Hamburger Beiträge
zur Angewandten Mathematik

Given's Transformation Applied to Quaternion Valued Vectors

D. Janovská and G. Opfer

This paper has been submitted to BIT
The final form may differ from this preprint

Reihe A
Preprint 171
February 2003

Hamburger Beiträge zur Angewandten Mathematik

Reihe A	Preprints
Reihe B	Berichte
Reihe C	Mathematische Modelle und Simulation
Reihe D	Elektrische Netzwerke und Bauelemente
Reihe E	Scientific Computing
Reihe F	Computational Fluid Dynamics and Data Analysis

BIT
2003, Vol. xx, No. x, pp. 001–012
0006-3835/00/4004-0001 $15.00

GIVENS' TRANSFORMATION APPLIED TO QUATERNION VALUED VECTORS *

D. JANOVSKÁ[1] and G. OPFER[2]

2003/2/9

[1] *Institute of Chemical Technology, Prague, Department of Mathematics, Technická 5 166 28 Prague 6, Czech Republic. email: janovskd@vscht.cz*

[2] *University of Hamburg, Department of Mathematics, Bundesstraße 55 20146 Hamburg, Germany. email: opfer@math.uni-hamburg.de*

Abstract.

Givens' transformation [1954] was originally applied to real matrices. We shall give an extension to complex and quaternion valued matrices. We observe that the classical Givens rotation in the real and in the complex case is itself a quaternion using an isomorphism between certain (2×2) matrices and $\mathbb{R}^4$ equipped with the quaternion multiplication. In the real and complex case Givens' (2×2) matrix is determined uniquely up to an arbitrary (real or complex) factor σ with $|\sigma| = 1$. However, because of the noncommutativity of quaternions we shall show that in the quaternion case such a factor must obey certain additional restrictions. There are two numerical examples.

AMS subject classification: 12E15 65F25 65F30.

Key words: Givens rotation, quaternion valued vectors, quaternion valued matrices.

1 Introduction.

The ordinary Givens transformation is an orthogonal, real (2×2) matrix

$$\mathbf{G} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix} \tag{1.1}$$

(cf. GOLUB & VAN LOAN[2, p. 215]). For a given vector $\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{R}^2 \backslash \{\mathbf{0}\}$, the elements c and s of $\mathbf{G}$ are chosen in such a way that $\mathbf{G}$ is orthogonal, i.e. $c^2 + s^2 = 1$, and

$$\mathbf{G}^{\mathrm{T}}\mathbf{x} = \mathbf{u} = u\,\mathbf{e}_1 \quad \text{where} \quad \mathbf{e}_1 = (1,0)^{\mathrm{T}} \in \mathbb{R}^2 \quad \text{and} \quad u \in \mathbb{R}\backslash\{0\}. \tag{1.2}$$

If we set $c = \cos\alpha$, $s = \sin\alpha$ and $\mathbf{x} = r(\cos\varphi, \sin\varphi)^{\mathrm{T}}$, then $\mathbf{G}^{\mathrm{T}}\mathbf{x} = \mathbf{u} = r(\cos(\varphi+\alpha), \sin(\varphi+\alpha))^{\mathrm{T}}$. Thus, we can see the rotational effect of $\mathbf{G}$. Therefore, $\mathbf{G}$ is also called a Givens *rotation*. In the literature (e. g. STOER[7, p. 252]) one also finds instead of $\mathbf{G}$ the matrix

$$\widetilde{\mathbf{G}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \mathbf{G} = \begin{pmatrix} c & s \\ s & -c \end{pmatrix},$$

*Received November 2002. Communicated by Zdenek Strakoš.

which is called a Givens *reflection*. This form has the advantage that $\widetilde{\mathbf{G}}$ is symmetric. But since, in general, $s\overline{c} - \overline{s}c \neq 0$ for $s, c \in \mathbb{C}$, the matrix $\widetilde{\mathbf{G}}$ is neither an orthogonal, nor a unitary matrix in a general complex case.

Our aim is to compute the matrix $\mathbf{G}$ in (1.1) in such a way that $\mathbf{G}^{\mathrm{T}}$ rotates the given vector $\mathbf{x} \in \mathbb{R}^2\backslash\{\mathbf{0}\}$ in the positive orientation into the vector $\mathbf{u}$ which is positioned on the x_1-axis, i.e. the rotation angle α lies in $[0, \pi[$. In the case that $\mathbf{x}$ is already on the x_1-axis, i.e. $x_2 = 0$, $\mathbf{x}$ will remain unchanged. If $x_1 = 0$, i.e. the vector $\mathbf{x}$ is on the x_2-axis, then we rotate the vector $\mathbf{x} = (0, x_2)^{\mathrm{T}}$ into the vector $\mathbf{u} = (-x_2, 0)^{\mathrm{T}}$. Taking into account all these requirements, we have to solve the following set of equations ($\|\mathbf{x}\| \neq 0$):

$$\begin{array}{rcl} cx_1 - sx_2 & = & u \neq 0, \\ sx_1 + cx_2 & = & 0, \end{array} \qquad c^2 + s^2 = 1, \tag{1.3}$$

$$\mathbf{G}^{\mathrm{T}} \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix}, \qquad \mathbf{G}^{\mathrm{T}} \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ 0 \end{pmatrix}.$$

For the matrix $\mathbf{G}$ in (1.1), we obtain the following formulae ($\|\mathbf{x}\| \neq 0$):

$$s = \frac{|x_2|}{||\mathbf{x}||}, \; c = \begin{cases} 1 & \text{if } x_2 = 0, \\ -\dfrac{(\operatorname{sgn} x_2)x_1}{||\mathbf{x}||} & \text{else,} \end{cases} \quad u = \begin{cases} x_1 & \text{if } x_2 = 0, \\ -||\mathbf{x}||\operatorname{sgn} x_2 & \text{else.} \end{cases} \tag{1.4}$$

