THE OCTONIONS

Michael Taylor

Contents

1. Introduction

There is a continuation of the construction leading from the real numbers to the complex numbers that proceeds for two more steps, yielding first the quaternions (of dimension 4 over $\mathbb{R}$) and then the octonions (of dimension 8 over $\mathbb{R}$). A quaternion $\xi \in \mathbb{H}$ is given by

$$\xi = a + bi + cj + dk, \quad a, b, c, d \in \mathbb{R}. \tag{1.1}$$

Addition is performed componentwise, and multiplication is an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$ in which 1 is a unit, products of distinct factors i, j, k behave like the cross product on $\mathbb{R}^3$, and $i^2 = j^2 = k^2 = -1$. This product is not commutative, but it is associative. An octonion $x \in \mathbb{O}$ is given by

$$x = (\xi, \eta), \quad \xi, \eta \in \mathbb{H}. \tag{1.2}$$

Addition is again defined componentwise, and multiplication is an $\mathbb{R}$-bilinear map $\mathbb{O} \times \mathbb{O} \to \mathbb{O}$, whose definition (given in §3) is a somewhat subtle modification of the definition of multiplication on $\mathbb{H}$. One major difference is that multiplication on $\mathbb{O}$ is no longer associative. Nevertheless, this non-associative algebra $\mathbb{O}$ has a very beautiful algebraic structure, whose study is the principal object of these notes.

We begin in §2 with a brief treatment of quaternions, giving results of use in our treatment of $\mathbb{O}$, which starts in §3. After defining the product xy of two elements of $\mathbb{O}$, we establish basic properties, introduce a norm and cross product, and study 4-dimensional subalgebras of $\mathbb{O}$ that are isomorphic to $\mathbb{H}$. These include $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_1u_2\}$, when u_1 and u_2 are orthonormal elements of $\mathrm{Im}(\mathbb{O})$. Analysis of the relationship of $\mathcal{A}$ and $\mathcal{A}^\perp$ gives rise to orthogonal linear maps on $\mathbb{O}$ that are shown to preserve the product, i.e., to automorphisms of $\mathbb{O}$, which form a group, denoted $\mathrm{Aut}(\mathbb{O})$.

Section 4 goes further into $\mathrm{Aut}(\mathbb{O})$, noting that it is a compact, connected Lie group of dimension 14, and analyzing some of its subgroups, including groups isomorphic to $SO(4)$ and groups isomorphic to $SU(3)$. Both of these types of subgroups contain two-dimensional tori. It is shown that $\mathrm{Aut}(\mathbb{O})$ contains no tori of larger dimension. These facts are used in §5 to show that $\mathrm{Aut}(\mathbb{O})$ is simple and to analyze its root system. The analysis reveals that $\mathrm{Aut}(\mathbb{O})$ falls into the classification of compact simple Lie groups as the group denoted G_2. In §6 we make further comments about the Lie algebra of $\mathrm{Aut}(\mathbb{O})$, including the fact that it has a $\mathbb{Z}/(3)$ grading.

The Lie group G_2 is the first in a list of 5 exceptional compact Lie groups, denoted G_2, F_4, E_6, E_7, and E_8, with complexified Lie algebras denoted $\mathfrak{G}_2, \mathfrak{F}_4, \mathfrak{E}_6, \mathfrak{E}_7$, and $\mathfrak{E}_8$. In Appendix A we briefly describe a unified construction of $\mathfrak{G}_2$ and $\mathfrak{E}_8$, due to Freudenthal, which for $\mathfrak{G}_2$ meshes nicely with the analysis in §6.

Further material on the algebra of octonions, its automorphism group, and other concepts arising here can be found in a number of sources, including the survey article [B] and the books [SV], [Por], and [H].

Acknowledgment. Thanks to Robert Bryant for valuable conversations on this topic, particularly for sharing his insights on the Moufang identities.

2. Quaternions

The space $\mathbb{H}$ of quaternions is a four-dimensional real vector space, identified with $\mathbb{R}^4$, with basis elements $1, i, j, k$, the element 1 identified with the real number 1. Elements of $\mathbb{H}$ are represented as follows:

$$\xi = a + bi + cj + dk, \tag{2.1}$$

with $a, b, c, d \in \mathbb{R}$. We call a the real part of ξ ($a = \operatorname{Re}\xi$) and $bi + cj + dk$ the vector part. We also have a multiplication on $\mathbb{H}$, an $\mathbb{R}$-bilinear map $\mathbb{H} \times \mathbb{H} \to \mathbb{H}$, such that $1 \cdot \xi = \xi \cdot 1 = \xi$, and otherwise governed by the rules

$$ij = k = -ji, \quad jk = i = -kj, \quad ki = j = -ik, \tag{2.2}$$

and

$$i^2 = j^2 = k^2 = -1. \tag{2.3}$$

Otherwise stated, if we write

$$\xi = a + u, \quad a \in \mathbb{R}, \quad u \in \mathbb{R}^3, \tag{2.4}$$

and similarly write $\eta = b + v,\ b \in \mathbb{R},\ v \in \mathbb{R}^3$, the product is given by

$$\xi\eta = (a + u)(b + v) = (ab - u \cdot v) + av + bu + u \times v. \tag{2.5}$$

Here $u \cdot v$ is the dot product in $\mathbb{R}^3$, and $u \times v$ is the cross product. The quantity $ab - u \cdot v$ is the real part of $\xi\eta$ and $av + bu + u \times v$ is the vector part. Note that

$$\xi\eta - \eta\xi = 2u \times v. \tag{2.6}$$

It is useful to take note of the following symmetries of $\mathbb{H}$.

Proposition 2.1. *Let $K : \mathbb{H} \to \mathbb{H}$ be an $\mathbb{R}$-linear transformation such that $K1 = 1$ and K cyclically permutes (i, j, k) (e.g., $Ki = j,\ Kj = k,\ Kk = i$). Then K preserves the product in $\mathbb{H}$, i.e.,*

$$K(\xi\eta) = K(\xi)K(\eta), \quad \forall\, \xi, \eta \in \mathbb{H}.$$

We say K is an automorphism of $\mathbb{H}$.

Proof. This is straightforward from the multiplication rules (2.2)–(2.3).

We move on to the following basic result.

Proposition 2.2. *Multiplication in $\mathbb{H}$ is associative, i.e.,*

$$\zeta(\xi\eta) = (\zeta\xi)\eta, \quad \forall\, \zeta, \xi, \eta \in \mathbb{H}. \tag{2.7}$$

Proof. Given the $\mathbb{R}$-bilinearity of the product, it suffices to check (2.7) when each ζ, ξ, and η is either $1, i, j$, or k. Since 1 is the multiplicative unit, the result (2.7) is easy when any factor is 1. Furthermore, one can use Proposition 2.1 to reduce the possibilities further; for example, one can take $\zeta = i$. We leave the final details to the reader.

Remark. In the case that $\xi = u, \eta = v$, and $\zeta = w$ are purely vectorial, we have

$$\begin{aligned} w(uv) &= w(-u \cdot v + u \times v) = -(u \cdot v)w - w \cdot (u \times v) + w \times (u \times v), \\ (wu)v &= (-w \cdot u + w \times u)v = -(w \cdot u)v - (w \times u) \cdot v + (w \times u) \times v. \end{aligned} \tag{2.8}$$

Then the identity of the two left sides is equivalent to the pair of identities

$$w \cdot (u \times v) = (w \times u) \cdot v, \tag{2.9}$$

$$w \times (u \times v) - (w \times u) \times v = (w \cdot u)v - (u \cdot v)w. \tag{2.10}$$

As for (2.10), it also follows from the pair of identities

$$w \times (u \times v) - (w \times u) \times v = (v \times w) \times u, \tag{2.11}$$

and

$$(v \times w) \times u = (w \cdot u)v - (u \cdot v)w, \tag{2.12}$$

which we leave as an exercise for the reader.

In addition to the product, we also have a conjugation operation on $\mathbb{H}$:

$$\overline{\xi} = a - bi - cj - dk = a - u. \tag{2.13}$$

A calculation gives

$$\xi\overline{\eta} = (ab + u \cdot v) - av + bu - u \times v. \tag{2.14}$$

In particular,

$$\operatorname{Re}(\xi\overline{\eta}) = \operatorname{Re}(\overline{\eta}\xi) = (\xi, \eta), \tag{2.15}$$

the right side denoting the Euclidean inner product on $\mathbb{R}^4$. Setting $\eta = \xi$ in (2.14) gives

$$\xi\overline{\xi} = |\xi|^2, \tag{2.16}$$

the Euclidean square-norm of ξ. In particular, whenever $\xi \in \mathbb{H}$ is nonzero, it has a multiplicative inverse,

$$\xi^{-1} = |\xi|^{-2}\overline{\xi}. \tag{2.17}$$

We say a ring $\mathcal{R}$ with unit 1 is a division ring if each nonzero $\xi \in \mathcal{R}$ has a multiplicative inverse. It follows from (F.17) that $\mathbb{H}$ is a division ring. It is not a field, since multiplication in $\mathbb{H}$ is not commutative. Sometimes $\mathbb{H}$ is called a "noncommutative field."

To continue with products and conjugation, a routine calculation gives

$$\overline{\xi\eta} = \overline{\eta}\,\overline{\xi}. \tag{2.18}$$

Hence, via the associative law,

$$|\xi\eta|^2 = (\xi\eta)(\overline{\xi\eta}) = \xi\eta\overline{\eta}\overline{\xi} = |\eta|^2\xi\overline{\xi} = |\xi|^2|\eta|^2, \tag{2.19}$$

or

$$|\xi\eta| = |\xi|\,|\eta|. \tag{2.20}$$

Note that $\mathbb{C} = \{a + bi : a, b \in \mathbb{R}\}$ sits in $\mathbb{H}$ as a commutative subring, for which the properties (2.16) and (2.20) are familiar.

Let us examine (2.20) when $\xi = u$ and $\eta = v$ are purely vectorial. We have

$$uv = -u \cdot v + u \times v. \tag{2.21}$$

Hence, directly,

$$|uv|^2 = (u \cdot v)^2 + |u \times v|^2, \tag{2.22}$$

while (2.20) implies

$$|uv|^2 = |u|^2|v|^2. \tag{2.23}$$

On the other hand, if θ is the angle between u and v in $\mathbb{R}^3$,

$$u \cdot v = |u|\,|v|\,\cos\theta.$$

Hence (2.22) implies

$$|u \times v|^2 = |u|^2|v|^2 \sin^2\theta. \tag{2.24}$$

We next consider the set of unit quaternions:

$$Sp(1) = \{\xi \in \mathbb{H} : |\xi| = 1\}. \tag{2.25}$$

Using (2.17) and (2.20), we see that $Sp(1)$ is a group under multiplication. It sits in $\mathbb{R}^4$ as the unit sphere S^3. We compare $Sp(1)$ with the group $SU(2)$, consisting of 2×2 complex matrices of the form

$$U = \begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix}, \quad \xi, \eta \in \mathbb{C}, \quad |\xi|^2 + |\eta|^2 = 1. \tag{2.26}$$

The group $SU(2)$ is also in natural one-to-one correspondence with S^3. Furthermore, we have:

Proposition 2.3. *The groups $SU(2)$ and $Sp(1)$ are isomorphic under the correspondence*

$$U \mapsto \xi + j\eta, \tag{2.27}$$

for U as in (2.26).

Proof. The correspondence (2.27) is clearly bijective. To see it is a homomorphism of groups, we calculate:

$$\begin{pmatrix} \xi & -\overline{\eta} \\ \eta & \overline{\xi} \end{pmatrix} \begin{pmatrix} \xi' & -\overline{\eta}' \\ \eta' & \overline{\xi}' \end{pmatrix} = \begin{pmatrix} \xi\xi' - \overline{\eta}\eta' & -\xi\overline{\eta}' - \overline{\eta}\overline{\xi}' \\ \eta\xi' + \overline{\xi}\eta' & -\eta\overline{\eta}' + \xi\overline{\xi}' \end{pmatrix}, \tag{2.28}$$

given $\xi, \eta \in \mathbb{C}$. Noting that, for $a, b \in \mathbb{R}$, $j(a + bi) = (a - bi)j$, we have

$$\begin{aligned} (\xi + j\eta)(\xi' + j\eta') &= \xi\xi' + \xi j\eta' + j\eta\xi' + j\eta j\eta' \\ &= \xi\xi' - \overline{\eta}\eta' + j(\eta\xi' + \overline{\xi}\eta'). \end{aligned} \tag{2.29}$$

Comparison of (2.28) and (2.29) verifies that (2.27) yields a homomorphism of groups.

