QUATERNIONS AND ROTATIONS

SVANTE JANSON

1. Introduction

The purpose of this note is to show some well-known relations between quaternions and the Lie groups $\mathsf{SO}(3)$ and $\mathsf{SO}(4)$ (rotations in $\mathbb{R}^3$ and $\mathbb{R}^4$) and $\mathsf{SU}(2)$ (unitary operators in $\mathbb{C}^2$ with determinant 1). In particular, this gives a simple description of the 2–1 covering homomorphisms $\mathsf{SU}(2) \to \mathsf{SO}(3)$ and $\mathsf{SU}(2) \times \mathsf{SU}(2) \to \mathsf{SO}(4)$.

We also include an explanation of the compact sympletic group $\mathsf{Sp}(n)$ as the group of norm-preserving $\mathbb{H}$-linear maps of the quaternionic vector space $\mathbb{H}^n$ into itself. (These could be thought of as rotations in $\mathbb{H}^n$.)

The quaternions were discovered by Sir William Rowan Hamilton in Dublin in 1843, and the results below were discovered rather shortly afterwards by him and others. In particular, the realization of $\mathsf{SO}(3)$ was found by Hamilton in 1844, and the realization of $\mathsf{SO}(4)$ by Cayley in 1855.

For more on quaternions, see e.g. [1, Chapter 7].

2. Quaternions

A *quaternion* is an element of the *quaternion algebra* $\mathbb{H}$, which is a four-dimensional algebra over $\mathbb{R}$ that as a vector space has a basis $\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$; i.e., a quaternion can uniquely be written as $w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ with $w, x, y, z \in \mathbb{R}$. The multiplication is bilinear, and is thus determined by the products of the basis elements; 1 is a unit and

$$\begin{aligned} &\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \\ &\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \qquad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \qquad \mathbf{ki} = -\mathbf{ik} = \mathbf{j}. \end{aligned} \tag{2.1}$$

The multiplication defined in this way is associative but not commutative. Moreover, $\mathbb{H}$ is a *division ring*, i.e., every quaternion $q \in \mathbb{H}$ with $q \neq 0$ has a (unique) inverse q^{-1}.

In fact, if we define the *conjugate* of a quaternion $q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ as

$$\bar{q} := w - x\mathbf{i} - y\mathbf{j} - z\mathbf{k}, \tag{2.2}$$

and the *norm* $\|q\|$ by

$$\|q\|^2 := q\bar{q}, \tag{2.3}$$

then a simple calculation shows that

$$\|q\|^2 := w^2 + x^2 + y^2 + z^2, \tag{2.4}$$

so $\|q\|$ is the usual Euclidean norm in $\mathbb{R}^4$, and in particular $\|q\| > 0$ if $q \neq 0$. Hence, when $q \neq 0$ we can define an inverse by

$$q^{-1} := \bar{q}/\|q\|^2. \tag{2.5}$$

Date: 27 January, 2017.

We regard $\mathbb{R}$ as the subspace of $\mathbb{H}$ spanned by 1; note that this is a subalgebra and that addition and multiplication in $\mathbb{R}$ and $\mathbb{H}$ agree for $x, y \in \mathbb{R}$.

Let $\operatorname{Im}\mathbb{H}$ be the three-dimensional subspace of $\mathbb{H}$ spanned by $\mathbf{i}, \mathbf{j}, \mathbf{k}$, i.e.,

$$\operatorname{Im}\mathbb{H} = \{x\mathbf{i} + y\mathbf{j} + z\mathbf{k} : x, y, z \in \mathbb{R}\} = \{q \in \mathbb{H} : \bar{q} = -q\}. \tag{2.6}$$

The elements of $\operatorname{Im}\mathbb{H}$ are called *purely imaginary quaternions.*

Note that $\mathbb{H}$ is a direct sum $\mathbb{R} \oplus \operatorname{Im}\mathbb{H}$ of the subspaces $\mathbb{R}$ and $\operatorname{Im}\mathbb{H}$. In other words, every quaternion q can be written (uniquely) as $r + p$, where $r \in \mathbb{R}$ and $p \in \operatorname{Im}\mathbb{H}$; we call r the *real part* and p the *imaginary part* or *pure part* or *vector part* of q.