Let us remark that as a consequence of $\alpha \in [0, \pi[$ we have $s \geq 0$. If we do not insist on positively oriented rotations and we prefer rotations $\alpha \in] -\pi/2, \pi/2]$ (i.e. $c \geq 0$), we have to modify the formulae (1.4) (cf. Schwarz[6, p. 291], for the reflection case Stoer[7, p. 253]):

$$s = \begin{cases} 1 & \text{if } x_1 = 0, \\ -\dfrac{(\operatorname{sgn} x_1)x_2}{||\mathbf{x}||} & \text{else,} \end{cases} \quad c = \frac{|x_1|}{||\mathbf{x}||}, \; u = \begin{cases} -x_2 & \text{if } x_1 = 0, \\ ||\mathbf{x}||\operatorname{sgn} x_1 & \text{else.} \end{cases} \tag{1.5}$$

2 Givens' transformation in the complex case.

In order that the Givens transformation $\mathbf{G}$ also works in the complex case some changes are necessary[1]. In the first place we define

$$\mathbf{G} = \begin{pmatrix} \overline{c} & s \\ -\overline{s} & c \end{pmatrix}. \tag{2.1}$$

A complex matrix $\mathbf{G}$ is called *unitary* if $\mathbf{G}^*\mathbf{G} = \mathbf{I}$, where $\mathbf{G}^*$ is formed from $\mathbf{G}$ by transposition and complex conjugation of the elements and $\mathbf{I}$ is the identity matrix. The matrix $\mathbf{G}$ in (2.1) will be unitary if and only if $|s|^2 + |c|^2 = 1$. Assume throughout this section, that $\mathbf{x} \in \mathbb{C}^2\backslash\{\mathbf{0}\}$ is a given vector. The requirement

$$\mathbf{G}^*\mathbf{x} = \mathbf{u} = u\mathbf{e}_1 \quad \text{where} \quad \mathbf{e}_1 = (1, 0)^{\mathrm{T}} \in \mathbb{C}^2 \quad \text{and} \quad u \in \mathbb{C}\backslash\{0\} \tag{2.2}$$

[1] A short treatment of the Givens' transformation in the complex case can be found in the book of Greenbaum[3, p. 40], the Householder transformation in the complex case is treated in Opfer[5, p. 250].

yields the equations corresponding to (1.3):

$$(1.3')\qquad \begin{array}{rcl} cx_1 - sx_2 & = & u \neq 0, \\ \overline{s}x_1 + \overline{c}x_2 & = & 0. \end{array}$$

Because $|s|^2 + |c|^2 = 1$, the second equation in (1.3′) implies

$$(2.3)\qquad |s| = \frac{|x_2|}{||\mathbf{x}||}, \quad |c| = \frac{|x_1|}{||\mathbf{x}||},$$

which yields the general solution of (1.3′)

$$(2.4)\qquad s = -\sigma\frac{\overline{x_2}}{||\mathbf{x}||}, \quad c = \gamma\frac{\overline{x_1}}{||\mathbf{x}||}, \quad \sigma, \gamma \in \mathbb{C}, \quad |\sigma| = |\gamma| = 1,$$

where the minus-sign in front of σ and the complex conjugations are taken for later convenience. In case $x_1x_2 = 0$, we can put the solution of (1.3′) into a similar form:

$$(2.5)\qquad \begin{cases} s = -\sigma\dfrac{\overline{x_2}}{|x_2|}, \; c = 0 & \text{if } x_1 = 0, \\ c = \sigma\dfrac{\overline{x_1}}{|x_1|}, \; s = 0 & \text{if } x_2 = 0, \end{cases} \qquad u = \sigma||\mathbf{x}||, \quad |\sigma| = 1.$$

LEMMA 2.1. *Let s, c be defined as in (2.4) and $x_1x_2 \neq 0$. Then, $\overline{s}x_1 + \overline{c}x_2 = 0$ implies*

$$(2.6)\qquad \gamma = \sigma, \quad u = \sigma||\mathbf{x}||.$$

PROOF. The given condition yields

$$\overline{s}x_1 + \overline{c}x_2 = -\overline{\sigma}\frac{x_2}{||\mathbf{x}||}x_1 + \overline{\gamma}\frac{x_1}{||\mathbf{x}||}x_2 = \frac{1}{||\mathbf{x}||}(-\overline{\sigma} + \overline{\gamma})x_1x_2 = 0. \qquad \square$$

With (2.4) and (2.6), all possible solutions of (2.2) are given. We shall use the following notation (also in the quaternion case $\mathbb{H} = \mathbb{R}^4$, to be treated later):

$$(2.7)\qquad \operatorname{sgn} x = \frac{\overline{x}}{|x|}, \quad x \in \mathbb{K}\backslash\{0\}, \quad \mathbb{K} \in \{\mathbb{R}, \mathbb{C}, \mathbb{H}\}.$$

