# A Quaternion Gradient Operator and Its Applications

D. P. Mandic, *Senior Member, IEEE*, C. Jahanchahi, and C. Cheong Took, *Member, IEEE*

*Abstract*—**Real functions of quaternion variables are typical cost functions in quaternion valued statistical signal processing, however, standard differentiability conditions in the quaternion domain do not permit direct calculation of their gradients. To this end, based on the isomorphism with real vectors and the use of quaternion involutions, we introduce the $\mathbb{HR}$ calculus as a convenient way to calculate derivatives of such functions. It is shown that the maximum change of the gradient is in the direction of the conjugate gradient, which conforms with the corresponding solution in the complex domain. Examples in some typical gradient based optimization settings support the result.**

*Index Terms*—**Conjugate gradient, HR calculus, quaternion gradient, quaternion LMS (QLMS), quaternion Wiener filter.**

## I. Introduction

**Q**UATERNION valued signal processing is a rapidly developing area, with applications across statistical signal processing, including adaptive filtering, neural networks, modelling of motion and tracking [1]–[3]. Recently introduced mathematical tools to support these developments include the quaternionic singular value decomposition (SVD) [4] and widely linear modelling [5], [6], however, in gradient based optimization the lack of analyticity of general quaternion valued functions remains an open issue. Existing generalizations of the Cauchy–Riemann equations to the quaternion domain—the Genaralised Cauchy–Riemann (GCR) and Cauchy–Riemann–Fueter (CRF) conditions [7], [8] are both too stringent and limit the class of quaternionic analytic functions to linear functions and constants, respectively.

In complex- and quaternion-valued statistical signal processing, a common optimization objective is to minimize a positive real function of complex/quaternion variables, typically in the form of error power, $J = |e|^2$. This poses a problem in both the complex, $\mathbb{C}$, and quaternion domain, $\mathbb{H}$, since the standard derivatives are not defined for functions dependent on both the variable $e$ and its conjugate $e^*$, such as the real valued cost function $J(e, e^*) = |e|^2 = ee^*$. The so called pseudogradient, that is, a sum of componentwise gradients is often employed to circumvent this problem; for complex valued gradients this is justified by Brandwood's result [9] and is further formalized within the $\mathbb{CR}$ calculus (for more detail see [10]), however, the corresponding result for quaternions is still missing.

The $\mathbb{CR}$ (or Wirtinger) calculus was developed based on the isomorphism between the fields $\mathbb{R}^2$ and $\mathbb{C}$ and the corresponding duality of total differentials of complex valued

functions $f(z, z^*) = u(x, y) + \jmath v(x, y) \in \mathbb{C}$ and real bivariate functions $g(x, y) \in \mathbb{R}^2$, and is proven to be a powerful framework comprising under the same umbrella both the standard Cauchy–Riemann differentiation (via the $\partial/\partial z$ derivative) and the differentiation of real functions of complex variables (via the $\partial/\partial z^*$ derivative, also called the $\mathbb{R}^*$ derivative) [10].

Our aim is to establish the duality between the derivatives of quaternion valued functions in $\mathbb{H}$ and the corresponding quadrivariate real functions in $\mathbb{R}^4$, in order to obtain a hypercomplex extension of the $\mathbb{CR}$ calculus termed the $\mathbb{HR}$ calculus. It is shown that this framework allows for calculating derivatives of both holomorphic quaternionic functions and nonholomorphic real functions of quaternion variables, typical cost functions. We also illustrate that in the context of stochastic gradient optimization the maximum change of a real function of quaternion vector variable is in the direction of conjugate gradient. This provides a generic extension of the corresponding result in the complex domain, and a rigorous framework for general gradient based optimization in $\mathbb{H}$.

## II. Quaternionic Linear Modelling

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

$$\mathbb{H} = \{q_a + \imath q_b + \jmath q_c + \kappa q_d \,|\, q_a, q_b, q_c, q_d \in \mathbb{R}\}$$

where the imaginary units $\imath, \jmath$ and $\kappa$ are also unit axis vectors, for which $\imath^2 = \jmath^2 = \kappa^2 = \imath\jmath\kappa = -1$. For any quaternion

$$q = q_a + \imath q_b + \jmath q_c + \kappa q_d = Sq + Vq \tag{1}$$

the scalar (real) part is denoted by $q_a = Sq = \Re(q)$, whereas the vector part (also called pure quaternion) $Vq = \Im(q) = \imath q_b + \jmath q_c + \kappa q_d$ comprises the three imaginary parts. The quaternion product for $q_1, q_2 \in \mathbb{H}$ is given by