We next define the map

$$\pi : Sp(1) \longrightarrow \mathcal{L}(\mathbb{R}^3) \tag{2.30}$$

by

$$\pi(\xi)u = \xi u \xi^{-1} = \xi u \overline{\xi}, \quad \xi \in Sp(1),\ u \in \mathbb{R}^3 \subset \mathbb{H}. \tag{2.31}$$

To justify (2.30), we need to show that if u is purely vectorial, so is $\xi u \overline{\xi}$. In fact, by (2.18),

$$\zeta = \xi u \overline{\xi} \Longrightarrow \overline{\zeta} = \overline{\overline{\xi}}\,\overline{u}\,\overline{\xi} = -\xi u \overline{\xi} = -\zeta, \tag{2.32}$$

so that is indeed the case. By (2.20),

$$|\pi(\xi)u| = |\xi|\,|u|\,|\overline{\xi}| = |u|, \quad \forall\, u \in \mathbb{R}^3,\ \xi \in Sp(1),$$

so in fact

$$\pi : Sp(1) \longrightarrow SO(3), \tag{2.33}$$

and it follows easily from the definition (2.31) that if also $\zeta \in Sp(1)$, then $\pi(\xi\zeta) = \pi(\xi)\pi(\zeta)$, so (2.33) is a group homomorphism. It is readily verified that

$$\operatorname{Ker} \pi = \{\pm 1\}. \tag{2.34}$$

Note that we can extend (2.30) to

$$\pi : Sp(1) \longrightarrow \mathcal{L}(\mathbb{H}), \quad \pi(\xi)\eta = \xi\eta\overline{\xi}, \quad \xi \in Sp(1),\ \eta \in \mathbb{H}, \tag{2.35}$$

and again $\pi(\xi\zeta) = \pi(\xi)\pi(\zeta)$ for $\xi, \zeta \in Sp(1)$. Furthermore, each map $\pi(\xi)$ is a ring homomorphism, i.e.,

$$\pi(\xi)(\alpha\beta) = (\pi(\xi)\alpha)(\pi(\xi)\beta), \quad \alpha, \beta \in \mathbb{H},\ \xi \in Sp(1). \tag{2.36}$$

Since $\pi(\xi)$ is invertible, this is a group of ring automorphisms of $\mathbb{H}$. The reader is invited to draw a parallel to the following situation. Define

$$\tilde{\pi} : SO(3) \longrightarrow \mathcal{L}(\mathbb{H}), \quad \tilde{\pi}(T)(a+u) = a + Tu, \tag{2.37}$$

given $a + u \in \mathbb{H},\ a \in \mathbb{R},\ u \in \mathbb{R}^3$. It is a consequence of the identity

$$T(u \times v) = Tu \times Tv, \text{ for } u, v \in \mathbb{R}^3,\ T \in SO(3),$$

that

$$\tilde{\pi}(T)(\alpha\beta) = (\tilde{\pi}(T)\alpha)(\tilde{\pi}(T)\beta), \quad \alpha, \beta \in \mathbb{H},\ T \in SO(3). \tag{2.38}$$

Thus $SO(3)$ acts as a group of automorphisms of $\mathbb{H}$. (Note that Proposition 2.1 is a special case of this.) We claim this is the same group of automorphisms as described in (2.35)–(2.36), via (2.33). This is a consequence of the fact that π in (2.33) is surjective. We mention that the automorphism K in Proposition 2.1 has the form (2.35) with

$$\xi = \frac{1}{2}(1 + i + j + k).$$

3. Octonions

The set of octonions (also known as Cayley numbers) is a special but intriguing example of a nonassociative algebra. This space is

$$\mathbb{O} = \mathbb{H} \oplus \mathbb{H}, \tag{3.1}$$

with product given by

$$(\alpha, \beta) \cdot (\gamma, \delta) = (\alpha\gamma - \overline{\delta}\beta, \delta\alpha + \beta\overline{\gamma}), \quad \alpha, \beta, \gamma, \delta \in \mathbb{H}, \tag{3.2}$$

with conjugation $\delta \mapsto \overline{\delta}$ on $\mathbb{H}$ defined as in §2. We mention that, with $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}$, the product in $\mathbb{H}$ is also given by (3.2), with $\alpha, \beta, \gamma, \delta \in \mathbb{C}$. Furthermore, with $\mathbb{C} = \mathbb{R} \oplus \mathbb{R}$, the product in $\mathbb{C}$ is given by (3.2), with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$. In the setting of $\mathbb{O} = \mathbb{H} \oplus \mathbb{H}$, the product in (3.2) is clearly $\mathbb{R}$-bilinear, but it is neither commutative nor associative. However, it does retain a vestige of associativity, namely

$$x(yz) = (xy)z \quad \text{whenever any two of } x, y, z \in \mathbb{O} \text{ coincide.} \tag{3.3}$$

We define a conjugation on $\mathbb{O}$:

$$x = (\alpha, \beta) \Longrightarrow \overline{x} = (\overline{\alpha}, -\beta). \tag{3.4}$$

We set $\operatorname{Re} x = (x + \overline{x})/2 = (\operatorname{Re} \alpha, 0)$. Note that $a = \operatorname{Re} x$ lies in the center of $\mathbb{O}$ (i.e., commutes with each element of $\mathbb{O}$), and $\overline{x} = 2a - x$. It is straightforward to check that

$$x, y \in \mathbb{O} \Longrightarrow \operatorname{Re} xy = \operatorname{Re} yx. \tag{3.5}$$

We have a decomposition

$$x = a + u, \quad a = \operatorname{Re} x, \ u = x - \operatorname{Re} x = \operatorname{Im} x, \tag{3.6}$$

parallel to (2.4). Again we call u the vector part of x, and we say that $u \in \operatorname{Im}(\mathbb{O})$. If also $y = b + v$, then

$$xy = ab + av + bu + uv, \tag{3.7}$$

with a similar formula for yx, yielding

$$xy - yx = uv - vu. \tag{3.8}$$

We now define the inner product

$$\langle x, y\rangle = \operatorname{Re}(x\overline{y}), \quad x, y \in \mathbb{O}. \tag{3.9}$$

To check symmetry, note that if $x = a + u,\ y = b + v$,

$$\langle x, y\rangle = ab - \operatorname{Re}(uv), \tag{3.10}$$

and (3.5) then implies

$$\langle x, y\rangle = \langle y, x\rangle. \tag{3.11}$$

In fact, (3.9) yields the standard Euclidean inner product on $\mathbb{O} \approx \mathbb{R}^8$, with square norm $|x|^2 = \sqrt{\langle x, x\rangle}$. We have

$$x = (\alpha, \beta) \Longrightarrow x\overline{x} = (\alpha\overline{\alpha} + \overline{\beta}\beta, 0) = (|x|^2, 0). \tag{3.12}$$

As a consequence, we see that

$$x \in \mathbb{O},\ x \neq 0,\ y = |x|^{-2}\overline{x} \Longrightarrow xy = yx = 1, \tag{3.13}$$

where $1 = (1, 0)$ is the multiplicative unit in $\mathbb{O}$.

Returning to conjugation on $\mathbb{O}$, we have, parallel to (2.18),

$$x, y \in \mathbb{O} \Longrightarrow \overline{xy} = \overline{y}\,\overline{x}, \tag{3.14}$$

via a calculation using the definition (3.2) of the product. Using the decomposition $x = a + u,\ y = b + v$, this is equivalent to $\overline{uv} = vu$, and since $\overline{uv} = 2\operatorname{Re}(uv) - uv = -2\langle u, v\rangle - uv$, this is equivalent to

$$u, v \in \operatorname{Im}(\mathbb{O}) \Longrightarrow uv + vu = -2\langle u, v\rangle. \tag{3.15}$$

In turn, (3.15) follows by expanding $(u+v)^2$ and using $w^2 = -|w|^2$ for $w \in \operatorname{Im}(\mathbb{O})$, with $w = u, v$, and $u + v$. We next establish the following parallel to (2.20).

Proposition 3.1. *Given $x, y \in \mathbb{O}$,*

$$|xy| = |x|\,|y|. \tag{3.16}$$

Proof. To begin, we bring in the following variant of (3.3),

$$x, y \in \mathbb{O} \Longrightarrow (xy)(yx) = ((xy)y)x, \tag{3.17}$$

which can be verified from the definition (3.2) of the product. Taking into account $\overline{x} = 2a - x,\ \overline{y} = 2b - y$, and (3.14), we have

$$\begin{aligned}(xy)(\overline{xy}) &= (xy)(\overline{y}\,\overline{x}) = ((xy)\overline{y})\overline{x} \\ &= (x|y|^2)\overline{x} = |x|^2|y|^2,\end{aligned} \tag{3.18}$$

which gives (3.16), since $|xy|^2 = (xy)(\overline{xy})$.

Continuing to pursue parallels with §2, we define a cross product on $\mathrm{Im}(\mathbb{O})$ as follows. Given $u, v \in \mathrm{Im}(\mathbb{O})$, set

$$u \times v = \frac{1}{2}(uv - vu). \tag{3.19}$$

By (3.5), this is an element of $\mathrm{Im}(\mathbb{O})$. Also, if $x = a + u, y = b + v$,

$$xy - yx = 2u \times v. \tag{3.20}$$

Compare (2.6). Putting together (3.15) and (3.19), we have

$$uv = -\langle u, v\rangle + u \times v, \quad u, v \in \mathrm{Im}(\mathbb{O}). \tag{3.21}$$

Hence

$$|uv|^2 = |\langle u, v\rangle|^2 + |u \times v|^2. \tag{3.22}$$

Now (3.16) implies $|uv|^2 = |u|^2|v|^2$, and of course $\langle u, v\rangle = |u|\,|v| \cos\theta$, where θ is the angle between u and v. Hence, parallel to (2.24),

$$|u \times v|^2 = |u|^2|v|^2|\sin\theta|^2, \quad \forall\, u, v \in \mathrm{Im}(\mathbb{O}). \tag{3.23}$$

We have the following complement.

Proposition 3.2. *If $u, v \in \mathrm{Im}(\mathbb{O})$, then*

$$w = u \times v \Longrightarrow \langle w, u\rangle = \langle w, v\rangle = 0. \tag{3.24}$$

Proof. We know that $w \in \mathrm{Im}(\mathbb{O})$. Hence, by (3.21),

$$\begin{aligned} \langle w, v\rangle &= \langle uv, v\rangle = \mathrm{Re}((uv)\overline{v}) \\ &= \mathrm{Re}(u(v\overline{v})) = |v|^2 \,\mathrm{Re}\, u = 0, \end{aligned} \tag{3.25}$$

the third identity by (3.3) (applicable since $\overline{v} = -v$). The proof that $\langle w, u\rangle = 0$ is similar.

Returning to basic observations about the product (3.2), we note that it is uniquely determined as the $\mathbb{R}$-bilinear map $\mathbb{O} \times \mathbb{O} \to \mathbb{O}$ satisfying

$$\begin{aligned} (\alpha, 0) \cdot (\gamma, 0) &= (\alpha\gamma, 0), \quad & (0, \beta) \cdot (\gamma, 0) &= (0, \beta\overline{\gamma}), \\ (\alpha, 0) \cdot (0, \delta) &= (0, \delta\alpha), \quad & (0, \beta) \cdot (0, \delta) &= (-\overline{\delta}\beta, 0), \end{aligned} \tag{3.26}$$

for $\alpha, \beta, \gamma, \delta \in \mathbb{H}$. In particular, $\mathbb{H} \oplus 0$ is a subalgebra of $\mathbb{O}$, isomorphic to $\mathbb{H}$. As we will see, $\mathbb{O}$ has lots of subalgebras isomorphic to $\mathbb{H}$. First, let us label the "standard" basis of $\mathbb{O}$ as

$$\begin{aligned} 1 &= (1,0), \quad e_1 = (i,0), \quad e_2 = (j,0), \quad e_3 = (k,0), \\ f_0 &= (0,1), \quad f_1 = (0,i), \quad f_2 = (0,j), \quad f_3 = (0,k), \end{aligned} \tag{3.27}$$

and describe the associated multiplication table. The mutiplication table for $1, e_1, e_2, e_3$ is the same as (2.2)–(2.3), of course. We also have $f_\ell^2 = -1$ and all the distinct e_ℓ and f_m anticommute. These results are special cases of the fact that

$$u, v \in \mathrm{Im}(\mathbb{O}),\ |u| = 1,\ \langle u, v\rangle = 0 \Longrightarrow u^2 = -1 \ \text{ and } \ uv = -vu, \tag{3.28}$$

which is a consequence of (3.15).

To proceed with the multiplication table for $\mathbb{O}$, note that (3.26) gives

$$(\alpha, 0) f_0 = (0, \alpha), \tag{3.29}$$

so

$$e_\ell f_0 = f_\ell, \quad 1 \le \ell \le 3. \tag{3.30}$$

By (3.28), $f_0 e_\ell = -f_\ell$. Using the notation $\varepsilon_1 = i, \varepsilon_2 = j, \varepsilon_3 = k \in \mathbb{H}$, we have

$$e_\ell f_m = (\varepsilon_\ell, 0) \cdot (0, \varepsilon_m) = (0, \varepsilon_m \varepsilon_\ell), \quad 1 \le \ell, m \le 3, \tag{3.31}$$

and the multiplication table (2.2)–(2.3) gives the result as $-f_0$ if $\ell = m$, and $\pm f_\mu$ if $\ell \neq m$, where $\{\ell, m, \mu\} = \{1, 2, 3\}$. Again by (3.28), $f_m e_\ell = -e_\ell f_m$. To complete the multiplication table, we have

$$f_0 f_m = (0,1) \cdot (0, \varepsilon_m) = (\varepsilon_m, 0) = e_m, \quad 1 \le m \le 3, \tag{3.32}$$

and

$$f_\ell f_m = (0, \varepsilon_\ell) \cdot (0, \varepsilon_m) = (\varepsilon_m \varepsilon_\ell, 0) = e_m e_\ell, \quad 1 \le \ell, m \le 3. \tag{3.33}$$

We turn to the task of constructing subalgebras of $\mathbb{O}$. To start, pick

$$u_1 \in \mathrm{Im}(\mathbb{O}), \quad \text{such that} \quad |u_1| = 1. \tag{3.34}$$

By (3.28), $u_1^2 = -1$, and we have the subalgebra of $\mathbb{O}$,

$$\mathrm{Span}\{1, u_1\} \approx \mathbb{C}. \tag{3.35}$$

To proceed, pick

$$u_2 \in \mathrm{Im}(\mathbb{O}), \quad \text{such that} \quad |u_2| = 1 \text{ and } \langle u_1, u_2\rangle = 0, \tag{3.36}$$

and set

$$u_3 = u_1 u_2. \tag{3.37}$$

By (3.28),

$$u_2^2 = -1, \text{ and } u_2 u_1 = -u_1 u_2 = -u_3. \tag{3.38}$$

Note that

$$\operatorname{Re} u_3 = \operatorname{Re}(u_1 u_2) = -\langle u_1, u_2 \rangle = 0. \tag{3.39}$$

Also, by (3.16), $|u_3| = 1$, so

$$1 = u_3 \overline{u}_3 = -u_3^2. \tag{3.40}$$

Furthermore, by (3.3),

$$\begin{aligned} u_1 u_3 &= u_1(u_1 u_2) = (u_1 u_1) u_2 = -u_2, \quad \text{and} \\ u_3 u_2 &= (u_1 u_2) u_2 = u_1(u_2 u_2) = -u_1. \end{aligned} \tag{3.41}$$

Let us also note that

$$u_3 = u_1 \times u_2. \tag{3.42}$$

Hence, by Proposition 3.2,

$$\langle u_3, u_1 \rangle = \langle u_3, u_2 \rangle = 0, \tag{3.43}$$

and, again by (3.28), $u_3 u_1 = -u_1 u_3$ and $u_2 u_3 = -u_3 u_2$. Thus we have for each such choice of u_1 and u_2 a subalgebra of $\mathbb{O}$,

$$\operatorname{Span}\{1, u_1, u_2, u_3\} \approx \mathbb{H}. \tag{3.44}$$

At this point we can make the following observation.