Consequently, $x \operatorname{sgn} x = |x|$. There are two favorite options for the choice of σ. One can make $s \geq 0$ by choosing $\sigma = -\overline{\operatorname{sgn} x_2}$ (in case $x_2 \neq 0$) which in the real case corresponds to a rotational angle $\alpha \in [0, \pi[$ or one can make $c \geq 0$ by choosing $\sigma = \overline{\operatorname{sgn} x_1}$ (in case $x_1 \neq 0$) which in the real case corresponds to a rotational angle $\alpha \in]-\pi/2, \pi/2]$. The first option yields

$$(2.8)\qquad s = \frac{|x_2|}{||\mathbf{x}||}, \; c = \begin{cases} 1 & \text{if } x_2 = 0, \\ -\dfrac{(\operatorname{sgn} \overline{x_2})\overline{x_1}}{||\mathbf{x}||} & \text{else}, \end{cases} \quad u = \begin{cases} x_1 & \text{if } x_2 = 0, \\ -||\mathbf{x}|| \operatorname{sgn} \overline{x_2} & \text{else}, \end{cases}$$

whereas the second option yields

$$s = \begin{cases} 1 & \text{if } x_1 = 0, \\ -\dfrac{(\operatorname{sgn}\overline{x_1})\overline{x_2}}{||\mathbf{x}||} & \text{else,} \end{cases} \quad c = \frac{|x_1|}{||\mathbf{x}||}, \quad u = \begin{cases} -x_2 & \text{if } x_1 = 0, \\ ||\mathbf{x}||\operatorname{sgn}\overline{x_1} & \text{else.} \end{cases} \tag{2.9}$$

Some more details with respect to these formulae are given in Section 4, see (4.17), (4.18).

3 Short review on the algebra of quaternions.

Let $\mathbb{H} = \mathbb{R}^4$ be equipped with the ordinary vector space structure and with an additional multiplicative operation $\mathbb{H} \times \mathbb{H} \longrightarrow \mathbb{H}$ which most easily can be defined by a multiplication table (see Table 3.1) for the four basis elements

$$\begin{aligned} (1,0,0,0) &= \mathbf{1}, \quad (0,1,0,0) = \mathbf{i}, \\ (0,0,1,0) &= \mathbf{j}, \quad (0,0,0,1) = \mathbf{k}\,. \end{aligned} \tag{3.1}$$

Table 3.1: Multiplication table for quaternions

	1	**i**	**j**	**k**
1	**1**	**i**	**j**	**k**
i	**i**	$-\mathbf{1}$	**k**	$-\mathbf{j}$
j	**j**	$-\mathbf{k}$	$-\mathbf{1}$	**i**
k	**k**	**j**	$-\mathbf{i}$	$-\mathbf{1}$

An element $x = (x_1, x_2, x_3, x_4) \in \mathbb{H}$ has the representation

$$x = x_1\mathbf{1} + x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k} = \Re x + \operatorname{Vec} x, \tag{3.2}$$

where $x_1, x_2, x_3, x_4 \in \mathbb{R}$, $\Re x = x_1$ is the *real part* of x, $\operatorname{Vec} x = x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}$ is the *vector part* of x. We will identify the quaternion $x = (x_1, 0, 0, 0)$ with the real number x_1, the quaternion $x = (x_1, x_2, 0, 0)$ will be identified with the complex number $x_1 + \mathrm{i}x_2$. If we denote $\mathbf{v} = (x_2, x_3, x_4) \in \mathbb{R}^3$ the vector part of x then the quaternion x has the representation:

$$x = (x_1, \mathbf{v}), \quad x_1 \in \mathbb{R}, \quad \mathbf{v} \in \mathbb{R}^3. \tag{3.3}$$

For $x = (x_1, x_2, x_3, x_4) = (x_1, \mathbf{v}) \in \mathbb{H}$, $y = (y_1, y_2, y_3, y_4) = (y_1, \mathbf{w}) \in \mathbb{H}$ it follows that

$$\begin{aligned} x\,y &= (x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4)\,\mathbf{1} + (x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3)\,\mathbf{i} + \\ &\quad +(x_1y_3 - x_2y_4 + x_3y_1 + x_4y_2)\,\mathbf{j} + (x_1y_4 + x_2y_3 - x_3y_2 + x_4y_1)\,\mathbf{k} \\ &= (x_1y_1 - \mathbf{v}\cdot\mathbf{w},\ x_1\mathbf{w} + y_1\mathbf{v} + \mathbf{v}\times\mathbf{w}), \end{aligned}$$

where $\cdot$, $\times$ are the scalar, vector products in $\mathbb{R}^3$, respectively. Obviously, in general, multiplication is not commutative here, but there are some classes of quaternions for which the product commutes (for example if one of the factors is real).

Given x according to (3.2), the *conjugate* $\overline{x}$ of x is defined to be

$$\overline{x} = (x_1, -x_2, -x_3, -x_4) = \Re x - \operatorname{Vec} x. \tag{3.4}$$

Let us note that conjugation obeys the rules

$$\overline{xy} = \overline{y}\,\overline{x}, \qquad \overline{\overline{x}} = x.$$

We define the *absolute value* of x by

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}. \tag{3.5}$$

We will use the following properties of $|x|$:

$$|x|^2 = x\,\overline{x} = \overline{x}\,x, \qquad |xy| = |yx| = |x||y|\,. \tag{3.6}$$

The space $\mathbb{H}$ is a normed vector space over $\mathbb{H}$, where the norm is introduced in (3.5).

