Chapter Two

The algebra of quaternions

In this chapter we introduce the quaternions and their algebra: multiplication, norm, automorphisms and antiautomorphisms, etc. We give matrix representations of various real linear maps associated with quaternion algebra. We also introduce representations of quaternions as real 4×4 matrices and as complex 2×2 matrices.

2.1 BASIC DEFINITIONS AND PROPERTIES

Fix an ordered basis $\{\mathsf{e}, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$ in a 4-dimensional real vector space H (we may take $\mathsf{H} = \mathsf{R}^4$, the vector space of columns consisting of four real components), and introduce multiplication in H by the formulas

$$\mathsf{ei} = \mathsf{ie} = \mathsf{i}, \quad \mathsf{ej} = \mathsf{je} = \mathsf{j}, \quad \mathsf{ek} = \mathsf{ke} = \mathsf{k},$$

$$\mathsf{i}^2 = \mathsf{j}^2 = \mathsf{k}^2 = -\mathsf{e}, \qquad \mathsf{e}^2 = \mathsf{e}, \qquad \mathsf{ij} = -\mathsf{ji} = \mathsf{k}, \quad \mathsf{jk} = -\mathsf{kj} = \mathsf{i}, \quad \mathsf{ki} = -\mathsf{ik} = \mathsf{j},$$

and by the requirement that the multiplication of elements of H is distributive with respect to addition and commutes with scalar multiplication:

$$x(y+z) = xy + xz, \quad (y+z)x = yx + yz, \quad x(\lambda y) = (\lambda x)y = \lambda(xy)$$

for all $x, y, z \in \mathsf{H}$ and all $\lambda \in \mathsf{R}$.

Definition 2.1.1. The elements of H with the algebraic operations of H as a real vector space and with the multiplication introduced as above are called the (real) *quaternions.*

The letter H stands for William Rowan Hamilton (1805–1865), inventor of quaternions. Clearly, the multiplication in the algebra H is noncommutative.

Proposition 2.1.2. H *is a unital associative algebra with the unity* e:

$$x(yz) = (xy)z, \quad \mathsf{e}x = x\mathsf{e} = x$$

for all $x, y, z \in \mathsf{H}$.

In the sequel we identify the real number λ with the quaternion $\lambda\mathsf{e}$; in particular, 1 stands for $1\mathsf{e}$. Also, it is easy to see that the real span of 1 and i is isomorphic (as a subalgebra of H) to C; thus, we identify, when convenient, C with the subalgebra of H spanned (as a real vector space) by 1 and i.

Definition 2.1.3. For a quaternion $x = x_0 + x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k}$, where $x_0, x_1, x_2, x_3 \in \mathsf{R}$, we define $\mathfrak{R}(x) = x_0$, the *real part* of x, and $\mathfrak{V}(x) = x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k}$, the *vector part* (or *imaginary part*) of x. The *conjugate quaternion* of x is defined by $x_0 - x_1\mathsf{i} - x_2\mathsf{j} - x_3\mathsf{k} = \mathfrak{R}(x) - \mathfrak{V}(x)$ and denoted $\overline{x}$ or x^*. The *norm* of x is $|x| = \sqrt{x^*x} = \sqrt{x_0^2 + x_1^2 + x_2^2 + x_3^2} \in \mathsf{R}$. We say that $x \in \mathsf{H}$ is a *unit quaternion* if $|x| = 1$.

Some elementary properties of the algebra of quaternions are listed below.

Proposition 2.1.4. *Let* $x, y \in \mathsf{H}$. *Then*:

1. $x^*x = xx^*$;
2. $|x| = |x^*|$;
3. $|\cdot|$ *is indeed a norm on* H; *in more detail, for all* $x, y \in \mathsf{H}$ *we have*:
$$|x| \geq 0 \quad \text{with equality if and only if } x = 0;$$
$$|x+y| \leq |x| + |y|; \quad |xy| = |yx| = |x| \cdot |y|;$$
4. $\mathsf{j}c\mathsf{j}^* = \mathsf{k}c\mathsf{k}^* = \overline{c}$ *for every* $c \in \mathsf{C}$;
5. $(xy)^* = y^*x^*$;
6. $x = x^*$ *if and only if* $x \in \mathsf{R}$;
7. *if* $a \in \mathsf{H}$, *then* $ax = xa$ *for every* $x \in \mathsf{H}$ *if and only if* $a \in \mathsf{R}$;
8. *every* $x \in \mathsf{H} \setminus \{0\}$ *has an inverse* $x^{-1} = x^*/|x|^2 \in \mathsf{H}$; *in more detail,*
$$x \cdot (x^*/|x|^2) = (x^*/|x|^2) \cdot x = 1;$$
9. $|x^{-1}| = |x|^{-1}$ *for every* $x \in \mathsf{H} \setminus \{0\}$;
10. $x \in \mathsf{H}$ *and* x^* *are solutions of the following quadratic equation with real coefficients*: $t^2 - 2\mathfrak{R}(x)t + |x|^2 = 0$;
11. *Cauchy-Schwarz-type inequality is* $\max\{|\mathfrak{R}(xy)|, |\mathfrak{V}(xy)|\} \leq |x| \cdot |y|$;
12. $\mathfrak{R}(xy) = \mathfrak{R}(yx)$ *for all* $x, y \in \mathsf{H}$;
13. *if* $\mathfrak{R}(x) = 0$, *then* $x^2 = -|x|^2$.

We indicate a proof of $|xy| = |x| \cdot |y|$:

$$|xy|^2 = xy(xy)^* = xyy^*x^* = yy^*xx^* = |y|^2\,|x|^2, \quad \text{for all } x, y \in \mathsf{H}.$$

Thus, H is a *division ring*, i.e., a unital ring in which every nonzero element has a multiplicative inverse, and also a 4-dimensional *algebra* over the field of real numbers R.

Note that the multiplication of quaternions with zero real parts can be expressed in terms of the usual inner product and cross product of vectors in R^3, namely, if $x = x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k}$, $y = y_1\mathsf{i} + y_2\mathsf{j} + y_3\mathsf{k} \in \mathsf{H}$, where $x_\ell, y_\ell \in \mathsf{R}$, then

$$xy = -p_x^T p_y + [\mathsf{i} \;\; \mathsf{j} \;\; \mathsf{k}](p_x \times p_y), \tag{2.1.1}$$

where

$$p_x = [x_1 \;\; x_2 \;\; x_3]^T, \quad p_y = [y_1 \;\; y_2 \;\; y_3]^T \in \mathsf{R}^{3\times 1}, \tag{2.1.2}$$

and where in the right-hand side of (2.1.1) $\times$ denotes the cross product (also known as vector product) of vectors in $\mathsf{R}^{3\times 1}$:

$$[x_1 \;\; x_2 \;\; x_3]^T \times [y_1 \;\; y_2 \;\; y_3]^T = (x_2y_3 - x_3y_2, -(x_1y_3 - x_3y_1), x_1y_2 - x_2y_1)^T.$$

The verification of (2.1.1) is straightforward. More generally, let

$$x = x_0 + x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k}, \quad y = y_0 + y_1\mathsf{i} + y_2\mathsf{j} + y_3\mathsf{k} \in \mathsf{H}, \qquad x_\ell, y_\ell \in \mathsf{R},$$

and define p_x, p_y by (2.1.2). Then

$$\mathfrak{R}(xy) = x_0y_0 - p_x^T p_y, \qquad \mathfrak{V}(xy) = x_0\mathfrak{V}(y) + y_0\mathfrak{V}(x) + [\mathsf{i} \;\; \mathsf{j} \;\; \mathsf{k}](p_x \times p_y).$$

2.2 REAL LINEAR TRANSFORMATIONS AND EQUATIONS

For fixed $a, b \in \mathsf{H}$, the map $x \mapsto axb$ is obviously a real linear transformation on H. We give the matrix form of this transformation with respect to the (ordered) basis $1, \mathsf{i}, \mathsf{j}, \mathsf{k}$ of H as a real vector space.

Theorem 2.2.1. *Let*

$$a = a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k}, \qquad b = b_0 + b_1\mathsf{i} + b_2\mathsf{j} + b_3\mathsf{k} \in \mathsf{H},$$

where $a_j, b_j \in \mathsf{R}$ for $j = 0, 1, 2, 3$. *Let*

$$T_{a,b}x = axb, \qquad x \in \mathsf{H}, \tag{2.2.1}$$

be a real linear transformation. Then $T_{a,b}$ is given by the following matrix with respect to the ordered real basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$ *in* H:

$$\begin{bmatrix} a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3 & -a_0b_1 - a_1b_0 + a_2b_3 - a_3b_2 \\ a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2 & a_0b_0 - a_1b_1 + a_2b_2 + a_3b_3 \\ a_0b_2 - a_1b_3 + a_2b_0 + a_3b_1 & -a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0 \\ a_0b_3 + a_1b_2 - a_2b_1 + a_3b_0 & a_0b_2 - a_1b_3 - a_2b_0 - a_3b_1 \end{matrix}$$

$$\begin{matrix} -a_0b_2 - a_1b_3 - a_2b_0 + a_3b_1 & -a_0b_3 + a_1b_2 - a_2b_1 - a_3b_0 \\ a_0b_3 - a_1b_2 - a_2b_1 - a_3b_0 & -a_0b_2 - a_1b_3 + a_2b_0 - a_3b_1 \\ a_0b_0 + a_1b_1 - a_2b_2 + a_3b_3 & a_0b_1 - a_1b_0 - a_2b_3 - a_3b_2 \\ -a_0b_1 + a_1b_0 - a_2b_3 - a_3b_2 & a_0b_0 + a_1b_1 + a_2b_2 - a_3b_3 \end{bmatrix}.$$

The proof is obtained by tedious but straightforward computation.

