(dq^* + dq^{i*} - dq^{j*} - dq^{k*})\\ &\quad + \frac{1}{4}\frac{\partial f}{\partial q_c} j(dq^* - dq^{i*} + dq^{j*} - dq^{k*}) + \frac{1}{4}\frac{\partial f}{\partial q_d} k(dq^* - dq^{i*} - dq^{j*} + dq^{k*})\\ &= \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right) dq^* + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right) dq^{i*}\\ &\quad + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i + \frac{\partial f}{\partial q_c} j - \frac{\partial f}{\partial q_d} k\right) dq^{j*} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b} i - \frac{\partial f}{\partial q_c} j + \frac{\partial f}{\partial q_d} k\right) dq^{k*}. \end{aligned} \tag{A 15}$$

We can now define the formal left derivatives $\partial f/\partial q^*$, $\partial f/\partial q^{i*}$, $\partial f/\partial q^{j*}$ and $\partial f/\partial q^{k*}$, so that

$$df = \frac{\partial f}{\partial q^*} dq^* + \frac{\partial f}{\partial q^{i*}} dq^{i*} + \frac{\partial f}{\partial q^{j*}} dq^{j*} + \frac{\partial f}{\partial q^k} dq^{k*} \tag{A 16}$$

holds. Upon comparing (A 16) with (A 15) and applying lemma A.2, the following left conjugate HR derivatives are obtained:

$$\begin{bmatrix} \dfrac{\partial f}{\partial q^*} \\ \dfrac{\partial f}{\partial q^{i*}} \\ \dfrac{\partial f}{\partial q^{j*}} \\ \dfrac{\partial f}{\partial q^{k*}} \end{bmatrix}^{\mathrm{T}} = \frac{1}{4} \begin{bmatrix} \dfrac{\partial f}{\partial q_a} \\ \dfrac{\partial f}{\partial q_b} \\ \dfrac{\partial f}{\partial q_c} \\ \dfrac{\partial f}{\partial q_d} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ i & i & -i & -i \\ j & -j & j & -j \\ k & -k & -k & k \end{bmatrix}. \tag{A 17}$$

The right case (A 9). There are two ways to link the real and quaternion differentials, the approach in (2.7) and its conjugate, which induce to the right HR derivatives or conjugate right HR derivatives, respectively.

A.3. The right HR derivatives

Applying a rotation transformation to both sides of (A 10), we have

$$\left.\begin{aligned} dq_a &= \tfrac{1}{4}(dq + dq^i + dq^j + dq^k), \quad dq_b = -\tfrac{1}{4}(dq + dq^i - dq^j - dq^k)i,\\ dq_c &= -\tfrac{1}{4}(dq - dq^i + dq^j - dq^k)j, \quad dq_d = -\tfrac{1}{4}(dq - dq^i - dq^j + dq^k)k. \end{aligned}\right\} \tag{A 18}$$

Then, by substituting (A 18) into (A 9), the differential of f becomes

$$\begin{aligned} df &= \frac{1}{4}(dq + dq^i + dq^j + dq^k)\frac{\partial f}{\partial q_a} - \frac{1}{4}(dq + dq^i - dq^j - dq^k)i\frac{\partial f}{\partial q_b}\\ &\quad - \frac{1}{4}(dq - dq^i + dq^j - dq^k)j\frac{\partial f}{\partial q_c} - \frac{1}{4}(dq - dq^i - dq^j + dq^k)k\frac{\partial f}{\partial q_d}\\ &= \frac{1}{4}dq\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right) + \frac{1}{4}dq^i\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right)\\ &\quad + \frac{1}{4}dq^j\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right) + \frac{1}{4}dq^k\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right). \end{aligned} \tag{A 19}$$

Now, define the formal right derivatives $\partial_r f/\partial q, \partial_r f/\partial q^i, \partial_r f/\partial q^j$ and $\partial_r f/\partial q^k$, so that

$$df = dq\frac{\partial_r f}{\partial q} + dq^i\frac{\partial_r f}{\partial q^i} + dq^j\frac{\partial_r f}{\partial q^j} + dq^k\frac{\partial_r f}{\partial q^k} \tag{A 20}$$

holds. Comparing (A 20) and (A 19) and using lemma A.1, the following right HR derivatives are obtained:

$$\begin{bmatrix} \dfrac{\partial_r f}{\partial q} \\ \dfrac{\partial_r f}{\partial q^i} \\ \dfrac{\partial_r f}{\partial q^j} \\ \dfrac{\partial_r f}{\partial q^k} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & -i & -j & -k \\ 1 & -i & j & k \\ 1 & i & -j & k \\ 1 & i & j & -k \end{bmatrix}\begin{bmatrix} \dfrac{\partial f}{\partial q_a} \\ \dfrac{\partial f}{\partial q_b} \\ \dfrac{\partial f}{\partial q_c} \\ \dfrac{\partial f}{\partial q_d} \end{bmatrix}. \tag{A 21}$$