Let us remark, that for any $x \in \mathbb{H}\backslash\{0\}$ an *inverse* quaternion x^{-1} is defined,

$$x^{-1} = \frac{\overline{x}}{|x|^2}. \tag{3.7}$$

Simple calculations give some more properties of operations on quaternions. Let $x = (x_1, x_2, x_3, x_4)$, $y = (y_1, y_2, y_3, y_4) \in \mathbb{H}$ be two quaternions. Then

$$\begin{aligned} x^2 &= x_1^2 - x_2^2 - x_3^2 - x_4^2 + 2x_1(x_2\mathbf{i} + x_3\mathbf{j} + x_4\mathbf{k}) = 2(\Re x)x - |x|^2, \\ (\operatorname{Vec} x)^2 &= -x_2^2 - x_3^2 - x_4^2 = -|\operatorname{Vec} x|^2, \\ \Re(xy) &= x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4 = \Re(yx). \end{aligned} \tag{3.8}$$

Two quaternions x and y are called *equivalent* if $y = \alpha^{-1}x\alpha$, for some $\alpha \in \mathbb{H}\backslash\{0\}$. Let us remark that x and y are equivalent if and only if $\Re x = \Re y$ and $|x| = |y|$.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)^{\mathrm{T}} \in \mathbb{H}^n$ and define

$$\|\mathbf{x}\| = \sqrt{\sum_{j=1}^{n} |x_j|^2}\,. \tag{3.9}$$

The space $\mathbb{H}^n$ becomes a normed vector space over $\mathbb{H}$ with the *norm* defined in (3.9). For $\mathbf{x} \in \mathbb{H}^n$, we denote by $\mathbf{x}^*$ the transpose of the entrywise conjugate of $\mathbf{x}$.

Let $\mathbf{B} \in \mathbb{H}^{m\times n}$ be a matrix with quaternion entries. The matrix $\mathbf{B}$ represents a linear mapping with respect to the multiplication from the right:

$$\begin{aligned} \mathbf{B}(\mathbf{x} + \mathbf{y}) &= \mathbf{B}\mathbf{x} + \mathbf{B}\mathbf{y} \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{H}^n, \\ \mathbf{B}(\mathbf{x}h) &= (\mathbf{B}\mathbf{x})h \quad \forall \mathbf{x} \in \mathbb{H}^n, \forall h \in \mathbb{H}. \end{aligned}$$

Let us note that for multiplication from the left we obtain

$$(\mathbf{B}(h\mathbf{x}))_j = \sum_k b_{jk}hx_k \neq h\sum_k b_{jk}x_k = (h\mathbf{B}\mathbf{x})_j.$$

We define the *conjugate transposition* of the matrix $\mathbf{B} = (b_{ij}) \in \mathbb{H}^{m\times n}$ as $\mathbf{B}^* = (\overline{b_{ji}}) \in \mathbb{H}^{n\times m}$. The square quaternion valued matrix $\mathbf{B} \in \mathbb{H}^{n\times n}$ is called *Hermitean* if $\mathbf{B} = \mathbf{B}^*$, and $\mathbf{B}$ is *positive definite* if it is Hermitean and

$$\mathbf{x}^*\,\mathbf{B}\,\mathbf{x} > 0 \qquad \forall\; \mathbf{x} \in \mathbb{H}^n\backslash\{\mathbf{0}\}\,.$$

A matrix $\mathbf{B} \in \mathbb{H}^{n\times n}$ is said to be *unitary* if $\mathbf{B}^*\mathbf{B} = \mathbf{I}$.

THEOREM 3.1. *A matrix $\mathbf{B} \in \mathbb{H}^{n\times n}$ is unitary if and only if $\|\mathbf{Bx}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{H}^n$.*

PROOF. Let $\mathbf{B} \in \mathbb{H}^{n\times n}$ be unitary, $\mathbf{x} \in \mathbb{H}^n$. Then

$$(\mathbf{Bx})^*\mathbf{Bx} = \mathbf{x}^*\mathbf{B}^*\mathbf{Bx} = \mathbf{x}^*\mathbf{x}, \quad \text{i.e.} \quad \|\mathbf{Bx}\| = \|\mathbf{x}\|.$$

To verify the converse, let us suppose that $\|\mathbf{Bx}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{H}^n$ and $\mathbf{B}^*\mathbf{B} = \mathbf{I} + \mathbf{A}$ for a matrix $\mathbf{A} \in \mathbb{H}^{n\times n}$. By assumption, $\mathbf{x}^*(\mathbf{I}+\mathbf{A})\mathbf{x} = \mathbf{x}^*\mathbf{x} + \mathbf{x}^*\mathbf{Ax} = \mathbf{x}^*\mathbf{x}$ which implies $\mathbf{x}^*\mathbf{Ax} = \mathbf{0}$ for all $\mathbf{x} \in \mathbb{H}^n$, i.e. $\mathbf{A} = \mathbf{0}$ and $\mathbf{B}$ is unitary. $\square$

DEFINITION 3.1. Let $\mathbf{B} \in \mathbb{H}^{n\times n}$. If there exist a vector $\mathbf{x} \in \mathbb{H}^n\backslash\{\mathbf{0}\}$ and a quaternion $\lambda \in \mathbb{H}$ such that

$$\mathbf{Bx} = \mathbf{x}\lambda, \tag{3.10}$$

we call λ an *eigenvalue* of $\mathbf{B}$ and $\mathbf{x}$ an *eigenvector corresponding to* λ.

Let us point out that in the above definition we have put λ as a right factor of $\mathbf{x}$, i.e. λ is the right eigenvalue of $\mathbf{B}$. This coincides with the fact that $\mathbf{B}$ represents a linear mapping with respect to multiplication from the right. In the literature, e.g. ZHANG[9], the left eigenvalues and the left eigenvectors are also defined, but these belong to a nonlinear theory.