Proposition 3.3. *Given any two elements $x_1, x_2 \in \mathbb{O}$, the algebra $\mathcal{A}$ generated by $1, x_1$, and x_2 is isomorphic to either $\mathbb{R}, \mathbb{C}$, or $\mathbb{H}$. In particular, it is associative.*

Proof. Consider $V = \operatorname{Span}\{1, x_1, x_2\}$. If $\dim V = 1$, then $\mathcal{A} \approx \mathbb{R}$. If $\dim V = 2$, the argument yielding (3.35) gives $\mathcal{A} \approx \mathbb{C}$. If $\dim V = 3$, then $\operatorname{Im} x_1$ and $\operatorname{Im} x_2$ are linearly independent. We can pick orthonormal elements u_1 and u_2 in their span. Then $\mathcal{A}$ is the algebra generated by $1, u_1$, and u_2, and the analysis (3.34)–(3.44) gives $\mathcal{A} \approx \mathbb{H}$.

The last assertion of Proposition 3.3 contains (3.3) and (3.17) as special cases. The failure of $\mathbb{O}$ to be associative is clearly illustrated by (3.31), which implies

$$e_\ell(e_m f_0) = (e_m e_\ell) f_0, \quad \text{for } \ 1 \le \ell, m \le 3, \tag{3.45}$$

so

$$e_\ell(e_m f_0) = -(e_\ell e_m) f_0, \quad \text{if } \ \ell \ne m.$$

Bringing in also (3.33) yields

$$f_\ell(e_m f_0) = e_m e_\ell, \quad \text{while } \ (f_\ell e_m) f_0 = e_\ell e_m. \tag{3.46}$$

We next explore how the subalgebra

$$\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_3\}, \tag{3.47}$$

from (3.44), interacts with its orthogonal complement $\mathcal{A}^\perp$. Pick

$$v_0 \in \mathcal{A}^\perp, \quad |v_0| = 1. \tag{3.48}$$

Note that $v_0 \in \mathrm{Im}(\mathbb{O})$. Taking a cue from (3.30), we set

$$v_\ell = u_\ell v_0, \quad 1 \le \ell \le 3. \tag{3.49}$$

Note that $\mathrm{Re}\, v_\ell = -\langle u_\ell, v_0 \rangle = 0$, so $v_\ell \in \mathrm{Im}(\mathbb{O})$. We claim that

$$\{v_0, v_1, v_2, v_3\} \ \text{ is an orthonormal set in } \ \mathbb{O}. \tag{3.50}$$

To show this, we bring in the following operators. Given $x \in \mathbb{O}$, define the $\mathbb{R}$-linear maps

$$L_x, R_x : \mathbb{O} \longrightarrow \mathbb{O}, \quad L_x y = xy, \ R_x y = yx. \tag{3.51}$$

By (3.16), for $y \in \mathbb{O}$,

$$|x| = 1 \Longrightarrow |L_x y| = |R_x y| = |y|. \tag{3.52}$$

Hence L_x and R_x are orthogonal transformations. Since the unit sphere in $\mathbb{O}$ is connected, $\det L_x$ and $\det R_x$ are $\equiv 1$ for such x, so

$$|x| = 1 \Longrightarrow L_x, R_x \in SO(\mathbb{O}). \tag{3.53}$$

Hence $R_{v_0} \in SO(\mathbb{O})$. Since

$$v_0 = R_{v_0} 1, \quad v_\ell = R_{v_0} u_\ell \ \text{ for } \ 1 \le \ell \le 3, \tag{3.54}$$

we have (3.50). We next claim that

$$v_\ell \perp u_m, \quad \forall\, \ell, m \in \{1,2,3\}. \tag{3.55}$$

In fact, since $L_{u_\ell} \in SO(\mathbb{O})$,

$$\begin{aligned} \langle v_\ell, u_m\rangle &= \langle u_\ell v_0, u_m\rangle = \langle u_\ell(u_\ell v_0), u_\ell u_m\rangle \\ &= \langle (u_\ell u_\ell) v_0, u_\ell u_m\rangle = -\langle v_0, u_\ell u_m\rangle = 0, \end{aligned} \tag{3.56}$$

the third identity by (3.3).

It follows that

$$\mathcal{A}^\perp = \operatorname{Span}\{v_0, v_1, v_2, v_3\}. \tag{3.57}$$

Consequently

$$\{1, u_1, u_2, u_3, v_0, v_1, v_2, v_3\} \ \text{ is an orthonormal basis of } \ \mathbb{O}. \tag{3.58}$$

Results above imply that

$$R_{v_0} : \mathcal{A} \xrightarrow{\approx} \mathcal{A}^\perp. \tag{3.59}$$

Such an argument applies to any unit length $v \perp \mathcal{A}$. Consequently

$$x \in \mathcal{A},\ y \in \mathcal{A}^\perp \Longrightarrow xy \in \mathcal{A}^\perp. \tag{3.60}$$

Noting that if also $x \in \operatorname{Im}(\mathbb{O})$ then $xy = -yx$, we readily deduce that

$$x \in \mathcal{A},\ y \in \mathcal{A}^\perp \Longrightarrow yx \in \mathcal{A}^\perp. \tag{3.61}$$

Furthermore, since $|x| = 1 \Rightarrow L_x, R_x \in SO(\mathbb{O})$, we have

$$x \in \mathcal{A}^\perp \Longrightarrow L_x, R_x : \mathcal{A}^\perp \longrightarrow \mathcal{A}, \tag{3.62}$$

hence

$$x, y \in \mathcal{A}^\perp \Longrightarrow xy \in \mathcal{A}. \tag{3.63}$$

Note that for the special case

$$\mathcal{H} = \mathbb{H} \oplus 0, \quad \mathcal{H}^\perp = 0 \oplus \mathbb{H}, \tag{3.64}$$

the results (3.60)–(3.63) follow immediately from (3.26).

We have the following important result about the correspondence between the bases (3.27) and (3.58) of $\mathbb{O}$.

Proposition 3.4. *Let $u_\ell, v_\ell \in \mathrm{Im}(\mathbb{O})$ be given as in (3.47)–(3.49). Then the orthogonal transformation $K : \mathbb{O} \to \mathbb{O}$, defined by*

$$K1 = 1, \quad Ke_\ell = u_\ell, \quad Kf_\ell = v_\ell, \tag{3.65}$$

preserves the product on $\mathbb{O}$:

$$K(xy) = K(x)K(y), \quad \forall\, x, y \in \mathbb{O}. \tag{3.66}$$

That is to say, K is an automorphism of $\mathbb{O}$.

Proof. What we need to show is that $\{u_1, u_2, u_3, v_0, v_1, v_2, v_3\}$ has the same multiplication table as $\{e_1, e_2, e_3, f_0, f_1, f_2, f_3\}$. That products involving only $\{u_\ell\}$ have such behavior follows from the arguments leading to (3.44). That $e_\ell f_0 = f_\ell$ is paralleled by $u_\ell v_0 = v_\ell$, for $1 \le \ell \le 3$, is the definition (3.49). It remains to show that the products $u_\ell v_m$ and $v_\ell v_m$ mirror the products $e_\ell f_m$ and $f_\ell f_m$, as given in (3.31)–(3.33).

First, we have, for $1 \le m \le 3$,

$$v_0 v_m = -v_m v_0 = -(u_m v_0) v_0 = -u_m (v_0 v_0) = u_m, \tag{3.67}$$

mirroring (3.32). Mirroring the case $\ell = m$ of (3.31), we have

$$u_\ell v_\ell = u_\ell (u_\ell v_0) = (u_\ell u_\ell) v_0 = -v_0. \tag{3.68}$$

The analogue of (3.31) for $\ell = m$ is simple, thanks to (3.15):

$$v_\ell v_\ell = -1. \tag{3.69}$$

It remains to establish the following:

$$u_\ell v_m = (u_m u_\ell) v_0, \quad v_\ell v_m = u_m u_\ell, \quad \text{for } \ 1 \le \ell, m \le 3, \ \ell \ne m. \tag{3.70}$$

Expanded out, the required identities are

$$u_\ell (u_m v_0) = (u_m u_\ell) v_0, \quad 1 \le \ell, m \le 3, \ \ell \ne m, \tag{3.71}$$

and

$$(u_\ell v_0)(u_m v_0) = u_m u_\ell, \quad 1 \le \ell, m \le 3, \ \ell \ne m. \tag{3.72}$$

Such identities as (3.71)–(3.72) are closely related to an important class of identities known as "Moufang identities," which we now introduce.

Proposition 3.5. *Given* $x, y, z \in \mathbb{O}$,

$$(xyx)z = x(y(xz)), \quad z(xyx) = ((zx)y)x, \tag{3.73}$$

and

$$(xy)(zx) = x(yz)x. \tag{3.74}$$

Regarding the paucity of parentheses here, we use the notation xwx to mean

$$xwx = (xw)x = x(wx), \tag{3.75}$$

the last identity by (3.3). Note also that the two identities in (3.73) are equivalent, respectively, to

$$L_{xyx} = L_x L_y L_x, \quad \text{and} \quad R_{xyx} = R_x R_y R_x. \tag{3.76}$$

A proof of Proposition 3.5 will be given later in this appendix. We now show how (3.73)–(3.74) can be used to establish (3.71)–(3.72).

We start with (3.72), which is equivalent to

$$(v_0 u_\ell)(u_m v_0) = u_\ell u_m. \tag{3.77}$$

In this case, (3.74) yields

$$\begin{aligned}(v_0 u_\ell)(u_m v_0) &= v_0(u_\ell u_m) v_0 \\ &= -(u_\ell u_m) v_0 v_0 \quad (\text{if } \ell \neq m) \\ &= u_\ell u_m,\end{aligned} \tag{3.78}$$

via a couple of applications of (3.15). This gives (3.72).
Moving on, applying L_{v_0}, we see that (3.71) is equivalent to

$$v_0(u_\ell(u_m v_0)) = v_0(u_m u_\ell) v_0, \tag{3.79}$$

hence to

$$v_0(u_\ell(v_0 u_m)) = v_0(u_\ell u_m) v_0. \tag{3.80}$$

Now the first identity in (3.73) implies that the left side of (3.80) is equal to

$$(v_0 u_\ell v_0) u_m = u_\ell u_m, \tag{3.81}$$

the latter identity because $v_0 u_\ell v_0 = -u_\ell v_0 v_0 = u_\ell$. On the other hand, if $\ell \neq m$, then

$$v_0(u_\ell u_m)v_0 = -(u_\ell u_m)v_0 v_0 = u_\ell u_m, \tag{3.82}$$

agreeing with the right side of (3.81). Thus we have (3.80), hence (3.71).

Rather than concluding that Proposition 3.4 is now proved, we must reveal that the proof of Proposition 3.5 given below actually uses Proposition 3.4. Therefore, it is necessary to produce an alternative endgame to the proof of Proposition 3.4.

We begin by noting that the approach to the proof of Proposition 3.4 described above uses the identities (3.73)–(3.74) with

$$x = v_0, \quad y = u_\ell, \quad z = u_m, \quad \ell \neq m, \tag{3.83}$$

hence $xy = -v_\ell, zx = v_m, yz = \pm u_h,\ \{h, \ell, m\} = \{1, 2, 3\}$. Thus the application of the first identity of (3.73) in (3.81) is justified by the following special case of (3.76):

Proposition 3.6. *If $\{u, v\} \in \mathrm{Im}(\mathbb{O})$ is an orthonormal set, then*

$$L_{uvu} = L_v = L_u L_v L_u. \tag{3.84}$$

Proof. Under these hypotheses, $u^2 = -1$ and $uv = -vu$. Bringing in (3.3), we have

$$uvu = -u^2 v = v, \tag{3.85}$$

which gives the first identity in (3.84). We also have

$$a \in \mathrm{Im}(\mathbb{O}) \Longrightarrow L_a^2 = L_{a^2} = -|a|^2 I, \tag{3.86}$$

the first identity by (3.3). Thus

$$\begin{aligned} -2I = L_{(u+v)}^2 &= (L_u + L_v)(L_u + L_v) \\ &= L_u^2 + L_v^2 + L_u L_v + L_v L_u, \end{aligned} \tag{3.87}$$

so

$$L_u L_v = -L_v L_u, \tag{3.88}$$

and hence

$$L_u L_v L_u = -L_v L_u^2 = L_v, \tag{3.89}$$

giving the second identity in (3.84).