We note the easily verified rule

$$\overline{q_1 q_2} = \bar{q}_2 \bar{q}_1, \qquad q_1, q_2 \in \mathbb{H}. \tag{2.7}$$

A simple consequence of this is that

$$\|q_1 q_2\|^2 = q_1 q_2 \bar{q}_2 \bar{q}_1 = q_1 \|q_2\|^2 \bar{q}_1 = q_1 \bar{q}_1 \|q_2\|^2 = \|q_1\|^2 \|q_2\|^2, \tag{2.8}$$

and thus

$$\|q_1 q_2\| = \|q_1\| \|q_2\| = \|q_2 q_1\|. \tag{2.9}$$

Furthermore,

$$\|\bar{q}\| = \|q\| \tag{2.10}$$

and, for every $q \neq 0$,

$$\|q^{-1}\| = \|q\|^{-1}. \tag{2.11}$$

3. Vector calculus

If we identify $\operatorname{Im}\mathbb{H}$ and $\mathbb{R}^3$ in the obvious way (identifying $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and the standard basis of $\mathbb{R}^3$), then it follows from (2.1) and bilinearity that for any two purely imaginary quaternions $u, v \in \operatorname{Im}\mathbb{H}$,

$$uv = -\langle u, v\rangle + u \times v, \tag{3.1}$$

where $\langle u, v\rangle$ is the scalar product and $u \times v$ the vector product in $\mathbb{R}^3$. (Thus, in (3.1), $-\langle u, v\rangle$ is the real part and $u \times v$ is the imaginary part of the quaternion uv.)

This connection with vector calculus in $\mathbb{R}^3$ was one important motivation for the interest in quaternions in the 19th century.

We use it to derive the following important symmetry property of the quaternion algebra.

Theorem 3.1. *Let $T \in \mathsf{SO}(3)$, i.e., a rotation of $\mathbb{R}^3$, and regard T as acting on $\operatorname{Im}\mathbb{H}$. Extend T to a linear operator $\widehat{T} : \mathbb{H} \to \mathbb{H}$ by $\widehat{T}(r + u) = r + Tu$ for $r \in \mathbb{R}$, $u \in \operatorname{Im}\mathbb{H}$. Then*

$$\widehat{T}(q_1 q_2) = \widehat{T}(q_1)\widehat{T}(q_2), \qquad q_1, q_2 \in \mathbb{H}, \tag{3.2}$$

and thus $\widehat{T}$ is an algebra automorphism of $\mathbb{H}$.

Proof. It is well-known that if $u, v \in \operatorname{Im}\mathbb{H} = \mathbb{R}^3$, then $\langle Tu, Tv\rangle = \langle u, v\rangle$ $T(u) \times T(v) = T(u \times v)$, and thus (3.1) yields $(\widehat{T}u)(\widehat{T}v) = (Tu)(Tv) = \widehat{T}(uv)$. Thus (3.2) holds when q_1 and q_2 are purely imaginary. Furthermore, (3.2) is trivial if $q_1 \in \mathbb{R}$ or $q_2 \in \mathbb{R}$, and the general case follows by bilinearity. ☐

Remark 3.2. Conversely, it is not difficult to show that every automorphism of $\mathbb{H}$ is of the form $\widehat{T}$ for some $T \in \mathsf{SO}(3)$. The automorphism group of $\mathbb{H}$ is thus isomorphic to $\mathsf{SO}(3)$ by the bijection $\widehat{T} \leftrightarrow T$.

4. Unit quaternions

Let

$$Q := \{q \in \mathbb{H} : \|q\| = 1\}. \tag{4.1}$$

Elements of Q are called *unit quaternions*.

If $q_1, q_2 \in Q$, then by (2.9), $q_1 q_2 \in Q$. Furthermore, if $q \in Q$, then $q^{-1} = \bar{q} \in Q$. Also, trivially, $1 \in Q$. These properties show that Q is a group.