The number of the eigenvalues of a quaternion valued matrix $\mathbf{B} \in \mathbb{H}^{n\times n}$ is, in general, not finite. If λ is an eigenvalue, one can easily show that the whole equivalence class $[\lambda]$ consists of eigenvalues, where

$$[\lambda] = \left\{\mu,\ \mu = h^{-1}\lambda h,\ h \in \mathbb{H}\backslash\{0\}\right\}.$$

We say that two eigenvalues $\lambda_1,\ \lambda_2$ are *equivalent*, denoted by $\lambda_1 \sim \lambda_2$, if they belong to the same equivalence class $[\lambda]$. As in the real and complex case one can show, that eigenvectors $\mathbf{x}_1, \mathbf{x}_2$ which correspond to two non equivalent eigenvalues λ_1, λ_2 are (right) linearly independent. Therefore, the number of equivalence classes is at most n. If the eigenvalue λ is real, then $h^{-1}\lambda h = \lambda$ for any $h \in \mathbb{H}\backslash\{0\}$, thus, $[\lambda] = \lambda$. It can be shown, that in $[\lambda]$ there is exactly one complex number with non negative real part. More precisely, one can show, that two quaternions λ_1, λ_2 are equivalent if and only if $\Re\lambda_1 = \Re\lambda_2$ and $|\lambda_1| = |\lambda_2|$. Thus, $\lambda = (l_1, l_2, l_3, l_4)$ and $\widetilde{\lambda} = (l_1, \sqrt{l_2^2 + l_3^2 + l_4^2}, 0, 0))$ are equivalent and $\widetilde{\lambda}$ is complex with non negative imaginary part and apparently the only equivalent quaternion with these properties.

The existence of a Schur canonical form guarantees the existence of n eigenvalues for any square quaternion matrix where the eigenvalues are either real or complex with positive imaginary part. Other canonical forms which are known to exist for quaternion valued matrices could do the same purpose. See ZHANG[9] for an overview.

THEOREM 3.2. *Let* $\mathbf{A} \in \mathbb{H}^{n\times n}$ *be Hermitean. Then* $\mathbf{A}$ *has only real eigenvalues and their number is* n.

PROOF. Let $\mathbf{x} \neq \mathbf{0}$, $\mathbf{Ax} = \mathbf{x}\lambda$. Then $\mathbf{x}^*\mathbf{Ax} = \|\mathbf{x}\|^2\lambda$. Since $\mathbf{x}^*\mathbf{A}^* = \overline{\lambda}\mathbf{x}^*$, we obtain $\mathbf{x}^*\mathbf{A}^*\mathbf{x} = \mathbf{x}^*\mathbf{Ax} = \|\mathbf{x}\|^2\overline{\lambda}$, i.e. λ is real. □

THEOREM 3.3. *For any unitary quaternion valued matrix* $\mathbf{A}$, *the eigenvalues* λ *satisfy* $|\lambda| = 1$.

PROOF. Let λ be an eigenvalue of a unitary quaternion valued matrix $\mathbf{A}$, $\mathbf{x} \neq \mathbf{0}$ a corresponding eigenvector. We multiply the equation $\mathbf{Ax} = \mathbf{x}\lambda$ from the left by $(\mathbf{Ax})^*$, i.e.

$$(\mathbf{Ax})^*\mathbf{Ax} = (\mathbf{Ax})^*\mathbf{x}\lambda. \tag{3.11}$$

On the left-hand side of the equation (3.11) we use the fact that $\mathbf{A}^*\mathbf{A} = \mathbf{I}$, on the right-hand side we substitute $\mathbf{x}^*\mathbf{A}^* = \overline{\lambda}\mathbf{x}^*$. We obtain $\|\mathbf{x}\|^2 = \|\mathbf{x}\|^2|\lambda|^2$ which implies $|\lambda| = 1$. □

4 Givens' transformation in the quaternion case.

Let $h \in \mathbb{H}$ have the representation

$$h = a_1\mathbf{1} + a_2\mathbf{i} + a_3\mathbf{j} + a_4\mathbf{k}. \tag{4.1}$$

By a well known isomorphism (cf. van der Waerden [1960, p. 55]) between $\mathbb{H}$ and complex 2×2 matrices of the form

$$\widetilde{\mathbf{h}} = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, \; \alpha = a_1 + \mathrm{i}a_2, \; \beta = a_3 + \mathrm{i}a_4, \tag{4.2}$$

we see that the Givens matrix $\mathbf{G}$, defined in (2.1) as well as $\mathbf{G}^*$ has exactly the quaternion form (4.2). Since we can extend any complex vector $\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{C}^2$ to a quaternion by defining $\widetilde{\mathbf{x}} = \begin{pmatrix} x_1 & -\overline{x_2} \\ x_2 & \overline{x_1} \end{pmatrix}$, the matrix-vector multiplication $\mathbf{u} = \mathbf{G}^*\mathbf{x}$ may be regarded as a multiplication of quaternions

$$\widetilde{\mathbf{y}} = \mathbf{G}^*\widetilde{\mathbf{x}},$$

where the first row of $\widetilde{\mathbf{y}}$ represents the product $\mathbf{u}$. We are not sure whether this interpretation has been given elsewhere in the literature.

We begin again with the following setting. A vector $\mathbf{x} \in \mathbb{H}^2\backslash\{\mathbf{0}\}$ is given and a unitary matrix $\mathbf{G}$ has to be constructed according to the following rules:

$$\mathbf{G} = \begin{pmatrix} \overline{c} & s \\ -\overline{s} & c \end{pmatrix} \in \mathbb{H}^{2\times 2}, \; \mathbf{G}^*\mathbf{x} = \mathbf{u} = u\mathbf{e}_1, \; \mathbf{e}_1 = (\mathbf{1}, 0)^{\mathrm{T}} \in \mathbb{H}^2, \; u \in \mathbb{H}\backslash\{0\}. \tag{4.3}$$

LEMMA 4.1. *The matrix* $\mathbf{G}$ *given in (4.3) is unitary if and only if*

$$|s|^2 + |c|^2 = 1, \tag{4.4}$$

$$s\overline{c} = \overline{c}s. \tag{4.5}$$

PROOF. Exercise. □

In order to find $\mathbf{G}$ of (4.3) we have to solve the following quaternion equations for s and c where x_1, x_2 are given quaternions, not both zero:

$$\begin{array}{rcl} cx_1 - sx_2 & = & u \neq 0, \\ \overline{s}x_1 + \overline{c}x_2 & = & 0, \end{array} \tag{4.6}$$

subject to (4.4), (4.5) mentioned in Lemma 4.1.