A.4. The right conjugate HR derivatives

By applying a rotation transformation to both sides of (A 14), we have

$$\left.\begin{aligned} dq_a &= \tfrac{1}{4}(dq^* + dq^{i*} + dq^{j*} + dq^{k*}), \quad dq_b = \tfrac{1}{4}(dq^* + dq^{i*} - dq^{j*} - dq^{k*})i, \\ dq_c &= \tfrac{1}{4}(dq^* - dq^{i*} + dq^{j*} - dq^{k*})j, \quad dq_d = \tfrac{1}{4}(dq^* - dq^{i*} - dq^{j*} + dq^{k*})k. \end{aligned}\right\} \tag{A 22}$$

Upon substituting (A 22) into (A 9), the differential of f can be written as

$$\begin{aligned} df &= \frac{1}{4}(dq^* + dq^{i*} + dq^{j*} + dq^{k*})\frac{\partial f}{\partial q_a} + \frac{1}{4}(dq^* + dq^{i*} - dq^{j*} - dq^{k*})i\frac{\partial f}{\partial q_b} \\ &\quad + \frac{1}{4}(dq^* - dq^{i*} + dq^{j*} - dq^{k*})j\frac{\partial f}{\partial q_c} + \frac{1}{4}(dq^* - dq^{i*} - dq^{j*} + dq^{k*})k\frac{\partial f}{\partial q_d} \\ &= \frac{1}{4}dq^*\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right) + \frac{1}{4}dq^{i*}\left(\frac{\partial f}{\partial q_a} + i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right) \\ &\quad + \frac{1}{4}dq^{j*}\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} + j\frac{\partial f}{\partial q_c} - k\frac{\partial f}{\partial q_d}\right) + \frac{1}{4}dq^{k*}\left(\frac{\partial f}{\partial q_a} - i\frac{\partial f}{\partial q_b} - j\frac{\partial f}{\partial q_c} + k\frac{\partial f}{\partial q_d}\right). \end{aligned} \tag{A 23}$$

Now, define the formal right derivatives $\partial_r f/\partial q^*, \partial_r f/\partial q^{i*}, \partial_r f/\partial q^{j*}$ and $\partial_r f/\partial q^{k*}$, so that

$$df = dq^*\frac{\partial_r f}{\partial q^*} + dq^{i*}\frac{\partial_r f}{\partial q^{i*}} + dq^{j*}\frac{\partial_r f}{\partial q^{j*}} + dq^{k*}\frac{\partial_r f}{\partial q^{k*}} \tag{A 24}$$

holds. Comparing (A 24) with (A 23) and using lemma A.2, we obtain the following right conjugate HR derivatives:

$$\begin{bmatrix} \dfrac{\partial_r f}{\partial q^*} \\ \dfrac{\partial_r f}{\partial q^{i*}} \\ \dfrac{\partial_r f}{\partial q^{j*}} \\ \dfrac{\partial_r f}{\partial q^{k*}} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix}\begin{bmatrix} \dfrac{\partial f}{\partial q_a} \\ \dfrac{\partial f}{\partial q_b} \\ \dfrac{\partial f}{\partial q_c} \\ \dfrac{\partial f}{\partial q_d} \end{bmatrix}. \tag{A 25}$$

Appendix B. The proof of the product rule

Within the GHR calculus, when a quaternion function is post-multiplied by a real function, the novel product rule degenerates into the traditional product rule. This is stated in the next lemma.

Lemma B.1. *If the functions $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{R}$ have the left GHR derivatives, then their product fg satisfies the traditional product rule*

$$\frac{\partial(fg)}{\partial q^\mu} = f\frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^\mu}g \quad \textit{and} \quad \frac{\partial(fg)}{\partial q^{\mu*}} = f\frac{\partial g}{\partial q^{\mu*}} + \frac{\partial f}{\partial q^{\mu*}}g, \tag{B 1}$$

where $\partial f/\partial q^\mu$ and $\partial f/\partial q^{\mu*}$ are the left GHR derivatives in definition 4.1.

