Hamilton's Quaternions

T. Y. Lam
University of California, Berkeley, Ca 94720

Contents

§1. A Spark Flashed Forth
§2. Great Expectations
§3. Matrix Models
§4. Related Mathematical Discoveries
§5. The Rotation Groups SO(3) and SO(4)
§6. Finite Groups of Quaternions
§7. Exponentials, De Moivre's Formula, and Other Information

Bibliography

§1. A Spark Flashed Forth

From a course in modern algebra, we learned that a division ring is a ring K with identity $1 \neq 0$ in which every nonzero element has an inverse. These rings deserve our close attention because in some sense they are the most "perfect" algebraic systems: in them, we can add, subtract, multiply, and divide (by nonzero elements), and with respect to these four basic operations, all of the usual rules of arithmetic hold — except possibly the commutativity of multiplication. If the commutative law for multiplication does hold in a division ring K, we say that K is a *field*. In other words, a division ring K is a field iff its multiplicative group K^* is abelian.

The first example of a division ring that is not a field was discovered by the great Irish mathematician Sir William Rowan Hamilton (1805-1865). This is the division ring of real quaternions generated by four basis elements $1, i, j, k$ over the real numbers $\mathbb{R}$, with the famous Hamiltonian relations

$$i^2 = j^2 = k^2 = ijk = -1. \tag{1.1}$$

Throughout this article, we shall denote this division ring by $\mathbb{H}$. It is easy to check from (1.1) that

$$ij = k = -ji, \quad jk = i = -kj, \quad \text{and} \quad ki = j = -ik. \tag{1.2}$$

In particular, the elements i, j, k pairwise anticommute. Using the rules (1.2), we can easily multiply out any pair of quaternions $a + bi + cj + dk$ and $a' + b'i + c'j + d'k$ by using the two distributive laws.

Hamilton's original motivation (coming from physics) was to find an algebraic formalism for the points (x_1, x_2, x_3) in 3-space, in generalization of the formalism of the complex numbers $\mathbb{C}$ as pairs of real numbers. His initial desire was thus to find a 3-dimensional "hypercomplex system" with all the right properties (e.g. containing $\mathbb{C}$, and having $x_1^2 + x_2^2 + x_3^2$ as a multiplicative norm function). Hamilton worked off and on without success on this problem in the period 1830-1843. Two decades later, reminiscing on this frustrating experience, he was to write in one of his letters to his son Archibald H. Hamilton:

> "*Every morning ..., on my coming down to breakfast, your (then) little brother William Edwin, and yourself, used to ask me, 'Well, Papa, can you multiply triplets?'. Whereto I was always obliged to reply, with a sad shake of the head: 'No, I can only add and subtract them'.*"

Of course, nowadays, any student of modern algebra would be able to see that Hamilton's effort was doomed from the start on several grounds. First, the system Hamilton tried to find was to contain properly the complex field $\mathbb{C}$. By the transitivity formula for dimensions, we know that such a system can only have even real dimension. Second, the existence of a composition formula for sums of three squares would quickly contradict the fact that such sums are not closed under multiplication in $\mathbb{Z}$ and in $\mathbb{Q}$. For a detailed discussion on the history of this problem, see K. O. May's article [Ma].

The breakthrough finally came on Monday, October 16, 1843. On that day, as Hamilton took a walk with his wife along the Royal Canal in Dublin, on way to a council meeting of the Royal Irish Academy, a sudden flash of genius led him to the *four*-dimensional system $\mathbb{H}$ with basis elements 1, i, j, k multiplied according to the laws (1.1) (and their consequences (1.2)). In Hamilton's own words (quoted from [Kl: p. 779]):

> "*I then and there felt the galvanic circuit of thought closed, and the sparks which fell from it were the fundamental equations between i, j and k, exactly such as I have used them ever since.*"

With Irish exuberance (borrowing a phrase of E.T. Bell), Hamilton took out his pocket knife, and carved his fundamental equations on a stone of the Brougham Bridge. The equations (1.1) and (1.2) have since appeared in countless books in mathematics, history of mathematics, as well as in many items commemorating the life and work of Sir William Hamilton, including a series of Irish postal stamps issued in 1983, one and a half century after Hamilton's discovery of the quaternions.

To see that $\mathbb{H}$ is a division ring, note that every quaternion $q = a + bi + cj + dk \in \mathbb{H}$ has a "norm"

$$\mathrm{N}(q) = a^2 + b^2 + c^2 + d^2 \in \mathbb{R}, \tag{1.3}$$

and a conjugate $\overline{q} = a - bi - cj - dk \in \mathbb{H}$. A quick calculation shows that $\mathrm{N}(q) = q\,\overline{q} = \overline{q}\,q$, so if $q \neq 0$, we have $\mathrm{N}(q) \neq 0$, and q^{-1} is given by $\overline{q}/\mathrm{N}(q)$. However, $\mathbb{H}$ is not a field since $ij \neq ji$ (and $jk \neq kj$, $ki \neq ik$). Thus, $\mathbb{H}$ gave historically the very first example of a *noncommutative* division ring.

Of Hamilton's monumental discovery of the quaternions, J.J. Sylvester wrote the following thoughtful passage:

> "*In Quaternions the example has been given of Algebra released from the yoke of the commutative principle of multiplication — an emancipation somewhat akin to Lobachevsky's of Geometry from Euclid's noted empirical axiom.*"

Sylvester's referral to Lobachevsky's noneuclidean geometry provided an interesting analogy for Hamilton's invention of the quaternions. This analogy becomes even somewhat uncanny if we also bring Gauss into the picture. It is well reported in history books in mathematics that, while Nikolai Ivanovich Lobachevsky discovered noneuclidean geometry in 1826 and published his findings between 1829-1837, and János Bolyai made similar discoveries in the period 1825-1833, Carl Friedrich Gauss (1777-1855) seemed to have been aware of the independence of Euclid's parallel postulate as early as 1799, and by 1813 he had arrived at the rudiments of a logically consistent "anti-euclidean geometry". However, always cautious about how his work would be received, and fearful of the ridicule of his revolutionary ideas by his contemporaries, Gauss never published his findings. As for the quaternions, a rather similar situation prevailed, although it is not as widely reported. There is no doubt that Hamilton deserved full credit for his discovery of the quaternions in 1843, but again Gauss had already a good "sighting" of the quaternion system around 1819-1823. In a short note [Ga] from his diary, in working with transformations of spaces, Gauss came up with a way of composing real quadruples. Given (a, b, c, d) and $(\alpha, \beta, \gamma, \delta)$ in $\mathbb{R}^4$, he associated the composite quadruple (A, B, C, D) given by

$$(1.4) \quad (a\alpha - b\beta - c\gamma - d\delta,\ a\beta + b\alpha - c\delta + d\gamma,\ a\gamma + b\delta + c\alpha - d\beta,\ a\delta - b\gamma + c\beta + d\alpha).$$

To see that this is no less than the composition of quaternions, one need only transpose the second and third coordinates in Gauss's notation. That is, if we express the quadruples (a, b, c, d) and $(\alpha, \beta, \gamma, \delta)$ as $a + bj + ci + dk$ and $\alpha + \beta\,j + \gamma\,i + \delta\,k$ in Hamilton's notation, Gauss's composition amounts *exactly* to the multiplication of two arbitrary quaternions! Gauss wrote, in an almost matter-of-fact fashion:

> "*Wir bezeichnen allgemein die Combination* a, b, c, d *durch* (a, b, c, d) *und schreiben*
> $$(a, b, c, d)\,(\alpha, \beta, \gamma, \delta) = (A, B, C, D).$$
> *Es ist also* $(a, b, c, d)\,(\alpha, \beta, \gamma, \delta)$ *nicht mit* $(\alpha, \beta, \gamma, \delta)\,(a, b, c, d)$ *zu verwechseln. ... Ferner bezeichne man die Combination* (a, b, c, d) *durch einen Buchstaben, z. B.* g, *und dann die Combination* $(a, -b, -c, -d)$ *durch* g'. *Es ist also*
> $$gg' = g'g = (aa + bb + cc + dd,\ 0, 0, 0)."$$

Although quadruples of the form (A, B, C, D) in (1.4) had certainly appeared earlier, notably in Euler's 4-square identity, Gauss took the unmistakable step of viewing (A, B, C, D) in (1.4) as a *composition* of (a, b, c, d) and $(\alpha, \beta, \gamma, \delta)$, and explicitly noted the failure of the commutative law. And of course, the last remark in the quotation above was tantamount to the realization that, under such a composition, the quadruples (a, b, c, d)'s form a division algebra over the real numbers. Although Gauss did not introduce the i, j, k notation as Hamilton did, it would be difficult to deny, on the basis of what Gauss had written (so succinctly!) in his diary, that he had essentially "come up" with the quaternion system as early as 1819. But again, Gauss never published anything about this discovery, so the "official" birth of the quaternions was to wait another quarter of a century, until Hamilton took his famous walk along the Royal Canal on that eventful October day in 1843.

§2. Great Expectations[1]

Hamilton noted his discovery of the quaternions on the Council Books of the Royal Academy on the very same October day, and obtained leave to read a paper on quaternions to a general meeting of the Academy on November 13, 1843. After this, he was to devote much of the remaining 22 years of his life to the study of quaternions and their applications. Besides the many papers he published on quaternions (one hundred and nine in all, according to Crowe [Cr: p. 41]), his main writings in the subject survived in his two books [H_1] and [H_2] (the latter published posthumously, first in 1866).

Hamilton had great expectations for the wide applicability of his quaternions to mathematics, physics, and astronomy. At one point, he even went so far as to compare his discovery to that of Newton on fluxions in the 17th century (see [Cr: p. 30]). In later years, his zeal in converting his contemporaries to the quaternionic thinking virtually verged on obsession. As it turned out, the controversy over the usefulness of quaternions was to last for several generations after Hamilton's death. Supporters of Hamilton touted the quaternions as one of the greatest inventions in the 19th century, "fitted to be of the greatest use in all parts of science", whereas antagonists of Hamilton's theory attacked the quaternions as "an unmixed evil to those who have touched them in anyway".[2] The only mathematical theory I can think of that has evoked among its practitioners such intense controversy and vigorous debate is perhaps, again, non-Euclidean geometry!

As is always the case, the truth is somewhere in between. In retrospect, Hamilton's original expectations were largely overblown, and indeed, a considerable part of the subsequent work on quaternions by his disciples and followers did not survive the time test.

[1] Our pilfering of the title of Charles Dickens's famous novel here is not purely accidental. Dickens, who lived from 1812 to 1870, was almost an exact contemporary of Hamilton: it seems remarkable that the life spans of these two great men differed only by a translation of about six years. *Great Expectations* was written in 1861 — four years before Hamilton's death. And, if the reader is curious as to what Charles Dickens was up to in 1843, the answer is that he wrote *A Christmas Carol*, which has also endured the times very well, and remained a delightful part of the Christmas tradition in countries around the world.

[2] For more complete original quotations and attributions to their sources, we refer the readers to [KR_1], [Al_2], and [Cr: Ch. 2].