$$q_1 q_2 = Sq_1 Sq_2 - Vq_1 \cdot Vq_2 + Sq_2 Vq_1 + Sq_1 Vq_2 + Vq_1 \times Vq_2$$

where the symbol '$\cdot$' denotes the scalar product and '$\times$' the vector product. Due to the presence of vector product, the quaternion product is noncommutative, that is, $q_1 q_2 \neq q_2 q_1$ and, e.g., $\imath\jmath = -\jmath\imath = \kappa, \jmath\kappa = -\kappa\jmath = \imath, \kappa\imath = -\imath\kappa = \jmath$, whereas the scalar product $q_1 \cdot q_2 = \langle q_1, q_2 \rangle$ is defined as

$$q_1 \cdot q_2 = q_{1a}q_{2a} + q_{1b}q_{2b} + q_{1c}q_{2c} + q_{1d}q_{2d} = \Re(q_1 q_2^*) = \Re(q_1^* q_2).$$

The quaternion conjugate is given by $q^* = Sq - Vq$, and the norm by $\| q \| = \sqrt{\langle q, q \rangle} = \sqrt{qq^*}$, and thus $q^{-1} = (q^*)/(qq^*) = (q^*)/(\| q \|^2) = (q^*)/(\langle q, q \rangle)$, and $\imath^{-1} = -\imath, \jmath^{-1} = -\jmath, \kappa^{-1} = -\kappa$.

### A. Equivalence Relations and Involutions

Similarity relation, denoted by "$\sim$" between quaternions $q_1$ and $q_2$ implies that $\exists \mu \in \mathbb{H}$, such that

$$q_1 \sim q_2 \quad \Leftrightarrow \quad q_1 = \mu q_2 \mu^{-1}, \quad \mu \neq 0. \tag{2}$$

It is an equivalence relation and $q_1 \sim q_2$ implies $\| q_1 \| = \| q_2 \|$ and $q_{1a} = q_{2a}$. Thus, for instance, the three imaginary units are similar, that is, $\imath \sim \jmath \sim \kappa$, and also $q \sim q^*$.

Equivalence relations of particular importance to this work are the quaternion involutions (self-inverse mappings) [11]

$$q^\imath = -\imath q \imath = q_a + \imath q_b - \jmath q_c - \kappa q_d$$
$$q^\jmath = -\jmath q \jmath = q_a - \imath q_b + \jmath q_c - \kappa q_d$$
$$q^\kappa = -\kappa q \kappa = q_a - \imath q_b - \jmath q_c + \kappa q_d. \qquad (3)$$

Notice that the quaternion conjugate is also an involution[1], that is, $(q^*)^* = q$. The four real components of a quaternion $q$ can now be expressed based on the involutions as

$$q_a = \frac{1}{4}[q + q^\imath + q^\jmath + q^\kappa] \qquad q_c = \frac{1}{4\jmath}[q - q^\imath + q^\jmath - q^\kappa]$$
$$q_b = \frac{1}{4\imath}[q + q^\imath - q^\jmath - q^\kappa] \qquad q_d = \frac{1}{4\kappa}[q - q^\imath - q^\jmath + q^\kappa] \qquad (4)$$

allowing any (either quadrivariate or quaternion-valued) function $g(q_a, q_b, q_c, q_d)$ of the four real variables $q_a, q_b, q_c, q_d$ to be expressed as a function of the quaternion variable $q$ and its perpendicular involutions[2] $q^\imath, q^\jmath, q^\kappa$ (or their conjugates) [11].

## III. THE $\mathbb{HR}$ DERIVATIVES

Similarly to the $\mathbb{CR}$ calculus [9], [10], in order to establish the duality between the derivatives of a quaternion valued function $f(q) \in \mathbb{H}$ and the derivatives of the corresponding 'composite' quadrivariate real function $g = g(q_a, q_b, q_c, q_d) \in \mathbb{R}^4$, we start from (since the fields $\mathbb{H}$ and $\mathbb{R}^4$ are isomorphic)

$$f(q) = f_a(q_a, q_b, q_c, q_d) + \imath f_b(q_a, q_b, q_c, q_d)$$
$$+ \jmath f_c(q_a, q_b, q_c, q_d) + \kappa f_d(q_a, q_b, q_c, q_d)$$
$$= g(q_a, q_b, q_c, q_d). \qquad (5)$$