The following particular cases are of interest.

Corollary 2.2.2. *The real linear transformations $T_{1,b}$, $T_{a,1}$, T_{a,a^*}, and $T_{a,a^{-1}}$ (in the latter case it is assumed $a \neq 0$) are given by the following matrices, respectively, with respect to the ordered real basis* $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$ *of* H *and using the notation of Theorem* 2.2.1:

$$\begin{bmatrix} b_0 & -b_1 & -b_2 & -b_3 \\ b_1 & b_0 & b_3 & -b_2 \\ b_2 & -b_3 & b_0 & b_1 \\ b_3 & b_2 & -b_1 & b_0 \end{bmatrix}, \quad \begin{bmatrix} a_0 & -a_1 & -a_2 & -a_3 \\ a_1 & a_0 & -a_3 & a_2 \\ a_2 & a_3 & a_0 & -a_1 \\ a_3 & -a_2 & a_1 & a_0 \end{bmatrix},$$

$$\begin{bmatrix} a_0^2 + a_1^2 + a_2^2 + a_3^2 & 0 \\ 0 & a_0^2 + a_1^2 - a_2^2 - a_3^2 \\ 0 & 2a_0a_3 + 2a_1a_2 \\ 0 & -2a_0a_2 + 2a_1a_3 \end{matrix}$$

$$\begin{matrix} 0 & 0 \\ -2a_0a_3 + 2a_1a_2 & 2a_0a_2 + 2a_1a_3 \\ a_0^2 - a_1^2 + a_2^2 - a_3^2 & -2a_0a_1 + 2a_2a_3 \\ 2a_0a_1 + 2a_2a_3 & a_0^2 - a_1^2 - a_2^2 + a_3^2 \end{bmatrix}, \tag{2.2.2}$$

and $(a_0^2 + a_1^2 + a_2^2 + a_3^2)^{-1}X$, *where X is the matrix* (2.2.2).

The statement of Corollary 2.2.2 concerning $T_{a,a^{-1}}$ follows from the observation that $T_{a,a^{-1}} = |a|^{-2}T_{a,a^*}$. Note that the matrices corresponding to $T_{b,1}$, resp. to $T_{1,a}$, are skewsymmetric if and only if $\mathfrak{R}(b) = 0$, resp. $\mathfrak{R}(a) = 0$, whereas the matrices corresponding to T_{a,a^*} and $T_{a,a^{-1}}$ are symmetric if and only if $\mathfrak{R}(a) = 0$ or $\mathfrak{V}(a) = 0$.

Corollary 2.2.3. *Let $a = a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k} \in \mathsf{H}$, $a_0, a_1, a_2, a_3 \in \mathsf{R}$. Then the real linear transformation $T_{1,a} - T_{a,1}$ that maps $x \in \mathsf{H}$ to $xa - ax$ is given by the skewsymmetric matrix*

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 2a_3 & -2a_2 \\ 0 & -2a_3 & 0 & 2a_1 \\ 0 & 2a_2 & -2a_1 & 0 \end{bmatrix}$$

with respect to the ordered real basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$.

Observe that for $a = a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k} \in \mathsf{H} \setminus \{0\}$, where a_0, a_1, a_2, a_3 are real, the matrix

$$U := \frac{1}{|a|^2} \begin{bmatrix} a_0^2 + a_1^2 - a_2^2 - a_3^2 & -2a_0a_3 + 2a_1a_2 & 2a_0a_2 + 2a_1a_3 \\ 2a_0a_3 + 2a_1a_2 & a_0^2 - a_1^2 + a_2^2 - a_3^2 & -2a_0a_1 + 2a_2a_3 \\ -2a_0a_2 + 2a_1a_3 & 2a_0a_1 + 2a_2a_3 & a_0^2 - a_1^2 - a_2^2 + a_3^2 \end{bmatrix}$$

is *orthogonal*, i.e., $U^TU = I$. A straightforward computation will verify this assertion. Moreover, $\det U = 1$. Indeed, the set of all nonzero quaternions is connected, and $\det U$ is a continuous function of the components of a. Therefore, the values of $\det U$ also form a connected set (Theorem 3.10.7). But determinants of real orthogonal matrices can be only 1 or -1. It follows that either $\det U = 1$ for all U, or $\det U = -1$ for all U. Since for $a = 1$ we have $U = I$, the second possibility cannot happen.

We obtain that 1 is an eigenvalue of U, and the corresponding eigenvector is unique up to scaling (apart from the case $U = I$). So, in a suitable orthonormal basis in R^3, the matrix U has the form

$$U = \begin{bmatrix} \cos\mu & -\sin\mu & 0 \\ \sin\mu & \cos\mu & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad 0 \le \mu < 2\pi.$$

Comparing with Corollary 2.2.2, the following geometric description of the transformation $T_{a,a^{-1}}$ is obtained. In this description, H_0 stands for the real vector space of all quaternions with zero real parts, and orthogonality in H_0 is understood in the sense of the real-valued inner product that has $\mathsf{i}, \mathsf{j}, \mathsf{k}$ as an orthonormal basis.

Corollary 2.2.4. *Let $a \in \mathsf{H} \setminus \{0\}$, and assume $T_{a,a^{-1}} \neq I$. Then $T_{a,a^{-1}}$ maps H_0 onto itself. Moreover, there is a unique (up to scaling) nonzero $x_0 \in \mathsf{H}_0$ such that $T_{a,a^{-1}}x_0 = x_0$, and denoting by $\mathsf{H}_0 \ominus \mathrm{Span}_{\mathsf{R}}\{x_0\}$ the 2-dimensional plane in H_0 orthogonal to x_0, we have that $T_{a,a^{-1}}$ acts as a rotation through a fixed angle μ, $0 < \mu < 2\pi$, in $\mathsf{H}_0 \ominus \mathrm{Span}_{\mathsf{R}}\{x_0\}$.*

Definition 2.2.5. We say that two quaternions x, y are *similar* if $axa^{-1} = y$ for some $a \in \mathsf{H} \setminus \{0\}$ and *congruent* if $axa^* = y$ for some $a \in \mathsf{H} \setminus \{0\}$.

Clearly, both similarity and congruence are equivalence relations. Denote by

$$\mathrm{Sim}\,(x) = \{y \in \mathsf{H} \,:\, y \text{ similar to } x\}$$

and

$$\mathrm{Con}\,(x) = \{y \in \mathsf{H} \,:\, y \text{ congruent to } x\}$$

the similarity orbit and the congruence orbit of $x \in \mathsf{H}$, respectively.

We have

$$\text{Con}\,(x) = \bigcup_{\lambda>0} \{\lambda \text{Sim}\,(x)\}.$$

Indeed, this follows from the formula $a^* = |a|^2 a^{-1}$, $a \in \mathsf{H} \setminus \{0\}$, with $\lambda = |a|^2$.

Theorem 2.2.6. *Fix* $x = x_0 + x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k} \in \mathsf{H}$, *where* $x_j \in \mathsf{R}$. *The following statements are equivalent for* $y = y_0 + y_1\mathsf{i} + y_2\mathsf{j} + y_3\mathsf{k} \in \mathsf{H}$, $y_j \in \mathsf{R}$:

(1) $y \in \text{Sim}\,(x)$;

(2) $y = axa^*$ *for some unit quaternion* a;

(3) $[y_0 \;\; y_1 \;\; y_2 \;\; y_3]^T = \begin{bmatrix} 1 & 0 \\ 0 & Q \end{bmatrix} [x_0 \;\; x_1 \;\; x_2 \;\; x_3]^T$ *for some* 3×3 *real orthogonal matrix* Q;

(4) $[y_0 \;\; y_1 \;\; y_2 \;\; y_3]^T = \begin{bmatrix} 1 & 0 \\ 0 & Q' \end{bmatrix} [x_0 \;\; x_1 \;\; x_2 \;\; x_3]^T$ *for some* 3×3 *real orthogonal matrix* Q' *having determinant* 1;

(5) $\mathfrak{R}(y) = \mathfrak{R}(x)$ *and* $|\mathfrak{V}(y)| = |\mathfrak{V}(x)|$.

Proof. (1) $\Longrightarrow$ (2): If $y = bxb^{-1}$, $b \in \mathsf{H} \setminus \{0\}$, then (2) holds with $a = b/|b|$.

(2) $\Longrightarrow$ (4): Follows from Corollary 2.2.2 and Ex. 2.7.7.

(4) $\Longrightarrow$ (3): Obvious.

(3) $\Longrightarrow$ (5): Follows from the isometric property of real orthogonal matrices, i.e., $\|Qu\| = \|u\|$ for all real orthogonal $Q \in \mathsf{R}^{n\times n}$ and all $u \in \mathsf{R}^{n\times 1}$.