Nevertheless, ever since that October day in 1843, the quaternions have assumed their permanent place in mathematics as the first noncommutative system to be studied. Their bold introduction affirmed the supreme freedom of mathematical thought, and opened the floodgate to the fruitful study of algebraic systems satisfying different sets of axioms, not all of which reflect the familiar properties of the real (or the complex) numbers. For instance, much of the work in hypercomplex systems (the theory of algebras) in late 19th century and the early part of the 20th century was rightfully a direct descendant of Hamilton's theory of quaternions. (For a survey on this line of work, see [Hap].) Theoretically, to every concept making sense for the reals and the complexes, there is a quaternionic analogue. Whether such a quaternionic analogue is truly worthwhile of study can only be seen in time. For instance, the efforts of Hamilton's followers in developing a quaternionic theory of functions (complete with differentiation, integration, and theorems of Gauss, Green and Stokes) were now largely forgotten. But in contemporary topology, there is certainly still a healthy interest in quaternionic manifolds. A recent search on the MathSciNet for papers with the word "quaternion" in their titles turned up 1972 entries. Quaternions did not become a main tool in physics as Hamilton had hoped. But late in the 20th century, there seemed to have been a revival of interest in the quaternions among some physicists, as they tried to formulate quaternionic generalizations of the postulates of quantum mechanics, quantum field theory, gravitational theory, and elementary particle symmetries. For some work in this area, see, for instance, [Ad].

§3. Matrix Models

The main purpose of this section is to discuss a matrix model of the quaternion algebra $\mathbb{H}$. This discussion will also help us to solidify our understanding of the quaternions before we return to look at other historical aspects pertaining to them.

Nowadays every college student in mathematics knows that matrices do not commute under multiplication, so the idea of a noncommutative system comes as no surprise. But matrices were not yet in the mathematical vocabulary in 1843. They were introduced only in 1855 by Arthur Cayley (1821-1895) who subsequently published his famous memoir [Ca] on the subject in 1858. For a detailed discussion on the invention of the theory of matrices, see [Kl: Ch. 33, §4].

From the viewpoint of matrices, the complex field $\mathbb{C}$ can be realized as a subalgebra of the matrix algebra $\mathbb{M}_2(\mathbb{R})$ as follows. We view $\mathbb{C}$ as a vector space over $\mathbb{R}$ with basis $\{1, -i\}$, and "identify" a complex number $a+bi$ $(a, b \in \mathbb{R})$ with the left multiplication map by $a+bi$ on $\mathbb{C}$ (the Cayley representation). With respect to the basis $\{1, -i\}$, this linear map has a matrix $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, so $\mathbb{C}$ may be identified with the algebra of all such matrices in $\mathbb{M}_2(\mathbb{R})$.

We can use a similar procedure to realize the quaternions as a certain real subalgebra of 2×2 *complex* matrices. To this end, let us identify the complex numbers $\mathbb{C}$ with $\mathbb{R} \oplus \mathbb{R}\, i$ in $\mathbb{H}$, and work with $\mathbb{H}$ as a 2-dimensional right vector space over $\mathbb{C}$ with basis $\{1, -j\}$. By the Cayley representation again, every quaternion q induces a $\mathbb{C}$-linear left multiplication map on $\mathbb{H}$, which can be expressed by a 2×2 complex matrix $L(q) \in \mathbb{M}_2(\mathbb{C})$. For instance, from $i \cdot 1 = i$ and $i \cdot (-j) = (-j)(-i)$, we have

$$(3.1) \qquad L(i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \text{and similarly, } L(j) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad L(k) = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}.$$

From these, we compute easily that, for a quaternion $q = a + bi + cj + dk$,

$$(3.2) \qquad L(q) = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}, \text{ where } \alpha = a + bi, \text{ and } \beta = c + di.$$

Since the Cayley representation is faithful ($L(v) = 0 \Rightarrow 0 = L(v)(1) = v$), we obtain an $\mathbb{R}$-algebra isomorphism

$$(3.3) \qquad \mathbb{H} \cong \left\{ \begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix} : \alpha, \beta \in \mathbb{C} \right\} \subseteq \mathbb{M}_2(\mathbb{C}),$$

which is given in many textbooks to describe (and sometimes to define) the division ring of the real quaternions. Note that this "Cayley model" of $\mathbb{H}$ makes it unnecessary to verify the associative law for multiplication (since matrix multiplication is always associative). Also, $L(i)$, $L(j)$ and $L(k)$ in (3.1) are now just three concrete anticommuting matrices over $\mathbb{C}$ (with squares $-I_2$ and product I_2): they no longer have the aura of mystery of the quaternions i, j and k.

The matrix model for $\mathbb{H}$ in (3.3) has several wonderful features. First, in this model, quaternionic conjugation $q \mapsto \overline{q}$ corresponds to the "conjugate transpose" operation * on $\mathbb{M}_2(\mathbb{C})$. More precisely,

$$(3.4) \qquad L(\overline{q}) = \begin{pmatrix} a-bi & -(c+di) \\ c-di & a+bi \end{pmatrix} = \begin{pmatrix} a+bi & c+di \\ -c+di & a-bi \end{pmatrix}^* = L(q)^*.$$

Secondly, the quaternionic norm corresponds to the determinant on $\mathbb{M}_2(\mathbb{C})$; namely,

$$(3.5) \qquad \mathrm{N}(q) = a^2 + b^2 + c^2 + d^2 = \det(L(q)).$$

The fact that "det" is multiplicative implies the same for the quaternion norm,[3] and this gives essentially the four-square identity of Euler. Thirdly, it is easy to verify that the matrices $L(1)$, $L(i)$, $L(j)$, $L(k)$ are linearly independent not only over $\mathbb{R}$, but also over $\mathbb{C}$. Thus, they form a $\mathbb{C}$-basis for $\mathbb{M}_2(\mathbb{C})$. From this, we get a $\mathbb{C}$-algebra isomorphism

$$(3.6) \qquad \mathbb{C} \otimes_{\mathbb{R}} \mathbb{H} \cong \mathbb{M}_2(\mathbb{C}).$$

[3] Of course, we don't really have to use the matrix model to prove the multiplicativity of N. Since conjugation on $\mathbb{H}$ is easily seen to be an involution, we have $\mathrm{N}(pq) = (pq)\,\overline{pq} = pq \cdot \overline{q}\,\overline{p} = p\,\mathrm{N}(q)\,\overline{p} = \mathrm{N}(p)\,\mathrm{N}(q)$.

Here, $\mathbb{C} \otimes_{\mathbb{R}} \mathbb{H}$ denotes the "scalar extension" of the quaternion algebra from $\mathbb{R}$ to $\mathbb{C}$. In other words, it is the quaternion algebra formed over the complexes; in Hamilton's work, this is called the *algebra of biquaternions.*[4] As a $\mathbb{C}$-algebra, it turns out to be isomorphic to $\mathbb{M}_2(\mathbb{C})$. (In modern parlance, we say that $\mathbb{H}$ "splits" over $\mathbb{C}$.)

It has been said that most of the applications of quaternions to physics are actually applications of biquaternions. There is perhaps considerable truth in this statement. Note that the matrices $L(i)$, $L(j)$, $L(k)$ in (3.1) are *unitary* matrices (that is, complex matrices U with $UU^* = I$). If we multiply them by the scalar $-i$, we get the following three *Hermitian* matrices (that is, complex matrices H such that $H^* = H$):

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{3.7}$$

These are the famous *Pauli spin matrices,* which are used by physicists in the study of the quantum mechanical motion of a spinning electron. The following relations among the Pauli spin matrices are familiar to all quantum physicists (see, e.g. [Me: p. 545]):

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I, \qquad \sigma_x\sigma_y\sigma_z = iI,$$
$$\sigma_x\sigma_y = -\sigma_y\sigma_x = i\sigma_z, \quad \sigma_y\sigma_z = -\sigma_z\sigma_y = i\sigma_x, \quad \sigma_z\sigma_x = -\sigma_x\sigma_z = i\sigma_y.$$

These are to be compared with the original Hamiltonian relations (1.1) and (1.2).

In the above, we have viewed $\mathbb{H}$ as a (right) $\mathbb{C}$-space so as to keep the matrices at size 2×2. If we simply view $\mathbb{H}$ as an $\mathbb{R}$-vector space with basis $\{1, i, j, k\}$, the Cayley representation would have given the 4×4 real matrix model:

$$\mathbb{H} \cong \left\{ \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} : \ a, b, c, d \in \mathbb{R} \right\} \subseteq \mathbb{M}_4(\mathbb{R}), \tag{3.8}$$

with the above matrix corresponding to the quaternion $q = a + bi + cj + dk$. Here, the quaternionic conjugate $q \mapsto \overline{q}$ corresponds to the ordinary transpose on $\mathbb{M}_4(\mathbb{R})$, while the determinant of the matrix above is now $\mathrm{N}(q)^2$.

§4. Related Mathematical Discoveries

To place Hamilton's discovery of the quaternions in the proper perspective of 19th (and 20th) century mathematics, it would be incumbent on us to describe several other important mathematical events which took place after 1843, and which turned out to have had a direct impact on the role of quaternions in mathematics (and physics). We itemize several of these relevant developments in the following.

[4] The word "biquaternion", however, may have a different meaning in other contexts. For instance, William Kingdon Clifford (1845-1879) used the same word for the elements in one of his hypercomplex number systems. In this paper, we have, unfortunately, no space to discuss Clifford's contributions.

• **Cayley Algebra.** Only a few months after Hamilton's discovery of the quaternions, his friend John T. Graves invented the 8-dimensional hypercomplex system $\mathbb{O}$, now known as the *algebra of octonions* (or *octaves*). Like $\mathbb{H}$, $\mathbb{O}$ is a real division algebra (every nonzero element has an inverse), and is not commutative. But unlike $\mathbb{H}$, $\mathbb{O}$ is no longer associative! Thus, a second bondage to the existing fundamental laws of algebra was quickly broken following on the heels of the first.

Although Graves made his discovery in 1843-44, his results were not published until 1848. In the mean time, Arthur Cayley made the same discovery and published it in 1845. From then on, the octonions also became known as the *Cayley numbers.* The quickest construction of the Graves-Cayley algebra results from defining the following multiplication on the additive group $\mathbb{O} := \mathbb{H} \times \mathbb{H}$:

$$(p_1,\, q_1)\,(p_2,\, q_2) = (p_1 p_2 - \overline{q}_2 q_1,\; q_2 p_1 + q_1 \overline{p}_2), \quad \text{where } (p_i,\, q_i) \in \mathbb{H} \times \mathbb{H}. \tag{4.1}$$

It is straightforward to check that, with this multiplication, $\mathbb{O}$ is indeed an $\mathbb{R}$-algebra, except for the lack of associativity, which can be seen from, say:

$$[(i,0)\,(j,0)]\,(0,k) = (k,0)\,(0,k) = (0,-1), \quad (i,0)\,[(j,0)\,(0,k)] = (i,0)\,(0,-i) = (0,1).$$

The definition of the multiplication in (4.1) is directly inspired by the following view of quaternion multiplications. Every quaternion

$$q = a + bi + cj + dk = (a+bi) + (c+di)j \in \mathbb{H}$$

can be "identified" with the pair $(\alpha,\, \beta) \in \mathbb{C}^2$, where $\alpha = a + bi$ and $\beta = c + di$ constitute the first row of the representing matrix $L(q)$ in (3.2). Taking the product of two such matrices and reading off the first row, we see that the multiplication of the quaternions as pairs of complex numbers is given by

$$(\alpha_1,\, \beta_1)\,(\alpha_2,\, \beta_2) = (\alpha_1 \alpha_2 - \beta_1 \overline{\beta}_2,\; \alpha_1 \beta_2 + \beta_1 \overline{\alpha}_2). \tag{4.2}$$

If we rewrite the RHS as $(\alpha_1\alpha_2 - \overline{\beta}_2 \beta_1,\; \beta_2 \alpha_1 + \beta_1 \overline{\alpha}_2)$ (here the α's and the β's commute), we get exactly the formula (4.1). The point is, however, that (4.1) *has to be carefully adapted from* (4.2), since the p's and the q's no longer commute. If one makes a wrong choice in the orders of the factors in (4.1), one will not get the Cayley algebra $\mathbb{O}$.