The differential of the function[3] $g = g(q_a, q_b, q_c, q_d)$ is

$$dg = \frac{\partial g}{\partial q_a}dq_a + \frac{\partial g}{\partial q_b}dq_b + \frac{\partial g}{\partial q_c}dq_c + \frac{\partial g}{\partial q_d}dq_d$$
$$= \frac{\partial f(q)}{\partial q_a}dq_a + \imath\frac{\partial f(q)}{\partial q_b}dq_b + \jmath\frac{\partial f(q)}{\partial q_c}dq_c + \kappa\frac{\partial f(q)}{\partial q_d}dq_d. \qquad (6)$$

From (4) the derivatives of the components of a quaternion

$$dq_a = \frac{1}{4}[dq + dq^\imath + dq^\jmath + dq^\kappa] \qquad dq_c = \frac{-\jmath}{4}[dq - dq^\imath + dq^\jmath - dq^\kappa]$$
$$dq_b = \frac{-\imath}{4}[dq + dq^\imath - dq^\jmath - dq^\kappa] \qquad dq_d = \frac{-\kappa}{4}[dq - dq^\imath - dq^\jmath + dq^\kappa]$$

[1]Involutions are generalizations of the complex conjugate operator, and are applied component-wise to every imaginary component of a quaternion.

[2]We can also use the relations $q_a = (1)/(2)(q + q^*)$, $q_b = (1)/(2\imath)(q - q^{\imath *})$, $q_c = (1)/(2\jmath)(q - q^{\jmath *})$, $q_d = (1)/(2\kappa)(q - q^{\kappa *})$, however, this way we lose two degrees of freedom, rendering this approach inappropriate for analyzing the duality with the gradients of general quadrivariate functions, addressed in Section III.

[3]Function $g \in \mathbb{R}^4$ is a function of four real variables $q_a, q_b, q_c, q_d \in \mathbb{R}$.

and since (also see the Appendix)

$$f(q) = f(q_a, q_b, q_c, q_d) = f(q, q^\imath, q^\jmath, q^\kappa) \qquad (7)$$

we have

$$df(q) = \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q}dq + \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\imath}dq^\imath$$
$$+ \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\jmath}dq^\jmath + \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\kappa}dq^\kappa \qquad (8)$$

For a real quadrivariate function $g(q_a, q_b, q_c, q_d) \in \mathbb{R}^4$ that is differentiable with respect to the real $q_a, \ldots, q_d$, from (6)–(8)

$$dg(q_a, q_b, q_c, q_d) = Adq + Bdq^\imath + Cdq^\jmath + Ddq^\kappa. \qquad (9)$$

Finally, solving for the factors $A, B, C, D$, from (6) we obtain the set of identities which we refer to as the $\mathbb{HR}$-derivatives. The derivative $\partial f(q, q^\imath, q^\jmath, q^\kappa)/\partial q$ is of particular interest and is given in (10) shown at the bottom of the page; the complete set of $\mathbb{HR}$ derivatives is

$$\begin{bmatrix} \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q} \\ \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\imath} \\ \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\jmath} \\ \frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q^\kappa} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & -\imath & -\jmath & -\kappa \\ 1 & -\imath & \jmath & \kappa \\ 1 & \imath & -\jmath & \kappa \\ 1 & \imath & \jmath & -\kappa \end{bmatrix}\begin{bmatrix} \partial f/\partial q_a \\ \partial f/\partial q_b \\ \partial f/\partial q_c \\ \partial f/\partial q_d \end{bmatrix} \qquad (12)$$

where we have used a shorthand notation $f = f(q_a, q_b, q_c, q_d)$.

### A. The $\mathbb{HR}^*$ Derivatives

In analogy to the $\mathbb{R}^*$ derivatives in the $\mathbb{CR}$ calculus, to arrive at the $\mathbb{HR}^*$ derivatives, we can express the real-valued components $q_a, q_b, q_c, q_d$ of a quaternion $q$ in terms of the quaternion conjugate $q^*$ and the conjugate involutions $q^{\imath *}, q^{\jmath *}, q^{\kappa *}$ (for an overview of the augmented bases, see [5])

$$q_a = \frac{1}{4}[q^* + q^{\imath *} + q^{\jmath *} + q^{\kappa *}] \qquad q_c = \frac{1}{4\jmath}[-q + q^{\imath *} - q^{\jmath *} + q^{\kappa *}]$$
$$q_b = \frac{1}{4\imath}[-q - q^{\imath *} + q^{\jmath *} + q^{\kappa *}] \qquad q_d = \frac{1}{4\kappa}[-q^* + q^{\imath *} + q^{\jmath *} - q^{\kappa *}].$$