As for the application of (3.74) to (3.78), we need the special case

$$(uv)(wu) = u(vw)u, \tag{3.90}$$

for $u = v_0, v = u_\ell, w = u_m, \ell \neq m, 1 \le \ell, m \le 3$ (so $uv = -v_\ell$), in which cases

$$\{u, v, w, uv\},\ \{u, vw\} \subset \mathrm{Im}(\mathbb{O}),\ \text{are orthonormal sets.} \tag{3.91}$$

In such a case, $u(vw)u = -(vw)u^2 = vw$, so it suffices to show that

$$(uv)(wu) = vw, \tag{3.92}$$

for

$$\{u, v, w, uv\} \subset \mathrm{Im}(\mathbb{O}), \quad \text{orthonormal.} \tag{3.93}$$

When (3.93) holds, we say $\{u, v, w\}$ is a *Cayley triangle.* The following takes care of our needs.

Proposition 3.7. *Assume $\{u, v, w\}$ is a Cayley triangle. Then*

$$v(uw) = -(vu)w, \tag{3.94}$$

$$\langle uv, uw\rangle = 0, \quad \text{so } \{u, v, uw\} \text{ is a Cayley triangle}, \tag{3.95}$$

and (3.92) holds.

Proof. To start, the hypotheses imply

$$vu = -uv, \quad vw = -wv, \quad uw = -wu, \quad (vu)w = -w(vu), \tag{3.96}$$

so

$$\begin{aligned} v(uw) + (vu)w &= -v(wu) - w(vu) \\ &= (v^2 + w^2)u - (v + w)(vu + wu) \\ &= (v + w)^2 u - (v + w)((v + w)u) \\ &= 0, \end{aligned} \tag{3.97}$$

and we have (3.94). Next,

$$\langle uv, uw\rangle = \langle L_u v, L_u w\rangle = \langle u, w\rangle = 0, \tag{3.98}$$

since $L_u \in SO(\mathbb{O})$. Thus $\{u, v, uw\}$ is a Cayley triangle. Applying (3.94) to this Cayley triangle (and bringing in (3.3)) then gives

$$\begin{aligned} (vu)(uw) &= -v(u(uw)) \\ &= -v(u^2 w) \\ &= vw, \end{aligned} \tag{3.99}$$

yielding (3.92).

At this point, we have a complete proof of Proposition 3.4. The proof of Proposition 3.5 will be given in the following section.

4. The automorphism group of $\mathbb{O}$

The set of automorphisms of $\mathbb{O}$ is denoted $\mathrm{Aut}(\mathbb{O})$. Note that $\mathrm{Aut}(\mathbb{O})$ is a group, i.e.,

$$K_j \in \mathrm{Aut}(\mathbb{O}) \Longrightarrow K_1K_2,\ K_j^{-1} \in \mathrm{Aut}(\mathbb{O}). \tag{4.1}$$

Clearly $K \in \mathrm{Aut}(\mathbb{O}) \Rightarrow K1 = 1$. The following result will allow us to establish a converse to Proposition 3.4.

Proposition 4.1. *Assume $K \in \mathrm{Aut}(\mathbb{O})$. Then*

$$K : \mathrm{Im}(\mathbb{O}) \longrightarrow \mathrm{Im}(\mathbb{O}). \tag{4.2}$$

Consequently

$$K\overline{x} = \overline{Kx}, \quad \forall\, x \in \mathbb{O}, \tag{4.3}$$

and

$$|Kx| = |x|, \quad \forall\, x \in \mathbb{O}, \tag{4.4}$$

so $K : \mathbb{O} \to \mathbb{O}$ is an orthogonal transformation.

Proof. To start, we note that, given $x \in \mathbb{O}$, x^2 is real if and only if either x is real or $x \in \mathrm{Im}(\mathbb{O})$. Now, given $u \in \mathrm{Im}(\mathbb{O})$,

$$(Ku)^2 = K(u^2) = -|u|^2K1 = -|u|^2 \quad \text{(real)}, \tag{4.5}$$

so either $Ku \in \mathrm{Im}(\mathbb{O})$ or $Ku = a$ is real. In the latter case, we have $K(a^{-1}u) = 1$, so $a^{-1}u = 1$, so $u = a$, contradicting the hypothesis that $u \in \mathrm{Im}(\mathbb{O})$. This gives (4.2). The result (4.3) is an immediate consequence. Thus, for $x \in \mathbb{O}$,

$$|Kx|^2 = (Kx)(\overline{Kx}) = (Kx)(K\overline{x}) = K(x\overline{x}) = |x|^2, \tag{4.6}$$

giving (4.4).

Now, given $K \in \mathrm{Aut}(\mathbb{O})$, define u_1, u_2, and v_0 by

$$u_1 = Ke_1, \quad u_2 = Ke_2, \quad v_0 = Kf_0. \tag{4.7}$$

By Proposition 4.1, these are orthonormal elements of $\mathrm{Im}(\mathbb{O})$. Also, $\mathcal{A} = K(\mathcal{H})$, spanned by $1, u_1, u_2$, and $u_1u_2 = u_1 \times u_2$, is a subalgebra of $\mathbb{O}$, and $v_0 \in \mathcal{A}^\perp$. These observations, together with Proposition 3.4, yield the following.

Proposition 4.2. *The formulas (4.7) provide a one-to-one correspondence between the set of automorphisms of* $\mathbb{O}$ *and*

$$\begin{aligned}&\text{the set of ordered orthonormal triples } (u_1, u_2, v_0) \text{ in } \mathrm{Im}(\mathbb{O})\\ &\text{such that } v_0 \text{ is also orthogonal to } u_1 \times u_2, \text{ that is,}\\ &\text{the set of Cayley triangles in } \mathrm{Im}(\mathbb{O}).\end{aligned} \tag{4.8}$$

It can be deduced from (4.8) that $\mathrm{Aut}(\mathbb{O})$ is a compact, connected Lie group of dimension 14.

We return to the Moufang identities and use the results on $\mathrm{Aut}(\mathbb{O})$ established above to prove them.

Proof of Proposition 3.5. Consider the first identity in (3.73), i.e.,

$$(xyx)z = x(y(xz)), \quad \forall\, x, y, z \in \mathbb{O}. \tag{4.9}$$

We begin with a few simple observations. First, (4.9) is clearly true if any one of x, y, z is scalar, or if any two of them coincide (thanks to Proposition 3.3). Also, both sides of (4.9) are linear in y and in z. Thus, it suffices to treat (4.9) for $y, z \in \mathrm{Im}(\mathbb{O})$. Meanwhile, multiplying by a real number and applying an element of $\mathrm{Aut}(\mathbb{O})$, we can assume $x = a + e_1$, for some $a \in \mathbb{R}$.

To proceed, (4.9) is clear for $y \in \mathrm{Span}(1, x)$, so, using the linearity in y, and applying Proposition 4.2 again, we can arrange that $y = e_2$. Given this, (4.9) is clear for $z \in \mathcal{H} = \mathrm{Span}(1, e_2, e_2, e_3 = e_1 e_2)$. Thus, using linearity of (4.9) in z, it suffices to treat $z \in \mathcal{H}^\perp$, and again applying an element of $\mathrm{Aut}(\mathbb{O})$, we can assume $z = f_1$.

At this point, we have reduced the task of proving (4.9) to checking it for

$$x = a + e_1, \quad y = e_2, \quad z = f_1, \quad a \in \mathbb{R}, \tag{4.10}$$

and this is straightforward. Similar arguments applied to the second identity in (3.73), and to (3.74), reduce their proofs to a check in the case (4.10).

We next look at some interesting subgroups of $\mathrm{Aut}(\mathbb{O})$. Taking $Sp(1)$ to be the group of unit quaternions, as in (2.25), we have group homomorphisms

$$\alpha, \beta : Sp(1) \longrightarrow \mathrm{Aut}(\mathbb{O}), \tag{4.11}$$

given by

$$\begin{aligned}\alpha(\xi)(\zeta, \eta) &= (\xi\zeta\overline{\xi}, \xi\eta\overline{\xi}),\\ \beta(\xi)(\zeta, \eta) &= (\zeta, \xi\eta),\end{aligned} \tag{4.12}$$

where $\zeta, \eta \in \mathbb{H}$ define $(\zeta, \eta) \in \mathbb{O}$. As in (2.31)–(2.36), for $\xi \in Sp(1)$, $\pi(\xi)\zeta = \xi\zeta\overline{\xi}$ gives an automorphism of $\mathbb{H}$, and it commutes with conjugation in $\mathbb{H}$, so the fact that $\alpha(\xi)$ is an automorphism of $\mathbb{O}$ follows from the definition (3.2) of the product in $\mathbb{O}$. The fact that $\beta(\xi)$ is an automorphism of $\mathbb{O}$ also follows directly from (3.2). Parallel to (2.34),

$$\operatorname{Ker}\alpha = \{\pm 1\} \subset Sp(1), \tag{4.13}$$

so the image of $Sp(1)$ under α is a subgroup of $\mathrm{Aut}(\mathbb{O})$ isomorphic to $SO(3)$. Clearly β is one-to-one, so it yields a subgroup of $\mathrm{Aut}(\mathbb{O})$ isomorphic to $Sp(1)$.

These two subgroups of $\mathrm{Aut}(\mathbb{O})$ do not commute with each other. In fact, we have, for $\xi_j \in Sp(1)$, $(\zeta,\eta) \in \mathbb{O}$,

$$\begin{aligned} \alpha(\xi_1)\beta(\xi_2)(\zeta,\eta) &= (\xi_1\zeta\overline{\xi}_1, \xi_1\xi_2\eta\overline{\xi}_1), \\ \beta(\xi_2)\alpha(\xi_1)(\zeta,\eta) &= (\xi_1\zeta\overline{\xi}_1, \xi_2\xi_1\eta\overline{\xi}_1). \end{aligned} \tag{4.14}$$

Note that, since $\xi_2\xi_1 = \xi_1(\overline{\xi}_1\xi_2\xi_1)$,

$$\beta(\xi_2)\alpha(\xi_1) = \alpha(\xi_1)\beta(\overline{\xi}_1\xi_2\xi_1). \tag{4.15}$$

It follows that

$$G_{\mathcal{H}} = \{\alpha(\xi_1)\beta(\xi_2) : \xi_j \in Sp(1)\} \tag{4.16}$$

is a subgroup of $\mathrm{Aut}(\mathbb{O})$. It is clear from (4.12) that each automorphism $\alpha(\xi_1), \beta(\xi_2)$, and hence each element of $G_{\mathcal{H}}$, preserves $\mathcal{H}$ (and also $\mathcal{H}^\perp$). The converse also holds:

Proposition 4.3. *The group $G_{\mathcal{H}}$ is the group of all automorphisms of $\mathbb{O}$ that preserve $\mathcal{H}$.*

Proof. Indeed, suppose $K \in \mathrm{Aut}(\mathbb{O})$ preserves $\mathcal{H}$. Then $K|_{\mathcal{H}}$ is an automorphism of $\mathcal{H} \approx \mathbb{H}$. Arguments in the paragraph containing (2.35)–(2.38) imply that there exists $\xi_1 \in Sp(1)$ such that $K|_{\mathcal{H}} = \alpha(\xi_1)|_{\mathcal{H}}$, so $K_0 = \alpha(\xi_1)^{-1}K \in \mathrm{Aut}(\mathbb{O})$ is the identity on $\mathcal{H}$. Now $K_0 f_1 = (0, \xi_2)$ for some $\xi_2 \in Sp(1)$, and it then follows from Proposition 4.2 that $K_0 = \beta(\xi_2)$. Hence $K = \alpha(\xi_1)\beta(\xi_2)$, as desired.

For another perspective on $G_{\mathcal{H}}$, we bring in

$$\tilde{\alpha} : Sp(1) \longrightarrow \mathrm{Aut}(\mathbb{O}), \quad \tilde{\alpha}(\xi) = \beta(\overline{\xi})\alpha(\xi). \tag{4.17}$$

Note that

$$\tilde{\alpha}(\xi)(\zeta,\eta) = (\xi\zeta\overline{\xi}, \eta\overline{\xi}), \tag{4.18}$$

so $\tilde{\alpha}$ is a group homomorphism. Another easy consequence of (4.18) is that $\tilde{\alpha}(\xi_1)$ and $\beta(\xi_2)$ commute, for each $\xi_j \in Sp(1)$. We have a surjective group homomorphism

$$\tilde{\alpha} \times \beta : Sp(1) \times Sp(1) \longrightarrow G_{\mathcal{H}}. \tag{4.19}$$

Note that $\mathrm{Ker}(\tilde{\alpha}\times\beta)=\{(1,1),(-1,-1)\}$, with 1 denoting the unit in $\mathbb{H}$. It follows that

$$G_{\mathcal{H}} \approx SO(4). \tag{4.20}$$

We now take a look at one-parameter families of automorphisms of $\mathbb{O}$, of the form

$$K(t)=e^{tA}, \quad A\in\mathcal{L}(\mathbb{O}), \tag{4.21}$$

where e^{tA} is the matrix exponential, studied in §25 of [T6], and in Chapter 3 of [T4]. To see when such linear transformations on $\mathbb{O}$ are automorphisms, we differentiate the identity

$$K(t)(xy)=(K(t)x)(K(t)y), \quad x,y\in\mathbb{O}, \tag{4.22}$$

obtaining

$$A(xy)=(Ax)y+x(Ay), \quad x,y\in\mathbb{O}. \tag{4.23}$$

When (4.23) holds, we say

$$A\in\mathrm{Der}(\mathbb{O}). \tag{4.24}$$

Proposition 4.4. *Given $A\in\mathcal{L}(\mathbb{O})$, $e^{tA}\in\mathrm{Aut}(\mathbb{O})$ for all $t\in\mathbb{R}$ if and only if $A\in\mathrm{Der}(\mathbb{O})$.*

Proof. The implication $\Rightarrow$ was established above. For the converse, suppose A satisfies (4.23). Take $x,y\in\mathbb{O}$, and set

$$X(t)=(e^{tA}x)(e^{tA}y). \tag{4.25}$$

Applying d/dt gives

$$\begin{aligned}\frac{dX}{dt}&=(Ae^{tA}x)(e^{tA}y)+(e^{tA}x)(Ae^{tA}y)\\&=A\big((e^{tA}x)(e^{tA}y)\big)\\&=AX(t),\end{aligned} \tag{4.26}$$

the second identity by (4.23). Since $X(0)=xy$, it follows from the standard uniqueness argument of ODE, cf. (25.11)–(25.16) of [T6], that

$$X(t)=e^{tA}(xy), \tag{4.27}$$

so indeed $e^{tA}\in\mathrm{Aut}(\mathbb{O})$.