Every Cayley number $x = (p,\, q) \in \mathbb{O}$ has a norm $\mathrm{N}(x) = \mathrm{N}(p) + \mathrm{N}(q)$, and a conjugate $\overline{x} = (\overline{p},\, -q)$. One has again $\mathrm{N}(x) = x\,\overline{x} = \overline{x}\,x$, so if $x \neq 0$, it has an inverse $\overline{x}/\mathrm{N}(x)$, just as in the case of quaternions.[5] This shows that $\mathbb{O}$ is a (nonassociative) division algebra. Unfortunately, the "duplication" process *cannot* be used on the Cayley numbers again to get a 16-dimensional division algebra. Our luck runs out at dimension 8!

[5] As the reader can perhaps guess, $\mathrm{N}(xy) = \mathrm{N}(x)\mathrm{N}(y)$ also holds for Cayley numbers, so that we get an 8-square identity. However, the proof for multiplicativity given for quaternions in Footnote 3 no longer works, since the associative law does not hold here! We invite the reader to supply a correct proof for the case of Cayley numbers.

Although $\mathbb{O}$ is not associative, it can be shown to satisfy the following two weaker versions of the associative law:

$$x(xy) = (xx)y, \quad (xy)y = x(yy), \tag{4.3}$$

for any $x, y \in \mathbb{O}$. An algebra satisfying these two laws is called an *alternative algebra.* Along with such algebras, various other kinds of nonassociative algebras have been studied, notably, the Lie algebras, and the Jordan algebras. Even within the framework of associative division algebras, one can try to get new objects by weakening some other axioms. For instance, if one keeps only one distributive law $x(y+z) = xy+xz$ and discards the other ($(y+z)x = yx+zx$), one gets a system called a *near field.* The investigations of these diverse systems were all made possible by Hamilton's pioneering discovery of the quaternions.

For more information on Cayley numbers, see [Ba], [Bl], and [W].

• **Frobenius' Theorem.** In a *tour de force* in 1877, F.G. Frobenius (1849-1917) proved that, up to isomorphisms, the only finite-dimensional real (associative) division algebras are $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$ [Fr]. The same theorem was also proved independently (if just a little bit later) by C.S. Peirce, the son of Benjamin Peirce, who published the result in an Appendix to his father's long memoir [Pe] on associative algebras in 1881. This is a truly marvelous theorem as far as the quaternions are concerned, since it showed that, beyond $\mathbb{R}$ and $\mathbb{C}$, $\mathbb{H}$ is in fact *the only* finite-dimensional real division algebra that can exist. This wonderful affirmation of the "cosmic" role of the quaternions in mathematics would seem to have at least partly fulfilled Hamilton's "great expectations", although unfortunately, it did not seem to have mattered very much in the great quaternion debate still raging on in the last decades of the 19th century. The anti-quaternionists were bent on arguing that quaternions were useless, whether they were shown to have any cosmic significance or not.

Nowadays, mathematicians certainly have a greater appreciation for uniqueness theorems. Frobenius' Theorem was perhaps one of the earliest classification results of its kind proved in algebra. This theorem has been imitated many times over in the 20th century in different contexts of classification. Most notable examples are: Zorn's Theorem (1933) on alternative real division algebras, the Gel'fand-Mazur Theorem (1938) for commutative Banach division algebras, and Hopf's Theorem (1940) on finite-dimensional commutative (but not necessarily associative) real division algebras. Hopf, a pioneer in applying topology to algebra, was the first one to see the possibility of proving results on division algebras from the geometry of spheres and projective spaces. His program in this direction was ultimately completed by Milnor and Kervaire. who proved independently in 1958 that *the unit sphere* S^{n-1} *is parallelizable*[6] *only if* $n = 1, 2, 4, 8$. An algebraic implication of this topological theorem is that the only possible finite dimensions of a (possibly noncommutative and nonassociative) real division algebra are $1, 2, 4$ or 8. This is a powerful expansion of Frobenius's 1877 theorem; however, up to this time, no purely

[6]A differentiable manifold is said to be *parallelizable* if it has a trivial tangent bundle.

algebraic proof of it is known. (This could very well be the most painful instance of the "topological thorn in the flesh of algebra" that Koecher and Remmert spoke about in the introduction to their article [KR_2: p. 223].) For an exposition on the topological methods used in the classification of division algebras, the reader may consult Hirzebruch's article [Hi]. For a modern formulation of a proof of Frobenius's Theorem, see [PS: (3.4)], or [La_2: (13.12)].

• **Vector Analysis.** The fast-paced progress in physics (especially in mechanics, electricity and magnetism) in the second half of the 19th century clearly called for a vectorial theory for the 3-space that could serve as a sound mathematical foundation for the study of the physical quantities. This was indeed very much on Hamilton's mind when he first set out to find a product operation for triplets. After Hamilton invented the quaternions $\mathbb{H}$, the space of "pure quaternions"

$$\mathbb{H}_0 := \mathbb{R}\, i \oplus \mathbb{R}\, j \oplus \mathbb{R}\, k \tag{4.4}$$

would seem to be a natural candidate for a model for the 3-space. However, for any $\mathbf{u} = u_1 i + u_2 j + u_3 k \in \mathbb{H}_0$, one has $\mathbf{u}^2 = -(u_1^2 + u_2^2 + u_3^2)$, which is a negative real number. It seemed counter-intuitive to physicists that the square of a vector should be a negative quantity, and this no doubt hindered the wide acceptance of the quaternion system by the physicists.

Only one year after Hamilton's discovery, Hermann Grassmann (1809-77) published his "Ausdehnungslehre 1844", which purported to provide not only a firm foundation for the linear algebra of space, but in fact a whole new discipline of mathematics. In retrospect, Grassmann's book was truly a ground-breaking masterpiece, and one of the most brilliant mathematical works of the 19th century. But his "linear extension theory" (complete with a myriad of all kinds of vector products) was couched in such general and abstract terms, and so laden with metaphysical overtones, that it came off as almost totally incomprehensible to his contemporaries. Gauss politely declined to read "Ausdehnungslehre", and Möbius openly admitted that he had not managed to get through more than "a few sheets" (see [Cr: p. 80]). The book essentially met with complete silence, and sold so poorly that by 1864 its publisher saw fit to shred about 600 of its remaining copies! A second edition of the book met with a somewhat better reception, but the ultimate recognition of Grassmann's genius would have to wait until the 20th century.

Grassmann's work certainly did not succeed in providing physicists with the vectorial system they felt comfortable to use in physical theories. This task was accomplished around the 1880s by Josiah Willard Gibbs (1839-1903) and Oliver Heaviside (1850-1925), whose system of "modern vector analysis" relatively quickly gained acceptance and popularity. Central among the concepts in this theory are the *inner product* and the *cross product* of vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, which we shall denote by $\langle \mathbf{u}, \mathbf{v} \rangle$ and $\mathbf{u} \times \mathbf{v}$. The former is a number, and the latter is a vector. The cross product is sometimes also called the outer product or the vector product. Grassmann himself had certainly defined inner and outer products before; however, his outer product for two vectors was not a vector, but an oriented area.

To a modern reader, the connection between quaternions and vector analysis is easy to discern. In view of (4.4), we can identify $\mathbb{H}_0$ with $\mathbb{R}^3$ (calling its elements "vectors"), and express $\mathbb{H}$ as a direct sum $\mathbb{R} \oplus \mathbb{H}_0$. Thus, every quaternion q is uniquely a sum $u_0 + \mathbf{u}$, where $u_0 \in \mathbb{R}$, and $\mathbf{u} \in \mathbb{H}_0 = \mathbb{R}^3$. Hamilton called u_0 the "scalar part", and $\mathbf{u}$ the "vector part", of the quaternion q, and explicitly used the notations: $S.q = u_0$, and $V.q = \mathbf{u}$. (This may well have been the very first usage of the words "scalar" and "vector" in the mathematical literature.) For any two "vectors" $\mathbf{u},\ \mathbf{v} \in \mathbb{H}_0$, the quaternion product of $\mathbf{u}$ and $\mathbf{v}$ is easily computed to be

$$\mathbf{u}\,\mathbf{v} = -\langle \mathbf{u},\, \mathbf{v}\rangle + \mathbf{u} \times \mathbf{v}. \tag{4.5}$$

Thus, in Hamilton's notation, $S.\mathbf{u}\,\mathbf{v} = -\langle \mathbf{u}, \mathbf{v}\rangle$, and $V.\mathbf{u}\,\mathbf{v} = \mathbf{u} \times \mathbf{v}$. From these, we see that $\mathbf{u}\,\mathbf{v}$ is in general not a "vector"; in fact, it is a vector iff $\mathbf{u},\ \mathbf{v}$ are *perpendicular* in $\mathbb{R}^3$, in which case we have $\mathbf{u}\,\mathbf{v} = \mathbf{u} \times \mathbf{v}$. From (4.5), we also see that the inner and outer products for vectors can be quickly retrieved from the quaternion products, namely:

$$\langle \mathbf{u},\, \mathbf{v}\rangle = -(\mathbf{u}\,\mathbf{v} + \mathbf{v}\,\mathbf{u})/2, \quad \text{and} \quad \mathbf{u} \times \mathbf{v} = (\mathbf{u}\,\mathbf{v} - \mathbf{v}\,\mathbf{u})/2. \tag{4.6}$$

The advantages of the inner and outer products are clear and many. First and foremost, they are well grounded in physics. Second, the mixed triple product

$$\langle \mathbf{u} \times \mathbf{v},\, \mathbf{w}\rangle = \langle \mathbf{u},\, \mathbf{v} \times \mathbf{w}\rangle \tag{4.7}$$

corresponds to $\det(\mathbf{u}, \mathbf{v}, \mathbf{w})$, and thus detects linear dependence. Third, the cross product is a closed binary operation on $\mathbb{R}^3$. so we do get a multiplication for triplets! Under the cross product, the elements $\{i, j, k\}$ multiply with each other exactly like the quaternions, but now

$$i \times i = j \times j = k \times k = 0. \tag{4.8}$$

It can be checked that the cross product turns $\mathbb{R}^3$ into a *Lie algebra.* From the Cayley model (3.3), one can further show that this is isomorphic to the Lie algebra of all 2×2 skew-Hermitian matrices with trace zero over $\mathbb{C}$ (under the Lie bracket $[A, B] = (AB - BA)/2$).