As a set, Q is the unit sphere in $\mathbb{H} = \mathbb{R}^4$, i.e., the three-dimensional sphere $\mathbb{S}^3$. We give Q the topology and manifold structure of $\mathbb{S}^3$, inherited from the Euclidean space $\mathbb{R}^4$. It is obvious that the group operations are continuous and differentiable, and thus Q is a Lie group. Furthermore, Q is compact, connected and simply connected, since $\mathbb{S}^3$ is.

5. Quaternions as matrices

The quaternion algebra $\mathbb{H}$ can be realized as a subalgebra of the matrix algebra $M_2(\mathbb{C})$ by identifying 1 and the unit matrix $I = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and setting

$$\mathbf{i} := \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad \mathbf{j} := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \mathbf{k} := \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}; \tag{5.1}$$

note that the matrices (5.1) satisfy (2.1). The quaternion algebra then is identified with the set $\widetilde{\mathbb{H}}$ of all complex 2×2 matrices $\left(\begin{smallmatrix} \alpha & \gamma \\ \beta & \delta \end{smallmatrix}\right)$ that are real linear combinations of I and the three matrices $\mathbf{i}, \mathbf{j}, \mathbf{k}$. This means that we make the identification

$$\mathbb{H} \cong \widetilde{\mathbb{H}} := \left\{ \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\}. \tag{5.2}$$

In this realization of the quaternions, the multiplication is ordinary matrix multiplication.

Moreover, if a quaternion $q = \left(\begin{smallmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{smallmatrix}\right)$, then $\bar{q} = \left(\begin{smallmatrix} \bar{\alpha} & \bar{\beta} \\ -\beta & \alpha \end{smallmatrix}\right) = \left(\begin{smallmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{smallmatrix}\right)^*$, and $\|q\|^2 = |\alpha|^2 + |\beta|^2$. In particular the group Q of unit quaternions becomes

$$Q = \left\{ \begin{pmatrix} \alpha & -\bar{\beta} \\ \beta & \bar{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C},\ |\alpha|^2 + |\beta|^2 = 1 \right\} = \mathsf{SU}(2), \tag{5.3}$$

the group of unitary 2×2 matrices, see [2, Chapter 1 and Exercise 1.5].

Remark 5.1. The matrices in (5.1) are not unique, and other choices can be used instead. (For example, a minor variation of the signs of some entries is used in [1].)

Remark 5.2. $M_2(\mathbb{C})$ is a 4-dimensional complex vector space, and therefore an 8-dimensional real vector space. The construction above identifies $\mathbb{H}$ with the 4-dimensional subspace $\widetilde{\mathbb{H}}$ of this real vector space.

6. Rotations

If $q \in Q$, let L_q and R_q be the linear operators $\mathbb{H} \to \mathbb{H}$ given by

$$L_q p = qp, \qquad R_q p = pq. \tag{6.1}$$

By (2.9), the operators L_q and R_q preserve the norm in $\mathbb{H} = \mathbb{R}^4$, and thus $L_q, R_q \in \mathsf{O}(4)$. Moreover, for $q = 1$ we obtain $L_1 = R_1 = I$ with determinant $+1$. Since the maps $q \mapsto L_q$ and $q \mapsto R_q$ are continuous maps $Q \to \mathsf{O}(4)$, and Q is connected, it follows that $L_q, R_q \in \mathsf{SO}(4)$ for every $q \in Q$.

Moreover, $L_{q_1 q_2} = L_{q_1} L_{q_2}$ and $R_{q_1 q_2} = R_{q_2} R_{q_1}$, and it follows that $q \mapsto L_q$ and $q \mapsto R_{q^{-1}}$ are homomorphisms $Q \to \mathsf{SO}(4)$. This mapping is obviously continuous and differentiable, and is thus a Lie group homomorphism.