(5) $\Longrightarrow$ (1): By hypothesis, $y_0 = x_0$ and $y_1^2 + y_2^{\stackrel{+}{2}} y_3^2 = x_1^2 + x_2^2 + x_3^2$. Consider the equation

$$(z_0 + z_1\mathsf{i} + z_2\mathsf{j} + z_3\mathsf{k})y = x(z_0 + z_1\mathsf{i} + z_2\mathsf{j} + z_3\mathsf{k}) \tag{2.2.3}$$

with real unknowns z_0, z_1, z_2, z_3. Equating the coefficients of each of $1, \mathsf{i}, \mathsf{j}, \mathsf{k}$ in the left and the right sides of (2.2.3), we see that (2.2.3), after some simple algebra, boils down to the system of equations

$$\begin{bmatrix} 0 & x_1 - y_1 & x_2 - y_2 & x_3 - y_3 \\ -x_1 + y_1 & 0 & x_3 + y_3 & -x_2 - y_2 \\ -x_2 + y_2 & -x_3 - y_3 & 0 & x_1 + y_1 \\ -x_3 + y_3 & x_2 + y_2 & -x_1 - y_1 & 0 \end{bmatrix} \begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ z_3 \end{bmatrix} = 0. \tag{2.2.4}$$

We claim that the matrix in the left-hand side of (2.2.4), call it X, is singular. Indeed, $X \begin{bmatrix} 0 \\ x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{bmatrix} = 0$, so, unless

$$x_1 + y_1 = x_2 + y_2 = x_3 + y_3 = 0, \tag{2.2.5}$$

the matrix X is singular. But if (2.2.5) holds, X is easily seen to be singular as well. Thus, (2.2.3) has a nontrivial solution, and (1) follows. ☐

Theorem 2.2.6 allows us to express the similarity orbits of quaternions in a more detailed way:

$$\text{Sim}\,(x) = \mathfrak{R}(x) + |\mathfrak{V}(x)|\mathsf{S},$$

where

$$\mathsf{S} := \{q \in \mathsf{H} \,:\, \mathfrak{R}(q) = 0, \quad |q| = 1\} = \{q \in \mathsf{H} \,:\, q^2 = -1\}. \tag{2.2.6}$$

Geometrically, S is the unit sphere in $\mathsf{R}^{3\times 1}$.

2.3 THE SYLVESTER EQUATION

Let $a, b \in \mathsf{H}$. In this section we study the *Sylvester equation*

$$ax - xb = y, \qquad x, y \in \mathsf{H},$$

and the corresponding real linear transformation

$$S_{a,b}(x) = ax - xb, \qquad x \in \mathsf{H}. \tag{2.3.1}$$

In what follows, we make use of the inner product.

Definition 2.3.1. The real-valued *inner product* of two quaternions is defined by

$$\backslash x_0 + x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k}, y_0 + y_1\mathsf{i} + y_2\mathsf{j} + y_3\mathsf{k}/ := x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3, \quad x_\ell, y_\ell \in \mathsf{R}.$$

Note that $\backslash x, x/ = |x|^2$ for every $x \in \mathsf{H}$. Also, for $x, y \in \mathsf{H}$ with zero real parts, we have $\backslash x, xy/ = \backslash y, xy/ = 0$, as can be easily verified using formula (2.1.1).

For given $a, b \in \mathsf{H}$, define the quaternions $\{x_+, y_+, x_-, y_-\}$ as follows:

(i) If $\mathfrak{V}(a)$ and $\mathfrak{V}(b)$ are linearly independent over R, we set

$$\begin{aligned} x_\pm &= (\pm|\mathfrak{V}(a)||\mathfrak{V}(b)| - (\mathfrak{V}(a))(\mathfrak{V}(b)))/n_\pm, \\ y_\pm &= (|\mathfrak{V}(a)|\mathfrak{V}(b) \pm |\mathfrak{V}(b)|\mathfrak{V}(a))/n_\pm, \end{aligned} \tag{2.3.2}$$

where

$$n_\pm = \sqrt{2|\mathfrak{V}(a)|\,|\mathfrak{V}(b)|\,(|\mathfrak{V}(a)||\mathfrak{V}(b)| \pm \backslash\mathfrak{V}(a), \mathfrak{V}(b)/)}.$$

Note that in view of the Cauchy-Schwarz inequality applied to $\mathfrak{V}(a)$ and $\mathfrak{V}(b)$ (interpreted as vectors in R^3) and the linear independence of $\mathfrak{V}(a)$ and $\mathfrak{V}(b)$, we have $|\mathfrak{V}(a)||\mathfrak{V}(b)| \pm \backslash\mathfrak{V}(a), \mathfrak{V}(b)/ > 0$.

(ii) Suppose $\mathfrak{V}(a)$ and $\mathfrak{V}(b)$ are linearly dependent over R. Then there exists $q \in \mathsf{H}$ with $\mathfrak{R}(q) = 0$, $|q| = 1$, $\mathfrak{V}(a) = |\mathfrak{V}(a)|q$, and $\mathfrak{V}(b) = |\mathfrak{V}(b)|q$ or $\mathfrak{V}(b) = -|\mathfrak{V}(b)|q$. Let $\widehat{q} \in \mathbb{H}$ be such that $\mathfrak{R}(\widehat{q}) = 0$, $|\widehat{q}| = 1$ and $\langle q, \widehat{q}\rangle = 0$. If $\mathfrak{V}(b) = |\mathfrak{V}(b)|q$, we define

$$x_+ = 1, \quad y_+ = q, \quad x_- = \widehat{q}, \quad y_- = q\widehat{q}.$$

If $\mathfrak{V}(b) = -|\mathfrak{V}(b)|q$, we define

$$x_+ = \widehat{q}, \quad y_+ = q\widehat{q}, \quad x_- = 1, \quad y_- = q.$$

Note that q and $\widehat{q}$ are not unique (for given a and b); more precisely, q is unique if $\mathfrak{V}(a) \neq 0$ and is unique up to negation if $\mathfrak{V}(b) \neq \mathfrak{V}(a) = 0$.

Furthermore, we define the subspaces

$$\mathcal{V}^+_{a,b} = \mathrm{Span}_{\mathsf{R}}\,\{x_+, y_+\}, \qquad \mathcal{V}^-_{a,b} = \mathrm{Span}_{\mathsf{R}}\,\{x_-, y_-\}.$$

The subspaces $\mathcal{V}^{\pm}_{a,b}$ are uniquely determined by a and b (unless $\mathfrak{V}(a) = \mathfrak{V}(b) = 0$), see Ex. 2.7.25. Moreover, we have

$$\mathcal{V}^{+}_{a,b} = \mathcal{V}^{-}_{a,b^*}, \quad \mathcal{V}^{-}_{a,b} = V^{+}_{a,b^*}$$

if at least one of a and b is nonreal.

Finally, we set

$$Q(\alpha, \beta) := \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}, \qquad \alpha, \beta \in \mathsf{R}.$$

Note the easily observed equality $\|Q(\alpha, \beta)\xi\| = \sqrt{\alpha^2 + \beta^2}\,\|\xi\|$ for all $\xi \in \mathsf{R}^{2\times 1}$, where we have used the euclidean norm in $\mathsf{R}^{2\times 1}$.

The main result of this section gives explicitly the orthonormal basis that reduces all three linear transformations $T_{a,1}$, $T_{1,b}$ (defined in (2.2.1)) and $S_{a,b}$ to a real Jordan form.

Theorem 2.3.2. (a) *The vectors x_+, y_+, x_-, y_- form an orthonormal basis* (*with respect to* $\langle \cdot, \cdot \rangle$) *of* H.

(b) *The equalities*

$$\begin{bmatrix} T_{a,1}(x_+) & T_{a,1}(y_+) & T_{a,1}(x_-) & T_{a,1}(y_-) \end{bmatrix} = \begin{bmatrix} x_+ & y_+ & x_- & y_- \end{bmatrix} \cdot (Q(\mathfrak{R}(a), |\mathfrak{V}(a)|) \oplus Q(\mathfrak{R}(a), |\mathfrak{V}(a)|))$$

and

$$\begin{bmatrix} T_{1,b}(x_+) & T_{1,b}(y_+) & T_{1,b}(x_-) & T_{1,b}(y_-) \end{bmatrix} = \begin{bmatrix} x_+ & y_+ & x_- & y_- \end{bmatrix} \cdot (Q(\mathfrak{R}(b), |\mathfrak{V}(b)|) \oplus Q(\mathfrak{R}(b), -|\mathfrak{V}(b)|))$$

hold true.

(c) *The subspaces $\mathcal{V}^{+}_{a,b}$ and $\mathcal{V}^{-}_{a,b}$ are both $T_{a,1}$- and $T_{1,b}$-invariant.*

(d) *The equality*

$$\begin{bmatrix} S_{a,b}(x_+) & S_{a,b}(y_+) & S_{a,b}(x_-) & S_{a,b}(y_-) \end{bmatrix} = \begin{bmatrix} x_+ & y_+ & x_- & y_- \end{bmatrix}$$
$$\cdot\, (Q(\mathfrak{R}(a) - \mathfrak{R}(b), |\mathfrak{V}(a)| - |\mathfrak{V}(b)|) \oplus Q(\mathfrak{R}(a) - \mathfrak{R}(b), |\mathfrak{V}(a)| + |\mathfrak{V}(b)|))$$

holds true.