Central to the Gibbs-Heaviside vector calculus were the three operators "grad", "div", and "curl". However, just like the inner and outer products, these operators were also more or less already "embedded" in Hamilton's quaternion system. Indeed, Hamilton was the first one to introduce (and to name) the "nabla" operator

$$\nabla = i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z}\,, \tag{4.9}$$

which, when applied to a differentiable scalar function $f(x, y, z)$, yields the gradient field $\operatorname{grad} f$. If we apply "nabla" instead to a vector field $\mathbf{F} = u(x, y, z)i + v(x, y, z)j + w(x, y, z)k$, then the use of quaternion multiplications leads to

$$\begin{aligned} \nabla\mathbf{F} &= -\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}\right) + \left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right) i + \left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) j + \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) k \\ &= -\operatorname{div}\mathbf{F} + \operatorname{curl}\mathbf{F}\,, \end{aligned}$$

so "div" and "curl" come into the picture simultaneously.[7] If we apply this formula to a gradient field $\mathbf{F} = \operatorname{grad} f$, then the i, j, k terms will all cancel out (because of the interchangeability of the order of partial differentiations), and we will be left with:

$$\nabla^2 f = -\left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}\right); \tag{4.10}$$

that is, we get the *negative* of the famous Laplace operator "$\triangle$" in potential theory.

On the basis of (4.5), (4.6), and the above discussions on "nabla", a liberal thinker today could very well say that vector algebra in $\mathbb{R}^3$ and the quaternions are "interchangeable systems". But they were definitely *not* the same in the eye of 19th century mathematicians, and it is a fact that Hamilton himself never explicitly introduced inner, outer products, or "div" and "curl". Gibbs had always maintained that these constructs were best studied on their own, and that it would not add anything to work with them as a part of the quaternionic structure. Heaviside held a similar view, priding himself on the fact that he never had to use a single quaternion in his own work. But ironically, it is perhaps still the following quotation from Heaviside [Cr: p. 192]) that best summed up the situation:

> "The invention of quaternions must be regarded as a most remarkable feat of human ingenuity. Vector analysis, without quaternions, could have been found by any mathematician by carefully examining the mechanics of the Cartesian mathematics; but to find out quaternions required a genius."

• **Composition of Sums of Squares.** We have pointed out before that the multiplicativity of the norm on the quaternions and on the Cayley numbers gave natural interpretations (and derivations) for the 4- and 8-square identities, just as the multiplicativity of complex number norms did for the 2-square identity. *How about n-square identities for other values of n?* Adolf Hurwitz (1859-1919) took up this question in 1888 and arrived at its complete solution [Hu], showing that, if there exists an identity

$$\left(x_1^2 + \cdots + x_n^2\right)\left(y_1^2 + \cdots + y_n^2\right) = z_1^2 + \cdots + z_n^2, \tag{4.11}$$

where $z_1, \ldots, z_n \in \mathbb{R}[\mathbf{x}, \mathbf{y}]$, then $n = 1, 2, 4, 8$! Later, Hurwitz and Radon considered the more general problem of finding, for a given n, the largest possible m for which there exists an identity

$$\left(x_1^2 + \cdots + x_m^2\right)\left(y_1^2 + \cdots + y_n^2\right) = z_1^2 + \cdots + z_n^2, \tag{4.12}$$

where $z_1, \ldots, z_n \in \mathbb{R}[\mathbf{x}, \mathbf{y}]$. The answer here is that $m = \rho(n)$, where ρ is the Hurwitz-Radon function, defined by

$$\rho(n) = 8a + 2^b \quad \text{for} \quad n = 2^{4a+b} n_0 \;\; (n_0 = \text{odd}, \;\; b \in \{0, 1, 2, 3\}). \tag{4.13}$$

[7] Note that, if we had used the cross products instead, then in view of (4.8), there would be no scalar term in this calculation, and we would have gotten the more familiar calculus formula $\nabla \times \mathbf{F} = \operatorname{curl} \mathbf{F}$.

By easy arithmetic, we have $\rho(n) \le n$, with equality iff $n = 1, 2, 4, 8$. Thus, the Hurwitz-Radon Theorem contains Hurwitz's original theorem in 1888.

The work of Hurwitz-Radon is of great significance to topologists, since the square identitities (4.12) for $m = \rho(n)$ can be used to construct readily $\rho(n) - 1$ linearly independent tangent vector fields on the unit sphere S^{n-1}. Much later (in 1962), using deep tools from topology, J. Frank Adams proved that this is indeed the best construction possible; that is, S^{n-1} *cannot admit* $\rho(n)$ *linearly independent tangent vector fields* [A: Thm. 1.1]. Adams's famous solution of the Vector Field Problem for Spheres implied, in particular, the aforementioned Milnor-Kervaire Theorem that the only parallelizable spheres are S^0, S^1, S^3, and S^7.

The study of sums-of-squares identities prompted by the existence of the quaternions and Cayley numbers have led to a lot of fruitful mathematics, both in algebra and in topology. However, the following original form of the sums-of-squares problem proposed by Hurwitz [Hu] in 1888 has remained unsolved to this date: *for a given integer* n*, what is the smallest* k *for which* $\left(x_1^2 + \cdots + x_n^2\right)\left(y_1^2 + \cdots + y_n^2\right)$ *is a sum of* k *squares in* $\mathbb{R}[\mathbf{x}, \mathbf{y}]$*?* For a thorough survey on this Hurwitz Problem and many other related problems, see Shapiro's monograph [Sh].

• **Fundamental Theorem of Algebra.** This important theorem says that $\mathbb{C}$ is an algebraically closed field; that is, any non-constant polynomial in $\mathbb{C}[x]$ has a zero in $\mathbb{C}$. Indeed, the main point of passing from $\mathbb{R}$ to $\mathbb{C}$ is exactly to get an "algebraic closure" of the reals. In this spirit, one may ask: *Is there an analogue of the Fundamental Theorem of Algebra for the quaternions?* To address this question properly, one must, however, first re-examine the notion of a "polynomial". The "usual" form of a polynomial over $\mathbb{H}$ is a (finite) expression of the kind $f(x) = \sum_i a_i x^i$, where $a_i \in \mathbb{H}$, and we can formally define the evaluation of f at $q \in \mathbb{H}$ to be $f(q) := \sum_i a_i q^i$. However, since $\mathbb{H}$ is not commutative, the above kind of polynomials is clearly not sufficiently general. For instance, one might legitimately consider a degree 2 expression of the shape $axbxc$ (where $a, b, c \in \mathbb{H}$), whose evaluation at $q \in \mathbb{H}$ would be $aqbqc$ (which *cannot* be reduced to $abcq^2$ due to the lack of commutativity).

To formulate a Fundamental Theorem of Algebra for $\mathbb{H}$, we shall, therefore, redefine a *monomial of degree* n to be an expression of the form

$$m(x) = a_0 x a_1 x \cdots a_{n-1} x a_n, \quad \text{where } a_i \in \mathbb{H}. \tag{4.13}$$

The evaluation $m(q)$ is then defined as $a_0 q a_1 q \cdots a_{n-1} q a_n$. A "polynomial" $f(x)$ is redefined to be a (finite) sum of monomials $m_i(x)$, and $f(q)$ is accordingly defined to be $\sum_i m_i(q)$. The "degree" of f is the highest degree of the monomials actually appearing in f. Notice that, if $\deg(f) = n$, the nth degree part of $f(x)$ is now a sum of degree n monomials, instead of a single monomial. (Of course, the same holds for other homogeneous parts of f.)

In 1944, building upon earlier work of Niven, Eilenberg and Niven [EN] proved the following beautiful result.

(4.14) Fundamental Theorem of Algebra for Quaternions. *Let $f(x)$ be a polynomial (in the generalized sense above) of degree $n \geq 1$ over $\mathbb{H}$, whose n-th degree part consists of a single monomial. Then the evaluation map $q \mapsto f(q)$ from $\mathbb{H}$ to $\mathbb{H}$ is surjective. In particular, f has a zero on $\mathbb{H}$.*

The only known proof of this theorem (so far) is again by an appeal to topology. We exploit the topology of the spheres, and think of S^4 as a compactification of $\mathbb{H} = \mathbb{R}^4$ by adding a "point at infinity". With f given as in (4.14), one observes that the evaluation map $q \mapsto f(q)$ can be extended to a *continuous* mapping $\hat{f} : S^4 \to S^4$ by taking $\hat{f}(\infty) = \infty$. This map can be shown to be of degree n in the *topological* sense; that is, its induced map in the 4-th homology multiplies a generator by n). Since $n \geq 1$, this map is onto, and therefore so is the original evaluation map $q \mapsto f(q)$ from $\mathbb{H}$ to $\mathbb{H}$.

In this theorem, the hypothesis on the highest degree term of $f(x)$ is a subtle one. Without this hypothesis, the theorem fails. For instance, for any nonzero $a \in \mathbb{H}$, any additive commutator $aq - qa$ has real part zero, and therefore the polynomial $f(x) = ax - xa + 1$ cannot have a zero. Here, the top degree part $ax - xa$ of f is not a single monomial.

It follows from (4.14), in particular, that any "ordinary" polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$ $(a_i \in \mathbb{H},\ a_n \neq 0)$ always has a zero in $\mathbb{H}$, provided that $n \geq 1$. This special case of (4.14) is known as the Niven-Jacobson Theorem. It can be proved without too much difficulty by pure algebra; see, for instance, [La$_2$: (16.14)]. From this theorem (and an appropriate Remainder Theorem), it follows that any polynomial $f \in \mathbb{H}[x]$ (as above) can be factored into a product of linear polynomials.

§5. The Rotation Groups $\mathbf{SO}(3)$ and $\mathbf{SO}(4)$

One of the most beautiful features of the quaternions is the role they play in the understanding and representation of the rotations of the low dimensional euclidean spaces. This piece of mathematics also goes back to Hamilton; we shall give an exposition of it here. Using modern notations, we write $\mathrm{O}(3)$ and $\mathrm{SO}(3)$ for the orthogonal group and the special orthogonal group of $\mathbb{R}^3 = \mathbb{H}_0$ (with respect to the euclidean norm), that is, the group of norm-preserving linear automorphisms and its (normal) subgroup of automorphisms of determinant 1. Let us consider the group homomorphism

(5.1) $$\varphi : \mathbb{H}^* \longrightarrow \mathrm{O}(3), \quad \text{defined by } \varphi(q)(\mathbf{v}) = q\,\mathbf{v}\,q^{-1} \quad (\forall\, \mathbf{v} \in \mathbb{H}_0).$$ [8]

Here, $\varphi(q) \in \mathrm{O}(3)$ since $\mathrm{N}(q\mathbf{v}q^{-1}) = \mathrm{N}(\mathbf{v})$ for all "vectors" $\mathbf{v} \in \mathbb{H}_0$. It is easy to see that $\mathbb{R}$ is the center of $\mathbb{H}$; this implies readily that the kernel of φ is $\mathbb{R}^*$. In the following, we'll show that the image of φ is $\mathrm{SO}(3)$.

