Hamilton's Quaternions

M. Koecher, R. Remmert

Love of fame moves and cheers great mathematicians (W.R. Hamilton).

INTRODUCTION

1. Sir William Rowan Hamilton was born in Dublin in 1805, and at the age of five was already reading Latin, Greek and Hebrew. He entered Trinity College Dublin in 1823, and while still an undergraduate was, in 1827, appointed Andrewes Professor of Astronomy at that university, and Director of the Dunsink Observatory with the title “Royal Astronomer of Ireland.” In that same year he began to develop geometric optics on extremal principles and in 1834/35 extended these ideas to dynamics, with the introduction of the principle of least action, the Hamiltonian function, and his canonical equations of motion. He was knighted in 1835 and was President of the Royal Irish Academy from 1837 to 1845. His great discovery of quaternions was made in 1843. He died in 1865 at Dunsink.

One of HAMILTON’s earlier achievements in 1835 had been to legitimize the traditional use of complex numbers in mathematics. He showed that calculating with complex numbers $x + iy$ was logically equivalent to performing operations on ordered pairs $(x, y)$ of real numbers in accordance with certain postulated rules (see 3.1.8). This was the origin of his interest in the question of whether the geometrical interpretation of addition, and more particularly of multiplication of complex numbers in the plane $\mathbb{R}^2$, might not somehow—through the creation of hypercomplex numbers—have an analogue in the three dimensional space $\mathbb{R}^3$ of our visual intuition.

HAMILTON had been hoping for many years to find a satisfactory form of multiplication for real number triples with the right properties. Shortly before his death in 1865 he wrote to his son (Math. Papers 3, p. XV): “Every morning, on my coming down to breakfast, you used to ask me: ‘Well, Papa, can you multiply triplets?’ Whereeto I was always obliged to reply, with a sad shake of the head: ‘No, I can only add and subtract them’.”

It is easy enough to see nowadays that there can be no $\mathbb{R}$-linear multiplication of all real number triples $(\alpha, \beta, \gamma)$ in $\mathbb{R}^3$ which simply extends the multiplication in $\mathbb{C} = \mathbb{R}^2 \subset \mathbb{R}^3$ of the pairs $(\alpha, \beta)$. For if $e := (1, 0, 0), i := (0, 1, 0), j := (0, 0, 1)$ be the canonical base of $\mathbb{R}^3$, then $ij$ would have
to be of the form $\rho e + \sigma i + \tau j$. It would then follow, if one assumes $i^2 = -e$ and $i(ij) = ii(j) = -j$, that

$$-j = \rho i - \sigma e + \tau ij = \rho i - \sigma e + \tau(\rho e + \sigma i + \tau j) = (\tau \rho - \sigma)e + (\tau \sigma + \rho)i + \tau^2 j,$$

and thus (since $e, i, j$ are linearly independent) that $\tau^2 = -1$, which would imply $\tau \notin \mathbb{R}$.\(^1\)

2. HAMILTON's efforts are at first unsuccessful: He is looking for a multiplication with triplets in which, as with number pairs, the usual rules would still apply (in other words he assumes a principle of permanence). He begins by trying

$$\alpha + \beta i + \gamma j \quad \text{with} \quad i^2 = j^2 = -1,$$

(in which the existence of a neutral element is already implied) and considers the simplest case

$$(*)(\alpha + \beta i + \gamma j)^2 = \alpha^2 - \beta^2 - \gamma^2 + 2i\alpha\beta + 2j\alpha\gamma + 2ij\beta\gamma,$$

where the expression on the right is calculated in the ordinary way using the commutative laws.

The “touchstone” which he uses to test the value of the product of two vectors is, as in the case of $\mathbb{C}$ (where we have the modulus law) the principle that the length of the “product” of two vectors should be equal to the product of their individual lengths; the length of $\alpha + \beta i + \gamma j$ being its “Euclidean” length $\sqrt{\alpha^2 + \beta^2 + \gamma^2}$. The sum of the squares of the coefficients of 1, $i$ and $j$ on the right hand side of (*) yields

$$(\alpha^2 - \beta^2 - \gamma^2)^2 + (2\alpha\beta)^2 + (2\alpha\gamma)^2 = (\alpha^2 + \beta^2 + \gamma^2)^2;$$

and thus HAMILTON has established the fact that the product rule will certainly hold provided $ij$ is made equal to zero. But he does not like this. And then he notices that the term on the right of (*) should really be $ij + ji$ rather than $2ij$. This has to vanish so that $ji = -ij$; and so he is led to sacrifice the commutative law. One can see all this very clearly from a letter which HAMILTON wrote to John GRAVES on the 17th October 1843 (Math. Papers 3, 106-110): “Behold me therefore tempted for a moment to

---

\(^1\)In fact, one can prove the better

**Theorem.** Every real division algebra $A$ of odd dimension with unit element $e$ is isomorphic to $\mathbb{R}$, and therefore has dimension 1.