Proof. Let $f = f_a + if_b + jf_c + kf_d$, where $f_a, f_b, f_c, f_d \in \mathbb{R}$, then $fg = f_a g + if_b g + jf_c g + kf_d g$. Using the property $\partial(\nu f)/\partial q^\mu = \nu(\partial f/\partial q^\mu)$ in (4.3), we have

$$\begin{aligned}\frac{\partial(fg)}{\partial q^\mu} &= \frac{\partial((f_a + if_b + jf_c + kf_d)g)}{\partial q^\mu}\\ &= \frac{\partial(f_a g)}{\partial q^\mu} + i\frac{\partial(f_b g)}{\partial q^\mu} + j\frac{\partial(f_c g)}{\partial q^\mu} + k\frac{\partial(f_d g)}{\partial q^\mu}\\ &= f_a\frac{\partial g}{\partial q^\mu} + if_b\frac{\partial g}{\partial q^\mu} + jf_c\frac{\partial g}{\partial q^\mu} + kf_d\frac{\partial g}{\partial q^\mu} + \frac{\partial f_a}{\partial q^\mu}g + i\frac{\partial f_b}{\partial q^\mu}g + j\frac{\partial f_c}{\partial q^\mu}g + k\frac{\partial f_d}{\partial q^\mu}g\\ &= f\frac{\partial g}{\partial q^\mu} + \left(\frac{\partial f_a}{\partial q^\mu} + i\frac{\partial f_b}{\partial q^\mu} + j\frac{\partial f_c}{\partial q^\mu} + k\frac{\partial f_d}{\partial q^\mu}\right)g = f\frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^\mu}g. \end{aligned} \tag{B 2}$$

Hence, the first part of the lemma is proved, and the rest follows in a similar way. ■

The proof of theorem 4.10. From (2.5), note that $\{1, i^\mu, j^\mu, k^\mu\}$ is another orthogonal basis of $\mathbb{H}$. Then, the quaternion-valued function g can be expressed in the following way:

$$g = g_a + i^\mu g_b + j^\mu g_c + k^\mu g_d, \quad g_a, g_b, g_c, g_d \in \mathbb{R}, \tag{B 3}$$

and $fg = fg_a + fi^\mu g_b + fj^\mu g_c + fk^\mu g_d, f \in \mathbb{H}$. Using the sum rule, it follows that

$$\frac{\partial(fg)}{\partial q^\mu} = \frac{\partial(fg_a)}{\partial q^\mu} + \frac{\partial(fi^\mu g_b)}{\partial q^\mu} + \frac{\partial(fi^\mu g_c)}{\partial q^\mu} + \frac{\partial(fi^\mu g_d)}{\partial q^\mu}. \tag{B 4}$$

Applying lemma B.1 to the right side of (B 4) yields

$$\begin{aligned}\frac{\partial(fg)}{\partial q^\mu} &= f\frac{\partial g_a}{\partial q^\mu} + fi^\mu\frac{\partial g_b}{\partial q^\mu} + fj^\mu\frac{\partial g_c}{\partial q^\mu} + fk^\mu\frac{\partial g_d}{\partial q^\mu} + \frac{\partial f}{\partial q^\mu}g_a + \frac{\partial(fi^\mu)}{\partial q^\mu}g_b + \frac{\partial(fj^\mu)}{\partial q^\mu}g_c + \frac{\partial(fk^\mu)}{\partial q^\mu}g_d\\ &= f\left(\frac{\partial g_a}{\partial q^\mu} + i^\mu\frac{\partial g_b}{\partial q^\mu} + j^\mu\frac{\partial g_c}{\partial q^\mu} + k^\mu\frac{\partial g_d}{\partial q^\mu}\right) + \frac{\partial f}{\partial q^\mu}g_a + \frac{\partial(fi^\mu)}{\partial q^\mu}g_b + \frac{\partial(fj^\mu)}{\partial q^\mu}g_c + \frac{\partial(fk^\mu)}{\partial q^\mu}g_d\\ &= f\frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^\mu}g_a + \frac{\partial(fi^\mu)}{\partial q^\mu}g_b + \frac{\partial(fj^\mu)}{\partial q^\mu}g_c + \frac{\partial(fk^\mu)}{\partial q^\mu}g_d. \end{aligned} \tag{B 5}$$