The second condition (4.5) requires that s and $\overline{c}$ commute in $\mathbb{H}$. Since real numbers commute with quaternions, a solution with real s (or real c) would satisfy (4.5). By using (4.4) the second equation of (4.6) yields

$$|s| = \frac{|x_2|}{||\mathbf{x}||}, \quad |c| = \frac{|x_1|}{||\mathbf{x}||}. \tag{4.7}$$

That implies, that the general solution of (4.6) must have the form

$$s = -\sigma\frac{\overline{x_2}}{||\mathbf{x}||}, \quad c = \gamma\frac{\overline{x_1}}{||\mathbf{x}||}, \quad \text{where} \quad \sigma, \gamma \in \mathbb{H} \text{ with } |\sigma| = |\gamma| = 1. \tag{4.8}$$

Since there are several forms for the general solution, we have chosen one which is similar to the complex case, Section 2, equation (2.4). The special case $x_1 x_2 = 0$ yields formally the same solution (2.5) as in the complex case.

LEMMA 4.2. *Let $\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{H}^2$ be given with $x_1 x_2 \neq 0$. Define s, c as in (4.8). Then, (i) $s\overline{c} = \overline{c}s$, (ii) $\overline{s}x_1 + \overline{c}x_2 = 0$, (iii) $u = cx_1 - sx_2 \neq 0$ imply*

$$\gamma = \sigma, \qquad u = \sigma||\mathbf{x}||. \tag{4.9}$$

PROOF. If we make use of formulae (4.8), we find from condition (i)

$$\underbrace{\sigma\overline{x_2}}_{b}\,\underbrace{x_1\overline{\gamma}}_{a} = \underbrace{x_1\overline{\gamma}}_{a}\,\underbrace{\sigma\overline{x_2}}_{b}, \tag{4.10}$$

and from condition (ii)

$$\underbrace{x_1\overline{\gamma}}_{a}\,x_2 = \underbrace{x_2\overline{\sigma}}_{\overline{b}}\,x_1. \tag{4.11}$$

If we multiply equation (4.11) from the left by b, we obtain $bax_2 = |b|^2 x_1 = |x_2|^2 x_1$. By using (4.10) we replace ba with ab,

$$|x_2|^2 x_1 = abx_2 = x_1\overline{\gamma}\sigma\overline{x_2}x_2 = x_1\overline{\gamma}\sigma|x_2|^2 \Rightarrow x_1 = x_1\overline{\gamma}\sigma \Rightarrow 1 = \overline{\gamma}\sigma \Rightarrow \sigma = \gamma.$$

The equation for u follows straightforwardly from (iii), by replacing γ with σ. □

In order to determine σ, we inspect (4.10), (4.11) again. If we set $\gamma = \sigma$ we see, that both equations are equivalent: one follows from the other. Thus, we have to determine the set

$$\Sigma := \{\sigma \in \mathbb{H}: \ \sigma\overline{x_2}x_1 = x_1\overline{x_2}\sigma, \ |\sigma| = 1\}. \tag{4.12}$$

The set Σ is not empty, since it contains the subset $\{\pm\operatorname{sgn}\overline{x_1}, \pm\operatorname{sgn}\overline{x_2}\}$. But it is different from the whole unit sphere, since ± 1 are in general not contained in Σ.

THEOREM 4.3. *Let* $\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{H}^2$ *be given with* $x_1 x_2 \neq 0$. *Set* $\xi = x_1\overline{x_2} = (\xi_1, \xi_2, \xi_3, \xi_4)$, $\eta = \overline{x_2}x_1 = (\eta_1, \eta_2, \eta_3, \eta_4)$. *Then* ξ *and* η *are equivalent (as quaternions) and the above* Σ *has the form*

$$\Sigma = \{\sigma \in \mathbb{H} : \ \xi\sigma = \sigma\eta, \ |\sigma| = 1\}. \tag{4.13}$$

In order to determine Σ *we shall distinguish the following two cases:*

a) *Let* x_1, x_2 *be linearly dependent, in the sense* $x_1 = \alpha x_2$ *where* $\alpha \in \mathbb{R}\backslash\{0\}$. *Then*
$$\Sigma = \{\sigma \in \mathbb{H} : |\sigma| = 1\}. \tag{4.14}$$

b) *Let* x_1, x_2 *not be linearly dependent in the above sense. In this case*

$$\Sigma = \left\{\sigma \in \mathbb{H} : \sigma = \frac{\alpha\operatorname{sgn}\overline{x_1} + \beta\operatorname{sgn}\overline{x_2}}{|\alpha\operatorname{sgn}\overline{x_1} + \beta\operatorname{sgn}\overline{x_2}|}, \ \alpha, \beta \in \mathbb{R}, \ |\alpha| + |\beta| > 0\right\}. \tag{4.15}$$

PROOF. Since by (3.6) and (3.8) $|\xi| = |\eta|$ and $\Re\xi = \Re\eta$, the quaternions ξ, η are equivalent and if they are real they coincide.

a) In case x_1, x_2 are linearly dependent, we have $\xi = \eta \in \mathbb{R}\backslash\{0\}$. The result follows from (4.13) immediately.