The set $\mathrm{Der}(\mathbb{O})$ has the following structure.

Proposition 4.5. Der($\mathbb{O}$) *is a linear subspace of* $\mathcal{L}(\mathbb{O})$ *satisfying*

$$A, B \in \mathrm{Der}(\mathbb{O}) \Longrightarrow [A, B] \in \mathrm{Der}(\mathbb{O}), \tag{4.28}$$

where $[A, B] = AB - BA$. *That is,* Der($\mathbb{O}$) *is a Lie algebra.*

Proof. That Der($\mathbb{O}$) is a linear space is clear from the defining property (4.23). Furthermore, if $A, B \in \mathrm{Der}(\mathbb{O})$, then, for all $x, y \in \mathbb{O}$,

$$\begin{aligned} AB(xy) &= A((Bx)y) + A(x(By)) \\ &= (ABx)y + (Bx)(Ay) + (Ax)(By) + x(ABy), \end{aligned} \tag{4.29}$$

and similarly

$$BA(xy) = (BAx)y + (Ax)(By) + (Bx)(Ay) + x(BAy), \tag{4.30}$$

so

$$[A, B](xy) = ([A, B]x)y + x([A, B]y), \tag{4.31}$$

and we have (4.28).

By Proposition 4.1, if $A \in \mathrm{Der}(\mathbb{O})$, then e^{tA} is an orthogonal transformation for each $t \in \mathbb{R}$. We have

$$(e^{tA})^* = e^{tA^*}, \tag{4.32}$$

so

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow A^* = -A, \tag{4.33}$$

i.e., A is skew-adjoint. It is clear that

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow A : \mathrm{Im}(\mathbb{O}) \to \mathrm{Im}(\mathbb{O}), \tag{4.34}$$

and since Im($\mathbb{O}$) is odd dimensional, the structural result Proposition 11.4 of [T6] implies

$$A \in \mathrm{Der}(\mathbb{O}) \Longrightarrow \mathcal{N}(A) \cap \mathrm{Im}(\mathbb{O}) \neq 0. \tag{4.35}$$

As long as $A \neq 0$, we can also deduce from Proposition 11.4 of [T6] that Im($\mathbb{O}$) contains a two-dimensional subspace with orthonormal basis $\{u_1, u_2\}$, invariant under A, and with repect to which A is represented by a 2×2 block

$$\begin{pmatrix} 0 & -\lambda \\ \lambda & 0 \end{pmatrix}. \tag{4.36}$$

Then, by (4.23),

$$\begin{aligned} A(u_1 u_2) &= (Au_1)u_2 + u_1(Au_2) \\ &= \lambda u_2^2 - \lambda u_1^2 \\ &= 0, \end{aligned} \tag{4.37}$$

so $u_1 u_2 = u_1 \times u_2 \in \mathcal{N}(A) \cap \mathrm{Im}(\mathbb{O})$. As in (3.36)–(3.44), $\mathrm{Span}\{1, u_1, u_2, u_3 = u_1 u_2\} = \mathcal{A}$ is a subalgebra of $\mathbb{O}$ isomorphic to $\mathbb{H}$. We see that A preserves $\mathcal{A}$, so the associated one-parameter group of automorphisms e^{tA} preserves $\mathcal{A}$.

Using Proposition 4.2, we can pick $K \in \mathrm{Aut}(\mathbb{O})$ taking $\mathcal{A}$ to $\mathcal{H}$, and deduce the following.

Proposition 4.6. *Given $A \in \mathrm{Der}(\mathbb{O})$, there exists $K \in \mathrm{Aut}(\mathbb{O})$ such that*

$$Ke^{tA}K^{-1} \in G_{\mathcal{H}}, \quad \forall\, t \in \mathbb{R}. \tag{4.38}$$

Note that then

$$Ke^{tA}K^{-1} = e^{t\widetilde{A}}, \quad \widetilde{A} = KAK^{-1} \in \mathrm{Der}(\mathbb{O}), \tag{4.39}$$

and (4.38) is equivalent to

$$\widetilde{A} : \mathcal{H} \longrightarrow \mathcal{H}, \quad \widetilde{A} \in \mathrm{Der}(\mathbb{O}), \tag{4.40}$$

which also entails $\widetilde{A} : \mathcal{H}^\perp \to \mathcal{H}^\perp$, since $\widetilde{A}$ is skew-adjoint. When (4.40) holds, we say

$$\widetilde{A} \in D_{\mathcal{H}}. \tag{4.41}$$

Going further, suppose we have d commuting elements of $\mathrm{Der}(\mathbb{O})$:

$$A_j \in \mathrm{Der}(\mathbb{O}), \quad A_j A_k = A_k A_j, \quad j, k \in \{1, \dots, d\}. \tag{4.42}$$

A modification of the arguments leading to Proposition 11.4 of [T6] yields a two-dimensional subspace of $\mathrm{Im}(\mathbb{O})$, with orthonormal basis $\{u_1, u_2\}$, invariant under each A_j, with respect to which each A_j is represented by a 2×2 block as in (4.36), with λ replaced by λ_j (possibly 0). As in (4.37),

$$A_j(u_1 u_2) = 0, \quad 1 \le j \le d, \tag{4.43}$$

so each A_j preserves $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_3 = u_1 u_2\}$, and so does each one-parameter group of automorphisms e^{tA_j}. Bringing in $K \in \mathrm{Aut}(\mathbb{O})$, taking $\mathcal{A}$ to $\mathcal{H}$, we have the following variant of Proposition 4.6.

Proposition 4.7. *Given commuting $A_j \in \mathrm{Der}(\mathbb{O})$, $1 \le j \le d$, there exists $K \in \mathrm{Aut}(\mathbb{O})$ such that*

$$Ke^{tA_j}K^{-1} \in G_{\mathcal{H}}, \quad \forall\, t \in \mathbb{R},\ j \in \{1, \dots, d\}. \tag{4.44}$$

As a consequence, we have

$$\widetilde{A}_j = KA_jK^{-1} \in D_{\mathcal{H}}, \quad \widetilde{A}_j \widetilde{A}_k = \widetilde{A}_k \widetilde{A}_j, \quad 1 \le j, k \le d. \tag{4.45}$$

Consequently, $e^{t\widetilde{A}_j}$ are mutually commuting one-parameter subgroups of $G_{\mathcal{H}}$, i.e.,

$$e^{t_j \widetilde{A}_j} \in G_{\mathcal{H}}, \quad e^{t_j \widetilde{A}_j} e^{t_k \widetilde{A}_k} = e^{t_k \widetilde{A}_k} e^{t_j \widetilde{A}_j}, \quad 1 \le j, k \le d. \tag{4.46}$$

One can produce *pairs* of such commuting groups, as follows. Take

$$\tilde{\alpha}(\xi_1(t_1)),\ \beta(\xi_2(t_2)) \in G_{\mathcal{H}}, \tag{4.47}$$

with β as in (4.11)–(4.12), $\tilde{\alpha}$ as in (4.17)–(4.18), and $\xi_\nu(t)$ one-parameter subgroups of $Sp(1)$, for example

$$\xi_\nu(t) = e^{t\omega_\nu}, \quad \omega_\nu \in \mathrm{Im}(\mathbb{H}) = \mathrm{Span}\{i, j, k\}. \tag{4.48}$$

The exponential $e^{t\omega_\nu}$ is amenable to a treatment parallel to that given in §25 of [T6]. Mutual commutativity in (4.47) follows from the general mutual commutativity of $\tilde{\alpha}$ and β. The following important structural information on $\mathrm{Aut}(\mathbb{O})$ says $d = 2$ is as high as one can go.

Proposition 4.8. *If $A_j \in \mathrm{Der}(\mathbb{O})$ are mutually commuting, for $j \in \{1, \dots, d\}$, and if $\{A_j\}$ is linearly independent in $\mathcal{L}(\mathbb{O})$, then $d \le 2$.*

Proof. To start, we obtain from A_j the mutually commuting one-parameter groups $Ke^{tA_j}K^{-1}$, subgroups of $G_{\mathcal{H}}$. Taking inverse images under the two-to-one surjective homomorphism (4.19), we get mutually commuting one-parameter subgroups $\gamma_j(t)$ of $Sp(1) \times Sp(1)$, which can be written

$$\gamma_j(t) = \begin{pmatrix} e^{\omega_j t} & \\ & e^{\sigma_j t} \end{pmatrix}, \quad \omega_j, \sigma_j \in \mathrm{Im}(\mathbb{H}), \quad 1 \le j \le d. \tag{4.49}$$

Parallel to Proposition 25.6 of [T6], this commutativity requires $\{\omega_j : 1 \le j \le d\}$ to commute in $\mathbb{H}$ and it also requires $\{\sigma_j : 1 \le j \le d\}$ to commute in $\mathbb{H}$. These conditions in turn require each ω_j to be a real multiple of some $\omega^\# \in \mathrm{Im}(\mathbb{H})$ and each σ_j to be a real multiple of some $\sigma^\# \in \mathrm{Im}(\mathbb{H})$.

Now the linear independence of $\{A_j : 1 \le j \le d\}$ in $\mathrm{Der}(\mathbb{O})$ implies the linear independence of $\{(\omega_j, \sigma_j) : 1 \le j \le d\}$ in $\mathrm{Im}(\mathbb{H}) \oplus \mathrm{Im}(\mathbb{H})$, and this implies $d \le 2$.

We turn to the introduction of another interesting subgroup of $\mathrm{Aut}(\mathbb{O})$. Note that, by Proposition I.7, given any unit $u_1 \in \mathrm{Im}(\mathbb{O})$, there exists $K \in \mathrm{Aut}(\mathbb{O})$ such that $Ke_1 = u_1$. Consequently, $\mathrm{Aut}(\mathbb{O})$, acting on $\mathrm{Im}(\mathbb{O})$ as a group of orthogonal transformations, acts *transitively* on the unit sphere S in $\mathrm{Im}(\mathbb{O}) \approx \mathbb{R}^7$, i.e., on $S \approx S^6$. We are hence interested in the group

$$\{K \in \mathrm{Aut}(\mathbb{O}) : Ke_1 = e_1\} = \mathcal{G}_{e_1}. \tag{4.50}$$

We claim that

$$\mathcal{G}_{e_1} \approx SU(3). \tag{4.51}$$

As preparation for the demonstration, note that each $K \in \mathcal{G}_{e_1}$ is an orthogonal linear transformation on $\mathbb{O}$ that leaves invariant $\mathrm{Span}\{1, e_1\}$, and hence it also leaves invariant the orthogonal complement

$$V = \mathrm{Span}\{1, e_1\}^\perp = \mathrm{Span}\{e_2, e_3, f_0, f_1, f_2, f_3\}, \tag{4.52}$$

a linear space of $\mathbb{R}$-dimension 6. We endow V with a complex structure. Generally, a complex structure on a real vector space V is an $\mathbb{R}$-linear map $J : V \to V$ such that $J^2 = -I_V$. One can check that this requires $\dim_{\mathbb{R}} V$ to be even, say $2k$. Then (V, J) has the structure of a complex vector space, with

$$(a + ib)v = av + bJv, \quad a, b \in \mathbb{R}, \ v \in V. \tag{4.53}$$

One has $\dim_{\mathbb{C}}(V, J) = k$. If V is a real inner product space, with inner product $\langle\ ,\ \rangle$, and if J is orthogonal (hence skew-adjoint) on V, then (V, J) gets a natural Hermitian inner product

$$(u, v) = \langle u, v\rangle + i\langle u, Jv\rangle. \tag{4.54}$$

If $T : V \to V$ preserves $\langle \ , \ \rangle$ and commutes with J, then it also preserves $(\ , \ )$, so it is a unitary transformation on (V, J).

We can apply this construction to V as in (4.52), with

$$Jv = L_{e_1} v = e_1 v, \tag{4.55}$$

noting that L_{e_1} is an orthogonal map on $\mathbb{O}$ that preserves $\mathrm{Span}\{1, e_1\}$, and hence also preserves V. To say that an $\mathbb{R}$-linear map $K : V \to V$ is $\mathbb{C}$-linear is to say that $K(e_1 v) = e_1 K(v)$, for all $v \in V$. Clearly this holds if $K \in \mathrm{Aut}(\mathbb{O})$ and $Ke_1 = e_1$. Thus each element of $\mathcal{G}_{e_1}$ defines a complex linear orthogonal (hence unitary) transformation on V, and we have an injective group homomorphism

$$\mathcal{G}_{e_1} \longrightarrow U(V, J). \tag{4.56}$$

Note that the 6 element real orthonormal basis of V in (4.52) yields the 3 element orthonormal basis of (V, J),

$$\{e_2, f_0, f_2\}, \tag{4.57}$$

since

$$e_3 = e_1 e_2, \quad f_1 = e_1 f_0, \quad f_3 = -e_1 f_2, \tag{4.58}$$

the latter two identities by (3.30)–(3.31). This choice of basis yields the isomorphism

$$U(V, J) \approx U(3). \tag{4.59}$$

We aim to identify the image of $\mathcal{G}_{e_1}$ in $U(3)$ that comes from (4.56) and (4.59).