Combining these two maps we get a homomorphism $\Psi : Q \times Q \to \mathsf{SO}(4)$ given by

$$(q_1, q_2) \mapsto \Psi_{q_1,q_2} := L_{q_1} R_{q_2^{-1}}, \tag{6.2}$$

i.e.,

$$\Psi_{q_1,q_2}(p) = q_1 p q_2^{-1}, \qquad p \in \mathbb{H}. \tag{6.3}$$

We specialize to the case $q_1 = q_2$ and write $\Phi_q := \Psi_{q,q} = L_q R_{q^{-1}}$, i.e.,

$$\Phi_q(p) = qpq^{-1}. \tag{6.4}$$

In particular, $\Phi_q 1 = 1$, and thus every element of $\mathbb{R} \subset \mathbb{H}$ is fixed by Φ_q. Since $\Phi_q \in \mathsf{SO}(4)$, also the orthogonal complement $\mathbb{R}^\perp = \operatorname{Im} \mathbb{H}$ is invariant under Φ_q. In other words, we can regard Φ_q as a linar operator $\operatorname{Im}\mathbb{H} \to \operatorname{Im}\mathbb{H}$, for every $q \in Q$. Since $\operatorname{Im}\mathbb{H} \cong \mathbb{R}^3$, we can thus regard Φ_q as a linear operator on $\mathbb{R}^3$. Since each Φ_q is an isometry, this means that $T_q \in \mathsf{O}(3)$, and thus $\Phi : q \mapsto \Phi_q$ is a homomorphism $Q \to \mathsf{O}(3)$; as above we see by continuity that in fact this homomorphism maps into $\mathsf{SO}(3)$, i.e., $\Phi_q \in \mathsf{SO}(3)$ for every $q \in Q$. This homomorphism is continuous and differentiable since Ψ is.

We have shown that the mapping $q \mapsto \Phi_q$ is a Lie group homomorphism $Q \to \mathsf{SO}(3)$. If we identify Q and $\mathsf{SU}(2)$ as in (5.3), then this can be seen as a homomorphism $\mathsf{SU}(2) \to \mathsf{SO}(3)$. In fact, this is the same as the homomorphism defined in [2, Section 1.4]. To see this, it suffices to note that the operator Φ_q defined there is the same as (6.4), and that the three-dimensional space V defined in [2] equals $i \operatorname{Im} \widetilde{\mathbb{H}}$, so $\operatorname{Im} \widetilde{\mathbb{H}} = iV$; obviously, Φ_q acts the same way in $\operatorname{Im} \widetilde{\mathbb{H}}$ and $i \operatorname{Im} \widetilde{\mathbb{H}}$. (To be formal, the representations in $\operatorname{Im}\mathbb{H}$ and $i\operatorname{Im}\mathbb{H}$ are isomorphic, with $p \mapsto ip$ as an intertwining operator.)

The rotation Φ_q can be described explicitly. Note that if a unit quaternion q is decomposed into real and imaginary parts as $q = r + p$, with $r \in \mathbb{R}$ and $p \in \operatorname{Im}\mathbb{H}$, then $r^2 + \|p\|^2 = \|q\|^2 = 1$, and thus there exists a real number θ such that $r = \cos\theta$ and $\|q\| = \sin\theta$. Consequently, if $q \in Q$, then there exists $\theta \in \mathbb{R}$ and a unit vector $u \in \operatorname{Im}\mathbb{H}$ such that

$$q = \cos\theta + \sin\theta \cdot u. \tag{6.5}$$

Lemma 6.1. *Let q be as in* (6.5). *Then Φ_q is the rotation by angle 2θ around the axis through u.*

Proof. First, it is clear that q and u commute, and thus $\Phi_q u = quq^{-1} = u$, so u is a fixed point of Φ_q, and thus Φ_q is a rotation around the axis through u.

To compute the angle of this rotation, consider first the case $u = \mathbf{i}$. Then a direct calculation using (2.1) shows that

$$\begin{aligned}\Phi_{\cos\theta+\mathbf{i}\sin\theta}\mathbf{j} &= (\cos\theta + \mathbf{i}\sin\theta)\mathbf{j}(\cos\theta - \mathbf{i}\sin\theta)\\ &= (\cos^2\theta - \sin^2\theta)\mathbf{j} + 2\sin\theta\cos\theta\,\mathbf{k} \qquad (6.6)\\ &= \cos(2\theta)\mathbf{j} + \sin(2\theta)\mathbf{k}.\end{aligned}$$

Since $\Phi_{\cos\theta+\mathbf{i}\sin\theta}$ has $\mathbf{i}$ as a fixed point, it is a rotation in the $\mathbf{jk}$ plane, and we see from (6.6) that it is a rotation by 2θ (using the standard orientation). This verifies the lemma when $u = \mathbf{i}$.