b) Let x_1, x_2 be linearly independent. This implies, that both, ξ and η are not real. The equation $\xi\sigma = \sigma\eta$ is equivalent to a homogeneous linear system of four equations of the form

$$\mathbf{A}\sigma = \mathbf{0}, \qquad \operatorname{rank}(\mathbf{A}) = 2, \qquad \mathbf{A} \in \mathbb{R}^{4\times 4}.$$

Thus, the solution set is a two-dimensional linear space. Since we already know two independent solutions, namely $\sigma_1 = \operatorname{sgn}\overline{x_1}$, $\sigma_2 = \operatorname{sgn}\overline{x_2}$, the space is spanned by these solutions. Since we are only interested in solutions with $|\sigma| = 1$, we have to add the given normalization. It remains to show, that $\operatorname{rank}(\mathbf{A}) = 2$. Let us set

$$a = \xi - \eta = (a_1, a_2, a_3, a_4), \quad b = \xi + \eta = (b_1, b_2, b_3, b_4).$$

Since $\Re\xi = \Re\eta$, $a_1 = 0$. Then, $\xi\sigma - \sigma\eta = 0$ is equivalent to $\mathbf{A}\sigma = \mathbf{0}$, where

$$\mathbf{A} = \begin{pmatrix} 0 & -a_2 & -a_3 & -a_4 \\ +a_2 & 0 & -b_4 & +b_3 \\ +a_3 & +b_4 & 0 & -b_2 \\ +a_4 & -b_3 & +b_2 & 0 \end{pmatrix}, \quad \mathbf{A} \in \mathbb{R}^{4\times 4}.$$

Observe, that $\mathbf{A} + \mathbf{A}^{\mathrm{T}} = \mathbf{0}$. Hence, all eigenvalues of $\mathbf{A}$ are located on the imaginary axis. Since purely imaginary eigenvalues always come in pairs of conjugate eigenvalues, the matrix $\mathbf{A}$ can have only even ranks: 0, 2 or 4. Let $\operatorname{rank}(\mathbf{A}) = 0$. Then $a_2 = a_3 = a_4 = b_2 = b_3 = b_4 = 0$. However, this implies that both, ξ and η are real, a contradiction. If the rank is four, the system has only the trivial solution $\sigma = 0$. This is again a contradiction, since we have already found some nontrivial solutions. The only remaining case is $\operatorname{rank}(\mathbf{A}) = 2$. ◻

We summarize the results in the following theorem.

THEOREM 4.4. *Let $\mathbf{x} = (x_1, x_2)^{\mathrm{T}} \in \mathbb{H}^2 \backslash \{\mathbf{0}\}$ be given. Define*

$$\mathbf{G} = \begin{pmatrix} \overline{c} & s \\ -\overline{s} & c \end{pmatrix} \in \mathbb{H}^{2\times 2} \quad \textit{with} \quad s = -\sigma \frac{\overline{x_2}}{||\mathbf{x}||}, \quad c = \sigma \frac{\overline{x_1}}{||\mathbf{x}||}, \quad |\sigma| = 1,$$

where σ is arbitrary in case x_1, x_2 are linearly dependent over $\mathbb{R}$. Otherwise, it must be chosen according to formula (4.15). Then $\mathbf{G}$ is a unitary matrix and $\mathbf{G}^*\mathbf{x} = \mathbf{u} = \sigma||\mathbf{x}||(\mathbf{1}, 0)^{\mathrm{T}}$.

PROOF. Contained in Theorem 4.3. ◻

Since for practical reasons we are interested in simple solutions, we could choose σ in such a way, that $s \geq 0$ or $c \geq 0$. This leads to $\sigma = -\operatorname{sgn} \overline{x_2}$ or $\sigma = \operatorname{sgn} \overline{x_1}$, respectively. The solution is formally the same as in the complex case, see Section 2, equations (2.8), (2.9). Therefore, we do not repeat it here.

When applying Givens' transformation, it is necessary to compute $\mathbf{v} = \mathbf{G}^*\mathbf{y}$ for many vectors $\mathbf{y}$ for one fixed matrix $\mathbf{G}$. Let $\mathbf{y} = (y_1, y_2)^{\mathrm{T}}$, $\mathbf{v} = (v_1, v_2)^{\mathrm{T}}$, then, explicitly, we have to compute (cf. (4.6))

$$\begin{aligned} v_1 &= c y_1 - s y_2, \\ v_2 &= \overline{s} y_1 + \overline{c} y_2. \end{aligned} \tag{4.16}$$

We can - for real s - slightly rearrange formulae (4.16) to obtain

$$\begin{aligned} v_2 &= \overline{s} y_1 + \overline{c} y_2, \\ v_1 &= \frac{c}{1+s}(y_1 + v_2) - y_2 \,. \end{aligned} \tag{4.16$'$}$$

If we choose $s \geq 0$ we can safely define a new quaternion constant $\mu = \dfrac{c}{1+s}$; for the operation count see Table 4.1.

Table 4.1: Number of operations

		formula (4.16)	formula (4.16$'$)
multiplication	real by quaternion	2	1
	quaternion by quaternion	2	2
addition of quaternions		2	3

As we can see from the Table 4.1, we exchange one multiplication real by quaternion by one addition of two quaternions. This is probably only a slight advantage. However, we can do a little more by recovering s and c from μ by

$$s = \frac{1-|\mu|^2}{1+|\mu|^2}, \quad c = (1+s)\mu = \frac{2}{1+|\mu|^2}\mu, \quad \text{where} \quad \mu = \frac{c}{1+s}. \tag{4.17}$$

Therefore, it is reasonable to compute only μ rather than s and c by the formula

$$\mu = \begin{cases} 1 & \text{if } x_2 = 0, \\ -\dfrac{\operatorname{sgn}\overline{x_2}\,\overline{x_1}}{||\mathbf{x}|| + |x_2|} & \text{else.} \end{cases} \tag{4.18}$$

The following Algorithm 4.1 (a MATLAB program) computes μ and on demand also u. It relies on several subprograms which handle quaternion arithmetic.