To accomplish this, we reason as follows. From Proposition 4.2 it follows that there is a natural one-to-one correspondence between the elements of $\mathcal{G}_{e_1}$ and

$$\begin{gathered} \text{the set of ordered orthonormal pairs } \{u_2, v_0\} \text{ in } V \\ \text{such that also } v_0 \perp e_1 u_2, \end{gathered} \tag{4.60}$$

or, equivalently,

$$\text{the set of ordered orthonormal pairs } \{u_2, v_0\} \text{ in } (V, J), \tag{4.61}$$

where (V, J) carries the Hermitian inner product (4.54). In fact, the correspondence associates to $K \in \mathcal{G}_{e_1}$ (i.e., $K \in \mathrm{Aut}(\mathbb{O})$ and $Ke_1 = e_1$) the pair

$$u_2 = Ke_2, \quad v_0 = Kf_0. \tag{4.62}$$

Then the image of $\mathcal{G}_{e_1}$ in $U(V,J)$ in (4.56) is uniquely determined by the action of K on the third basis element in (4.57), as

$$Kf_2 = K(e_2f_0) = K(e_2)K(f_0) = u_2v_0 = u_2 \times v_0, \tag{4.63}$$

where we recall from (3.30) that $f_2 = e_2f_0$, and the last identity in (4.63) follows from (3.21).

From (4.60)–(4.61), it can be deduced that $\mathcal{G}_{e_1}$ is a compact, connected Lie group of dimension 8. Then (4.55) and (4.58) present $\mathcal{G}_{e_1}$ as isomorphic to a subgroup (call it $\widetilde{\mathcal{G}}$) of $U(3)$ that is a compact, connected Lie group of dimension 8. Meanwhile, $\dim U(3) = 9$, so $\widetilde{\mathcal{G}}$ has codimension 1. We claim that this implies

$$\widetilde{\mathcal{G}} = SU(3). \tag{4.64}$$

We sketch a proof of (4.64), using some elements of Lie group theory.

To start, one can show that a connected, codimension-one subgroup of a compact, connected Lie group must be *normal*. Hence $\widetilde{\mathcal{G}}$ is a normal subgroup of $U(3)$. This implies $U(3)/\widetilde{\mathcal{G}}$ is a group. This quotient is a compact Lie group of dimension 1, hence isomorphic to $S^1 = \{z \in \mathbb{C} : |z| = 1\}$, and the projection $U(3) \to U(3)/\widetilde{\mathcal{G}}$ produces a continuous, surjective group homomorphism

$$\vartheta : U(3) \longrightarrow S^1, \quad \operatorname{Ker}\vartheta = \widetilde{\mathcal{G}}. \tag{4.65}$$

Now a complete list of such homomorphisms is given by

$$\vartheta_j(K) = (\det K)^j, \quad j \in \mathbb{Z} \setminus 0, \tag{4.66}$$

and in such a case, $\operatorname{Ker}\vartheta_j$ has $|j|$ connected components. Then connectivity of $\widetilde{\mathcal{G}}$ forces $\vartheta = \vartheta_{\pm 1}$ in (4.65), which in turn gives (4.64).

It is useful to take account of various subgroups of $\operatorname{Aut}(\mathbb{O})$ that are conjugate to $G_{\mathcal{H}}$ (given by (4.16)) or to $\mathcal{G}_{e_1}$ (given by (4.50)). In particular, when $\mathcal{A} \subset \mathbb{O}$ is a four-dimensional subalgebra, we set

$$G_{\mathcal{A}} = \{K \in \operatorname{Aut}(\mathbb{O}) : K(\mathcal{A}) \subset \mathcal{A}\}, \tag{4.67}$$

and if $u \in \operatorname{Im}(\mathbb{O}), |u| = 1$, we set

$$\mathcal{G}_u = \{K \in \operatorname{Aut}(\mathbb{O}) : Ku = u\}. \tag{4.68}$$

We see that each group $G_{\mathcal{A}}$ is conjugate to $G_{\mathcal{H}}$, and isomorphic to $SO(4)$, and each group $\mathcal{G}_u$ is conjugate to $\mathcal{G}_{e_1}$, and isomorphic to $SU(3)$.

It is of interest to look at $\mathcal{G}_u \cap \mathcal{G}_v$, where u and v are unit elements of $\operatorname{Im}(\mathbb{O})$ that are not collinear. Then

$$\mathcal{G}_u \cap \mathcal{G}_v = \{K \in \operatorname{Aut}(\mathbb{O}) : K = I \text{ on } \operatorname{Span}\{u,v\}\}. \tag{4.69}$$

Now we can write $\mathrm{Span}\{u, v\} = \mathrm{Span}\{u_1, u_2\}$, with $u_1 = u, u_2 \perp u_1$, and note that $Ku_j = u_j \Rightarrow K(u_1u_2) = u_1u_2$, so (4.69) is equal to

$$\mathcal{G}_{\mathcal{A}} = \{K \in \mathrm{Aut}(\mathbb{O}) : K = I \text{ on } \mathcal{A}\}, \tag{4.70}$$

where $\mathcal{A} = \mathrm{Span}\{1, u_1, u_2, u_1u_2\}$ is a four-dimensional subalgebra of $\mathbb{O}$. Clearly

$$\mathcal{G}_{\mathcal{A}} \subset G_{\mathcal{A}}, \quad \text{and} \quad \mathcal{G}_{\mathcal{A}} \approx Sp(1) \approx SU(2). \tag{4.71}$$

In fact, $\mathcal{G}_{\mathcal{A}}$ is conjugate to $\mathcal{G}_{\mathcal{H}} = \beta(Sp(1))$, with β as in (4.11)–(4.12).

Extending (4.52), we have associated to each unit $u \in \mathrm{Im}(\mathbb{O})$ the space

$$V_u = \mathrm{Span}\{1, u\}^{\perp}, \tag{4.72}$$

and $L_u : V_u \to V_u$ gives a complex structure $J_u = L_u|_{V_u}$, so (V_u, J_u) is a three-dimensional complex vector space. Parallel to (4.56), we have an injective group homomorphism

$$\mathcal{G}_u \longrightarrow U(V_u, J_u), \tag{4.73}$$

whose image is a codimension-one subgroup isomorphic to $SU(3)$. Associated to the family (V_u, J_u) is the following interesting geometrical structure. Consider the unit sphere $S \approx S^6$ in $\mathrm{Im}(\mathbb{O})$. There is a natural identification of V_u with the tangent space T_uS to S at u:

$$T_uS = V_u, \tag{4.74}$$

and the collection of complex structures J_u gives S what is called an *almost complex structure*. Now an element $K \in \mathrm{Aut}(\mathbb{O})$ acts on S, thanks to Proposition 4.1. Furthermore, for each $u \in S$,

$$K : V_u \longrightarrow V_{Ku} \tag{4.75}$$

is an isomtery, and it is $\mathbb{C}$-linear, since

$$v \in V_u \Longrightarrow K(uv) = K(u)K(v). \tag{4.76}$$

Thus $\mathrm{Aut}(\mathbb{O})$ acts as a group of rotations on S that preserve its almost complex structure. In fact, this property characterizes $\mathrm{Aut}(\mathbb{O})$. To state this precisely, we bring in the following notation. Set

$$\iota : \mathrm{Aut}(\mathbb{O}) \longrightarrow SO(\mathrm{Im}(\mathbb{O})), \quad \iota(K) = K\big|_{\mathrm{Im}(\mathbb{O})}. \tag{4.77}$$

This is an injective group homomorphism, whose image we denote

$$A^b(\mathbb{O}) = \iota\, \mathrm{Aut}(\mathbb{O}). \tag{4.78}$$

The inverse of the isomorphism $\iota : \mathrm{Aut}(\mathbb{O}) \to A^b(\mathbb{O})$ is given by

$$\begin{gathered} j\Big|_{A^b(\mathbb{O})}, \quad j : SO(\mathrm{Im}(\mathbb{O})) \to SO(\mathbb{O}), \\ J(K_0)(a + u) = a + K_0u. \end{gathered} \tag{4.79}$$

Our result can be stated as follows.

Proposition 4.9. *The group* Γ *of rotations on* $\mathrm{Im}(\mathbb{O})$ *that preserve the almost complex structure of* S *is equal to* $A^b(\mathbb{O})$.

Proof. We have seen that $A^b(\mathbb{O}) \subset \Gamma$. It remains to prove that $\Gamma \subset A^b(\mathbb{O})$, so take $K_0 \in \Gamma$, and set $K = j(K_0)$, as in (4.79). We need to show that $K \in \mathrm{Aut}(\mathbb{O})$. First, one readily checks that, if $K = j(K_0)$, then

$$K \in \mathrm{Aut}(\mathbb{O}) \Longleftrightarrow K(uv) = K(u)K(v), \ \forall\, u, v \in \mathrm{Im}(\mathbb{O}), \tag{4.80}$$

and furthermore we can take $|u| = 1$. Now the condition $K_0 \in \Gamma$ implies

$$K_0(uv) = K_0(u)K_0(v), \quad \forall\, u \in \mathrm{Im}(\mathbb{O}),\ v \in V_u. \tag{4.81}$$

To finish the argument, we simply note that if $K_0 \in \Gamma$ and $K = j(K_0)$, and if u is a unit element of $\mathrm{Im}(\mathbb{O})$ and $v \in V_u$, then for all $a \in \mathbb{R}$,

$$\begin{aligned} K(u(au+v)) &= K(-a+uv) \\ &= -a + K_0(uv) \\ &= -a + K_0(u)K_0(v), \end{aligned} \tag{4.82}$$

while

$$\begin{aligned} (Ku)(K(au+v)) &= (K_0u)(aK_0u + K_0v) \\ &= a(K_0u)^2 + (K_0u)(K_0v) \\ &= -a + K_0(u)K_0(v). \end{aligned} \tag{4.83}$$

This finishes the proof.

Further results on almost complex 6-dimensional submanifolds, including submanifolds of $\mathbb{O}$, can be found in [Br1] and [Br2].

5. Simplicity and root structure of $\mathrm{Aut}(\mathbb{O})$

Our first goal in this section is to establish the following.

Proposition 5.1. *The group* $\mathrm{Aut}(\mathbb{O})$ *is simple.*

We will deduce this from the facts that $\mathrm{Aut}(\mathbb{O})$ is a compact, connected Lie group of dimension 14 and that it has rank 2. We recall from basic Lie group theory that if G is a compact Lie group, it has a maximal torus, and any two such are conjugate. The dimension of such a maximal torus is the rank of G. That $\mathrm{Aut}(\mathbb{O})$ has rank 2 follows from Proposition 4.8. The following general result basically does the trick.

Proposition 5.2. *Let G be a compact Lie group of rank 2. If its Lie algebra $\mathfrak{g}$ has a non-trivial ideal, $\mathfrak{h}$, then* $\dim G \le 6$.

Proof. Give $\mathfrak{g}$ an ad-invariant inner product. If $\mathfrak{h} \subset \mathfrak{g}$ is an ideal, then ad $\mathfrak{g}$ preserves both $\mathfrak{h}$ and $\mathfrak{h}^\perp$, so $\mathfrak{h}^\perp$ is also an ideal, and each $X \in \mathfrak{h}$ commutes with each $Y \in \mathfrak{h}^\perp$. Now if $\mathfrak{h}$ and $\mathfrak{h}^\perp$ are both nonzero,

$$\text{Rank } \mathfrak{g} = 2 \Longrightarrow \text{Rank } \mathfrak{h} = \ \text{Rank } \mathfrak{h}^\perp = 1.$$

But, as is well known,

$$\text{Rank } \mathfrak{h} = 1 \Longrightarrow \dim \mathfrak{h} = 1 \ \text{ or } \ 3,$$

so we have the conclusion that $\dim G \le 6$.

It follows from Proposition 5.2 that the Lie algebra $\mathrm{Der}(\mathbb{O})$ of $\mathrm{Aut}(\mathbb{O})$ has no nontrivial ideals. A connected Lie group with this property is typically said to be simple. However, we can establish the more precise result that $\mathrm{Aut}(\mathbb{O})$ contains no nontrivial normal subgroups. Indeed, if H were such a subgroup, so would be its closure, so it suffices to consider the case when H is closed. (The reader can show that a proper *dense* subgroup of a noncommutative, connected Lie group cannot be normal.) Then H is a Lie group, and Proposition 5.2 implies that either $H = \mathrm{Aut}(\mathbb{O})$ or H is discrete, hence finite. In such a case, H normal implies H is the center of $\mathrm{Aut}(\mathbb{O})$, so our task is reduced to showing

$$\mathrm{Aut}(\mathbb{O}) \ \text{ has trivial center.} \tag{5.1}$$

Indeed, suppose K_0 belongs to the center of $\mathrm{Aut}(\mathbb{O})$. Then K_0 belongs to a one-parameter subgroup e^{tA}, and (4.35) applies, to yield $u \in S \subset \mathrm{Im}(\mathbb{O})$, fixed under the action of e^{tA}, hence fixed by K_0. Then, for each $K \in \mathrm{Aut}(\mathbb{O})$, $KK_0K^{-1} = K_0$ fixes Ku, and since $\mathrm{Aut}(\mathbb{O})$ acts transitively on the unit sphere $S \subset \mathrm{Im}(\mathbb{O})$, K_0 must fix each point of $\mathrm{Im}(\mathbb{O})$, so $K_0 = I$, and we have (5.1).