Proof. Part (a) is established by a straightforward but tedious computation. Parts (c) and (d) are immediate from (b) as $S_{a,b} = T_{a,1} - T_{1,b}$. The verification of (b) is straightforward (Proposition 2.1.4(13) is used repeatedly). □

We can read off many important properties of $S_{a,b}$ from Theorem 2.3.2, such as the following.

Theorem 2.3.3. *Let $a, b \in \mathsf{H}$, and $S_{a,b}$ defined by* (2.3.1). *Then:*

(1) *the four singular values of $S_{a,b}$ are*

$$\sigma_1 = \sigma_2 = \sqrt{(\mathfrak{R}(a) - \mathfrak{R}(b))^2 + (|\mathfrak{V}(a)| + |\mathfrak{V}(b)|)^2},$$
$$\sigma_3 = \sigma_4 = \sqrt{(\mathfrak{R}(a) - \mathfrak{R}(b))^2 + (|\mathfrak{V}(a)| - |\mathfrak{V}(b)|)^2};$$

moreover, $|S_{a,b}(x)| = \sigma_4|x|$ for $x \in \mathcal{V}^{+}_{a,b}$, and $|S_{a,b}(x)| = \sigma_1|x|$ for $x \in \mathcal{V}^{-}_{a,b}$;

(2) *$S_{a,b}$ is singular if and only if $\mathfrak{R}(a) = \mathfrak{R}(b)$ and $|\mathfrak{V}(a)| = |\mathfrak{V}(b)|$. If these conditions hold and $a, b \notin \mathsf{R}$, then*

$$\operatorname{Ker} S_{a,b} = \mathcal{V}^+_{a,b} = \mathcal{V}^-_{a,b^*} = \operatorname{Ran} S_{a,b^*}, \quad \operatorname{Ran} S_{a,b} = \mathcal{V}^-_{a,b} = \mathcal{V}^+_{a,b^*} = \operatorname{Ker} S_{a,b^*};$$

(3) *$S_{a,b}$ has a real eigenvalue (which then is $\mathfrak{R}(a) - \mathfrak{R}(b)$) if and only if $|\mathfrak{V}(a)| = |\mathfrak{V}(b)|$ and the associated eigenspace is $\mathcal{V}^+_{a,b}$;*

(4) *the centralizer of $a \in \mathsf{H}$ is*

$$\operatorname{Cen}(a) := \{x \in \mathsf{H} : ax = xa\} = \operatorname{Ker} S_{a,a};$$

we have

$$\operatorname{Cen}(a) = \begin{cases} \mathsf{H} & \text{if } a \in \mathsf{R}, \\ \mathcal{V}^+_{a,a} = \operatorname{Span}_{\mathsf{R}}\{1, a\} & \text{otherwise.} \end{cases}$$

In the case a and b are similar (by Theorem 2.2.6 this happens if and only if $\mathfrak{R}(a) = \mathfrak{R}(b)$ and $|\mathfrak{V}(a)| = |\mathfrak{V}(b)|$), the kernel and image of $S_{a,b}$ have alternative descriptions.

Theorem 2.3.4. *Assume $a, b \in \mathsf{H}\backslash\mathsf{R}$ are similar, so that $b = z^{-1}az$, $z \in \mathsf{H}\backslash\{0\}$. Then:*

(a) $\operatorname{Ran} S_{a,b} = \operatorname{Ker} S_{a,b^*}$. *In other words, the equation $ax - xb = y$ has a solution x if and only if $ay = yb^*$;*

(b) $\operatorname{Ker} S_{a,b} = \operatorname{Cen}(a)\, z = \operatorname{Span}_{\mathsf{R}}\{z, az\}$.

Proof. Part (a) follows from Theorem 2.3.3(2). Part (b) is a consequence of the identity

$$ax - x(z^{-1}az) = (a(xz^{-1}) - (xz^{-1})a)z.$$

□

We conclude this section with formulas for the unique solution of the Sylvester equation, provided $S_{a,b}$ is invertible. For $a, b \in \mathsf{H}$, define

$$f_1(a, b) = b^2 - 2\mathfrak{R}(a)b + |a|^2, \quad f_2(a, b) = a^2 - 2\mathfrak{R}(b)a + |b|^2.$$

Proposition 2.3.5. *The following statements are equivalent:*

(1) $f_1(a, b) = 0$;

(2) $f_2(a, b) = 0$;

(3) *a and b are similar;*

(4) *$\mathfrak{R}(a) = \mathfrak{R}(b)$ and $|a| = |b|$.*

Proof. Equivalence of (3) and (4) follows from equivalence of (1) and (5) in Theorem 2.2.6. The implications (3) ⇒ (1) and (3) ⇒ (2) follow easily from Proposition 2.1.4(10). Suppose (1) holds true. Subtracting the equality $b^2 - 2\mathfrak{R}(b)b + |b|^2 = 0$ from $f_1(a, b) = 0$, we obtain $2(\mathfrak{R}(a) - \mathfrak{R}(b))b = |a|^2 - |b|^2$. If $\mathfrak{R}(a) \neq \mathfrak{R}(b)$, then b must be real. Subtracting $a^2 - 2\mathfrak{R}(a)a + |a|^2 = 0$ from $f_1(a, b) = 0$ yields $a = b$. If $\mathfrak{R}(a) = \mathfrak{R}(b)$, then $|a|^2 - |b|^2 = 0$, and (4) follows. Analogously, one proves that (2) implies (4). □

Theorem 2.3.6. *If $S_{a,b}$ is nonsingular, then the unique solution to the equation $S_{a,b}(x) = y$ satisfies*

$$x = a^*y(f_1(a,b))^{-1} - y(f_1(a,b))^{-1}b = a(f_2(a,b))^{-1}y - (f_2(a,b))^{-1}yb^*.$$

The proof follows from the equalities

$$S_{a,b}(a^*z - zb) = zf_1(a,b), \qquad S_{a,b}(az - zb^*) = f_2(a,b)z,$$

for all $z \in \mathsf{H}$, which can be verified without difficulty.

2.4 AUTOMORPHISMS AND INVOLUTIONS

Definition 2.4.1. An ordered triple of quaternions (q_1, q_2, q_3) is said to be a *units triple* if

$$\begin{aligned} q_1^2 = q_2^2 = q_3^2 = -1, \qquad & q_1q_2 = -q_2q_1 = q_3, \\ q_2q_3 = -q_3q_2 = q_1, \qquad & q_3q_1 = -q_1q_3 = q_2. \end{aligned} \tag{2.4.1}$$

For example, $\{\mathsf{i}, \mathsf{j}, \mathsf{k}\}$ is a units triple.

Proposition 2.4.2. *An ordered triple (q_1, q_2, q_3), $q_j \in \mathsf{H}$, is a units triple if and only if there exists a 3×3 real orthogonal matrix $P = [p_{\alpha,\beta}]_{\alpha,\beta=1}^3$ with determinant 1 such that*

$$q_\alpha = p_{1,\alpha}\mathsf{i} + p_{2,\alpha}\mathsf{j} + p_{3,\alpha}\mathsf{k}, \quad \alpha = 1, 2, 3. \tag{2.4.2}$$

Proof. A straightforward computation verifies that $x \in \mathsf{H}$ satisfies $x^2 = -1$ if and only if

$$x = a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k},$$

where $a_1, a_2, a_3 \in \mathsf{R}$ and $a_1^2 + a_2^2 + a_3^2 = 1$. Thus, we may assume that q_α are given by (2.4.2) with the vectors $p_\alpha := (p_{1,\alpha}, p_{2,\alpha}, p_{3,\alpha})^T \in \mathsf{R}^{3\times 1}$ having euclidean norm 1, for $\alpha = 1, 2, 3$. Next, in view of (2.1.1) we have

$$q_uq_v = -p_u^T p_v + \begin{bmatrix} \mathsf{i} & \mathsf{j} & \mathsf{k} \end{bmatrix} (p_u \times p_v), \quad u, v \in \{1, 2, 3\}. \tag{2.4.3}$$

The result of Proposition 2.4.2 now follows easily. □

In particular, for every units triple (q_1, q_2, q_3) the quaternions $1, q_1, q_2, q_3$ form a basis of the real vector space H.

Next, we consider endomorphisms and antiendomorphisms of quaternions.

Definition 2.4.3. A map $\phi : \mathsf{H} \longrightarrow \mathsf{H}$ is called an *endomorphism*, resp. an *antiendomorphism*, if $\phi(xy) = \phi(x)\phi(y)$, resp., $\phi(xy) = \phi(y)\phi(x)$ for all $x, y \in \mathsf{H}$, and $\phi(x + y) = \phi(x) + \phi(y)$ for all $x, y \in \mathsf{H}$. An antiendomorphism ϕ is called an *involution* if $\phi(\phi(x)) = x$ for every $x \in \mathsf{H}$.

An involution is necessarily one-to-one and onto.