In general, note first that if T is any rotation of $\operatorname{Im}\mathbb{H}$, $q \in Q$ and $h \in \mathbb{H}$, then, using the notation of Theorem 3.1,

$$T(\Phi_q h) = \widehat{T}(\Phi_q h) = \widehat{T}(qhq^{-1}) = \widehat{T}(q)\widehat{T}(h)\widehat{T}(q)^{-1} = \Phi_{\widehat{T}(q)}(Th), \qquad (6.7)$$

and thus, replacing h by $T^{-1}h$, $\Phi_{\widehat{T}(q)}(h) = T(\Phi_q(T^{-1}(h)))$, i.e.,

$$\Phi_{\widehat{T}(q)} = T \circ \Phi_q \circ T^{-1}. \qquad (6.8)$$

Now let q be an arbitrary unit quaternion given by (6.5). Let T be a rotation of $\operatorname{Im}\mathbb{H}$ such that $T\mathbf{i} = u$, and let $q_1 := \cos\theta + \mathbf{i}\sin\theta$. Then, $q = \widehat{T}(q_1)$ and thus, by (6.8),

$$\Phi_q = T \circ \Phi_{q_1} \circ T^{-1}, \qquad (6.9)$$

where Φ_{q_1} is a rotation by 2θ around the axis through $\mathbf{i}$. The composition (6.9) is thus a rotation by 2θ around $T\mathbf{i} = u$. □

Remark 6.2. Note that the representation (6.5) of a quaternion is not unique. Of course, θ can be replaced by $\theta + 2n\pi$ for any integer n. Moreover, (θ, u) can be replaced by $(-\theta, -u)$, and thus by $(-\theta + 2n\pi, -u)$. (If $q \neq \pm 1$, these are the only possibilities.) Note that both representations yield the same rotation Φ_q in Lemma 6.1, as they must, since a rotation by 2θ around u is the same as a rotation by -2θ around $-u$.

Theorem 6.3 (Hamilton (1844))**.** *The Lie group homomorphism* $\Phi : Q \to \mathsf{SO}(3)$ *is onto and* 2–1*, with kernel* $\{\pm 1\}$.

Proof. Let $q \in Q$. Then $T_q = I$ if and only if $qpq^{-1} = p$ for every $p \in \operatorname{Im}\mathbb{H}$, i.e., if $qp = pq$ for every $p \in \operatorname{Im}\mathbb{H}$. Using this with $p = \mathbf{i}, \mathbf{j}, \mathbf{k}$ shows by a simple calculation that $q \in \mathbb{R} = \{r \cdot 1\} \subset \mathbb{H}$. Since also $q \in Q$, it follows that $q = \pm 1$. Conversely, if $q = \pm 1$, then $\Phi_q = I$.

Lemma 6.1 shows that Φ is onto $\mathsf{SO}(3)$. □

Alternatively, it follows from Lemma 6.1 that $\Phi_{q_1} = \Phi_{q_2} \iff q_1 = \pm q_2$.

Theorem 6.4 (Cayley (1855))**.** *The Lie group homomorphism* $\Psi : Q \times Q \to \mathsf{SO}(3)$ *is onto and* 2–1*, with kernel* $\{\pm(1,1)\}$.

Proof. First, suppose that $\Psi_{q_1,q_2} = I$. This means, by (6.3), that $q_1 p q_2^{-1} = p$ for every $p \in \mathbb{H}$. In particular, taking $p = 1$ yields $q_1 q_2^{-1} = 1$ and thus

$q_1 = q_2$. Hence $\Phi_{q_1} = \Psi_{q_1,q_1} = I$ and by Theorem 6.3, $q_1 = \pm 1$. Conversely, if $q_1 = q_2 = \pm 1$, then $\Psi_{q_1,q_2} = I$.