In this conjugate basis, the derivative of the function $f(q) = f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})$ can be written as

$$df(q) = \frac{\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})}{\partial q^*}dq^* + \frac{\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})}{\partial q^{\imath *}}dq^{\imath *}$$
$$+ \frac{\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})}{\partial q^{\jmath *}}dq^{\jmath *} + \frac{\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})}{\partial q^{\kappa *}}dq^{\kappa *}$$

and thus the derivative of the quadrivariate $g(q_a, q_b, q_c, q_d)$ is

$$dg(q_a, q_b, q_c, q_d) = \bar{A}dq^* + \bar{B}dq^{\imath *} + \bar{C}dq^{\jmath *} + \bar{D}dq^{\kappa *}. \qquad (13)$$

In the same spirit as above, solving for factors $\bar{A}, \ldots, \bar{D}$ yields the $\mathbb{HR}^*$ derivatives; the derivative $\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})/\partial q^*$

$$\frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q} = \frac{1}{4}\left[\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_a} - \imath\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_b} - \jmath\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_c} - \kappa\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_d}\right] \qquad (10)$$

$$\frac{\partial f(q^*, q^{\imath *}, q^{\jmath *}, q^{\kappa *})}{\partial q^*} = \frac{1}{4}\left[\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_a} + \imath\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_b} + \jmath\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_c} + \kappa\frac{\partial f(q_a, q_b, q_c, q_d)}{\partial q_d}\right] \qquad (11)$$

is given in (11), and the complete set of $\mathbb{HR}^*$ derivatives in a vector-matrix form becomes

$$
\begin{bmatrix}
\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})/\partial q^* \\
\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})/\partial q^{\imath*} \\
\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})/\partial q^{\jmath*} \\
\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})/\partial q^{\kappa*}
\end{bmatrix}
= \frac{1}{4}
\begin{bmatrix}
1 & \imath & \jmath & \kappa \\
1 & \imath & -\jmath & -\kappa \\
1 & -\imath & \jmath & -\kappa \\
1 & -\imath & -\jmath & \kappa
\end{bmatrix}
\nabla \mathbf{f}
$$

where $\nabla \mathbf{f} = [\partial f/\partial q_a, \partial f/\partial q_b, \partial f/\partial q_c, \partial f/\partial q_d]^T$. (14)

### B. Comparison With the $\mathbb{CR}$-Calculus

The derivative $(\partial f(q, q^\imath, q^\jmath, q^\kappa))/(\partial q) \in \mathbb{H}$ in (10) can be regarded as a generalization of the derivative $(\partial f(z, z^*))/(\partial z) \in \mathbb{C}$, however, there are significant differences:

○ *Functions $f(q, q^*)$*. The $\mathbb{HR}$ derivative $\partial f/\partial q$ can be directly applied to functions $f(q, q^*)$, since the substitution $q^* = (1/2)(q^\imath + q^\jmath + q^\kappa - q)$ shows that $q$ and $q^*$ and not independent and thus $\partial f(q, q^*)/\partial q \neq 0$.

○ *Placement of unit vectors $\imath, \jmath, \kappa$*. Notice that both the quaternion forms $\imath(\partial f(q_a, q_b, q_c, q_d))/(\partial q_a)$ and $(\partial f(q_a, q_b, q_c, q_d))/(\partial q_a)\imath$ are valid, since the quadarivariate function $g(q_a, q_b, q_c, q_d) \in \mathbb{R}^4$ from (6) is real and thus e.g., $(\partial g(q_a, q_b, q_c, q_d))/(\partial q_a)dq_a = dq_a(\partial g(q_a, q_b, q_c, q_d))/(\partial q_a)$. These terms, however, do not commute when replaced in the quaternion valued function $f$ in (5), and so we need to take care of the positioning.

○ *The use of involutions*. Owing to the similarity relation in (2), for real functions of quaternion variables it is advantageous to differentiate with respect to $q, q^\imath, q^\jmath, q^\kappa$ rather than directly with respect to $q_a, q_b, q_c, q_d$ (see the Appendix); this may also help with the uncertainty regarding the placement of the imaginary units.