Theorem 2.4.4. *Let ϕ be an endomorphism or an antiendomorphism of H. Assume that ϕ does not map H into zero. Then ϕ is one-to-one and onto H; thus, ϕ is in fact an automorphism or an antiautomorphism. Moreover, ϕ is real linear, and representing ϕ as a 4×4 real matrix with respect to the basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$, we have:*

(a) *ϕ is an automorphism if and only if*

$$\phi = \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix}, \tag{2.4.4}$$

where T is a 3×3 real orthogonal matrix of determinant 1;

(b) *ϕ is an antiautomorphism if and only if ϕ has the form* (2.4.4), *where T is a 3×3 real orthogonal matrix of determinant* -1;

(c) *ϕ is an involution if and only if*

$$\phi = \begin{bmatrix} 1 & 0 \\ 0 & T \end{bmatrix},$$

where either $T = -I_3$ or T is a 3×3 real orthogonal symmetric matrix with eigenvalues $1,\ 1,\ -1$.

Proof. Clearly, ϕ is one-to-one (indeed, if $\phi(x) = 0$ for some nonzero x, then for all $y \in \mathsf{H}$ we have $\phi(y) = \phi(yx^{-1})\phi(x) = 0$, a contradiction to the hypotheses of Theorem 2.4.4). Also, $\phi(x) = x$ for every real rational x.

We will use an observation which can be easily checked by straightforward algebra: *if $x \in \mathsf{H}$ is nonreal, then the commutant of x, namely, the set of all $y \in \mathsf{H}$ such that $xy = yx$ coincides with the set of all quaternions of the form $a + bx$, where $a, b \in \mathsf{R}$.*

Next, we prove that ϕ maps reals into reals. Arguing by contradiction, assume that $\phi(x)$ is nonreal for some real x. Since $xy = yx$ for every $y \in \mathsf{H}$, we have that $\phi(\mathsf{H})$ is contained in the commutant of $\phi(x)$, i.e., by the above observation,

$$\phi(\mathsf{H}) \subseteq \mathsf{R} + \mathsf{R}\phi(x).$$

However, the set $\mathsf{R} + \mathsf{R}\phi(x)$ contains only two square roots of -1, namely,

$$\pm \frac{\mathfrak{V}(\phi(x))}{|\mathfrak{V}(\phi(x))|}.$$

On the other hand, H contains a continuum of square roots of -1. Since ϕ maps square roots of -1 onto square roots of -1, ϕ cannot be one-to-one, a contradiction. Thus, ϕ maps reals into reals, and the restriction of ϕ to R is a nonzero endomorphism of the field of real numbers. Now R has no nontrivial endomorphisms (indeed, any nonzero endomorphism of R fixes every rational number, and since only nonnegative real numbers have real square roots, any such endomorphism is also order preserving, and these properties easily imply that any such endomorphism must be the identity). Therefore, we must have $\phi(x) = x$ for all $x \in \mathsf{R}$. Now, clearly, ϕ is a real linear map. Representing ϕ as a 4×4 matrix with respect to the basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$, we obtain the result of Part (a) from Proposition 2.4.2. For Part (b), note that if ϕ is an antiautomorphism, then a composition of ϕ with any fixed antiautomorphism is an automorphism. Taking a composition of ϕ with the antiautomorphism of standard conjugation $\mathsf{i} \longrightarrow -\mathsf{i}$, $\mathsf{j} \longrightarrow -\mathsf{j}$, $\mathsf{k} \longrightarrow -\mathsf{k}$, we see by Part (a) that the composition has the form (2.4.4) with T a real orthogonal matrix of determinant 1. Since the standard conjugation has the form (2.4.4) with $T = -I$, we obtain the result of Part (b). Finally, clearly ϕ is involutory if and only if the matrix T of (2.4.4) has eigenvalues ± 1, and (c) follows at once from (a) and (b). $\square$

Definition 2.4.5. If the former case of (c) holds true, then ϕ is the standard conjugation, and we say that ϕ is *standard*. If the latter case of (c) holds true, we say that these involutions are *nonstandard*.

Thus, the nonstandard involutions are parameterized by 1-dimensional real subspaces (representing eigenvectors of T corresponding to the eigenvalue -1) in R^3. In other words, the set of nonstandard involutions can be identified (as a topological space) with the 2-dimensional real projective space.

Here is another useful property of nonstandard involutions.

Lemma 2.4.6. *Let ϕ_1 and ϕ_2 be two distinct nonstandard involutions. Then for any $\alpha \in \mathsf{H}$, the equality $\phi_1(\alpha) = \phi_2(\alpha)$ holds if and only if $\phi_1(\alpha) = \phi_2(\alpha) = \alpha$.*

Proof. The "if" part is obvious. To prove the "only if" part, let T_1 and T_2 be the 3×3 real orthogonal matrices with eigenvalues $1, 1, -1$ such that

$$\phi_j = \begin{bmatrix} 1 & 0 \\ 0 & T_j \end{bmatrix}, \quad j = 1, 2,$$

as in Theorem 2.4.4(c). Then the "only if" part amounts to the following: if $T_1 \neq T_2$ and $T_1 x = T_2 x$ for some $x \in \mathsf{R}^{3\times 1}$, then $T_1 x = T_2 x = x$. Considering the orthogonal complement of a common eigenvector of T_1 and T_2 corresponding to the eigenvalue 1, the proof reduces to the statement that if $\widehat{T_1} \neq \widehat{T_2}$ are 2×2 real orthogonal symmetric matrices with determinants -1, then $\det(\widehat{T_1} - \widehat{T_2}) \neq 0$. This can be verified by elementary matrix manipulations, taking $\widehat{T_j}$ in the form $\begin{bmatrix} \cos\tau_j & \sin\tau_j \\ -\sin\tau_j & -\cos\tau_j \end{bmatrix}$, where $0 \leq \tau_j < 2\pi$, for $j = 1, 2$. □

Theorem 2.4.4 allows us to prove easily, using elementary linear algebra, the following well-known fact.

Proposition 2.4.7. *Every automorphism of H is inner—i.e., if $\phi : \mathsf{H} \to \mathsf{H}$ is an automorphism, then there exists $\alpha \in \mathsf{H} \setminus \{0\}$ such that*

$$\phi(x) = \alpha^{-1} x \alpha, \qquad \text{for all } x \in \mathsf{H}. \tag{2.4.5}$$

Proof. An elementary (but tedious) calculation shows that for

$$\alpha = a + b\mathsf{i} + c\mathsf{j} + d\mathsf{k} \in \mathsf{H} \setminus \{0\},$$

the 4×4 matrix representing the R-linear transformation $x \mapsto \alpha^{-1} x \alpha$ in the standard basis $\{1, \mathsf{i}, \mathsf{j}, \mathsf{k}\}$ is equal to

$$\begin{bmatrix} 1 & 0 \\ 0 & S \end{bmatrix},$$

where

$$S := \frac{1}{|\alpha|^2} \left(\begin{bmatrix} b \\ c \\ d \end{bmatrix} \begin{bmatrix} b & c & d \end{bmatrix} + \begin{bmatrix} a & d & -c \\ -d & a & b \\ c & -b & a \end{bmatrix}^2 \right) \tag{2.4.6}$$

(cf. Corollary 2.2.2). In view of Theorem 2.4.4, the matrix S is orthogonal with determinant 1, because the transformation $x \mapsto \alpha^{-1} x \alpha$ is an automorphism. Conversely, every 3×3 real orthogonal matrix T with determinant 1 has the form of the right-hand side of (2.4.6): Indeed, if $T = I$, choose $a \neq 0$ and $b = c = d = 0$. If $T \neq I$, choose $(b, c, d)^T$ as a unit length eigenvector of T corresponding to the eigenvalue 1. Observe that $(b, c, d)^T$ is also an eigenvector of S corresponding to the eigenvalue 1. Since the trace of S is equal to

$$1 + \frac{2}{a^2 + b^2 + c^2 + d^2}(a^2 - b^2 - c^2 - d^2),$$

it remains to choose a so that

$$\frac{a^2 - b^2 - c^2 - d^2}{a^2 + b^2 + c^2 + d^2}$$

coincides with the real part of those eigenvalues of T which are different from 1. The choice of $a \in \mathsf{R}$ is always possible because the real part of the eigenvalues of T other than 1 is between -1 and 1 and is not equal to 1. $\square$

One can write down concrete formulas for the automorphism ϕ given by (2.4.5). Namely, without loss of generality, we can assume that $\alpha \in \mathsf{H} \setminus \{0\}$ has the form

$$\alpha = \cos\left(\frac{\theta}{2}\right) - \sin\left(\frac{\theta}{2}\right) q_3, \quad \text{for some } \theta, \quad 0 \leq \theta < 2\pi,$$

where (q_1, q_2, q_3) is a suitable units triple. Indeed, if α is real, we take $\theta = 0$, and if $\alpha \notin \mathsf{R}$, we take $q_3 = \mathfrak{V}(\alpha)/|\mathfrak{V}(\alpha)|$ and, in both cases, divide α by $|\alpha|$. Then

$$\begin{array}{rclrcl} \phi(1) & = & 1, & \phi(q_1) & = & \cos(\theta)q_1 + \sin(\theta)q_2, \\ \phi(q_2) & = & -\sin(\theta)q_1 + \cos(\theta)q_2, & \phi(q_3) & = & q_3. \end{array}$$