To show that Ψ maps onto $\mathsf{SO}(4)$, let $T \in \mathsf{SO}(4)$, as above regarded as a group of rotations of $\mathbb{H}$. Let $q_1 := T1$. Then $\|q_1\| = \|1\| = 1$, and thus $q_1 \in Q$. Moreover, $L_{q_1^{-1}}(T1) = 1$, and thus $T_1 := L_{q_1^{-1}} \circ T \in \mathsf{SO}(4)$ is a rotation that preserves 1, and thus T_1 preserves also the orthogonal complement $\mathbb{R}^\perp = \operatorname{Im}\mathbb{H}$. The restriction of T_1 to $\operatorname{Im}\mathbb{H}$ is thus a rotation of $\operatorname{Im}\mathbb{H}$, i.e., an element of $\mathsf{SO}(3)$, and Theorem 6.3 shows that there exists a quaternion $q \in Q$ such that $T_1 = \Phi_q$ on $\operatorname{Im}\mathbb{H}$, and thus $T = L_{q_1}T_1 = L_{q_1}\Phi_q$ on $\operatorname{Im}\mathbb{H}$. Explicitly, this means that for every $p \in \operatorname{Im}\mathbb{H}$,

$$Tp = q_1\Phi_q(p) = q_1 q p q^{-1}. \tag{6.10}$$

Furthermore, $T1 = q_1 = q_1 q 1 q^{-1}$, and thus (6.10) holds also for $p = 1$, and by linearity thus for all $p \in \mathbb{H}$. This shows that $T = \Psi_{q_1 q, q}$, and thus Ψ is onto $\mathsf{SO}(4)$. □

Using the isomorphism $Q \cong \mathsf{SU}(2)$ in Section 5, the homomorphisms in Theorems 6.3 and 6.4 can also be seen as 2–1 Lie group homomorphisms $\mathsf{SU}(2) \to \mathsf{SO}(3)$ and $\mathsf{SU}(2) \times \mathsf{SU}(2) \to \mathsf{SO}(4)$.

Remark 6.5. These homomorphisms $Q \to \mathsf{SO}(3)$ and $Q \times Q \to \mathsf{SO}(4)$ (or $\mathsf{SU}(2) \to \mathsf{SO}(3)$ and $\mathsf{SU}(2) \times \mathsf{SU}(2) \to \mathsf{SO}(4)$) are covering maps, and since $Q \cong \mathbb{S}^3$ is simply connected, universal covering maps, as is easily seen from the definition [2, Definition 13.1].

7. $\mathsf{Sp}(n)$ and quaternionic vector spaces

The quaternion algebra $\mathbb{H}$ is a non-commutative division ring and not a field. Nevertheless, much of the theory of vector spaces can (with some care) be extended to (left) modules over a division rings, and we can thus talk about *quaternionic vector spaces*, with quaternions as scalars. In particular, the space $\mathbb{H}^n$ is an n-dimensional quaternionic vector space with the multiplication $q(q_1, \dots, q_n) = (qq_1, \dots, qq_n)$.

Note that $\mathbb{H}^n$ is a $4n$-dimensional real vector space, and that an map $T : \mathbb{H}^n \to \mathbb{H}^n$ is $\mathbb{H}$-linear if and only if T is a real linear map such that $T(qv) = qT(v)$ for every $q \in \mathbb{H}$ and $v \in \mathbb{H}^n$. By (real) linearity, this is equivalent to the three conditions

$$T(\mathbf{i}v) = \mathbf{i}T(v), \qquad T(\mathbf{j}v) = \mathbf{j}T(v), \qquad T(\mathbf{k}v) = \mathbf{k}T(v) \tag{7.1}$$

for all $v \in \mathbb{H}^n$. Since $\mathbf{ij} = \mathbf{k}$, we see that the first two conditions (for all v) imply the third, so it suffices to verify the first two.

We can identify $\mathbb{H}^n$ with $\mathbb{R}^{4n}$ and give $\mathbb{H}^n$ the corrresponding Euclidean norm.