To accomplish our goal, we'll need some notations. For the rest of this paper, we'll write $\mathbb{H}_1$ for the multiplicative group of quaternions of norm 1. (We'll call these the *unit*

[8] We leave to the reader the easy task of checking that $\mathbf{v} \in \mathbb{H}_0 \Longrightarrow q\,\mathbf{v}\,q^{-1} \in \mathbb{H}_0$.

quaternions.) For any vector $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$, we denote by $\tau_{\mathbf{u}}$ the reflection of $\mathbb{H}_0$ with respect to the plane (through the origin) with unit normal $\mathbf{u}$, and by $\rho_{\mathbf{u}}^{\theta}$ the rotation of the 3-space $\mathbb{H}_0$ about the vector $\mathbf{u}$ (anti-clockwise) by the angle θ.

(5.2) Theorem. (1) *If* $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$, *then* $\varphi(\mathbf{u}) = -\tau_{\mathbf{u}} = \rho_{\mathbf{u}}^{\pi}$.
(2) *Let* $\mathbf{u}_1, \mathbf{u}_2$ *be vectors in* $\mathbb{H}_0 \cap \mathbb{H}_1$, *making an angle* $\theta \in (0, \pi)$, *and let* $\mathbf{u}$ *be the unit vector in the direction of the cross product* $\mathbf{u}_1 \times \mathbf{u}_2$. *Then* $\tau_{\mathbf{u}_2} \tau_{\mathbf{u}_1} = \rho_{\mathbf{u}}^{2\theta}$.
(3) *For any* $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$ *and any angle* θ, $\varphi(\cos\,\theta + (\sin\,\theta)\,\mathbf{u}) = \rho_{\mathbf{u}}^{2\theta}$.
(4) *(Rodrigues' Formula for Rotations) For* $\mathbf{u}$ *as in* (3), *and any vector* $\mathbf{v} \in \mathbb{H}_0$:

$$\rho_{\mathbf{u}}^{2\theta}(\mathbf{v}) = (\cos\,2\theta)\,\mathbf{v} + (\sin\,2\theta)\,(\mathbf{u} \times \mathbf{v}) + (1 - \cos\,2\theta)\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u}.$$

(5) $\varphi(\mathbb{H}^*) = \varphi(\mathbb{H}_1) = \mathrm{SO}(3)$. *Thus, we have an exact sequence*

$$1 \longrightarrow \mathbb{R}^* \longrightarrow \mathbb{H}^* \xrightarrow{\varphi} \mathrm{SO}(3) \longrightarrow 1,$$

though φ *is not a split epimorphism.*

Proof. (1) Here, $\mathbf{u}^2 = -\mathbf{u}\,\overline{\mathbf{u}} = -\mathrm{N}(\mathbf{u}) = -1$, so (4.6) yields

$$\mathbf{u}\,\mathbf{v}\,\mathbf{u} = -\big(\mathbf{v}\,\mathbf{u} + 2\,\langle \mathbf{v},\, \mathbf{u} \rangle\big)\,\mathbf{u} = \mathbf{v} - 2\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u} \qquad (\forall\,\mathbf{v} \in \mathbb{H}_0). \tag{5.3}$$

Thus, $\varphi(\mathbf{u})(\mathbf{v}) = -\mathbf{u}\,\mathbf{v}\,\mathbf{u} = 2\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u} - \mathbf{v} = -\tau_{\mathbf{u}}(\mathbf{v}) = \rho_{\mathbf{u}}^{\pi}(\mathbf{v})$, as claimed.

(2) This is a well-known geometric fact, for which we'll only give a brief proof. Note that $\mathbf{u}$ is on the line of intersection of the two planes (through the origin) normal to the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$. With this observation, the proof can be then reduced to checking that, in the plane, the composition of reflections with respect to two lines intersecting at an angle θ is a rotation (anti-clockwise) by 2θ about the point of intersection. This is easily verified by drawing a picture.

(3) Without loss of generality, we may assume that the vector $\mathbf{u}$ is as in (2) above. We can thus carry out the proof here using the notations in (2). By the definition of inner products and cross products, we have $\langle \mathbf{u}_1,\, \mathbf{u}_2 \rangle = \cos\,\theta$, and $\mathbf{u}_1 \times \mathbf{u}_2 = (\sin\,\theta)\,\mathbf{u} = -\mathbf{u}_2 \times \mathbf{u}_1$. Therefore, by (4.5),

$$-\mathbf{u}_2\,\mathbf{u}_1 = \langle \mathbf{u}_1,\, \mathbf{u}_2 \rangle - \mathbf{u}_2 \times \mathbf{u}_1 = \cos\,\theta + (\sin\,\theta)\,\mathbf{u}. \tag{5.4}$$

From this, and from (1), (2) above, it follows that

$$\begin{aligned} \varphi\big(\cos\,\theta + (\sin\,\theta)\,\mathbf{u}\big) &= \varphi\,(-\mathbf{u}_2\,\mathbf{u}_1) = \varphi\,(\mathbf{u}_2)\,\varphi\,(\mathbf{u}_1) \\ &= (-\tau_{\mathbf{u}_2})(-\tau_{\mathbf{u}_1}) = \tau_{\mathbf{u}_2}\,\tau_{\mathbf{u}_1} = \rho_{\mathbf{u}}^{2\theta}. \end{aligned}$$

(4) Using (3), we can give the following quaternionic derivation of Rodrigues' Formula:

$$\begin{aligned} \rho_{\mathbf{u}}^{2\theta}(\mathbf{v}) = &= (\cos\,\theta + \mathbf{u}\,\sin\,\theta)\,\mathbf{v}\,(\cos\,\theta - \mathbf{u}\,\sin\,\theta) \\ &= (\cos^2\theta)\,\mathbf{v} + \cos\,\theta\,\sin\,\theta\,(\mathbf{u}\,\mathbf{v} - \mathbf{v}\,\mathbf{u}) - (\sin^2\theta)\,\mathbf{u}\,\mathbf{v}\,\mathbf{u} \\ &= (\cos^2\theta)\,\mathbf{v} + 2\sin\,\theta\,\cos\,\theta\,(\mathbf{u} \times \mathbf{v}) - (\sin^2\theta)\,\big(\mathbf{v} - 2\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u}\,\big) \quad \text{(by (5.3))} \\ &= (\cos\,2\theta)\,\mathbf{v} + (\sin\,2\theta)\,(\mathbf{u} \times \mathbf{v}) + 2\sin^2\theta\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u} \\ &= (\cos\,2\theta)\,\mathbf{v} + (\sin\,2\theta)\,(\mathbf{u} \times \mathbf{v}) + (1 - \cos\,2\theta)\,\langle \mathbf{v},\, \mathbf{u} \rangle\,\mathbf{u}. \end{aligned}$$

(5) Note that, for any $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$, $\cos\theta + (\sin\theta)\,\mathbf{u}$ is a unit quaternion. Since any element in $\mathrm{SO}(3)$ is a rotation $\rho_{\mathbf{u}}^{2\theta}$, (3) above shows that $\mathrm{SO}(3) \subseteq \varphi(\mathbb{H}_1)$, and of course, we have $\varphi(\mathbb{H}_1) = \varphi(\mathbb{H}^*)$. The inclusion above will be an equality if we can show that *every* unit quaternion $q = a + bi + cj + dk$ can be expressed in the form $\cos\theta + (\sin\theta)\mathbf{u}$, where $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$, and $\theta \in [0, \pi]$. Indeed, since $a^2 + b^2 + c^2 + d^2 = 1$, there exists a unique $\theta \in [0, \pi]$ such that $\cos\theta = a$. If $q \neq \pm 1$, then $\theta \in (0, \pi)$, so $b^2 + c^2 + d^2 = 1 - a^2 = \sin^2\theta > 0$, and we have

$$q = \cos\theta + \mathbf{u}\sin\theta, \quad \text{uniquely for } \mathbf{u} = (\sin\theta)^{-1}(bi + cj + dk) \in \mathbb{H}_0 \cap \mathbb{H}_1. \tag{5.5}$$

This expression is called the *polar form* of the unit quaternion q. It is uniquely determined if $q \neq \pm 1$. (If $q = \pm 1$, we take $\theta = 0$ or π, and take $\mathbf{u}$ to be any unit vector.)

If φ splits, there would exist a subgroup $G \subset \mathbb{H}^*$ giving a complement to the central subgroup $\mathbb{R}^*$, and hence $\mathbb{H}^* = \mathbb{R}^* \times G$. This implies that G contains all commutators in $\mathbb{H}^*$; but then $-1 = i^{-1}j^{-1}ij \in G$, a contradiction. □

For a little bit of history: Olinde Rodrigues (1795-1851) was of Spanish ancestry, though he was born in Bordeaux. An amateur mathematician, he spent his life in France as an economist, banker, social reformer, and railroad tycoon. Besides the rotation formula in (4) above, his claim to fame was the "other" Rodrigues formula expressing the Legendre polynomials in terms of the higher derivatives of $(x^2 - 1)^n$, which apparently came from his doctoral dissertation. For more information on Rodrigues' life and work (especially his contribution to the theory of rotations), see [Al$_1$], [Al$_2$].

Coming back to (5.2), it is of interest to note that the last part of this theorem can be used to give a rational parametrization of $\mathrm{SO}(3)$. In fact, if $q = a + bi + cj + dk \in \mathbb{H}^*$, then $q^{-1} = (a - bi - cj - dk)/n$ with $n = a^2 + b^2 + c^2 + d^2$. The orthogonal transformation $\varphi(q)$ on $\mathbb{H}_0$ given by

$$\mathbf{v} \mapsto q\,\mathbf{v}\,q^{-1} = n^{-1}(a + bi + cj + dk)\,\mathbf{v}\,(a - bi - cj - dk) \tag{5.6}$$

has thus a matrix with entries of the form r_{ij}/n, where r_{ij} are real quadratic forms in a, b, c, d. After the calculation is carried out, we see that this matrix is

$$n^{-1} \cdot \begin{pmatrix} a^2 + b^2 - c^2 - d^2 & -2ad + 2bc & 2ac + 2bd \\ 2ad + 2bc & a^2 - b^2 + c^2 - d^2 & -2ab + 2cd \\ -2ac + 2bd & 2ab + 2cd & a^2 - b^2 - c^2 + d^2 \end{pmatrix}. \tag{5.7}$$

Since $\varphi(\mathbb{H}^*) = \mathrm{SO}(3)$, (5.7) gives a rational parametrization of $\mathrm{SO}(3)$. This rational parametrization of $\mathrm{SO}(3)$ is exactly that discovered by Euler (using Euler angles) in 1770. (It is true that Euler did not work with matrices *per se*; however, the three columns of an $\mathrm{SO}(3)$ matrix are nothing more than a triad of unit vectors with a left-hand orientation.)