Our next goal is to analyze the root structure of $\mathrm{Aut}(\mathbb{O})$. We start with a general definition. Let G be a compact, connected Lie group, with maximal torus $\mathbb{T}$, having associated Lie algebras $\mathfrak{t} \subset \mathfrak{g}$. Give $\mathfrak{g}$ an Ad-invariant inner product. The adjoint representation Ad of G on $\mathfrak{g}_{\mathbb{C}}$ has derived Lie algebra representation ad of $\mathfrak{g}$ by skew-adjoint transformations on $\mathfrak{g}_{\mathbb{C}}$, which simultaneously diagonalize when restricted to $\mathfrak{t}$. We have the root space decomposition

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{\alpha} \mathfrak{g}_{\alpha}, \tag{5.2}$$

where, given $\alpha \in \mathfrak{t}'$, $\alpha \neq 0$,

$$\mathfrak{g}_{\alpha} = \{z \in \mathfrak{g}_{\mathbb{C}} : [x, z] = i\alpha(x)z,\ \forall\, x \in \mathfrak{t}\}. \tag{5.3}$$

If $\mathfrak{g}_{\alpha} \neq 0$, we call α a root, and nonzero elements of $\mathfrak{g}_{\alpha}$ are called root vectors. It is a fact (cf. [T2], §35), that

$$[\mathfrak{g}_{\alpha}, \mathfrak{g}_{\beta}] \subset \mathfrak{g}_{\alpha+\beta}, \tag{5.4}$$

that

$$\alpha \ \text{ root } \Longrightarrow \dim \mathfrak{g}_{\alpha} = 1, \tag{5.5}$$

and that if $\mathfrak{z}$ denotes the center of $\mathfrak{g}$,

$$\mathfrak{z} = 0 \Longrightarrow \ \text{the set of roots spans } \ \mathfrak{t}'. \tag{5.6}$$

Before tackling the particulars for $G = \mathrm{Aut}(\mathbb{O})$, we describe the most classical case $SU(n)$. Details can be found in §§19–22 of [T2].

The group $SU(n)$ has maximal torus

$$\mathbb{T} = \left\{ \begin{pmatrix} e^{ix_1} & & \\ & \ddots & \\ & & e^{ix_n} \end{pmatrix} : x_j \in \mathbb{R},\ \sum_j x_j = 0 \right\}, \tag{5.7}$$

leading to the identification

$$\mathfrak{t} = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_1 + \cdots + x_n = 0\}. \tag{5.8}$$

Then the set Δ of roots of $SU(n)$ is given by

$$\Delta = \{\omega_{jk} : j \neq k, 1 \leq j, k \leq n\}, \tag{5.9}$$

where $\omega_{jk} \in \mathfrak{t}'$ is given by

$$\omega_{jk}(x) = x_j - x_k, \quad x \in \mathfrak{t}. \tag{5.10}$$

See [T2], (19.13).

Such results, for $n = 3$, actually yield half the roots of $\mathrm{Aut}(\mathbb{O})$, as we now explain. As seen in §4, we have the subgroup

$$\mathcal{G}_{e_1} = \{K \in \mathrm{Aut}(\mathbb{O}) : Ke_1 = e_1\} \approx SU(3). \tag{5.11}$$

Thus a maximal torus $\mathbb{T}$ of $\mathcal{G}_{e_1}$ is two-dimensional, and, by Proposition 4.8, this must also be a maximal torus of $\mathrm{Aut}(\mathbb{O})$. The adjoint action of $\mathfrak{t}$ on $\mathrm{Der}(\mathbb{O})$ leaves invariant the Lie algebra $\mathfrak{g}_{e_1}$ of $\mathcal{G}_{e_1}$, so, with the identification (5.8), we see that

$$\{\omega_{jk} : 1 \le j, k \le 3, j \ne k\}, \quad \omega_{jk}(x) = x_j - x_k, \tag{5.12}$$

are roots of $\mathrm{Aut}(\mathbb{O})$. This gives six roots. Since $\dim \mathrm{Aut}(\mathbb{O}) = 14$ and $\mathfrak{t}$ has dimension 2, it follows from (5.2)–(5.5) that $\mathrm{Aut}(\mathbb{O})$ has 12 roots. It remains to find the other six.

Let us abstract the setting. Let G be a compact, connected Lie group, $H \subset G$ a compact, connected subgroup, and assume that a maximal torus $\mathbb{T}$ of G is contained in H, i.e., $\mathbb{T} \subset H$. Then the adjoint action of G on $\mathfrak{g}_{\mathbb{C}}$, restricted to $\mathbb{T}$, is also the restriction to $\mathbb{T}$ of the action of H on $\mathfrak{g}_{\mathbb{C}}$, obtained by restricting Ad from G to H. This latter is a unitary representation of H on $\mathfrak{g}_{\mathbb{C}}$, which we will denote by π. Thus the roots of G coincide with the weights of π.

We recall the definition of weights. Let H be as above, with maximal torus $\mathbb{T}$, whose Lie algebra is $\mathfrak{t}$, and let π be a unitary representation of H on a finite-dimensional complex inner-product space V. Then there is an orthogonal decomposition

$$V = \bigoplus_{\lambda} V_\lambda, \tag{5.13}$$

where, for $\lambda \in \mathfrak{t}'$,

$$V_\lambda = \{v \in V : d\pi(x)v = i\lambda(x)v, \ \forall\, x \in \mathfrak{t}\}. \tag{5.14}$$

If $V_\lambda \ne 0$, we call λ a weight, and any nonzero $v \in V_\lambda$ a weight vector. Generally, if π is a representation of H on V, we define the *contragredient* representation $\overline{\pi}$ of H on V' by $\overline{\pi}(g) = \pi(g^{-1})^t$. It is readily verified that $\lambda \in \mathfrak{t}'$ is a weight of π if and only if $-\lambda$ is a weight of $\overline{\pi}$.

To take an example, let $H = SU(n)$, with maximal torus given by (5.7) and $\mathfrak{t}$ as in (5.8), and let π_0 be the standard representation of $SU(n)$ on $\mathbb{C}^n$. Then the weights of π_0 are

$$\{\lambda_j : 1 \le j \le n\}, \quad \lambda_j(x) = x_j, \tag{5.15}$$

with asociated weight spaces $V_{\lambda_j} = \mathrm{Span}\{e_j\}$, where $\{e_1, \dots, e_n\}$ is the standard basis of $\mathbb{C}^n$. The weights of the contragredient representation $\overline{\pi}_0$ of $SU(n)$ on $\mathbb{C}^n$ are given by

$$\{-\lambda_j : 1 \le j \le n\}. \tag{5.16}$$

We return to the situation introduced three paragraphs above, with $\mathbb{T} \subset H \subset G$, and π the restriction to H of the adjoint representation of G on $\mathfrak{g}$ (and on its complexification $\mathfrak{g}_{\mathbb{C}}$). Taking an Ad-invariant inner product on $\mathfrak{g}$, we can write

$$\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{h}^{\perp}, \tag{5.17}$$

and both pieces are invariant under π, say

$$\pi = \pi_{\mathfrak{h}} \oplus \pi_1. \tag{5.18}$$

Of course, $\pi_{\mathfrak{h}}$ is simply the adjoint action of H on $\mathfrak{h}$. We need to analyze π_1.

To do this, it is convenient to look at the homogeneous space $M = G/H$, on which G acts transitively. Then H is the subgroup of elements of G that fix the point $p = eH \in M$. This gives rise to an action of H on T_pM, i.e., a real representation ρ of H on T_pM. Furthermore, we have natural equivalences

$$\mathfrak{h}^{\perp} \approx T_pM, \quad \pi_1 \approx \rho. \tag{5.19}$$

We now apply this to

$$G = \mathrm{Aut}(\mathbb{O}), \quad H = \mathcal{G}_{e_1}, \quad M = S \subset \mathrm{Im}(\mathbb{O}), \quad p = e_1. \tag{5.20}$$

Then, as seen in §4, T_pS carries a complex structure, with respect to which, via the isomorphism $\mathcal{G}_{e_1} \approx SU(3)$ set up in §4, ρ becomes the standard representation π_0 of $SU(3)$ on $\mathbb{C}^3$.

However, we need to regard ρ as a real representation on T_pS, and then complexify this vector space. When this is done, the resulting representation on $(\mathfrak{h}^{\perp})_{\mathbb{C}}$ is seen to be

$$\pi_0 \oplus \overline{\pi}_0, \tag{5.21}$$

with weights

$$\{\lambda_j, -\lambda_j : 1 \le j \le 3\}, \quad \lambda_j(x) = x_j. \tag{5.22}$$

We have the following conclusion.

Proposition 5.3. *The roots of* $\mathrm{Aut}(\mathbb{O})$ *are the linear functionals on*

$$\mathfrak{t} = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\} \tag{5.23}$$

given by (5.12) and (5.22).

We want to investigate the Weyl group of $\mathrm{Aut}(\mathbb{O})$. Generally, if G is a compact, connected Lie group with maximal torus $\mathbb{T}$, the Weyl group of G is

$$W(G) = N(\mathbb{T})/\mathbb{T}, \quad N(\mathbb{T}) = \{g \in G : g^{-1}\mathbb{T}g = \mathbb{T}\}. \tag{5.24}$$

We define the representation $\mathcal{W}$ of $N(\mathbb{T})$ on $\mathfrak{t}$ by

$$\mathcal{W}(g) = \operatorname{Ad}(g)\Big|_{\mathfrak{t}}, \quad \text{for } \ g \in N(\mathbb{T}), \tag{5.25}$$

and denote by $\overline{\mathcal{W}}$ the contragredient representation on $\mathfrak{t}'$ (and its complexification). Of course, these two representations are equivalent via the isomorphism $\mathfrak{t} \approx \mathfrak{t}'$ induced by the Ad-invariant inner product we use on $\mathfrak{g}$.

A key example is

$$W(SU(n)) \approx S_n, \tag{5.26}$$

the symmetric group on n symbols, which arises as follows (if n is odd). For $\sigma \in S_n$, the permutation matrix $E_\sigma \in U(n)$, defined on the standard basis $\{u_1, \dots, u_n\}$ of $\mathbb{C}^n$ by $E_\sigma u_k = u_{\sigma(k)}$, has the property that

$$C(E_\sigma) : \mathbb{T} \longrightarrow \mathbb{T}, \quad C(E_\sigma)V = E_\sigma^{-1} V E_\sigma, \tag{5.27}$$

with $\mathbb{T}$ as in (5.7). Since $\det E_\sigma = \operatorname{sgn} \sigma$, we need to alter (5.27) to get an element of $N(\mathbb{T}) \subset SU(n)$. For n odd, we can just replace E_σ in (5.27) by

$$\widetilde{E}_\sigma = (\operatorname{sgn} \sigma) E_\sigma. \tag{5.28}$$

For n even, see (37.35)–(37.37) of [T2]. Of immediate interest here is the case $n = 3$. Note that $\sigma \mapsto \widetilde{E}_\sigma$ gives a group homomorphism

$$S_3 \longrightarrow N(\mathbb{T}) \subset SU(3), \tag{5.29}$$

whose composition with $\widetilde{E}_\sigma \mapsto C(\widetilde{E}_\sigma) : \mathbb{T} \to \mathbb{T}$ yields an isomorphism of S_3 with the image of $W(SU(3))$ under the map $\overline{\mathcal{W}}$. In connection with these facts, we mention the following general results regarding $W(G)$, for an arbitrary compact, connected, semisimple Lie group G. For details, see §37 of [T2] and Chapter 8 of [Si].

Proposition A. *Let π be a unitary representation of G on V, with weight space decomposition $V = \oplus V_\lambda$. Then*

$$g \in N(\mathbb{T}) \Longrightarrow \pi(g) : V_\lambda \to V_{\overline{\mathcal{W}}(g)\lambda}. \tag{5.30}$$

Proposition B. *If $g \in G$ and $g^{-1} u g = u$ for each $u \in \mathbb{T}$, then $g \in \mathbb{T}$. Hence if $g \in N(\mathbb{T})$ and $\mathcal{W}(g) = I$ on $\mathfrak{t}$, then $g \in \mathbb{T}$. Consequently, we can identify $W(G)$ with its image under $\mathcal{W}$ in $G\ell(\mathfrak{t})$, and therefore also with its image under $\overline{\mathcal{W}}$ in $G\ell(\mathfrak{t}')$.*

Proposition C. *The image of $W(G)$ under $\overline{\mathcal{W}}$ in $G\ell(\mathfrak{t}')$ is generated by the set of reflections S_α across hyperplanes in $\mathfrak{t}'$ orthogonal to α, as α runs over the set of roots of G.*

It is straightforward to verify these results for $G = SU(3)$. Note that, under $\overline{\mathcal{W}}$, the Weyl group $W(SU(3))$ acts transitively on each of the sets

$$(5.31) \qquad \{\omega_{jk} : j \neq k, 1 \leq j,k \leq 3\}, \quad \{\lambda_j : 1 \leq j \leq 3\}, \quad \{-\lambda_j : 1 \leq j \leq 3\},$$

defined as in (5.12), (5.15), and (5.16). The first set is the set of roots for $SU(3)$, and the last two sets are, respectively, the sets of weights for π_0 and $\overline{\pi}_0$.

Composing the map $\sigma \mapsto \widetilde{E}_\sigma$ in (5.29) with the inclusion $SU(3) \approx \mathcal{G}_{e_1} \subset \mathrm{Aut}(\mathbb{O})$ yields the injective group homomorphism

$$(5.32) \qquad W(SU(3)) \longrightarrow W(\mathrm{Aut}(\mathbb{O})).$$

However, $W(\mathrm{Aut}(\mathbb{O}))$ is bigger than $W(SU(3))$. By Proposition C, the image under $\overline{\mathcal{W}}$ of $W(SU(3))$ is generated by the reflections in $\mathfrak{t}'$ across lines orthogonal to ω_{12}, ω_{23}, and ω_{31}, respectively. The image under $\overline{\mathcal{W}}$ of $W(\mathrm{Aut}(\mathbb{O})$ is generated by these 3 reflections plus 3 more: reflections in $\mathfrak{t}'$ across lines orthogonal to λ_1, λ_2, and λ_3, rspectively. In particular, the image under $\overline{\mathcal{W}}$ of $W(\mathrm{Aut}(\mathbb{O}))$ acts transitively on each of the sets

$$(5.33) \qquad \{\omega_{jk} : j \neq k, 1 \leq j,k \leq 3\}, \quad \{\lambda_j, -\lambda_j : 1 \leq j \leq 3\},$$

which together give all the roots of $\mathrm{Aut}(\mathbb{O})$. We see that $W(\mathrm{Aut}(\mathbb{O}))$ is isomorphic to the group of isometries of a regular hexagon.