### C. Examples—Quaternion Valued Functions

1) *Derivatives of the holomorphic $f(q) = q$*. Using the $\mathbb{HR}$ derivatives in (12) we obtain the partial derivatives for $f(q) = q$, as $\partial f(q)/\partial q = 1, \partial f(q)/\partial q^\imath = 0, \partial f(q)/\partial q^\jmath = 0, \partial f(q)/\partial q^\kappa = 0$. This is equivalent to the standard Cauchy–Riemann–Fueter (CRF) derivative, which gives $f'(q) = 1$.

2) *Derivatives of the nonholomorphic $f(q, q^*) = q^*$*. The four partial $\mathbb{HR}$ derivatives for $f(q) = q^*$ are $\partial f(q)/\partial q = -(1/2), \partial f(q)/\partial q^\imath = (1/2), \partial f(q)/\partial q^\jmath = (1/2), \partial f(q)/\partial q^\kappa = (1/2)$.
This is the main difference from the $\mathbb{CR}$ calculus, where the derivative $\partial z^*/\partial z = 0$; since $q^* = (1/2)(q^\imath + q^\jmath + q^\kappa - q)$, the $\mathbb{HR}$ derivative in (10) gives $\partial q^*/\partial q = -(1/2)$.

3) *Derivatives of involutions $f(q) = q^\eta, \eta = \{\imath, \jmath, \kappa\}$*. As desired, the $\mathbb{HR}$ derivatives give

$$
\frac{\partial q^\eta}{\partial q^\theta} = \begin{cases} 1, & \eta = \theta, \ \eta, \theta \in \{\imath, \jmath, \kappa\} \\ 0, & \eta \neq \theta, \ \eta, \theta \in \{\imath, \jmath, \kappa\}. \end{cases}
\quad (15)
$$

4) *The $\mathbb{HR}^*$ derivatives of $f(q) = q^*$*. In the same spirit as above, based on the $\mathbb{HR}^*$ derivative in (11) we have $(\partial q^*)/(\partial q^*) = 1$, since for the involutions $(\partial q^*)/(\partial q^{\eta*}) = 0, \eta \in \{\imath, \jmath, \kappa\}$.

## IV. EXTENSION OF BRANDWOOD'S RESULTS IN $\mathbb{H}$

In the complex domain, as shown in [9], [10], stationary points of a real function of complex variables $f(z, z^*)$ are defined by both $\partial f/\partial z = 0$ and $\partial f/\partial z^* = 0$. It was also

shown that the conjugate gradient provides the direction of the maximum steepness on the optimization surface, and so for instance, $\nabla_{\mathbf{w}^*} J$ is used for calculating the maximum gradient of the cost function $J$ with respect to the coefficient vector $\mathbf{w}$; this coincides with the so-called pseudogradient $\nabla_{\mathbf{w}} = \nabla_{\mathbf{w_r}} + \jmath \nabla_{\mathbf{w_i}}$ which was introduced without formal justification.

In the quaternion domain, as shown in (3), the role of the complex self-inverse $(\cdot)^*$ is played by the involutions, $q^\imath, q^\jmath, q^\kappa$, allowing for more than one equivalent augmented basis [5]. Of particular interest to this work are the bases

$$
\begin{aligned}
f(q) &= f(q_a, q_b, q_c, q_d) = f(q, q^\imath, q^\jmath, q^\kappa) \\
&= f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*}).
\end{aligned}
$$

Based on the $\mathbb{HR}$ and $\mathbb{HR}^*$ derivatives in (10) and (11), the stationary points of function $f(q)$ are obtained for $\partial f(q)/\partial q_a = \partial f(q)/\partial q_b = \partial f(q)/\partial q_c = \partial f(q)/\partial q_d = 0$, giving

$$
\frac{\partial f(q, q^\imath, q^\jmath, q^\kappa)}{\partial q} = 0 \text{ and } \frac{\partial f(q^*, q^{\imath*}, q^{\jmath*}, q^{\kappa*})}{\partial q^*} = 0. \quad (16)
$$

Thus, either $(\partial f)/(\partial q) = 0$ or $(\partial f)/(\partial q^*) = 0$ is a necessary and sufficient condition for the stationary points of a real function of quaternion variables.