A nice way to think of the group $\mathbb{H}_1$ is the following. Using the matrix model (3.3), we can interpret $\mathbb{H}_1$ as the group of complex matrices $\begin{pmatrix} \alpha & \beta \\ -\overline{\beta} & \overline{\alpha} \end{pmatrix}$ of determinant 1. This is exactly the special unitary group $\mathrm{SU}(2)$. Making this identification, and restricting the homomorphism φ to $\mathbb{H}_1$, we get a new exact sequence

$$(5.8)\qquad 1 \longrightarrow \{\pm I_2\} \longrightarrow \mathrm{SU}(2) \xrightarrow{\varphi} \mathrm{SO}(3) \longrightarrow 1,$$

where, again, φ is nonsplit. Here, we have a 2-fold fibration of compact Lie groups. The best way to understand the topological structure of $\mathrm{SU}(2) \approx \mathbb{H}_1$ is to think of it as the 3-sphere S^3 in the euclidean 4-space $\mathbb{R}^4 = \mathbb{H}$. It is well known that S^3 is simply connected (that is, it has a trivial fundamental group), so the same holds for $\mathrm{SU}(2)$. Thus, via the epimorphism φ, $\mathrm{SU}(2)$ provides a (2-fold) *universal covering* of $\mathrm{SO}(3)$. This is called the *spin covering* of $\mathrm{SO}(3)$, and for this reason $\mathrm{SU}(2)$ is often called a "spin group" (and denoted by $\mathrm{Spin}(3)$). The spin covering is useful for many purposes; for instance, a continuous representation of $\mathrm{Spin}(3)$ that is not trivial on $\{\pm 1\}$ can be considered as a double-valued "spin representation" of $\mathrm{SO}(3)$ (and all of this has generalizations to the higher dimensions).

Using the topological model S^3 for $\mathrm{SU}(2)$, we see, in addition, that the rotation group $\mathrm{SO}(3)$ can be identified with the real projective space $\mathbb{RP}^3$. (Note that, under the covering φ in (5.8), we are identifying a matrix in $\mathrm{SU}(2)$ with its negative. This corresponds to identifying a unit quaternion with its negative, which translates into identifying antipodal points on S^3.) The physicists, however, have always preferred working with $\mathrm{SU}(2)$ (to working with S^3) since they also have to deal with the higher unitary groups $\mathrm{SU}(3)$ and $\mathrm{SU}(4)$ in elementary particle physics.

As it turns out that, the quaternions can also be used to describe the group $\mathrm{SO}(4)$. The possibility of this was already observed by Cayley. We use $\mathbb{H}$ as a model for the 4-space, and the main point is that the usual inner product on $\mathbb{R}^4$ can be expressed by the pairing

$$(5.9)\qquad (p,\, q) \mapsto B(p,\, q) := \big(p\,\overline{q} + q\,\overline{p}\big)/2 \in \mathbb{R}$$

on $\mathbb{H}$. (This fact can be checked by an easy direct calculation.) Given this, we can already use quaternion multiplications to construct a large family of isometries of $\mathbb{R}^4$. For, if $x,\, y \in S^3 = \mathbb{H}_1$, the map $q \mapsto x\,q\,\overline{y}$ is an isometry on $(\mathbb{H},\, B)$, as

$$\mathrm{N}(x\,q\,\overline{y}) = \mathrm{N}(x)\,\mathrm{N}(q)\,\mathrm{N}(\overline{y}) = \mathrm{N}(q).$$

Moreover, $q \mapsto x\,q$ and $q \mapsto q\,\overline{y}$ both have determinant 1, so the isometry constructed above is in $\mathrm{SO}(4)$. Cayley's beautiful result says that *all* isometries of $\mathbb{R}^4 = (\mathbb{H},\, B)$ arise in this way.

(5.10) Theorem. *Let* $\psi : S^3 \times S^3 \to \mathrm{SO}(4)$ *be defined by* $\psi(x,y)(q) = x\,q\,\overline{y}$ *(for all* $q \in \mathbb{H}$*). Then* ψ *is a group homomorphism, and we have an exact sequence*

$$(5.11)\qquad 1 \longrightarrow \{\pm(1,1)\} \longrightarrow S^3 \times S^3 \xrightarrow{\psi} \mathrm{SO}(4) \longrightarrow 1.$$

Proof. Here, we continue to identify $\mathbb{H}_1$ with S^3, so the latter becomes a group (and so does $S^3 \times S^3$). It is routine to check that ψ is a group homomorphism. To compute its kernel, suppose $\psi(x, y) = 1$, where $x, y \in \mathbb{H}_1$. Then $1 = \psi(x, y)(1) = x\,\overline{y}$ implies that $x = \overline{y}^{-1} = y$, and we need to have

$$q = \psi(x, y)(q) = x\,q\,\overline{y} = x\,q\,\overline{x} = x\,q\,x^{-1} \quad \text{for all } q \in \mathbb{H}.$$

Since the center of $\mathbb{H}$ is $\mathbb{R}$, this means that $x \in \mathbb{R} \cap \mathbb{H}_1 = \{\pm 1\}$. Thus, $(x, y) = \pm(1, 1)$, as desired.

It only remains to prove that ψ is *onto*. The proof of this becomes quite easy if we assume the Cartan-Dieudonné Theorem on isometries. According to (a weak version of) this theorem, any isometry on a finite-dimensional nonsingular symmetric bilinear space (over a field of characteristic $\neq 2$) is a product of hyperplane reflections (see [La$_1$: p. 27]). Thus, our group SO(4) is generated by "pair-products" of hyperplane reflections[9] $\tau_a\,\tau_b$, where a, b range over $\mathbb{H}_1 = S^3$. Therefore, it suffices to show that $\tau_a\,\tau_b \in \operatorname{im}(\psi)$. Now, using the pairing B in (5.9), we can explicitly compute the effect of τ_b as follows. For any $q \in \mathbb{H}$:

$$\tau_b(q) = q - 2B(q, b)\,b = q - \left(q\,\overline{b} + b\,\overline{q}\right) b = -b\,\overline{q}\,b. \tag{5.12}$$

Therefore, $\tau_a\,\tau_b\,(q) = \tau_a\big(-b\,\overline{q}\,b\big) = -a\,\overline{(-b\,\overline{q}\,b)}\,a = a\,\overline{b}\,q\,\overline{b}\,a$. Since this holds for all $q \in \mathbb{H}$, and $\overline{b}\,a = \overline{\overline{a}\,b}$, we have $\tau_a\tau_b = \psi\big(a\,\overline{b},\,\overline{a}\,b\big)$, as desired. ☐

In many ways, the group SO(4) turns out to be exceptional in the series of the orthogonal groups SO(n) ($n \geq 3$). In this series, SO(n) is a simple group when n is odd, and has a unique nontrivial normal subgroup given by its center $\{\pm I\}$ when n is even and $\neq 4$ (see Appendix II in [Di]). But for SO(4), the exact sequence (5.11) implies, remarkably, that SO(4) has a pair of nontrivial normal subgroups $\psi(S^3 \times \{1\})$ and $\psi(\{1\} \times S^3)$ (both isomorphic to S^3) intersecting at its center $\{\pm I\}$. This exceptional behavior of SO(4) (exhibited above by quaternion constructions) may have been one of the sources of the rather peculiar role played by the 4-space in mathematics. As Dieudonné wrote in [Di: p. 172], "The claim that four-dimensional spaces are quite exceptional is no idle talk."

§6. Finite Groups of Quaternions

For any field K, it is well known that any finite subgroup of the multiplicative group K^* is cyclic, so nothing much remains to be said about these finite subgroups. In the case of a *division ring* K, finding the finite subgroups of K^* is a much more interesting problem. Since $\mathbb{H}$ was the first known noncommutative division ring, it would seem particularly natural to try to find all finite subgroups of $\mathbb{H}^*$. However, this problem was not satisfactorily solved until 1940, when Coxeter came to the scene.

[9] As before, we use the notation τ_b for hyperplane reflections. Thus, here, τ_b denotes the reflection of $\mathbb{R}^4$ with respect to the 3-dimensional subspace orthogonal to b in the usual inner product.

In his classical paper [Co], Coxeter not only determined the finite subgroups of $\mathbb{H}^*$, but also put these groups in the general context of what he called the *binary polyhedral groups.* Before stating this classification result, let us first introduce two general families of groups after Coxeter.

For integers ℓ, m, n such that $2 \le \ell \le m \le n$, let (ℓ, m, n) and $\langle \ell, m, n\rangle$ denote, respectively, the following two groups defined by generators and relations:

$$\langle R, S, T \mid R^\ell = S^m = T^n = RST = 1 \rangle; \tag{6.1}$$

$$\langle R, S, T \mid R^\ell = S^m = T^n = RST \rangle. \tag{6.2}$$

Note that the element RST is central in $\langle \ell, m, n\rangle$ (being a power of each of the generators), and so $\langle \ell, m, n\rangle/\langle RST\rangle \cong (\ell, m, n)$. Coxeter has proved that (ℓ, m, n) is finite exactly in the following four cases:

$$(a)\ (2,2,n), \quad (b)\ (2,3,3), \quad (c)\ (2,3,4), \quad \text{and} \quad (d)\ (2,3,5). \tag{6.3}$$

These groups are, respectively:

(a) the *dihedral groups* of order $2n$;
(b) the *tetrahedral group* of order 12, isomorphic to the alternating group A_4;
(c) the *octahedral group* of order 24, isomorphic to the symmetric group S_4; and
(d) the *icosahedral group* of order 60, isomorphic to the alternating group A_5.

These are the symmetry groups of the regular n-gon and the regular tetrahedron, octahedron and icosahedron, respectively, so they are subgroups of SO(3). It is well known that, up to conjugation, these "(regular) polyhedral groups" and the cyclic groups (generated by rotations by the angle $2\pi/n$, $n \ge 1$) are, in fact, *all finite subgroups* of SO(3).

In the four cases in (6.3), Coxeter showed that the element RST has order 2 in $\langle \ell, m, n\rangle$, so the latter groups are also finite, with order twice that of (ℓ, m, n). The resulting finite groups $\langle \ell, m, n\rangle$ are called the *binary polyhedral groups*, since they are 2-fold coverings of the (regular) polyhedral groups. Note, for instance,

$$\begin{aligned} \langle 2,2,n\rangle &= \langle R, S, T \mid R^2 = S^2 = T^n = RST \rangle \\ &= \langle S, T \mid (ST)^2 = S^2 = T^n \rangle \\ &= \langle S, T \mid T^{2n} = 1,\ S^2 = T^n,\ S^{-1}TS = T^{-1} \rangle; \end{aligned}$$

this is called the *generalized quaternion* (or *dicyclic*) *group* of order $4n$.

The name "generalized quaternion group" suggests that these groups (for $n \ge 2$) are strongly related to the quaternions. Indeed, a straightforward calculation shows that the two unit quaternions

$$s = j, \quad \text{and} \quad t = \cos(\pi/n) + i\sin(\pi/n) \tag{6.4}$$

generate a subgroup of $\mathbb{H}_1$ that is isomorphic to $\langle 2,2,n\rangle$ (by the isomorphism $s \leftrightarrow S,\ t \leftrightarrow T$). In the simplest case $n = 2$, we get the group generated by $\{i,\ j\}$, which is the ordinary quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$, familiar to all students in elementary group theory!