Next, by using the result $\partial(f\nu)/\partial q^\mu = (\partial f/\partial q^{\nu\mu})\nu$ in (4.3), we have

$$\begin{aligned}\frac{\partial(fg)}{\partial q^\mu} &= f\frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^\mu}g_a + \frac{\partial f}{\partial q^{\mu i}}i^\mu g_b + \frac{\partial f}{\partial q^{\mu j}}j^\mu g_c + \frac{\partial f}{\partial q^{\mu k}}k^\mu g_d\\ &= f\frac{\partial g}{\partial q^\mu} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}i^\mu - \frac{\partial f}{\partial q_c}j^\mu - \frac{\partial f}{\partial q_d}k^\mu\right)g_a + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}i^\mu + \frac{\partial f}{\partial q_c}j^\mu + \frac{\partial f}{\partial q_d}k^\mu\right)i^\mu g_b\\ &\quad + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b}i^\mu - \frac{\partial f}{\partial q_c}j^\mu + \frac{\partial f}{\partial q_d}k^\mu\right)j^\mu g_c + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} + \frac{\partial f}{\partial q_b}i^\mu + \frac{\partial f}{\partial q_c}j^\mu - \frac{\partial f}{\partial q_d}k^\mu\right)k^\mu g_d, \end{aligned} \tag{B 6}$$

where definition 4.1 and (2.4) were used in the last equality above. Grouping together $\partial f/\partial q_a$, $\partial f/\partial q_b$, $\partial f/\partial q_c$ and $\partial f/\partial q_d$ in (B 7) finally yields

$$\begin{aligned}\frac{\partial(fg)}{\partial q^\mu} &= f\frac{\partial g}{\partial q^\mu} + \frac{1}{4}\frac{\partial f}{\partial q_a}\left(g_a + i^\mu g_b + j^\mu g_c + k^\mu g_d\right) - \frac{1}{4}\frac{\partial f}{\partial q_b}\left(g_a + i^\mu g_b + j^\mu g_c + k^\mu g_d\right)i^\mu\\ &\quad - \frac{1}{4}\frac{\partial f}{\partial q_c}\left(g_a + i^\mu g_b + j^\mu g_c + k^\mu g_d\right)j^\mu - \frac{1}{4}\frac{\partial f}{\partial q_d}\left(g_a + i^\mu g_b + j^\mu g_c + k^\mu g_d\right)k^\mu\\ &= f\frac{\partial g}{\partial q^\mu} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a}g - \frac{\partial f}{\partial q_b}gi^\mu - \frac{\partial f}{\partial q_c}gj^\mu - \frac{\partial f}{\partial q_d}gk^\mu\right)\\ &= f\frac{\partial g}{\partial q^\mu} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}gi^\mu g^{-1} - \frac{\partial f}{\partial q_c}gj^\mu g^{-1} - \frac{\partial f}{\partial q_d}gk^\mu g^{-1}\right)g\\ &= f\frac{\partial g}{\partial q^\mu} + \frac{1}{4}\left(\frac{\partial f}{\partial q_a} - \frac{\partial f}{\partial q_b}i^{g\mu} - \frac{\partial f}{\partial q_c}j^{g\mu} - \frac{\partial f}{\partial q_d}k^{g\mu}\right)g = f\frac{\partial g}{\partial q^\mu} + \frac{\partial f}{\partial q^{g\mu}}g, \end{aligned} \tag{B 7}$$

where (2.4) was used in the second to last equality above, and definition 4.1 was used in the last equality. Hence, the first part of the theorem follows, and the second part can be proved in an analogous manner. ■

Appendix C. The proof of the chain rule

To prove the chain rule, we shall use the following lemma.

Lemma C.1. *Let $q = q_a + iq_b + jq_c + kq_d$, where $q_a, q_b, q_c, q_d \in \mathbb{R}$. Then, the partial derivatives of the quaternion composite function $f(g(q))$ satisfy the following chain rule:*

$$\left.\begin{aligned}\frac{\partial f(g(q))}{\partial \xi} &= \frac{\partial f}{\partial g^{\nu}}\frac{\partial g^{\nu}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial \xi}\\ \text{and}\quad \frac{\partial f(g(q))}{\partial \xi} &= \frac{\partial f}{\partial g^{\nu *}}\frac{\partial g^{\nu *}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu i *}}\frac{\partial g^{\nu i *}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu j *}}\frac{\partial g^{\nu j *}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu k *}}\frac{\partial g^{\nu k *}}{\partial \xi},\end{aligned}\right\} \tag{C 1}$$

where $\xi \in \{q_a, q_b, q_c, q_d\}$ and $\nu \in \mathbb{H}$, $\nu \neq 0$.