**Proof.** Let $a \in A$. The “left-multiplication” $L_a: A \to A, x \mapsto ax$ is a vector space endomorphism. Since $\dim A$ is odd, $L_a$ has a real eigenvalue (by the Bolzano-Cauchy intermediate value theorem). If $v \neq 0$ is an associated eigenvector, then $av = \lambda v$, that is $(a - \lambda e)v = 0$. Since $A$ is a division algebra, it follows that $a = \lambda e$, or in other words $a \in \mathbb{R}e$, from which we see that $A = \mathbb{R}e$. 


fancy that \(ij = 0\). But this seemed odd and uncomfortable, and I perceived that the same suppression of the term which was \textit{de trop} might be attained by assuming what seemed to me less harsh, namely that \(ji = -ij\). I made therefore \(ij = k, ji = -k\), reserving to myself to inquire whether \(k\) was 0 or not.”

And now HAMILTON hit upon the ingenious idea that gave a new and decisive direction to the whole problem: he “jumped with \(k\) into a fourth dimension.” In other words, he took \(k\) to be linearly independent of 1, \(i\) and \(j\). In his letter to GRAVES he wrote (loc. cit.) “and there dawned on me the notion that we must admit, in some sense, a fourth dimension of space for the purpose of calculating with triplets.”

HAMILTON now carefully investigates what \(k^2\) should be. If one were to use the associative law it would be immediately apparent that

\[k^2 = (ij)(ij) = i(ji)j = -i(ij)j = -i^2j^2 = -1;\]

but he does not use this argument, because he is not sure whether his multiplication is associative (his notes on this point are to be found in \textit{Math. Papers 3}, 103–105).

Later on he brings out clearly the validity of the associative law; thus he writes (\textit{Math. Papers 3}, p. 114): “... the commutative character is lost .... However it will be found that another important property of the old multiplication is preserved, or extended to the new, namely, that which may be called the \textit{associative} character of the operation ....” This could well be the first introduction of the word “associative” in Mathematics.

3. The breakthrough came to HAMILTON on the 16th October 1843 on his way to a meeting of the Royal Irish Academy; during that meeting he announced his discovery of quaternions. He devoted the remainder of his life exclusively to their further exploration. He himself described in 1858 the moment of discovery in the following words (\textit{North British Rev. 14}, 1858): “...Tomorrow will be the fifteenth birthday of the Quaternions. They started into life, or light, full grown, on the 16th of October, 1843, as I was walking with Lady Hamilton to Dublin, and came up to Brougham Bridge. That is to say, I then and there felt the galvanic circuit of thought closed, and the sparks which fell from it were the fundamental equations between \(i, j, k\) \textit{exactly such} as I have used them ever since. I pulled out, on the spot, a pocketbook, which still exists, and made an entry, on which, \textit{at the very moment}, I felt that it might be worth my while to expend the labour of at least ten (or it might be fifteen) years to come. But then it is fair to say that this was because I felt a \textit{problem} to have been at that moment \textit{solved}, an intellectual \textit{want relieved}, which had \textit{haunted} me for at least \textit{fifteen years} before...” And in the letter, which we have already mentioned, to his son, he says, referring to that memorable October day: “Nor could I resist the impulse—unphilosophical as it may have been—to
cut with a knife on a stone of Brougham Bridge the fundamental formula with the symbols $i, j, k$:

$$i^2 = j^2 = k^2 = ijk = -1.$$ 

With great delight HAMILTON verifies the validity of the product rule for his quaternion multiplication, and writes (Math. Papers 3, p. 108): “But I considered it essential to try whether [my] equations were consistent with the law of moduli, . . . , without which consistence being verified, I should have regarded the whole speculation as a failure.”