The next important step is to observe that the other three finite groups in the family $\langle \ell, m, n\rangle$, that is, the binary tetrahedral, octahedral, and icosahedral groups, can likewise be realized as subgroups of $\mathbb{H}_1$. For instance, we can check that the following set of 24 quaternions

$$\{\pm 1,\ \pm i,\ \pm j,\ \pm k,\ (\pm 1 \pm i \pm j \pm k)/2\} \subset \mathbb{H}_1 \subset \mathbb{H}^* \tag{6.5}$$

(where the signs are arbitrarily chosen) constitute a model of the binary tetrahedral group $\langle 2,3,3\rangle$, with an isomorphism given by

$$R \leftrightarrow i, \quad S \leftrightarrow (1+i+j-k)/2, \quad \text{and} \quad T \leftrightarrow (1+i+j+k)/2. \tag{6.6}$$

In the same spirit, we can also construct triplets of elements in $\mathbb{H}_1$ that generate subgroups of $\mathbb{H}_1$ providing models for the binary octahedral and binary icosahedral groups. We shall only give the "answers" here:

$$R \leftrightarrow (i+j)/\sqrt{2}, \quad S \leftrightarrow (1+i+j+k)/2, \quad \text{and} \quad T \leftrightarrow (1+i)/\sqrt{2}, \tag{6.7}$$

$$R \leftrightarrow i, \quad S \leftrightarrow (1+\tau i-\tau^{-1}k)/2, \quad \text{and} \quad T \leftrightarrow (\tau+i+\tau^{-1}j)/2, \tag{6.8}$$

where, in (6.8), τ denotes the "golden ratio" $(1+\sqrt{5})/2$. These explicit constructions came from [Co], except that, curiously enough, there was a mistake in [Co: p. 371] in the modeling of the binary icosahedral group! Our proposed correction is based on (7.6)(3) below.[10]

A more conceptual way to see that the binary polyhedral groups occur as subgroups of $\mathbb{H}_1$ is to use the spin covering $\varphi : \mathbb{H}_1 \to \mathrm{SO}(3)$ with kernel $\{\pm 1\}$ (constructed in §5). Fixing copies of the finite polyhedral groups in $\mathrm{SO}(3)$, we can form the preimages of these groups under the covering φ. The resulting groups in $\mathbb{H}_1$, twice the size of the polyhedral groups, turn out to be isomorphic to the binary polyhedral groups, though this statement would require a careful proof.

Given the above constructions, and assuming the fact stated earlier on the classification of finite subgroups in $\mathrm{SO}(3)$, it is now relatively easy to prove the following beautiful result of Coxeter.

(6.9) Coxeter's Theorem. *Up to conjugation, the models of the binary polyhedral groups constructed above, and the cyclic groups* $\langle \cos(2\pi/n) + i\sin(2\pi/n)\rangle$ *(for* $n \geq 1$*), are all the finite subgroups of* $\mathbb{H}^*$ *.*

[10]For nice "geometric pictures" of these constructions via the regular solids, see [Vi: pp. 15-16].

Proof (Sketch). Let G be a finite subgroup of $\mathbb{H}^*$. Since every element in G has finite order, G must lie in $\mathbb{H}_1$. If G is cyclic, it is not hard to show that G is conjugate to $\langle \cos(2\pi/n) + i\sin(2\pi/n) \rangle$ (for some n). If G is not cyclic, then $|G|$ must be even. (If $|G|$ was odd, we would have $G \cong \varphi(G) \subseteq \mathrm{SO}(3)$. But the only odd-order subgroups of $\mathrm{SO}(3)$ are cyclic.) By considering an element of order 2 in G, we see that $G \supseteq \{\pm 1\}$. Then $\varphi(G)$ is conjugate to a polyhehral group in $\mathrm{SO}(3)$. Taking preimages under the spin covering $\varphi : \mathbb{H}_1 \to \mathrm{SO}(3)$, we can argue easily that G is conjugate to a (fixed) binary polyhehral group in $\mathbb{H}_1$. $\square$

With this theorem proven, one ambitious question to ask is: *how about finite subgroups of division rings in general?* It turned out that this difficult problem also has a satisfactory answer: in his paper [Am] in 1955, S.A. Amitsur furnished a complete classification for the finite groups that are embeddable in the multiplicative groups of division rings. There is no need for us to reproduce Amitsur's list of groups here (which, of course, includes all finite cyclic groups and binary polyhedral groups). Let us just content ourselves by mentioning the rather remarkable fact that all finite groups in Amitsur's list turned out to be *solvable* — with the sole exception of the binary icosahedral group! (See [Am: p. 384]; also, cf. [Br: p. 102].) This group has order 120, and is isomorphic to the special linear group $\mathrm{SL}_2(\mathbb{F}_5)$. The latter group is a double covering of $\mathrm{PSL}(\mathbb{F}_5) = \mathrm{SL}_2(\mathbb{F}_5)/\{\pm 1\}$, which is the icosahedral group A_5 (the smallest nonabelian simple group). It is interesting that the *only* nonsolvable finite group embeddable in a division ring appears already in Hamilton's system of real quaternions.

§7. Exponentials, De Moivre's Formula, and Other Information

In this final section, we shall assemble a few additional facts which are pertinent to understanding the structure of the division ring of quaternions. This will be followed by some parting pointers to the literature.

First, we note that every quaternion $q = a + bi + cj + dk$ satisfies a quadratic equation over $\mathbb{R}$. For, if we write $\mathrm{Tr}(q) = q + \overline{q} = 2a$ (the trace of q), then

$$(7.1) \qquad q^2 - \mathrm{Tr}(q)\, q + \mathrm{N}(q) = q^2 - (q + \overline{q})\, q - q\,\overline{q} = 0.$$

Thus, q satisfies $q^2 - 2a\, q + (a^2 + b^2 + c^2 + d^2) = 0$. If $q \notin \mathbb{R}$, this is clearly the minimal equation of q over $\mathbb{R}$, so the quadratic extension $\mathbb{R}\,[q]$ must be isomorphic to the complex field $\mathbb{C}$. It is easy to check further that: (1) after identifying $\mathbb{R}\,[q]$ with $\mathbb{C}$, the norm of q is exactly the squared modulus of q as a complex number, and (2) the centralizer of q in $\mathbb{H}$ is exactly the field $\mathbb{R}\,[q]$.

One thing we can do with the above information is to define an exponential function on the quaternions. This can be done exactly in the usual way: for any $q \in \mathbb{H}$, we simply define

$$(7.2) \qquad e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!} = 1 + \frac{q}{1!} + \frac{q^2}{2!} + \frac{q^3}{3!} + \cdots \in \mathbb{H}.$$

By the standard methods, it can be shown that this series is (absolutely) convergent (with respect to the euclidean metric on $\mathbb{H}$), so its sum lies in $\mathbb{H}$ as indicated. But actually, this verification is not strictly necessary. For, if x is not a real number, then $\mathbb{R}[q]$ is just a copy of $\mathbb{C}$ as we have pointed out above. Therefore, we know that the series in (7.2) is (absolutely) convergent in any case, and hence already $e^q \in \mathbb{R}[q]$.

One of the principal properties of the ordinary exponential function is $e^p \cdot e^q = e^{p+q}$. If $pq = qp \in \mathbb{H}$, then they are contained in a common quadratic extension of $\mathbb{R}$, so this equality will indeed hold. But if $pq \neq qp$, we can no longer prove (and thus should not expect) the same equality. For instance, $e^{\pi i} \cdot e^{\pi j} = (-1)(-1) = 1$, but $e^{\pi(i+j)}$ is definitely not 1. The failure of the equality $e^p \cdot e^q = e^{p+q}$ on $\mathbb{H}$ is a serious drawback, and may have been the principal roadblock to the development of a really useful theory of functions on the quaternions.

Nevertheless, the exponential function on $\mathbb{H}$ leads to nice ways of representing quaternions. For any $\mathbf{u} \in \mathbb{H}_0 \cap \mathbb{H}_1$, we have $1 = \mathbf{u}\,\overline{\mathbf{u}} = -\mathbf{u}^2$, so $\mathbf{u}^2 = -1$. Thus, for any real angle θ:

$$\begin{aligned} e^{\theta\,\mathbf{u}} &= 1 + \frac{\theta\,\mathbf{u}}{1!} - \frac{\theta^2}{2!} - \frac{\theta^3\,\mathbf{u}}{3!} + \frac{\theta^4}{4!} + \frac{\theta^5\,\mathbf{u}}{5!} - \frac{\theta^6}{6!} - \frac{\theta^7\,\mathbf{u}}{7!} + \cdots \\ &= \left(1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots\right) + \mathbf{u}\left(\theta - \frac{\theta^3}{3!} + \frac{\theta^5}{5!} - \cdots\right) \\ &= \cos\,\theta + \mathbf{u}\,\sin\,\theta. \end{aligned}$$

This is entirely to be expected. For, we may identify $\mathbb{R}[\mathbf{u}]$ with $\mathbb{C}$ by identifying $\mathbf{u}$ with the complex number i, so the above equation could have been obtained from Euler's Formula $e^{i\theta} = \cos\,\theta + i\,\sin\,\theta$ for the complex numbers. Using this method, we also obtain without further proof the following extension of the usual "De Moivre Formula" to unit quaternions:

$$(7.3) \qquad \big(\cos\,\theta + \mathbf{u}\,\sin\,\theta\big)^n = \cos\,n\theta + \mathbf{u}\,\sin\,n\theta \quad \text{(for any integer } n\text{)}.$$

Any quaternion $q \in \mathbb{H}^*$ can be written in the form[11] $\rho \cdot (\cos\,\theta + \mathbf{u}\,\sin\,\theta)$, where $\rho = \mathrm{N}(q)^{1/2}$, and $\theta \in [0,\,\pi]$ is called the *polar angle* of q. (Hamilton, who had a penchant for inventing strange names, called $\cos\,\theta + \mathbf{u}\,\sin\,\theta$ the *versor* of q.) Thus, we can represent q in the exponential form $\rho\,e^{\theta\,\mathbf{u}}$, in exactly the same way as we represent complex numbers. We can then compute e^q as follows. Writing $a = \rho\,\cos\,\theta$ and $b = \rho\,\sin\,\theta$, we have $q = a + b\,\mathbf{u}$, and thus,

$$(7.4) \qquad e^q = e^{a+b\,\mathbf{u}} = e^a\,e^{b\,\mathbf{u}} = e^a\,\big(\cos\,b + \mathbf{u}\,\sin\,b\big).$$

From this, it follows, for instance, that e^q has norm e^{2a}, and "versor" $\cos\,b + \mathbf{u}\,\sin\,b$.

[11] The case when $q \in \mathbb{R}^*$ is a little exceptional. In this case, as we have noted earlier, we take θ to be 0 or π depending on whether q is positive or negative, and take $\mathbf{u}$ to be any unit vector.

The (generalized) De Moivre Formula (7.3) is extremely useful in analyzing the multiplicative structure of $\mathbb{H}^*$. From this formula for q^n, we can, for instance, deduce the following nice fact, which enables us to compute quickly the multiplicative order of any nonzero quaternion q via $\mathrm{N}(q)$ and the real part of q.

(7.5) Theorem. *A nonzero quaternion q has finite multiplicative order iff $\mathrm{N}(q) = 1$ and the polar angle θ of q (in radian measure) is a rational multiple of 2π. In this case, the order of q is the smallest positive integer n such that $n\theta \in 2\pi\,\mathbb{Z}$.*

The explicit examples below will illustrate the considerable power of (7.5).

(7.6) Examples.

(1) Let $\alpha = (\pm i \pm j \pm k)/\sqrt{3}$ (where the signs are arbitrary). Then α has norm 1 and polar angle $90^o = \pi/2$. Thus, $\alpha^2 = -1$, and α has order 4. (The same holds for any unit pure quaternion!)