Proof. Let $g(q) = g_a + ig_b + jg_c + kg_d$, where $g_a, g_b, g_c, g_d \in \mathbb{R}$. Then, the function $f(g(q))$ can be seen as a function of the four real-valued variables g_a, g_b, g_c and g_d, and the partial derivative of $f(g(q))$ can be expressed as

$$\frac{\partial f}{\partial \xi} = \frac{\partial f}{\partial g_a}\frac{\partial g_a}{\partial \xi} + \frac{\partial f}{\partial g_b}\frac{\partial g_b}{\partial \xi} + \frac{\partial f}{\partial g_c}\frac{\partial g_c}{\partial \xi} + \frac{\partial f}{\partial g_d}\frac{\partial g_d}{\partial \xi}. \tag{C 2}$$

By definition 4.1, the partial derivatives $\partial f/\partial g_a$, $\partial f/\partial g_b$, $\partial f/\partial g_c$ and $\partial f/\partial g_d$ are given by

$$\left.\begin{aligned}\frac{\partial f}{\partial g_a} &= \frac{\partial f}{\partial g^{\nu}} + \frac{\partial f}{\partial g^{\nu i}} + \frac{\partial f}{\partial g^{\nu j}} + \frac{\partial f}{\partial g^{\nu k}},\\ \frac{\partial f}{\partial g_b} &= \left(\frac{\partial f}{\partial g^{\nu}} + \frac{\partial f}{\partial g^{\nu i}} - \frac{\partial f}{\partial g^{\nu j}} - \frac{\partial f}{\partial g^{\nu k}}\right) i^{\nu},\\ \frac{\partial f}{\partial g_c} &= \left(\frac{\partial f}{\partial g^{\nu}} - \frac{\partial f}{\partial g^{\nu i}} + \frac{\partial f}{\partial g^{\nu j}} - \frac{\partial f}{\partial g^{\nu k}}\right) j^{\nu}\\ \text{and}\quad \frac{\partial f}{\partial g_d} &= \left(\frac{\partial f}{\partial g^{\nu}} - \frac{\partial f}{\partial g^{\nu i}} - \frac{\partial f}{\partial g^{\nu j}} + \frac{\partial f}{\partial g^{\nu k}}\right) k^{\nu}.\end{aligned}\right\} \tag{C 3}$$

Upon applying the quaternion rotation transform $(\cdot)^{\nu}$ to both sides of (2.7) and replacing q with g, the real-valued components g_a, g_b, g_c and g_d can be expressed as

$$\left.\begin{aligned} g_a &= \frac{1}{4}(g^{\nu} + g^{\nu i} + g^{\nu j} + g^{\nu k}), \quad g_b = -\frac{i^{\nu}}{4}(g^{\nu} + g^{\nu i} - g^{\nu j} - g^{\nu k}),\\ g_c &= -\frac{j^{\nu}}{4}(g^{\nu} - g^{\nu i} + g^{\nu j} - g^{\nu k}), \quad g_d = -\frac{k^{\nu}}{4}(g^{\nu} - g^{\nu i} - g^{\nu j} + g^{\nu k}).\end{aligned}\right\} \tag{C 4}$$

Taking the partial derivatives of both sides of the above equations yields

$$\left.\begin{aligned}\frac{\partial g_a}{\partial \xi} &= \frac{1}{4}\left(\frac{\partial g^{\nu}}{\partial \xi} + \frac{\partial g^{\nu i}}{\partial \xi} + \frac{\partial g^{\nu j}}{\partial \xi} + \frac{\partial g^{\nu k}}{\partial \xi}\right),\\ \frac{\partial g_b}{\partial \xi} &= -\frac{i^{\nu}}{4}\left(\frac{\partial g^{\nu}}{\partial \xi} + \frac{\partial g^{\nu i}}{\partial \xi} - \frac{\partial g^{\nu j}}{\partial \xi} - \frac{\partial g^{\nu k}}{\partial \xi}\right),\\ \frac{\partial g_c}{\partial \xi} &= -\frac{j^{\nu}}{4}\left(\frac{\partial g^{\nu}}{\partial \xi} - \frac{\partial g^{\nu i}}{\partial \xi} + \frac{\partial g^{\nu j}}{\partial \xi} - \frac{\partial g^{\nu k}}{\partial \xi}\right)\\ \text{and}\quad \frac{\partial g_d}{\partial \xi} &= -\frac{k^{\nu}}{4}\left(\frac{\partial g^{\nu}}{\partial \xi} - \frac{\partial g^{\nu i}}{\partial \xi} - \frac{\partial g^{\nu j}}{\partial \xi} + \frac{\partial g^{\nu k}}{\partial \xi}\right).\end{aligned}\right\} \tag{C 5}$$