### A. Maximum Change of the Quaternion Gradient

The first order multivariate Taylor series expansion of the function $f(q) = f(q, q^\imath, q^\jmath, q^\kappa)$ gives

$$
df(q) = \frac{\partial f(q)}{\partial q}dq + \frac{\partial f(q)}{\partial q^\imath}dq^\imath + \frac{\partial f(q)}{\partial q^\jmath}dq^\jmath + \frac{\partial f(q)}{\partial q^\kappa}dq^\kappa. \quad (17)
$$

Observe that $q^* = (1/2)(q^\imath + q^\jmath + q^\kappa - q)$ and thus

$$
2\frac{\partial f}{\partial q^*}dq^* + \frac{\partial f}{\partial q}dq = \frac{\partial f}{\partial q^\imath}dq^\imath + \frac{\partial f}{\partial q^\jmath}dq^\jmath + \frac{\partial f}{\partial q^\kappa}dq^\kappa
$$

allowing us to simplify the Taylor series expansion in (17) as[4]

$$
df = \frac{\partial f}{\partial q}dq + 2\left(\frac{\partial f}{\partial q}dq\right)^* + \frac{\partial f}{\partial q}dq = 4\Re\left(\frac{\partial f}{\partial q}dq\right). \quad (18)
$$

This result is readily applied to vectors $\mathbf{q} = [q_1, \ldots, q_N]^T$, as

$$
df(\mathbf{q}) = 2\nabla_{\mathbf{q}}^T f(\mathbf{q})d\mathbf{q} + 2d\mathbf{q}^* \nabla_{\mathbf{q}^*}^T f(\mathbf{q}) = 4\Re[\nabla_{\mathbf{q}}^T f(\mathbf{q})d\mathbf{q}]
$$

where $\nabla_{\mathbf{q}} f = [(\partial f)/(\partial q_1), \ldots, (\partial f)/(\partial q_N)]^T, \nabla_{\mathbf{q}^*} f = [(\partial f)/(\partial q_1^*), \ldots, (\partial f)/(\partial q_N^*)]^T$.

Since quaternions form a Hilbert space, for real functions of quaternion variables $(\partial f)/(\partial \mathbf{q}^*) = ((\partial f)/(\partial \mathbf{q}))^*$, and (18) becomes

$$
df(\mathbf{q}) = 4\Re[(\nabla_{\mathbf{q}^*} f(\mathbf{q}))^H d\mathbf{q}] = 4\langle \nabla_{\mathbf{q}^*} f(\mathbf{q}), d\mathbf{q}\rangle. \quad (19)
$$

According to the Schwarz inequality, the inner product of two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{H}$ satisfies $\langle \mathbf{a}, \mathbf{b}\rangle \leq \| \mathbf{a} \| \| \mathbf{b} \|$. The equality holds only when $\mathbf{a}$ is collinear with $\mathbf{b}$, that is, $\mathbf{a} = \gamma\mathbf{b}, \gamma \in \mathbb{R}$, and thus the maximum of the derivative $df(\mathbf{q})$ in (19) occurs when $d\mathbf{q}$ is collinear with the conjugate gradient $\nabla_{\mathbf{q}^*} f(\mathbf{q})$. This shows that the conjugate gradient $\nabla_{\mathbf{q}^*} f(\mathbf{q})$ defines the direction of the maximum rate of change of the function $f(\mathbf{q})$, thus making it a natural choice in gradient type optimization of real functions of quaternion variables.

[4]In the process, for a real $f(q)$ we have employed the similarity relation in (2) to obtain the identities $(\partial f(q))/(\partial q) = (\partial f^\eta(q))/(\partial q) = ((\partial f(q))/(\partial q^\eta))^\eta$ and $((\partial f(q))/(\partial q))^\eta = (\partial f(q))/(\partial q^\eta)$, for $\eta \in \{\imath, \jmath, \kappa\}$. The properties of the involution product $(q_1 q_2)^\eta = q_1^\eta q_2^\eta$, and the conjugate $(q_1 q_2)^* = q_2^* q_1^*$ were also used.

## V. Some Signal Processing Applications

In gradient based optimization problems in adaptive filtering we desire to minimize the real cost function $J(e, e^*) = E[ee^*]$, where the error $e = d - y$, $d$ is the desired signal, $y = \mathbf{w}^H \mathbf{q}$ the filter output, $E[\cdot]$ the statistical expectation operator, $\mathbf{w} = [w_1, \ldots, w_N]^T$ the vector of filter coefficients, $\mathbf{q}$ the input vector, and $(\cdot)^H$ the Hermitian transpose operator.