(2) Let $\beta := (1 \pm i \pm j \pm k)/2$, which has norm 1. Its polar angle is $60^o = \pi/3$. Thus, $\beta^3 = -1$, and β has order 6. On the other hand, any quaternion $(-1 \pm i \pm j \pm k)/2$ has norm 1 and polar angle $120^o = 2\pi/3$, so it has order 3.

(3) Let τ denote the golden ratio $(1+\sqrt{5}\,)/2$, which is the positive root of $\tau^2 = \tau + 1$. From $\tau - \tau^{-1} = 1$, one gets $\tau^2 + \tau^{-2} = 3$. Thus, any quaternion of the form $\gamma := (\tau \pm i \pm \tau^{-1} j)/2$ has norm 1. Furthermore, since $\tau = 2\cos 36^o$, γ has polar angle $36^o = \pi/5$. From this, we see that $\gamma^5 = -1$, and the order of γ is 10. On the other hand, if, say, $\delta := (1 \pm \tau i \pm \tau^{-1} k)/2$, then $\delta^3 = -1$, and δ has order 6.

The observations above enable us to check quickly that the constructions in (6.6), (6.7) and (6.8) give us well-defined quaternion models for the binary tetrahedral, octahedral, and icosahedral groups.

We record below, largely without proof, some of the other properties of quaternions that are useful to know.

- A quaternion $q \notin \mathbb{R}$ is a pure quaternion iff $q^2 \in \mathbb{R}$.

- Two quaternions in $\mathbb{H}$ are conjugate iff they have the same norm and the same trace, iff they have the same minimal polynomial over $\mathbb{R}$. In particular, two pure quaternions q_1, q_2 are conjugate iff $q_1^2 = q_2^2$, so a typical conjugacy class of $\mathbb{H}^*$ in the pure quaternions is the set of zeros (in $\mathbb{H}$) of a polynomial $x^2 + r$ (where $r \in \mathbb{R}_+^*$). Each of these zero sets is *uncountable* (contrary to the case over fields where any polynomial of degree n has at most n roots).

- Every $q \in \mathbb{H}_1$ can be written as a *single* commutator $x^{-1}y^{-1}xy$, where $x, y \in \mathbb{H}_1$. In particular, the commutator subgroups of $\mathbb{H}^*$ and $\mathbb{H}_1$ are both equal to $\mathbb{H}_1$.

- (Stereographic Projection) Taking 1 as the "north pole" of the sphere $S^3 = \mathbb{H}_1$, we can use a stereographic projection to map $\mathbb{H}_1 \setminus \{1\}$ onto the space of pure quaternions $\mathbb{H}_0$.

This map works out to be $q \mapsto \frac{1+q}{1-q}$, with inverse map $p \mapsto \frac{p-1}{p+1}$ for any pure quaternion $p \in \mathbb{H}_0$. Both of these are conformal mappings.

• Any quaternion is a product of two pure quaternions. (This can be easily deduced from the proof of (5.2)(3).)

• Every linear endomorphism of $\mathbb{H}$ as a real vector space is expressible in the form $q \mapsto \sum_i x_i\, q\, y_i$, for suitable quaternions $x_i,\, y_i \in \mathbb{H}$.

• Every automorphism σ of $\mathbb{H}$ as a ring is an inner automorphism. (This is proved by first observing that $\sigma(\mathbb{R}) = \mathbb{R}$, since σ must take the center of $\mathbb{H}$ onto itself. Thus, σ is an $\mathbb{R}$-automorphism, and the Skolem-Noether Theorem gives the desired conclusion.)[12]

We close by mentioning some additional references on quaternions, especially on topics that we have had no space to touch upon in this article. On the life and work of Sir William Rowan Hamilton, a good modern reference is [Ha]. For more historical backgrounds related to the discovery of quaternions, see [Cr], [Kl], [Wa$_1$] and [Wa$_2$]. For a survey on determinants of quaternionic matrices, see [As]. For general discussions on the use of quaternions in physics, see [AJ], [Va], and [W]. For more information on rotations as related to quaternions, with applications to astronomy and aerospace sciences, see [Ku].

References

[A] J. F. Adams: *Vector fields on spheres,* Annals of Math. **72**(1962), 603-632.

[Ad] S. L. Adler: *Quaternion Quantum Mechanics and Quantum Fields,* Oxford Univ. Press, New York-Oxford, 1995.

[Al$_1$] S. L. Altmann: *Rotations, Quaternions, and Double Groups,* Oxford Univ. Press, 1986.

[Al$_2$] S. L. Altmann: *Hamilton, Rodrigues, and the quaternion scandal,* Math. Mag. **62**(1989), , 291-308.

[Am] S. A. Amitsur: *Finite subgroups of division rings,* Trans. Amer. Math. Soc. **80**(1955), 361-386.

[AJ] R. Anderson and G. C. Joshi: *Quaternions and the heuristic role of mathematical structures in physics,* Physical Essays **6**(1993), 308-319.

[As] H. Aslaksen: *Quaternionic determinants,* Math. Intelligencer **18**(1996), 57-65.

[12] In contrast to this, however, the automorphism group of $\mathbb{C}$ is more complicated. Contrary to a claim once made by Dedekind, this automorphism group contains more elements than just the identity map and the complex conjugation. The point is that an automorphism of $\mathbb{C}$ need not take $\mathbb{R}$ to $\mathbb{R}$: this shows a subtle difference between the complex numbers and the quaternions. Dedekind's claim would have been correct if we consider only $\mathbb{R}$-automorphisms on $\mathbb{C}$.

[Ba] J. C. Baez: *The octonions,* Bull. Amer. Math. Soc. (N.S.) **39**(2002), 145-205.

[Bl] F, van der Blij: *History of the octaves,* Simon Stevin **34**(1961), 106-125.

[Br] R. Brown: *Frobenius groups and classical maximal orders,* Mem. Amer. Math. Soc., No. **717**(2000), AMS, Providence, R.I.

[Ca] A. Cayley: *A memoir on the theory of matrices,* London Phil. Trans. **148**(1858), 17-37. (Collected Works, Vol. 2, 475-496.)

[Co] H.S.M. Coxeter: *The binary polyhehral groups, and other generalizations of the quaternion groups,* Duke Math. Journal **7**(1940), 367-379.

[Cr] M. J. Crowe: *A History of Vector Analysis,* Univ. of Notre Dame Press, 1967. (Reprinted by Dover Publications, Inc., New York, 1985.)

[Di] J. Dieudonné: *Linear Algebra and Geometry,* Hermann, Houghton Mifflin Co., New York, 1969.

[EN] S. Eilenberg and I. Niven: *The "fundamental theorem of algebra" for quaternions,* Bull. Amer. Math. Soc. **50**(1944), 246-248.

[Fr] F. G. Frobenius: *Über lineare Substitutionen und bilineare Formen,* Crelle Journal **84**(1878), 1-63. (Ges. Abhandl. **1**, 343-405, Springer-Verlag, 1968.)

[Ga] C. F. Gauss: *Mutation des Raumes,* Carl Friedrich Gauss Werke, Achter Band, pp. 357-361, König. Gesell. Wissen. Göttingen, 1900. (This note appeared to be an updated version of an entry in Gauss's diary; editor Stäckl of the volume placed the date of the entry at around 1819.)

[H_1] W. R. Hamilton: *Lectures on Quaternions,* Hodges and Smith, Dublin, 1853.

[H_2] W. R. Hamilton: *Elements of Quaternions,* 1866. [Second Ed. (2 volumes): C.J. Jolly, editor, Longman, Green & Co., London, 1899-1901; reprinted by Chelsea, 1969.]

[Ha] T. L. Hankins: *Sir William Rowan Hamilton,* Johns Hopkins Univ. Press, Baltimore, 1980.

[Hap] D. Happel: *Klassifikationstheorie endlich-dimensionaler Algebren in der Zeit von* 1880 *bis* 1920*,* L'Enseig. Math. **26**(1980), 91-102.

[Hi] F. Hirzebruch: *Division algebras and topology,* in *Numbers,* (H.-D. Ebbinghaus et al., eds.), pp. 281-302, Readings in Math., Springer-Verlag, Berlin-Heidelberg-New York, 1991.

[Hu] A. Hurwitz: *Über die Komposition der quadratischen Formen von beliebig vielen Variablen,* Nach. der könig. Gesell. Wissen. Göttingen, Math.-Physik. Klasse, 1898, 309-316. (Math. Werke **2**, 565-571, Basel, 1933.)

[Kl] M. Kline: *Mathematical Thought from Ancient to Modern Times,* Oxford Univ. Press, Oxford-New York-Toronto, 1972.

[KR_1] M. Koecher and R. Remmert: *Hamilton's quaternions,* in *Numbers,* (H.-D. Ebbinghaus et al., eds.), Graduate Texts in Math. **123** (Readings in Mathematics), pp. 189-220, Springer-Verlag, Berlin-Heidelberg-New York, 1991.

[KR_2] M. Koecher and R. Remmert: *The isomorphism theorems of Frobenius, Hopf and Gelfand-Mazur, ibid.* pp. 221-247.

[Ku] J. Kuipers: *Quaternions and Rotation Sequences,* Princeton Univ. Press, Princeton, N.J., 1999.

[La_1] T. Y. Lam: *The Algebraic Theory of Quadratic Forms,* (second printing, with revisions), Benjamin-Addison Wesley, Reading, Mass., 1980.

[La_2] T. Y. Lam: *A First Course in Noncommutative Rings*, Second Edition, Graduate Texts in Math., Vol. **131**, Springer-Verlag, Berlin-Heidelberg-New York, 2001.

[Ma] K. O. May: *The impossibility of a division algebra of vectors in three-dimensional space,* Amer. Math. Monthly **73**(1966), 289-291.

[Me] A. Messiah: *Quantum Mechanics,* Vol. 2, North-Holland, J. Wiley & Sons, Inc., New York, 1965.

[Pe] B. Peirce: *Linear associative algebras*, Amer. J. Math **4**(1881), 97-229.

[PS] G. Pickert and H.-G. Steiner: *Komplexe Zahlen und Quaternionen,* Kapitel 8 in *Grundzüge der Mathematik I,* (Behnke, Bachmann, Fladt, Süss, eds.), Vandenhoeck & Ruprecht, Göttingen, 1962.

[Sh] D. Shapiro: *Compositions of Quadratic Forms,* de Gruyter Expos. in Math., Vol. 33, Berlin, 2000.

[Va] M. A. Vandyck: *Quaternions in physics: Victorian curiosities or enlightening language? New insights 150 years after Hamilton's discovery,* Rev. Questions Sci. **172**(2002), 293-307.

[Vi] M.-F. Vigneras: *Arithmétique des Algèbres de Quaternions,* Lecture Notes in Math., Vol. **800**, Springer-Verlag, Berlin-Heidelberg-New York, 1980.

[Wa_1] B.L. van der Waerden: *Hamiltons Entdeckung der Quaternionen,* erweitierte Fassung eines Vortrags 26 Juni, 1973, Vandenhoeck & Ruprecht, Göttingen, 1973.

[Wa_2] B.L. van der Waerden: *Hamilton's discovery of quaternions,* Math. Magazine **49**(1976), 227-234.

[W] J. P. Ward: *Quaternions and Cayley Numbers,* Mathematics and its Applications, **403**. Kluwer Academic Publishers Group, Dordrecht, 1997.

University of California
Berkeley, Ca 94720

lam@math.berkeley.edu