Upon substituting the results from (C 3) and (C 5) into (C 2), and using the distributive law and merging similar terms, we finally have

$$\frac{\partial f(g(q))}{\partial \xi} = \frac{\partial f}{\partial g^{\nu}}\frac{\partial g^{\nu}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial \xi} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial \xi}. \tag{C 6}$$

Hence, the first equality of the lemma follows, the second equality following in a similar way. ■

The proof of theorem 4.16. Using definition 4.1, the left HR derivative of the product fg can be expressed as

$$\frac{\partial f(g(q))}{\partial q^\mu} = \frac{1}{4}\left(\frac{\partial f(g(q))}{\partial q_a} - \frac{\partial f(g(q))}{\partial q_b} i^\mu - \frac{\partial f(g(q))}{\partial q_c} j^\mu - \frac{\partial f(g(q))}{\partial q_d} k^\mu\right). \tag{C 7}$$

Upon substituting the first equality of lemma C.1 into (C 7), it then follows that

$$\begin{aligned}\frac{\partial f(g(q))}{\partial q^\mu} = {} & \frac{1}{4}\left(\frac{\partial f}{\partial g^\nu}\frac{\partial g^\nu}{\partial q_a} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q_a} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q_a} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q_a}\right) \\ & - \frac{1}{4}\left(\frac{\partial f}{\partial g^\nu}\frac{\partial g^\nu}{\partial q_b} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q_b} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q_b} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q_b}\right) i^\mu \\ & - \frac{1}{4}\left(\frac{\partial f}{\partial g^\nu}\frac{\partial g^\nu}{\partial q_c} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q_c} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q_c} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q_c}\right) j^\mu \\ & - \frac{1}{4}\left(\frac{\partial f}{\partial g^\nu}\frac{\partial g^\nu}{\partial q_d} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q_d} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q_d} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q_d}\right) k^\mu. \end{aligned} \tag{C 8}$$

Grouping the terms $\partial f/\partial g^\nu$, $\partial f/\partial g^{\nu i}$, $\partial f/\partial g^{\nu j}$ and $\partial f/\partial g^{\nu k}$ in (C 8) together, we have

$$\begin{aligned}\frac{\partial f(g(q))}{\partial q^\mu} = {} & \frac{\partial f}{\partial g^\nu}\frac{1}{4}\left(\frac{\partial g^\nu}{\partial q_a} - \frac{\partial g^\nu}{\partial q_b} i^\mu - \frac{\partial g^\nu}{\partial q_c} j^\mu - \frac{\partial g^\nu}{\partial q_d} k^\mu\right) \\ & + \frac{\partial f}{\partial g^{\nu i}}\frac{1}{4}\left(\frac{\partial g^{\nu i}}{\partial q_a} - \frac{\partial g^{\nu i}}{\partial q_b} i^\mu - \frac{\partial g^{\nu i}}{\partial q_c} j^\mu - \frac{\partial g^{\nu i}}{\partial q_d} k^\mu\right) \\ & + \frac{\partial f}{\partial g^{\nu j}}\frac{1}{4}\left(\frac{\partial g^{\nu j}}{\partial q_a} - \frac{\partial g^{\nu j}}{\partial q_b} i^\mu - \frac{\partial g^{\nu j}}{\partial q_c} j^\mu - \frac{\partial g^{\nu j}}{\partial q_d} k^\mu\right) \\ & + \frac{\partial f}{\partial g^{\nu k}}\frac{1}{4}\left(\frac{\partial g^{\nu k}}{\partial q_a} - \frac{\partial g^{\nu k}}{\partial q_b} i^\mu - \frac{\partial g^{\nu k}}{\partial q_c} j^\mu - \frac{\partial g^{\nu k}}{\partial q_d} k^\mu\right) \\ = {} & \frac{\partial f}{\partial g^\nu}\frac{\partial g^\nu}{\partial q^\mu} + \frac{\partial f}{\partial g^{\nu i}}\frac{\partial g^{\nu i}}{\partial q^\mu} + \frac{\partial f}{\partial g^{\nu j}}\frac{\partial g^{\nu j}}{\partial q^\mu} + \frac{\partial f}{\partial g^{\nu k}}\frac{\partial g^{\nu k}}{\partial q^\mu}. \end{aligned} \tag{C 9}$$