These formulas can be verified by direct computation. One also verifies that in terms of formula (2.4.4), the automorphism ϕ is given with

$$T = P \cdot \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot P^T, \tag{2.4.7}$$

where the real orthogonal matrix $P \in \mathsf{R}^{3\times 3}$ is defined by

$$[q_1 \;\; q_2 \;\; q_3] = [\mathsf{i} \;\; \mathsf{j} \;\; \mathsf{k}]P.$$

Since every antiautomorphism of H is a composition of a fixed antiautomorphism (such as the conjugation) and a suitable automorphism, Proposition 2.4.7 that every antiautomorphism ϕ of H has the form $\phi(x) = \beta^{-1}x^*\beta$ for some $\beta \in \mathsf{H}$ with $|\beta| = 1$. One easily verifies that ϕ is an involution if and only if β^2 is real. Also, ϕ is nonstandard if and only if $\beta^2 = -1$, and then $\phi(\beta) = -\beta$. It follows from Theorem 2.4.4(c) that for every nonstandard involution ϕ, there is a unique (up to negation) $\beta \in \mathsf{H}$ such that $\phi(\beta) = -\beta$ and $\beta^2 = -1$. We select one of the two quaternions β with these properties and denote $\beta(\phi) = \beta$. Letting (q_1, q_2, q_3) be a units triple with $q_1 = \beta(\phi)$, we see that for every $x_0, x_1, x_2, x_3 \in \mathsf{R}$,

$$\phi(x_0 + x_1q_1 + x_2q_2 + x_3q_3) = x_0 - x_1q_1 + x_2q_2 + x_3q_3. \tag{2.4.8}$$

In particular, $\phi(x) = -x$, $x \in \mathsf{H}$, if and only if $x \in \mathsf{R}\beta$. We denote by $\text{Inv}\,(\phi)$ the set of quaternions left invariant by a nonstandard involution ϕ:

$$\text{Inv}\,(\phi) := \{x \in \mathsf{H} \,:\, \phi(x) = x\} = \text{Span}_{\mathsf{R}}\,\{1, q_2, q_3\}.$$

2.5 QUADRATIC MAPS

For a fixed involution ϕ consider quadratic maps of the form $x \mapsto \phi(x)\alpha x$, where $\alpha \in \mathsf{H} \setminus \{0\}$ is such that either $\phi(\alpha) = \alpha$ or $\phi(\alpha) = -\alpha$. (We exclude the trivial case $\alpha = 0$.) It is useful to find information about ranges of these maps. For example, if $\phi(\alpha) = \alpha$, is it true that the range of the quadratic map coincides with the set of quaternions that are fixed by ϕ?

We use the notation $\mathrm{Quad}_\phi(\alpha)$ for the map $x \mapsto \phi(x)\alpha x$, $x \in \mathsf{H}$.

If ϕ is a nonstandard involution, then there exists a unique (up to negation) $\beta \in \mathsf{H}$ such that $\phi(\beta) = -\beta$ and $|\beta| = 1$; note that β has zero real part. When working with nonstandard involutions ϕ, we often fix one such β and write $\beta(\phi)$ for β.

Theorem 2.5.1. (a) *If ϕ is a nonstandard involution, then*

$$\phi(x)\beta(\phi)x = \beta(\phi)|x|^2, \qquad \text{for every } x \in \mathsf{H}. \tag{2.5.1}$$

In particular, the range of $\mathrm{Quad}_\phi(\alpha)$ is the half-line

$$\{a\alpha \, : \, a \in \mathsf{R}, \; a \geq 0\} \tag{2.5.2}$$

for every $\alpha \neq 0$ such that $\phi(\alpha) = -\alpha$.

(b) *If ϕ is a nonstandard involution, then for every $\alpha \neq 0$ such that $\phi(\alpha) = \alpha$, the range of $\mathrm{Quad}_\phi(\alpha)$ coincides with the set (actually, a real subspace of H) $\mathrm{Inv}\,(\phi) := \{x \in \mathsf{H} \, : \, \phi(x) = x\}$; moreover, for every $\lambda \in \mathrm{Inv}\,(\phi)$ there exists $x \in \mathrm{Inv}\,(\phi)$ such that $\phi(x)\alpha x = \lambda$.*

(c) *If ϕ is the conjugation, then for every $\alpha \neq 0$ such that $\phi(\alpha) = -\alpha$ (in other words, $\mathfrak{R}(\alpha) = 0$, $\mathfrak{V}(\alpha) \neq 0$), the range of $\mathrm{Quad}_\phi(\alpha)$ coincides with the set (real subspace of H) of quaternions with zero real parts.*

(d) *If ϕ is the conjugation, then for every $\alpha \neq 0$ such that $\phi(\alpha) = \alpha$ (in other words, $\mathfrak{R}(\alpha) \neq 0$, $\mathfrak{V}(\alpha) = 0$), we have $\phi(x) = \alpha|x|^2$ for all $x \in \mathsf{H}$. In particular, the range of $\mathrm{Quad}_\phi(\alpha)$ is the half-line $\{x \in \mathsf{R} : x \geq 0\}$ if $\alpha > 0$ and the half-line $\{x \in \mathsf{R} : x \leq 0\}$ if $\alpha < 0$.*

Proof. The verification of (2.5.1) is straightforward, and Part (d) follows immediately from the basic properties of Proposition 2.1.4. Part (c) follows from the description of the conjugate orbit of α given in Ex. 2.7.24; indeed,

$$\mathrm{Quad}_\phi(\alpha) = \mathrm{Con}\,(\alpha) \cup \{0\}$$

if ϕ is the conjugation. Part (b) will be proved later. □

For a proof of Theorem 2.5.1(b), we investigate the square root function on quaternions, which is of independent interest. In the next theorem, we use the notation

$$\mathsf{R}_- := \{x \in \mathsf{H} \, : \, \mathfrak{R}(x) \leq 0, \, \mathfrak{V}(x) = 0\},$$

and $\sqrt{\cdot}$ stands for the nonnegative square root of a nonnegative number.

Theorem 2.5.2. (a) *Let $\lambda \in \mathsf{R}_-$. Then $x^2 = \lambda$ if and only if $x = \sqrt{|\lambda|}q$ for some $q \in \mathsf{H}$ with $\mathfrak{R}(q) = 0$, $|q| = 1$.*

(b) *Let* $\lambda \in \mathsf{H} \setminus \mathsf{R}_-$. *Then* $x^2 = \lambda$ *if and only if* $x = x_\lambda$ *or* $x = -x_\lambda$, *where*

$$x_\lambda := \frac{|\lambda| + \lambda}{\sqrt{2(|\lambda| + \Re(\lambda))}}.$$

If $\lambda \notin \mathsf{R}$ *then* x_λ *can be written in the form*

$$x_\lambda = \sqrt{\frac{|\lambda| + \Re(\lambda)}{2}} + \sqrt{\frac{|\lambda| - \Re(\lambda)}{2}}\,\frac{\mathfrak{V}(\lambda)}{|\mathfrak{V}(\lambda)|}.$$

Since x_λ in Part (b) is the unique square root of λ with positive real part, we introduce the notation

$$\sqrt{\lambda} := x_\lambda, \qquad \text{for all } \lambda \in \mathsf{H} \setminus \mathsf{R}_-. \tag{2.5.3}$$

Proof. We leave aside the trivial case $\lambda = 0$. Part (a) follows from the fact (easily proved by straightforward verification using the representation $x = x_0 + x_1\mathsf{i} + x_2\mathsf{j} + x_3\mathsf{k} \in \mathsf{H}$, where $x_0, x_1, x_2, x_3 \in \mathsf{R}$) that $x^2 = -1$ if and only if $\Re(x) = 0$ and $|x| = 1$.

Part (b). The equation $x^2 = \lambda$ implies

$$\lambda = x^2 = (2\Re(x) - x^*)x = 2\Re(x)x - |x|^2 = 2\Re(x)x - |\lambda|.$$

Hence, $|\lambda| + \lambda = 2\Re(x)x$, and (taking the real part in this equality) $|\lambda| + \Re(\lambda) = 2(\Re(x))^2$. If $\lambda \notin \mathsf{R}_-$ then $\Re(\lambda) + |\lambda| > 0$, and it follows that

$$x = \frac{|\lambda| + \lambda}{2\Re(x)} = \frac{|\lambda| + \lambda}{2\sqrt{(|\lambda| + \Re(\lambda))/2}} = \frac{|\lambda| + \lambda}{\sqrt{2(|\lambda| + \Re(\lambda))}}.$$

Thus,

$$\Re(x) = \frac{|\lambda| + \Re(\lambda)}{\sqrt{2(\Re(\lambda) + |\lambda|)}}$$

and, if $\mathfrak{V}(\lambda) \neq 0$, then

$$\mathfrak{V}(x) = \frac{\mathfrak{V}(\lambda)}{\sqrt{2(|\lambda| + \Re(\lambda))}} = \frac{\Re(\lambda)\sqrt{|\lambda| - \Re(\lambda)}}{\sqrt{2(|\lambda|^2 - \Re(\lambda)^2)}} = \sqrt{(|\lambda| - \Re(\lambda))/2} \cdot \frac{\mathfrak{V}(\lambda)}{|\mathfrak{V}(\lambda)|},$$

and the theorem is proved. □

Following from Theorem 2.5.2, the function $\lambda \mapsto \sqrt{\lambda}$ is continuous (even real analytic) on $\mathsf{H} \setminus \{0\}$. However, there is no continuous square root function on H as the following corollary shows.