Let $\mathsf{Sp}(n)$ be the group of $\mathbb{H}$-linear maps $\mathbb{H}^n \to \mathbb{H}^n$ that preserve the norm. (This is thus the analogue of $\mathsf{O}(n)$ and $\mathsf{U}(n)$ in the quaternionic case.) We will show that this is isomorphic to the compact symplectic group as defined in [2, Section 1.2.4], using a minor variation of the argument in [2, Section 1.2.8].

In order to do this, we use again the realization of the quaternions as the matrix algebra $\widetilde{\mathbb{H}}$ in Section 5. Each matrix in $M_2(\mathbb{C})$ acts on $\mathbb{C}^2$ by the usual multiplication of a matrix and a vector, and it follows that this defines a scalar multiplication $\widetilde{\mathbb{H}} \times \mathbb{C}^2 \to \mathbb{C}^2$ that makes $\mathbb{C}^2$ into a 1-dimensional quaternionic vector space.

Actually, we will use a different construction of $\mathbb{C}^2$ as a quaternionic vector space, which is equivalent to the one just given by the (real linear) involution $\iota : (z_1, z_2) \mapsto (z_1, -\bar{z}_2)$. If $q \in \mathbb{H}$, $\tilde{q} \in \widetilde{\mathbb{H}}$ is the corresponding matrix defined in Section 5 and $v = (z, w) \in \mathbb{C}^2$, we thus thus define (regarding the vectors as column vectors)

$$q \cdot v := \iota(\tilde{q}\iota(v)). \tag{7.2}$$

In particular, by (5.1), this yields

$$\mathbf{i} \cdot \begin{pmatrix} z \\ w \end{pmatrix} = \iota\left(\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}\begin{pmatrix} z \\ -\bar{w} \end{pmatrix}\right) = \iota\begin{pmatrix} iz \\ i\bar{w} \end{pmatrix} = \begin{pmatrix} iz \\ iw \end{pmatrix} = i\begin{pmatrix} z \\ w \end{pmatrix}, \tag{7.3}$$

so $\mathbf{i}$ acts by usual complex multiplication, and

$$\mathbf{j} \cdot \begin{pmatrix} z \\ w \end{pmatrix} = \iota\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} z \\ -\bar{w} \end{pmatrix}\right) = \iota\begin{pmatrix} -\bar{w} \\ -z \end{pmatrix} = \begin{pmatrix} -\bar{w} \\ \bar{z} \end{pmatrix}. \tag{7.4}$$

(We do not have to write explicitly the action of $\mathbf{k}$, since $\mathbf{k} = \mathbf{ij}$.)

Taking the direct sum of n copies of this $\mathbb{H}$-vector space $\mathbb{C}^2$, we obtain an n-dimensional quaternionic vector space $\mathbb{C}^{2n}$, where (if we order the coordinates suitably), $\mathbf{i}$ acts by ordinary complex multiplication by i and $\mathbf{j}$ by

$$\mathbf{j} \cdot (\alpha, \beta) = J(\alpha, \beta) := (-\bar{\beta}, \bar{\alpha}), \qquad \alpha, \beta \in \mathbb{C}^n. \tag{7.5}$$

This is an n-dimensional quaternionic vector space, and therefore it is isomorphic to $\mathbb{H}^n$. Consider now the $\mathbb{H}$-linear maps of this vector space into itself. By (7.1) and the comment after it, and the description above of multiplication by quaternions in $\mathbb{C}^{2n}$, a real linear map $T : \mathbb{C}^{2n} \to \mathbb{C}^{2n}$ is $\mathbb{H}$-linear if and only if $T(iv) = iT(v)$ and $T(J(v)) = J(T(v))$ for all $v \in \mathbb{C}^{2n}$. However, $T(iv) = iT(v)$ if and only if T is complex linear. Thus the $\mathbb{H}$-linear maps $\mathbb{C}^{2n} \to \mathbb{C}^{2n}$ are precisely the complex linear maps that commute with J.