Hence, the first equality of the theorem follows, the other equalities following in a similar fashion. ■

Appendix D. Fundamental results based on the generalized HR derivatives

For convenience, several of the most important results for the left GHR derivatives are summarized in table 1. The symbols ν, ω and λ denote quaternion constants, q is a quaternion-valued variable and μ any quaternion constant or expression.

References

1. Pei SC, Cheng CM. 1999 Color image processing by using binary quaternion-moment-preserving thresholding technique. *IEEE Trans. Image Process.* **8**, 614–628. (doi:10.1109/83.760310)
2. Bulow T, Sommer G. 2001 Hypercomplex signals: a novel extension of the analytic signal to the multidimensional case. *IEEE Trans. Signal Process.* **49**, 2844–2852. (doi:10.1109/78.960432)
3. Fortuna L, Muscato G, Xibilia M. 2001 A comparison between HMLP and HRBF for attitude control. *IEEE Trans. Neural Netw.* **12**, 318–328. (doi:10.1109/72.914526)
4. Took CC, Mandic DP. 2009 The quaternion LMS algorithm for adaptive filtering of hypercomplex processes. *IEEE Trans. Signal Process.* **57**, 1316–1327. (doi:10.1109/TSP.2008.2010600)
5. Le Bihan N, Mars J. 2004 Singular value decomposition of quaternion matrices: a new tool for vector-sensor signal processing. *Signal Process.* **84**, 1177–1199. (doi:10.1016/j.sigpro.2004.04.001)
6. Buchholz S, Le Bihan N. 2008 Polarized signal classification by complex and quaternionic multi-layer perceptrons. *Int. J. Neural Syst.* **18**, 75–85. (doi:10.1142/S0129065708001403)
7. Hamilton WR. 1847 On quaternions. *Proc. R. Irish Acad.* **4**, 1–16.
8. Xia Y, Jahanchahi C, Mandic DP. 2015 Quaternion-valued echo state networks. *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 663–673. (doi:10.1109/TNNLS.2014.2320715)
9. Mizoguchi T, Yamada I. 2014 An algebraic translation of Cayley-Dickson linear systems and its applications to online learning. *IEEE Trans. Signal Process.* **62**, 1438–1453. (doi:10.1109/TSP.2013.2296881)
10. Sudbery A. 1979 Quaternionic analysis. *Math. Proc. Camb. Phil. Soc.* **85**, 199–225. (doi:10.1017/S0305004100055638)
11. Deavours CA. 1973 The quaternion calculus. *Amer. Math. Mon.* **80**, 995–1008. (doi:10.2307/2318774)
12. De Leo S, Rotelli PP. 2003 Quaternionic analyticity. *App. Math. Lett.* **16**, 1077–1081. (doi:10.1016/S0893-9659(03)90097-8)
13. Mandic DP, Jahanchahi C, Took CC. 2011 A quaternion gradient operator and its applications. *IEEE Signal Process. Lett.* **18**, 47–50. (doi:10.1109/LSP.2010.2091126)
14. Gentili G, Struppa DC. 2006 A new approach to Cullen-regular functions of a quaternionic variable. *C. R. Math. Acad. Sci. Paris* **342**, 741–744. (doi:10.1016/j.crma.2006.03.015)
15. Brandwood D. 1983 A complex gradient operator and its application in adaptive array theory. *IEEE*

Commun. Radar Signal Process. **130**, 11–16. (doi:10.1049/ip-f-1.1983.0003)