Neither HAMILTON nor anyone else at the time was aware that EULER had already been in possession of the characteristic laws applying to quaternions, as early as 1748. In a letter to GOLDBACH on the 4th of May he gives the product rule in the form of the “four squares theorem” (see 2.3 on this point). GAUSS also knew about the rules for calculating with quaternions; he wrote in 1819 a short note (not published at the time) on “Mutations of space,” in which the quaternion formulae appear (Werke 8, 357–362).

4. HAMILTON regarded the creation of his quaternions as being on a par with the creation of the infinitesimal calculus. He acknowledged no contemporary mathematicians other than GAUSS and GRASSMANN as having played any part. F. ENGEL, on page 208 of his very readable account of GRASSMANN’s life, wrote: “Grassmann teilt sich mit Gauß in die Ehre, daß Hamilton ihm zutraut, er könne die Quaternionen gefunden haben, und sich immer von Neuem freut, daß es allem Anscheine nach doch nicht der Fall ist” (“Grassmanns Leben,” Teubner Verlag, Leipzig 1911). [GRASSMANN shares with GAUSS the honor, accorded to him by HAMILTON, that he (GRASSMANN) could have discovered quaternions, and that he (HAMILTON) is always delighted with the news that to all appearances this is not the case.]

HAMILTON believed that his quaternions would play a key role in physics. With missionary zeal he strove to get them accepted by the mathematical world. Thus in Dublin, quaternions became an official examination subject; a “cosmic” significance was attributed to them. Felix KLEIN in his well-known Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert (Vol. 1, p. 184) [Lectures on the development of mathematics in the 19th century] gave a very harsh judgement when he wrote: “Hamilton selbst gestaltete sie [= Quaternionen] für sich zu einer Art orthodoxer Lehre des mathematischen Credo, in die er alle seine geometrischen und sonstigen Interessen hineinzwang, je mehr sich gegen Ende seines Lebens sein Geist vereinseitigte und . . . .” [Hamilton himself regarded quaternions as a kind of orthodox doctrine of the mathematical Credo, into which all his geometrical and other interests were forced, and this tendency became
more pronounced as towards the end of his life his mind set and he became obsessed with a single idea...]

5. In Ireland and England, HAMILTON became the figurehead of a school of “quaternionists” who “outdid their master in intolerance and rigidity.” At the center stood a mystic formalism treated with due reverence by the initiated. One dreamt of a quaternionistic theory of functions and expected to gain new and profound insights into the whole realm of mathematics. To promote these utopian aims there was even founded, in 1895, an “International Association for promoting the study of quaternions,” at Yale University in New Haven, Connecticut. Even now there are still faint echoes from the great days of the quaternionists in Ireland. Thus Eamon de VALERA, the President of Ireland from 1959 to 1973, during his period of office, would occasionally attend a mathematical colloquium in Dublin, whenever the announcement of the discourse contained the word “quaternions.”

The history of algebra has shown that the significance of quaternions was vastly overestimated in the last century. Nowadays it has become clear that the quaternion algebra is only a particular algebra of complex $2 \times 2$ matrices (see §1). It was not the discovery of quaternions which was the great achievement, but rather the recognition which came about as a result of that discovery, of the great freedom which one has available, to construct hypercomplex systems. Lord KELVIN (1824–1907) the famous Scottish physicist and writer on thermodynamics, commented caustically: “Quaternions came from Hamilton after his really good work had been done; and though beautifully ingenious, have been an unmixed evil to those who have touched them in any way.”

In contrast to this opinion is a well-known saying by Thomas HILL (who was a student of B. PEIRCE, the President of Harvard in 1862): “In the great mathematical birth of 1843, the Quaternions of HAMILTON, there is as much real promise of benefit to mankind as in any event of Victoria’s reign.”

We refer readers who would like further historical details to:


§1. The Quaternion Algebra $\mathbb{H}$

We introduce quaternions in §1.1, following Hamilton's example, by means of the multiplication table for the natural basis. In §1.2 quaternions are represented as special complex $2 \times 2$ matrices. A subalgebra $\mathcal{H}$ of $\text{Mat}(2, \mathbb{C})$ and a natural isomorphism $F: \mathbb{H} \to \mathcal{H}$ of the quaternion algebra $\mathbb{H}$ onto $\mathcal{H}$ is constructed, which was already known to Cayley in 1858. With this isomorphism it becomes obvious among other things that $\mathbb{H}$ is an associative division algebra over $\mathbb{R}$. Hamilton had to find a direct verification of the associativity of $\mathbb{H}$, because in the year of the discovery 1843, matrices were as yet unknown. It was not until 1858, that Cayley introduced matrices and the matrix calculus in his "A memoir on the theory of matrices" (Math. Papers 2, 475–496), which includes the quaternion calculus as a special case. The algebra $\mathcal{H}$ and the isomorphism $F$ can be usefully applied throughout the whole of this chapter.