ALGORITHM 4.1. *Givens' code for quaternion valued 2-vectors* $\mathbf{x}$

```
%function [mu,u]=quat_givens(x);
%Givens code for quaternion-valued 2-vector x.
%The result will be quaternions mu, u.
%Matrices are all cells.
%The following seven subprograms have to be provided:
%1. function [no,noinf]=quat_normvector(x);
        %computes euclidean and max-norm
%2. function result=quat_iszero(M);
        %if M is a zero quaternion matrix, result=1, else 0.
%3. function q=quat_realmult(p,a);
        % p*a: p is quaternion (may be matrix), a is real
%4. function a=quat_sign(x);
        %elementwise sign(h)=h/|h|, for h\not=0,
        %sign([0 0 0 0])=[1 0 0 0].
%5. function s=quat_mult(h1,h2);
        %elementwise quaternion product of h1, h2
%6. function Astar=quat_conj(A);
        %Astar=A^* (A must not be square)
%7. function a=quat_abs(x);
        %elementwise absolute value of quaternion x, a is real,
        %same size as x

function [mu,u]=quat_givens(x);
  b1{1}=x{1};     b2{1}=x{2};
  h=quat_normvector(x); u=b1;
  if h==0 | quat_iszero(b2)
    mu{1}=[1,0,0,0]'; return
  end; %if
  a=quat_realmult(quat_sign(b2),-1);
  mu=quat_realmult(quat_mult(a,quat_conj(b1)),1/(h+quat_abs(b2)));
  if nargout==2
    u=quat_realmult(a,h);
  end; %if
return
```

EXAMPLE 4.1. For $\mathbf{x} = ([1,2,3,4],[-4,-3,-2,-1])^{\mathrm{T}}$ (in this example we write quaternions in square brackets) the above program yields
$\mu = [0.27614237491540, -0.13807118745770, 0, -0.27614237491540]$,
$u = [5.65685424949238, 4.24264068711928, 2.82842712474619, 1.41421356237309]$.
By hand computation we obtain $\mu = \frac{\sqrt{2}-1}{3}[2,-1,0,-2]$, $u = \sqrt{2}[4,3,2,1]$, $s = 0.5\sqrt{2}$, $c = \frac{\sqrt{2}}{6}[2,-1,0,-2]$, which coincides with the above given numerical solution in the given precision. The MATLAB flop count for this example is 114.

Let us repeat this example with the strategy

$$\sigma = \frac{\operatorname{sgn} \overline{x_1} + \operatorname{sgn} \overline{x_2}}{|\operatorname{sgn} \overline{x_1} + \operatorname{sgn} \overline{x_2}|}.$$

We obtain:

$$\sigma = \begin{pmatrix} -0.67082039324994 \\ -0.22360679774998 \\ 0.22360679774998 \\ 0.67082039324994 \end{pmatrix}, \quad u = \begin{pmatrix} -5.19615242270663 \\ -1.73205080756888 \\ 1.73205080756888 \\ 5.19615242270663 \end{pmatrix},$$

$$s = \begin{pmatrix} -0.28867513459481 \\ 0.28867513459481 \\ -0.00000000000000 \\ 0.57735026918963 \end{pmatrix}, \quad c = \begin{pmatrix} 0.28867513459481 \\ 0.28867513459481 \\ 0.00000000000000 \\ 0.57735026918963 \end{pmatrix}.$$

We also see by examples, that a choice of σ with $|\sigma| = 1$ but $\sigma \notin \Sigma$ where Σ is defined in (4.15), p. 9 fails to produce a unitary matrix $\mathbf{G}$.

General information on quaternions may be obtained from KUIPERS[4]. A survey on older and recent quaternion literature was given by ZHANG[9].

Acknowledgement.

The authors acknowledge with pleasure the support of the Grant Agency of the Czech Republic (grant # 201/02/0844, grant # 201/02/0595) and support of the project MSM 223400007 of the Czech Ministry of Education. The authors also thank Professor Ron Guenther, Oregon State University, Oregon for carefully reading this manuscript and giving suggestions for improvement.

REFERENCES

1. J. W. Givens, *Numerical computation of the characteristic values of a real symmetric matrix*, Oak Ridge Natl. Lab. Report: ORNL 1574, 1954.
2. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD., 1997, 694 p.
3. A. Greenbaum, *Iterative Methods for Solving Linear Systems*, siam, Philadelphia, 1997, 220 p.
4. J. B. Kuipers, *Quaternions and Rotation Sequences, a Primer with Applications to Orbits, Aerospace, and Virtual Reality*, Princeton University Press, Princeton, NJ, 1999, 371 p.
5. G. Opfer, *Numerische Mathematik für Anfänger*, 4. Aufl., Vieweg, Braunschweig, Wiesbaden, 2002, 357 p.
6. H. R. Schwarz, *Numerische Mathematik*, 4. Aufl. Teubner, Stuttgart, 1997, 653 p.
7. J. Stoer, *Numerische Mathematik 1*, 8. Aufl., Springer, Berlin, 1999, 371 p.
8. B. L. van der Waerden, *Algebra I*, 5. Aufl., Springer, Berlin, Göttingen, Heidelberg, 1960, 292. p.
9. F. Zhang, *Quaternions and matrices of quaternions*, Linear Algebra Appl. **251**, (1997), 21-57.