Corollary 2.5.3. *Let* S *be a nonempty and connected subset of* $\mathsf{H} \setminus \mathsf{R}_-$, *and let* $u : S \to \mathsf{H}$ *be a function such that* $u(\lambda)^2 = \lambda$ *for all* $\lambda \in S$.

(a) *The function* $u(\cdot)$ *is continuous if and only if either* $u(\lambda) = \sqrt{\lambda}$ *for all* $\lambda \in S$ *or* $u(\lambda) = -\sqrt{\lambda}$ *for all* $\lambda \in S$.

(b) *Suppose that the function* $u(\cdot)$ *is continuous. Let* $\mu \in \mathsf{R}_-.$, $\mu \neq 0$. *Suppose further that there is a sequence* $\lambda_k \in S \setminus \mathsf{R}$ *such that* $\lim_{k\to\infty} \lambda_k = \mu$ *and* $\lim_{k\to\infty} \mathfrak{V}(\lambda_k)/|\mathfrak{V}(\lambda_k)|$ *does not exist. Then the function* $u(\cdot)$ *admits no continuous extension to* $S \cup \{\mu\}$.

Proof. Part (a): For $\lambda \notin \mathsf{R}_-$ the equation $u(\lambda)^2 = \lambda$ implies $u(\lambda) = s(\lambda)\sqrt{\lambda}$, where $s(\lambda) \in \{-1, 1\}$. Continuity of u and connectedness of S yield that $s(\lambda)$ is independent of λ.

Part (b): Without loss of generality, we may suppose $u(\lambda) = \sqrt{\lambda}$. Thus, for a sequence λ_k as in (b), we have

$$u(\lambda_k) = \underbrace{\sqrt{\frac{|\lambda_k| + \mathfrak{R}(\lambda_k)}{2}}}_{=a_k} + \underbrace{\sqrt{\frac{|\lambda_k| - \mathfrak{R}(\lambda_k)}{2}}}_{=b_k} \frac{\mathfrak{V}(\lambda_k)}{|\mathfrak{V}(\lambda_k)|}.$$

As k tends to infinity, a_k tends to 0 and b_k tends to $\sqrt{|\mu|}$, while the sequence $(\mathfrak{V}\lambda_k)/|\mathfrak{V}(\lambda_k)|$ does not converge. Thus each extension of $u(\cdot)$ to $S \cup \{\mu\}$ is discontinuous at μ. $\square$

Proof of Theorem 2.5.1(b). Clearly, the range of $\mathrm{Quad}_\phi\,(\alpha)$ is contained in $\mathrm{Inv}\,(\phi)$. We show that for all $\lambda \in \mathrm{Inv}\,(\phi)$, the equation $x_\phi \alpha x = \lambda$ has a solution $x \in \mathrm{Inv}\,(\phi)$. Suppose first that $\alpha\lambda \notin \mathsf{R}_-$. Then

$$x = \alpha^{-1}\sqrt{\alpha\lambda} = \alpha^{-1}\frac{|\alpha\lambda| + \alpha\lambda}{\sqrt{2(|\alpha\lambda| + \mathfrak{R}(\alpha\lambda))}} = \frac{\alpha^* \frac{|\lambda|}{|\alpha|} + \lambda}{\sqrt{2(|\alpha\lambda| + \mathfrak{R}(\alpha\lambda))}}$$

is a solution. To see this, observe that

$$x \in \mathrm{Span}_{\mathsf{R}}\,\{\alpha^*, \lambda\} \subseteq \mathrm{Inv}\,(\phi).$$

Thus,

$$x_\phi \alpha x = x\alpha x = \alpha^{-1}(\alpha x)^2 = \lambda.$$

Now suppose that $\alpha\lambda \in \mathsf{R}_-$. Then for any $q \in \mathsf{H}$ with $\mathfrak{R}(q) = 0$ and $|q| = 1$ the quaternion $x = \alpha^{-1}\sqrt{|\alpha\lambda|}\,q$ satisfies $(\alpha x)^2 = \alpha\lambda$. Hence, $x\alpha x = \lambda$. If $\alpha \notin \mathsf{R}$, choose $q = \mathfrak{V}(\alpha)/|\mathfrak{V}(\alpha)|$. Then $x \in \mathrm{Inv}\,(\phi)$. If $\alpha \in \mathsf{R}$, then $x \in \mathrm{Inv}\,(\phi)$ for any $q \in \mathrm{Inv}\,(\phi)$. $\square$

2.6 REAL AND COMPLEX MATRIX REPRESENTATIONS

In the sequel it will be often useful to represent quaternions as 4×4 real matrices or 2×2 complex matrices. These representations are described as follows.

Define the map

$$\chi : \mathsf{H} \to \mathsf{R}^{4\times 4}, \quad \chi(a_0 + a_1\mathsf{i} + a_2\mathsf{j} + a_3\mathsf{k}) = \begin{bmatrix} a_0 & -a_1 & a_3 & -a_2 \\ a_1 & a_0 & -a_2 & -a_3 \\ -a_3 & a_2 & a_0 & -a_1 \\ a_2 & a_3 & a_1 & a_0 \end{bmatrix},$$

where $a_0, a_1, a_2, a_3 \in \mathsf{R}$.

Proposition 2.6.1. *The map χ is a unital (i.e., $\chi(I) = I$) isomorphism of H onto the algebra of all 4×4 real matrices of the form $\lambda I + S$, where $\lambda \in \mathsf{R}$ and $S \in \mathsf{R}^{4\times 4}$ is a skew symmetric matrix of the form*

$$S = \left\{ \begin{bmatrix} 0 & -a_1 & a_3 & -a_2 \\ a_1 & 0 & -a_2 & -a_3 \\ -a_3 & a_2 & 0 & -a_1 \\ a_2 & a_3 & a_1 & 0 \end{bmatrix} : a_1, a_2, a_3 \in \mathsf{R} \right\}. \tag{2.6.1}$$

Proposition 2.6.1 can be proved by straightforward verification that $\chi(x+y) = \chi(x) + \chi(y)$, $\chi(xy) = \chi(x)\chi(y)$ for all $x, y \in \mathsf{H}$, $\chi(1) = I$, and χ is one-to-one and onto map on the indicated algebra of 4×4 real matrices (Ex. 2.7.10).

For $\alpha = a_0 + \mathsf{i}a_1 + \mathsf{j}a_2 + \mathsf{k}a_3 \in \mathsf{H}$, $a_0, a_1, a_2, a_3 \in \mathsf{R}$, define

$$\omega(\alpha) = \begin{bmatrix} a_0 + \mathsf{i}a_1 & a_2 + \mathsf{i}a_3 \\ -a_2 + \mathsf{i}a_3 & a_0 - \mathsf{i}a_1 \end{bmatrix} \in \mathsf{C}^{2\times 2}.$$

Proposition 2.6.2. *The map ω is a unital isomorphism of H onto the algebra of all 2×2 complex matrices of the form $\begin{bmatrix} z & u \\ -\overline{u} & \overline{z} \end{bmatrix}$, where $z, u \in \mathsf{C}$ are arbitrary.*

The proof is again by straightforward verification of the required properties (Ex. 2.7.11).

The ordered triple of matrices

$$-\mathsf{i}\omega(\mathsf{k}) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad -\mathsf{i}\omega(\mathsf{j}) = \begin{bmatrix} 0 & -\mathsf{i} \\ \mathsf{i} & 0 \end{bmatrix}, \quad -\mathsf{i}\omega(\mathsf{i}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

is known as *Pauli spin matrices*, of importance of studies of spin in quantum mechanics. Note that the Pauli spin matrices are hermitian and unitary. Together with I_2, the Pauli spin matrices form a basis in the real vector space of 2×2 hermitian matrices.

2.7 EXERCISES

Ex. 2.7.1. Verify Proposition 2.1.2.

Ex. 2.7.2. Verify Proposition 2.1.4.

Ex. 2.7.3. Prove that $|x + y| = |x| + |y|$, $x, y \in \mathsf{H}$, holds if and only if either at least one of x, y is zero or $x \neq 0$ and $y \neq 0$ are positive real multiples of each other.

Ex. 2.7.4. Solve the following equations:

(1) $x^4 + 1 = 0$, $x \in \mathsf{H}$;

(2) $x^m + 1 = 0$, $x \in \mathsf{H}$, where m is a fixed even positive integer.

Hint: For a fixed $q \in \mathsf{H}$, with $\mathfrak{R}(q) = 0$ and $|q| = 1$, consider solutions in $\text{Span}_{\mathsf{R}}\,\{1, q\}$ which is isomorphic to C.

Ex. 2.7.5. Verify the following equality for $x \in \mathsf{H}$:

$$x^2 = |\mathfrak{R}(x)|^2 - |\mathfrak{V}(x)|^2 + 2\mathfrak{R}(x)\mathfrak{V}(x).$$

Ex. 2.7.6. Let $a \in \mathsf{H}$ be such that $ax = xa$ for every $x \in \mathsf{H}_0$, where H_0 is a real 3-dimensional subspace of H. Show that $a \in \mathsf{R}$. Show by example that this statement is generally not true if H_0 is taken to be a real 2-dimensional subspace of H.