*Wiener filter.* To find the optimum weights $\mathbf{w}_o$, we expand

$$J(e, e^*) = E[dd^*] + E[yy^*] - E[yd^*] - E[dy^*]. \quad (20)$$

Based on the involutions in (3) and Example 2 in Section III-C, note that $\partial \mathbf{w}/\partial \mathbf{w}^* = -(1/2)\mathbf{I} \neq \mathbf{0}$, and thus

$$\begin{aligned}
\nabla_{\mathbf{w}^*} J &= E[\nabla_{\mathbf{w}^*}(dd^*)] + E[\nabla_{\mathbf{w}^*}(\mathbf{w}^H \mathbf{q}\mathbf{q}^H \mathbf{w})] \\
&\quad - E[\nabla_{\mathbf{w}^*}(\mathbf{w}^H \mathbf{q}d^*)] - E[\nabla_{\mathbf{w}^*}(d\mathbf{q}^H \mathbf{w})] \\
&= \mathbf{R}_{qq}\mathbf{w} - \frac{1}{2}\mathbf{w}^H \mathbf{R}_{qq} - \mathbf{r}_{qd} + \frac{1}{2}\mathbf{r}_{dq}^T \quad (21)
\end{aligned}$$

where $\mathbf{R}_{qq} = E[\mathbf{q}\mathbf{q}^H]$ and $\mathbf{r}_{qd} = E[\mathbf{q}d^*], \mathbf{r}_{dq} = E[d\mathbf{q}^*]$. Upon setting to zero, owing to limited algebraic manipulation due to the noncommutativity of the quaternion product, we obtain the Wiener solution in two parts, as in [5]. As a simplification, using the 'complex form' of quaternions, $q^* = Sq - Vq$, and thus, like in $\mathbb{C}$, assuming $q \perp q^*$, the two terms multiplied by $1/2$ in (21) vanish, giving the generic Wiener solution $\mathbf{w}_o = \mathbf{R}_{qq}^{-1}\mathbf{r}_{qd}$.

*The Quaternion LMS (QLMS).* For consistency with the QLMS algorithm in [1], consider the cost function $J(k) = e(k)e^*(k)$, and the filter output in the form $y(k) = \mathbf{w}^T(k)\mathbf{q}(k)$. The weight update is $\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \nabla J(k)$, and the $\mathbb{HR}$ calculus gives $\nabla J(k) = \nabla_{\mathbf{w}^*} J(k)$, to yield

$$\nabla_{\mathbf{w}^*} J(k) = e(k)\frac{\partial e^*(k)}{\partial \mathbf{w}^*(k)} + \frac{\partial e(k)}{\partial \mathbf{w}^*(k)}e^*(k). \quad (22)$$

The errors are given by $e(k) = d(k) - \mathbf{w}^T(k)\mathbf{q}(k)$ and $e^*(k) = d^*(k) - \mathbf{q}^H(k)\mathbf{w}^*(k)$, and thus the error gradients become

$$\frac{\partial e^*(k)}{\partial \mathbf{w}^*(k)} = -\mathbf{q}^*(k) \qquad \frac{\partial e(k)}{\partial \mathbf{w}^*(k)} = \frac{1}{2}\mathbf{q}(k).$$

As desired, this gives the QLMS update in the original form [1] (with a scalar constant absorbed in $\mu$)

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\left(e(k)\mathbf{q}^*(k) - \frac{1}{2}\mathbf{q}(k)e^*(k)\right). \quad (23)$$

As a simplification, using the complex form of a quaternion and thus formally assuming $\partial \mathbf{w}/\partial \mathbf{w}^* = \mathbf{0}$, we obtain the same generic form[5] for QLMS as that of the complex LMS (CLMS)

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e(k)\mathbf{q}^*(k)$$

*The widely linear QLMS (WLQLMS).* The WLQLMS [12] is based on the augmented basis in (25). Using the gradient defined as $\nabla J = \nabla_{\mathbf{w}^*} J$, we arrive directly at the WQLMS update in its augmented form (see the Appendix)

$$\mathbf{w}^a(k+1) = \mathbf{w}^a(k) + \mu\left(e(k)\mathbf{x}^*(k) - \frac{1}{2}\mathbf{x}(k)e^*(k)\right)$$

where $\mathbf{w}^a = [\mathbf{u}, \mathbf{v}, \mathbf{g}, \mathbf{h}]^T$ and $\mathbf{x} = [\mathbf{q}, \mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa]^T$.