6. More on the Lie algebra of $\mathrm{Aut}(\mathbb{O})$

As seen in §§4-5, the Lie algebra $\mathrm{Der}(\mathbb{O})$ of $\mathrm{Aut}(\mathbb{O})$ can be written as a vector space sum

$$\mathrm{Der}(\mathbb{O}) = \mathfrak{su}(3) \oplus V, \tag{6.1}$$

where $\mathfrak{su}(3)$ is the Lie algebra of $\mathcal{G}_{e_1} \approx SU(3)$ and V, the orthogonal complement of $\mathfrak{su}(3)$, is isomorphic to $T_{e_1}S$, a vector space of $\mathbb{R}$-dimension 6, with a complex structure J, so $(V, J) \approx \mathbb{C}^3$, and the natural action ρ of $\mathcal{G}_{e_1}$ on V is equivalent to the standard action of $SU(3)$ on $\mathbb{C}^3$. Thus an element of $\mathrm{Der}(\mathbb{O})$ can be represented as a pair (X, v), with $X \in \mathfrak{su}(3),\ v \in V$. If also $(Y, w) \in \mathrm{Der}(\mathbb{O})$, we want to look at the Lie bracket

$$[(X, v), (Y, w)] = [X, Y] + [X, w] + [v, Y] + [v, w]. \tag{6.2}$$

Of course, $[X, Y]$ is the standard bracket on $\mathfrak{su}(3)$. Meanwhile, by (5.18)–(5.19),

$$[X, w] = d\rho(X)w \in V, \tag{6.3}$$

and similarly $[v, Y] = -[Y, v] = -d\rho(Y)v$.

It remains to examine $[v, w]$, which will typically have a component in $\mathfrak{su}(3)$ and a component in V. The component in $\mathfrak{su}(3)$ is specified by

$$\begin{aligned} \langle X, [v, w]\rangle &= \langle X, \mathrm{ad}(v), w\rangle \\ &= -\langle \mathrm{ad}(v)X, w\rangle \\ &= \langle d\rho(X)v, w\rangle. \end{aligned} \tag{6.4}$$

For further analysis of $[v, w]$, it is convenient to bring in the complexification

$$V_{\mathbb{C}} = V_1 \oplus V_{-1}, \tag{6.5}$$

where

$$V_\mu = \{v \in V_{\mathbb{C}} : Jv = \mu i v\}, \quad \mu = \pm 1. \tag{6.6}$$

Since $d\rho(X)$ commutes with J, we have, for $v, w \in V_{\mathbb{C}}$,

$$\begin{aligned} \langle X, [Jv, w]\rangle &= \langle d\rho(X)Jv, w\rangle \\ &= \langle J d\rho(X)v, w\rangle \\ &= -\langle d\rho(X)v, Jw\rangle \\ &= -\langle X, [v, Jw]\rangle, \end{aligned} \tag{6.7}$$

and hence

$$\langle X, [Jv, Jw] \rangle = -\langle X, [v, J^2 w] \rangle \\ = \langle X, [v, w] \rangle. \tag{6.8}$$

Meanwhile, for each $\mu = \pm 1$,

$$v, w \in V_\mu \Longrightarrow [Jv, Jw] = -[v, w], \tag{6.9}$$

so

$$v, w \in V_\mu \Longrightarrow [v, w] \perp X, \quad \forall\, X \in \mathfrak{su}(3). \tag{6.10}$$

More precisely, we can show that

$$v, w \in V_\mu \Longrightarrow [v, w] \in V_{-\mu}. \tag{6.11}$$

This can be seen from the root space decomposition, established in §5. With $\mathfrak{g}_{\mathbb{C}}$ denoting the complexification of $\mathrm{Der}(\mathbb{O})$, we have

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_{-1}, \\ \mathfrak{g}_0 = \mathfrak{su}(3)_{\mathbb{C}}, \quad \mathfrak{g}_1 = V_1, \quad \mathfrak{g}_{-1} = V_{-1}, \tag{6.12}$$

and

$$\mathfrak{g}_0 = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{j \neq k} \mathfrak{g}_{\omega_{jk}}, \quad \mathfrak{g}_1 = \bigoplus_j \mathfrak{g}_{\lambda_j}, \quad \mathfrak{g}_{-1} = \bigoplus_j \mathfrak{g}_{-\lambda_j}, \tag{6.13}$$

with $\{\omega_{jk}\}$ and $\{\pm\lambda_j\}$ as in (5.10), (5.15). It follows from (5.4) that

$$[\mathfrak{g}_j, \mathfrak{g}_k] \subset \mathfrak{g}_\ell, \quad \ell = j + k \bmod 3. \tag{6.14}$$

In particular,

$$[\mathfrak{g}_1, \mathfrak{g}_{-1}] \subset \mathfrak{g}_0, \tag{6.15}$$

so this bracket action is completely determined by (6.4). It remains to analyze

$$[\mathfrak{g}_1, \mathfrak{g}_1] \to \mathfrak{g}_{-1}, \quad \text{and} \quad [\mathfrak{g}_{-1}, \mathfrak{g}_{-1}] \to \mathfrak{g}_1, \tag{6.16}$$

or equivalently

$$[V_1, V_1] \to V_{-1}, \quad [V_{-1}, V_{-1}] \to V_1, \tag{6.17}$$

with $V_{\pm 1}$ as in (6.5)–(6.6). The following observation is useful.

Lemma 6.1. *If (V, J) is a vector space with complex structure J, equipped with a Hermitian inner product $(\ ,\ )$, and $V_{\mathbb{C}} = V_1 \oplus V_{-1}$, as in (6.5)–(6.6), then there are natural $\mathbb{C}$-linear isomorphisms*

$$V_1' \approx V_{-1} \quad \textit{and} \quad V_{-1}' \approx V_1. \tag{6.18}$$

Proof. The inner product $(\ ,\ )$ on V extends to a $\mathbb{C}$-bilinear form on $V_{\mathbb{C}}$. If $u - iJu \in V_1$ and $v + iJv \in V_{-1}$ (with $u, v \in V$), then

$$\begin{aligned}(u - iJu, v + iJv) &= (u, v) - i(Ju, v) + i(u, Jv) + (Ju, Jv) \\ &= 2(u, v),\end{aligned} \tag{6.19}$$

so the left side yields a $\mathbb{C}$-linear dual pairing of V_1 and V_{-1}. Note that $(i(u - iJu), v + iJv) = 2(Ju, v) = 2i(u, v)$ and $(u - iJu, i(v + iJv)) = -2(u, Jv) = 2i(u, v)$.

It follows that the bilinear maps in (6.17) yield tri-linear maps

$$\varphi : V_1 \times V_1 \times V_1 \to \mathbb{C}, \quad \psi : V_{-1} \times V_{-1} \times V_{-1} \to \mathbb{C}, \tag{6.20}$$

via

$$\varphi(u, v, w) = ([u, v], w), \quad u, v, w \in V_1, \tag{6.21}$$

and analogously for ψ. Note that

$$\begin{aligned}([u, v], w) &= (\operatorname{ad} u(v), w) \\ &= -(v, \operatorname{ad} u(w)) \\ &= -(v, [u, w]) \\ &= -([u, w], v),\end{aligned}$$

so φ is anti-symmetric in its arguments. On the other hand,

$$\dim_{\mathbb{C}} V_1 = 3 \Longrightarrow \Lambda^3_{\mathbb{C}} V_1 \approx \mathbb{C}, \tag{6.22}$$

so φ is uniquely determined, up to a scalar multiple, by the anti-symmetry property. Let us note that φ in (6.20) is not zero, i.e., the bracket $[V_1, V_1] \hookrightarrow V_{-1}$ is not identically zero. In fact, for example, $[\mathfrak{g}_{\lambda_1}, \mathfrak{g}_{\lambda_3}]$ has nonzero image in $\mathfrak{g}_{-\lambda_2}$ (cf. Proposition 35.6 of [T2]).

A. From G_2 to E_8

The complexification of Der($\mathbb{O}$), analyzed in §6, is the first of 5 exceptional complex simple Lie algebras, introduced by Killing and Cartan, denoted $\mathfrak{G}_2, \mathfrak{F}_4, \mathfrak{E}_6, \mathfrak{E}_7$, and $\mathfrak{E}_8$. We describe a uniform construction of $\mathfrak{G}_2$ and $\mathfrak{E}_8$, due to Freudenthal. In each case, the complex Lie algebra has a $\mathbb{Z}/(3)$ grading:

$$\mathfrak{g} = \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1, \quad \mathbb{Z}/(3) = \{-1, 0, 1\}. \tag{A.1}$$

We will have $[\mathfrak{g}_j, \mathfrak{g}_k] \subset \mathfrak{g}_{j+k}$, with $j+k$ computed mod 3. In each case, the complex Lie algebra $\mathfrak{g}_0$ has a representation ρ on a complex vector space, with contragredient representation ρ' on V'. We set

$$\mathfrak{g}_1 = V, \quad \mathfrak{g}_{-1} = V', \tag{A.2}$$

and define the actions $[\mathfrak{g}_0, \mathfrak{g}_j] \to \mathfrak{g}_j$ via these representations. In the cases $\mathfrak{g} = \mathfrak{G}_2$ or $\mathfrak{E}_8$, we take respectively

$$\mathfrak{g}_0 = s\ell(3, \mathbb{C}), \quad \mathfrak{g}_0 = s\ell(9, \mathbb{C}), \tag{A.3}$$

and, respectively,

$$V = \mathbb{C}^3 \ \text{ and } \ V = \Lambda^3\mathbb{C}^9. \tag{A.4}$$

There is a natural representation ρ of $\mathfrak{g}_0$ on V in each case. In the first case, we have $\Lambda^3 V = \Lambda^3\mathbb{C}^3 \approx \mathbb{C}$, via an invariant complex volume element, and in the second case $\Lambda^3 V \to \Lambda^9\mathbb{C}^9 \approx \mathbb{C}$. Thus we have natural bilinear maps

$$V \times V \longrightarrow V', \quad V' \times V' \longrightarrow V, \tag{A.5}$$

which are anti-symmetric. These define Lie brackets

$$[\mathfrak{g}_1, \mathfrak{g}_1] \to \mathfrak{g}_{-1}, \quad [\mathfrak{g}_{-1}, \mathfrak{g}_{-1}] \to \mathfrak{g}_1. \tag{A.6}$$

It remains to specify

$$[\mathfrak{g}_1, \mathfrak{g}_{-1}] \to \mathfrak{g}_0. \tag{A.7}$$

This is done as follows. Given $v \in V, v' \in V'$, we define $[v, v'] \in \mathfrak{g}_0$ by

$$-B(\lambda, [v, v']) = \langle \rho(\lambda)v, v' \rangle, \quad \lambda \in \mathfrak{g}_0, \tag{A.8}$$

where B is the Killing form on the simple Lie algebra $\mathfrak{g}_0$.

In this fashion, the Lie algebras are constructed. For $\mathfrak{G}_2$, the construction outlined here is consistent with the analysis of the complexification of Der($\mathbb{O}$) done in §6. For $\mathfrak{E}_8$, one needs to verify that the "products" defined above satisfy the Jacobi identity. For details on this, and the analysis of the root system for $\mathfrak{E}_8$, see [Ad], [SV].

References

[Ad] J. Adams, Lectures on Exceptional Lie Groups, Univ. Chicago Press, 1996.

[Art] M. Artin, Algebra, Prentice-Hall, Englewood Cliffs, NJ, 1991.

[B] J. Baez, The octonions, Bull. AMS 39 (2002), 145–205.

[Br1] R. Bryant, Submanifolds and special structures on the octonions, J. Diff. Geom. 17 (1982), 185–232.

[Br2] R. Bryant, On the geometry of almost-complex 6-manifolds, Asian J. Math. 10 (2006), 561–605.

[H] F. R. Harvey, Spinors and Calibrations, Academic Press, New York, 1990.

[LM] H. B. Lawson and M. L. Michelson, Spin Geometry, Princeton Univ. Press, Princeton NJ, 1989.

[Mcc] K. McCrimmon, Jordan algebras and their applications, Bull. AMS 84 (1978), 612–627.

[Por] I. Porteous, Clifford Algebras and Classical Groups, Cambridge Univ. Press, 1995.

[P] C. Procesi, Lie Groups – an Approach through Invariants and Representations, Springer, New York, 2007.

[Si] B. Simon, Representations of finite and compact groups, AMS, Providence, RI, 1996.

[SV] T. Springer and F. Veldkamp, Octonions, Jordan Algebras, and Exceptional Groups, Springer, Berlin, 2000.

[T1] M. Taylor, Introduction to Analysis in Several Variables (Advanced Calculus), Lecture Notes, available at
http://www.unc.edu/math/Faculty/met/math521.html

[T2] M. Taylor, Lectures on Lie Groups, Lecture Notes, available at
http://www.unc.edu/math/Faculty/met/lieg.html

[T3] M. Taylor, Partial Differential Equations, Vols. 1–3, Springer, New York, 1996 (2nd ed., 2011).

[T4] M. Taylor, Introduction to Differential Equations, Amer. Math. Soc., Providence RI, 2011.

[T5] M. Taylor, Differential Geometry, Lecture Notes, available at
http://www.unc.edu/math/Faculty/met/diffg.html

[T6] M. Taylor, Linear Algebra, Lecture Notes, available at
http://www.unc.edu/math/Faculty/met/linalg.html