In particular, if we define the group $\mathsf{GL}(n; \mathbb{H})$ to be the set of all $\mathbb{H}$-linear maps of $\mathbb{C}^{2n}$ (or $\mathbb{H}^n$) into itself, this yields

$$\mathsf{GL}(n; \mathbb{H}) = \{T \in \mathsf{GL}(2n; \mathbb{C}) : TJ = JT\}. \tag{7.6}$$

Moreover, by definition, a complex linear map $T : \mathbb{C}^{2n} \to \mathbb{C}^{2n}$ preserves the (Euclidean) norm if and only $T \in \mathsf{U}(2n)$. Consequently, with the definition of $\mathsf{Sp}(n)$ above, we have

$$\mathsf{Sp}(n) = \mathsf{GL}(n; \mathbb{H}) \cap \mathsf{U}(2n) = \{T \in \mathsf{U}(2n) : TJ = JT\}. \tag{7.7}$$

Finally, let $\langle z, w \rangle = \sum_{i=1}^{2n} \bar{z}_i w_i$ denote the usual inner product in $\mathbb{C}^{2n}$ (with the less usual conjugation convention in [2]), and define

$$\omega(z, w) := \langle Jz, w \rangle, \qquad z, w \in \mathbb{C}^n. \tag{7.8}$$

Since J is conjugate-linear; ω is a bilinear form on $\mathbb{C}^{2n}$, and it is easily seen to be the standard skewsymmetric (sympletic) form defined in [2, (1.7)].

Suppose that $T \in \mathsf{U}(2n)$. By (7.8),

$$\omega(Tz, Tw) = \langle JTz, Tw \rangle = \langle T^*JTz, w \rangle = \langle T^{-1}JTz, w \rangle, \tag{7.9}$$

and by comparing with (7.8), we see that $\omega(Tz, Tw) = \omega(z, w)$ for all $z, w \in \mathbb{C}^{2n}$ if and only if $T^{-1}JT = J$, which is equivalent to $JT = TJ$.

Thus, recalling that $\mathsf{Sp}(n;\mathbb{C})$ is the group of all complex-linear maps $T : \mathbb{C}^{2n} \to \mathbb{C}^{2n}$ such that $\omega(Tz, Tw) = \omega(z, w)$ for all $z, w \in \mathbb{C}^{2n}$, we see by (7.7) that

$$\mathsf{Sp}(n) = \{T \in \mathsf{U}(2n) : T \in \mathsf{Sp}(n;\mathbb{C})\} = \mathsf{U}(2n) \cap \mathsf{Sp}(n;\mathbb{C}). \tag{7.10}$$

This is the definition used in [2].

Example 7.1. Consider the simple case $n = 1$. Let $T : \mathbb{H}^1 \to \mathbb{H}^1$ be an $\mathbb{H}$-linear map. Let $q := T1$; then, by $\mathbb{H}$-linearity, for every $p \in \mathbb{H}$,

$$Tp = T(p \cdot 1) = pT(1) = pq, \tag{7.11}$$

and thus $T = R_q$, defined in (6.1). Conversely, every $R_q : \mathbb{H} \to \mathbb{H}$ is $\mathbb{H}$-linear. Consequently, the set of $\mathbb{H}$-linear maps $\mathbb{H} \to \mathbb{H}$ equals $\{R_q : q \in \mathbb{H}\}$. (Note that L_q is in general *not* $\mathbb{H}$-linear.)

Moreover, R_q preserves the norm if and only if $\|q\| = 1$. It follows that $q \mapsto R_{q^{-1}}$ is an isomorphism $Q \to \mathsf{Sp}(1)$. Thus, there are Lie group isomorphisms

$$\mathsf{Sp}(1) \cong Q \cong \mathsf{SU}(2). \tag{7.12}$$

References

[1] H.-D. Ebbinghaus et al. *Numbers*. Springer-Verlag, New York, 1991.

[2] Brian Hall. *Lie Groups, Lie Algebras, and Representations*. 2nd ed. Springer, 2015. (Section and Exercise numbers are different in the 1st ed.)

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden

E-mail address: svante.janson@math.uu.se

URL: http://www.math.uu.se/svante-janson