16. Mandic DP, Goh SL. 2009 *Complex valued nonlinear adaptive filters: noncircularity, widely linear and neural models*. London, UK: John Wiley and Sons Ltd.
17. Kreutz-Delgado K. 2009 The complex gradient operator and the CR Calculus. (http://arxiv.org/abs/0906.4835)
18. Schwartz C. 2009 Calculus with a quaternionic variable. *J. Math. Phys.* **50**, 013523. (doi:10.1063/1.3058642)
19. Ward JP. 1997 *Quaternions and Cayley numbers: algebra and applications*. London, UK: Kluwer Academic Publishers.
20. Xu D, Mandic DP. 2015 The theory of quaternion matrix derivatives. *IEEE Trans. Signal Process*. **63**, 1543–1556. (doi:10.1109/TSP.2015.2399865)
21. Ell TA, Sangwine SJ. 2007 Quaternion involutions and anti-involutions. *Comput. Math. Appl.* **53**, 137–143. (doi:10.1016/j.camwa.2006.10.029)
22. Navarro-Moreno J, Fernández-Alcalá RM, Ruiz-Molina JC. 2014 A quaternion widely linear model for nonlinear Gaussian estimation. *IEEE Trans. Signal Process*. **62**, 6414–6424. (doi:10.1109/TSP.2014.2364790)
23. Luna-Elizarrarás ME, Shapiro M. 2011 A survey on the (hyper-) derivatives in complex, quaternionic and Clifford analysis. *Milan J. Math*. **79**, 521–542. (doi:10.1007/s00032-011-0169-0)
24. Fueter R. 1934/5 Die Funktionentheorie der Differentialgleichungen $\Delta u = 0$ und $\Delta\Delta u = 0$ mit vier reellen Variablen. *Comment. Math. Helv.* **7**, 307–330. (doi:10.1007/BF01292723)
25. Fueter R. 1935/6 Über die analytische Darstellung der regulären Funktionen einer Quaternionenvariablen. *Math. Helv.* **8**, 371–378. (doi:10.1007/BF01199562)
26. Gentili G, Struppa DC. 2007 A new theory of regular functions of a quaternionic variable. *Adv. Math*. **216**, 279–301. (doi:10.1016/j.aim.2007.05.010)
27. Colombo F, Sabadini I, Struppa DC. 2009 Slice monogenic functions. *Israel J. Math*. **171**, 385–403. (doi:10.1007/s11856-009-0055-4)
28. Ujang BC, Took CC, Mandic DP. 2011 Quaternion valued nonlinear adaptive filtering. *IEEE Trans. Neural Netw.* **22**, 1193–1206. (doi:10.1109/TNN.2011.2157358)
29. Colombo F, Sabadini I, Struppa DC. 2011 *Noncommutative functional calculus: theory and applications of slice hyperholomorphic functions*. Basel, Switzerland: Birkhäuser.
30. Gentili G, Stoppato C, Struppa DC. 2013 *Regular functions of a quaternionic variable*. Heidelberg, Germany: Springer.
31. Wirtinger W. 1927 Zur formalen theorie der funktionen von mehr komplexen veränderlichen. *Math. Ann.* **97**, 357–375. (doi:10.1007/BF01447872)
32. Vía J, Ramírez D, Santamaría I. 2010 Properness and widely linear processing of quaternion random vectors. *IEEE Trans. Inform. Theory* **56**, 3502–3515. (doi:10.1109/TIT.2010.2048440)
33. Isokawa T, Nishimura H, Matsui N. 2012 Quaternionic multilayer perceptron with local analyticity. *Information* **3**, 756–770. (doi:10.3390/info3040756)
34. Apostol TM 1967 *Calculus*, vol. I, 2nd edn. London, UK: John Wiley and Sons Ltd.
35. Took CC, Mandic DP. 2010 A quaternion widely linear adaptive filter. *IEEE Trans. Signal Process*. **58**, 4427–4431. (doi:10.1109/TSP.2010.2048323)
36. Xu D, Xia Y, Mandic DP. In press. Optimization in quaternion dynamic systems: gradient, Hessian, and learning algorithms. *IEEE Trans. Neural Netw. Learn. Syst.* (doi:10.1109/TNNLS.2015.2440473)
37. Ujang BC, Took CC, Mandic DP. 2010 Split quaternion nonlinear adaptive filtering. *Neural Netw.* **23**, 426–434. (doi:10.1016/j.neunet.2009.10.006)
38. Goh SL, Mandic DP. 2007 An augmented CRTRL for complex-valued recurrent neural networks. *Neural Netw.* **20**, 1061–1066. (doi:10.1016/j.neunet.2007.09.015)
39. Widrow B, McCool J, Ball M. 1975 The complex LMS algorithm. *Proc. IEEE* **63**, 719–720. (doi:10.1109/PROC.1975.9807)
40. Jahanchahi C, Mandic DP. 2013 A unifying approach to quaternion adaptive filtering: addressing the gradient and convergence. (http://arxiv.org/abs/1310.5612)