Ex. 2.7.7. Verify that the matrix $Y := (a_0^2 + a_1^2 + a_2^2 + a_3^2)^{-1}X$ of Corollary 2.2.2 is orthogonal and has determinant 1. Hint: The determinant of a real orthogonal matrix can be only ± 1, and the case that the determinant is equal to -1 is impossible here.

Ex. 2.7.8. Prove that H has no divisors of zero: If $x_1, \ldots, x_k \in \mathsf{H}$ are such that $x_1 \cdots x_k = 0$, then at least one of the x_j's is equal to zero.

Ex. 2.7.9. Find the kernel and the range of the real linear transformation $x \mapsto xa - ax$, $x \in \mathsf{H}$, where $a \in \mathsf{H} \setminus \{0\}$ is fixed.

Ex. 2.7.10. Prove Proposition 2.6.1.

Ex. 2.7.11. Prove Proposition 2.6.2.

Ex. 2.7.12. Show that any two involutions ϕ_1 and ϕ_2 of H different from the conjugation are similar: There exists $\alpha \in \mathsf{H} \setminus \{0\}$ (which depends on ϕ_1 and ϕ_2) such that $\phi_2(x) = \alpha^{-1}\phi_1(x)\alpha$, $x \in \mathsf{H}$.

Ex. 2.7.13. Let (q_1, q_2, q_3) be a units triple. Show that $xq_1 = -q_1 x$, $x \in \mathsf{H}$ if and only if x is a real linear combination of q_2 and q_3.

Ex. 2.7.14. Let α and β be two quaternions with zero real parts. Prove that $\alpha\beta = \beta^*\alpha$ if and only if α and β are *orthogonal,* i.e., writing

$$\alpha = \mathsf{i}\alpha_1 + \mathsf{j}\alpha_2 + \mathsf{k}\alpha_3, \quad \beta = \mathsf{i}\beta_1 + \mathsf{j}\beta_2 + \mathsf{k}\beta_3, \qquad \alpha_\ell, \beta_\ell \in \mathsf{R} \ \text{ for } \ell = 1, 2, 3,$$

we have

$$\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 = 0. \tag{2.7.1}$$

Ex. 2.7.15. Let u, v and u', v' be two pairs of unit length quaternions with zero real parts. Prove that there exists an automorphism ϕ of H such that $\phi(u) = u'$, $\phi(v) = v'$ if and only if the angle θ $(0 \le \theta \le \pi)$ between u and v is equal to that between u' and v'. Is ϕ unique?

Ex. 2.7.16. Let u, v and u', v' be as in Ex. 2.7.15. Prove that there exists an antiautomorphism ψ of H such that $\psi(u) = u'$, $\psi(v) = v'$ if and only if the angle between u and v is equal to that between u' and v'. Is ψ unique?

Ex. 2.7.17. Identify all pairs (x, y) of quaternions in the set $\{1 + \mathsf{i}, 1 + 2\mathsf{j}, 2 + 2\mathsf{j}, 1 + 3\mathsf{k}\}$ that are similar, resp. congruent, to each other. Find $a \in \mathsf{H}$ such that $axa^{-1} = y$ in case x and y are similar or $axa^* = y$ in case x and y are congruent.

Ex. 2.7.18. Find:

(a) all $x \in \mathsf{H}$ such that $\mathsf{i}x = x\mathsf{j}$;

(b) all $y \in \mathsf{H}$ such that $2\mathsf{i}y = y(\mathsf{j} + \mathsf{k})$;

(c) all $z \in \mathsf{H}$ such that $3\mathsf{i}z = \sqrt{3}z(\mathsf{i} + \mathsf{j} + \mathsf{k})$.

Ex. 2.7.19. Show that all solutions z of the equation $z^2 = p + \mathsf{i}q$, where $p, q \in \mathsf{R}$ and $q \neq 0$ and $z \in \mathsf{H}$ is to be found, belong to $\text{Span}_{\mathsf{R}}\{1, \mathsf{i}\}$.

Ex. 2.7.20. Prove Bohr's inequality:

$$|z + w|^2 \le p|z|^2 + q|w|^2, \qquad z, w \in \mathsf{H}, \tag{2.7.2}$$

where p, q are positive real numbers such that $(1/p) + (1/q) = 1$, with the equality holding in (2.7.2) if and only if $w = (p - 1)z$.

Ex. 2.7.21. Show that every quaternion can be written in infinitely many ways as product of two quaternions with zero real parts. Hint: Show that for all $x \in \text{Span}_{\mathsf{R}}\{1, \mathsf{i}\}$, $x_1 \in \text{Span}_{\mathsf{R}}\{\mathsf{j}, \mathsf{k}\}$ with $x_1 \neq 0$, there exists a $y \in \text{Span}_{\mathsf{R}}\{\mathsf{j}, \mathsf{k}\}$ such that $x = x_1 y$.

Ex. 2.7.22. Let $S := \{x \in \mathsf{H} : |x| = 1\}$ be the set of quaternions of norm one.

(1) Show that S is a (multiplicative) subgroup of $\mathsf{H} \setminus \{0\}$, i.e., $x, y \in S$ implies $xy \in S$ and $x^{-1} \in S$.

(2) Prove that the function $f : S \times S \to S$ defined by $f(x, y) = y^{-1}x^{-1}yx$ maps onto S. Hint: Use the result of Ex. 2.7.21.

Ex. 2.7.23. Define S as in Ex. 2.7.22.

(a) Show that every automorphism or antiautomorphism of H maps S onto itself.

(b) Show that every automorphism ϕ of H has at least four fixed points x in S, i.e., $x \in S$ such that $\phi(x) = x$. Hint: Every 3×3 real orthogonal matrix with determinant one has eigenvalue 1.

Ex. 2.7.24. Verify the following formula for the congruence orbit of $x \in \mathsf{H}$:

$$\text{Con}\,(x) = \{(\mathfrak{R}(x))r + \mathsf{S}|\mathfrak{V}x|r \,:\, r > 0\},$$

where S is defined in (2.2.6).

Ex. 2.7.25. Verify that if at least one of $a, b \in \mathsf{H}$ is nonreal, then the subspaces $\mathcal{V}^{\pm}_{a,b}$ (defined in Section 2.3) are uniquely determined by a and b.

Ex. 2.7.26. Verify formula (2.4.7).

2.8 NOTES

Most of the material in this chapter is standard (except Section 2.3), although some of it, e.g., Theorem 2.2.1, is difficult to locate in the literature. The representation of quaternion multiplication as rotations in 3-dimensional real vector space, as stated in Corollary 2.2.4, as well as connection with cross products (2.1.1), is a key to many applications in geometry, mechanics, and engineering, which is written up in many books; see, e.g., Ward [157]. For applications in feedback regulator problems and stability analysis, see, e.g., Wie et al. [159].

Proposition 2.1.4 contains basic elementary properties of H; a large set of such properties is found in Zhang [164]. The result of Theorem 2.3.3(2), as well as that of Theorem 2.3.4(a), is found in Johnson [74]. Ex. 2.7.21 is taken from Koecher and Remmert [81].

The material of Section 2.3 is taken from Karow [79]. Theorem 2.5.2 and Corollary 2.5.3, as well as the proof of Theorem 2.5.1(b), were suggested by Karow [80].

For some material on antiautomorphisms of quaternions, see von Randow [127].

We mention the following “fundamental theorem of algebra” concerning polynomial equations in quaternions:

Theorem 2.8.1. *Let $f(x)$ be a function of the form*

$$f(x) = a_0xa_1xa_2 \cdots a_{n-1}xa_n + g(x), \qquad x \in \mathsf{H},$$

where $a_0, a_1, \ldots a_n \in \mathsf{H} \setminus \{0\}$, and $g(x)$ is a finite sum of monomials of the form $b_0xb_1x \cdots b_{k-1}xb_k$ with $k < n$. Then there exists $x_0 \in \mathsf{H}$ such that $f(x_0) = 0$.

Theorem 2.8.1 was proved by Eilenberg and Niven [36]. The complete proof is based on topological methods and is beyond the scope of this book.

Quaternions were a starting point in many important developments in modern algebra: octonions, division algebras, and Clifford algebras, to name a few. While we cannot here go in depth into any of these developments, some references for further reading are provided: Conway and Smith [30], Lounesto [103], and Cohn [29]. In particular, note that the division ring of quaternions is a special case of a finite dimensional central simple algebra over a field. Many well-known results for such algebras apply to H, for example, the Skolem-Noether theorem that asserts (in a basic formulation) that every automorphism of such algebras is inner (cf. Proposition 2.4.7); see, e.g., Berhuy and Oggier [14] or Farb and Dennis [37] for more details.

Rotations transforms in 3- and 4-dimensional real vector space via quaternions (cf. Corollary 2.2.4) and related topics are studied in depth in many books, e.g., Ward [157], Vince [153], and Kuipers [84].

For remarks and further references concerning historical developments of quaternions, consult Koecher and Remmert [81] or Vince [153].