---

[5] The same form is obtained for a gradient defined via involutions (3), that is, $\nabla J = \nabla_{\mathbf{w}^\eta} J$, $\eta \in \{\imath, \jmath, \kappa\}$, and assuming $\mathbf{w}$ and $\mathbf{w}^*$ are independent.

## VI. Conclusion

We have introduced the $\mathbb{HR}$ caluclus, a unifying framework for the calculation of gradients of both quaternion holomorphic functions and real functions of quaternion variables. This has been achieved by making use of perpendicular involutions, augmented bases, and isomorphism with real vectors. The analysis has shown that the direction of maximum change of the gradient of real functions of quaternion variables is in the direction of the conjugate gradient, conforming with the corresponding solution in the complex domain. The usefulness of the $\mathbb{HR}$ calculus has been illustrated on some typical gradient based optimization problems in signal processing.

## Appendix
## Quaternion Widely Linear Model

Consider the real-valued minimum mean square error (MMSE) estimator $\hat{y} = E[y \,|\, x]$ which estimates the values of signal $y$ in terms of another observation $x$. For zero mean, jointly normal $y$ and $x$, the solution is a linear estimator (where $\mathbf{x} = [x_1, \ldots, x_N]^T$ and $\mathbf{w}$ is the coefficient vector) given by

$$\hat{y} = E[y \,|\, x] \quad \rightarrow \quad \hat{y} = \mathbf{w}^T \mathbf{x}. \quad (24)$$

In the same spirit, the standard conditional MSE estimator in the quaternion domain is given by $\hat{y} = E[y \,|\, q], \forall q, y \in \mathbb{H}$. However, observe that $\hat{y}_{\{m\}} = E[y_{\{m\}} \,|\, q_a, q_b, q_c, q_d], \quad m \in \{a, b, c, d\}$. The *real valued* components $q_a, q_b, q_c, q_d$ can now be expressed via their *quaternion valued* perpendicular involutions in (4), to give $\hat{y} = E[y \,|\, q, q^\imath, q^\jmath, q^\kappa]$, thus leading to the *quaternion widely linear model* [6], [12]

$$\hat{y} = \mathbf{u}^T \mathbf{q} + \mathbf{v}^T \mathbf{q}^\imath + \mathbf{g}^T \mathbf{q}^\jmath + \mathbf{h}^T \mathbf{q}^\kappa = \mathbf{w}^{aT}\mathbf{x} \quad (25)$$

where $\mathbf{u}, \mathbf{v}, \mathbf{g}, \mathbf{h}$ are coefficient vectors, $\mathbf{q}^\eta, \eta \in \{\imath, \jmath, \kappa\}$ input involution vectors, $\mathbf{x} = [\mathbf{q}, \mathbf{q}^\imath, \mathbf{q}^\jmath, \mathbf{q}^\kappa]^T$, $\mathbf{w}^a = [\mathbf{u}, \mathbf{v}, \mathbf{g}, \mathbf{h}]^T$.

## References

[1] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex real world processes," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, 2009.

[2] P. Arena, L. Fortuna, G. Muscato, and M. G. Xibilia, *Neural Networks in Multidimensional Domains*. Berlin, Germany: Springer-Verlag, 1998.

[3] A. M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 7, pp. 1346–1356, 2006.

[4] N. L. Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing," *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, 2004.

[5] C. C. Took and D. P. Mandic, "Augmented second order statistics of quaternion random signals," *Signal Process.*, vol. 91, pp. 214–224, 2011.

[6] J. Via, D. Ramirez, and I. Santamaria, "Properness and widely linear processing of quaternion random vectors," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3502–3515, 2010.

[7] A. Sudberry, "Quaternionic analysis," *Math. Proc. Cambridge Philos. Soc.*, vol. 85, pp. 199–225, 1979.

[8] I. Frenkel and M. Libine, "Quaternionic analysis, representation theory, and physics," *Adv. Math.*, vol. 218, pp. 1806–1877, 2008.

[9] D. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Proc. Inst. Elect. Eng. F, Commun., Radar, Signal Processing*, vol. 130, no. 1, pp. 11–16, 1983.

[10] K. Kreutz-Delgado, The Complex Gradient Operator and the $\mathbb{CR}$ Calculus pp. 1–74, 2006, Lecture Supplement ECE275A.

[11] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Comput, Math. Applicat.*, vol. 53, pp. 137–143, 2007.

[12] C. C. Took and D. P. Mandic, "A quaternion widely linear adaptive filter," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4427–4431, 2010.