In paragraphs §1.3 to §1.7 the basic algebraic properties of the quaternions will be discussed.

1. The Algebra $\mathbb{H}$ of the Quaternions. In the four-dimensional $\mathbb{R}$-vector space $\mathbb{R}^4$ of ordered real number quadruples, we choose the standard basis

$$e_1 := (1,0,0,0), \quad e_2 := (0,1,0,0), \quad e_3 := (0,0,1,0), \quad e_4 := (0,0,0,1).$$

We now introduce the so-called Hamiltonian multiplication. Let $e_1$, be the unit element; then the nine products $e_\mu e_\nu$, $2 \leq \mu, \nu \leq 4$, still have to be specified, and we define them by the following relations

$$
\begin{align*}
    e_2 e_2 &:= -e_1, & e_2 e_3 &:= e_4, & e_2 e_4 &:= -e_3 \\
    e_3 e_2 &:= -e_4, & e_3 e_3 &:= -e_1, & e_3 e_4 &:= e_2 \\
    e_4 e_2 &:= e_3, & e_4 e_3 &:= -e_2, & e_4 e_4 &:= -e_1
\end{align*}
$$

(HAMILTON relations)

This is often set out in the form of a multiplication table

<table>
<thead>
<tr>
<th></th>
<th>$e_2$</th>
<th>$e_3$</th>
<th>$e_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_2$</td>
<td>$-e_1$</td>
<td>$e_4$</td>
<td>$-e_3$</td>
</tr>
<tr>
<td>$e_3$</td>
<td>$-e_4$</td>
<td>$-e_1$</td>
<td>$e_2$</td>
</tr>
<tr>
<td>$e_4$</td>
<td>$e_3$</td>
<td>$-e_2$</td>
<td>$-e_1$</td>
</tr>
</tbody>
</table>

The four-dimensional real $\mathbb{R}$-algebra constructed in this way is called the quaternion algebra and denoted by $\mathbb{H}$. The elements of $\mathbb{H}$ were given the name of quaternions by Hamilton. Since $e_2 e_3 \neq e_3 e_2$ it is clear that the quaternion algebra $\mathbb{H}$ is not commutative.

---

2 The word means any group of four persons or things, and was used, for example, in the New Testament to describe the four groups of four soldiers used by King Herod to guard Peter. (Acts of the apostles, 12, 4): "... he put him in prison, and delivered him to four quaternions of soldiers to keep him" (see Temple 100 years of mathematics, London, Duckworth, 1981, p. 46).
§1. The Quaternion Algebra \( \mathbb{H} \)

The validity of the 27 equations \((e_\lambda e_\mu)e_\nu = e_\lambda(e_\mu e_\nu), \ 2 \leq \lambda, \mu, \nu \leq 4,\) can be checked directly from the multiplication table, thus verifying that the quaternion algebra is associative. We refrain from doing this because associativity and more will emerge in the next paragraph in a more elegant way. Traditionally \(e_1, e_2, e_3, e_4\) are denoted by \(e, i, j, k\) respectively so that

\[
i^2 = j^2 = k^2 = ij = k, \quad ij = -ji = k.
\]

The other products are derived from these by cyclic interchange of \(i, j, k\).

Using the distributive law we thus obtain the

**Product formula:**

\[
(\alpha e + \beta i + \gamma j + \delta k)(\alpha' e + \beta' i + \gamma' j + \delta' k) = (\alpha \alpha' - \beta \beta' - \gamma \gamma' - \delta \delta')e + (\alpha \beta' + \beta \alpha' + \gamma \delta' - \delta \gamma')i + (\alpha \gamma' - \beta \delta' + \gamma \alpha' + \delta \beta')j + (\alpha \delta' + \beta \gamma' - \gamma \beta' + \delta \alpha')k.
\]

The classical method of writing quaternions with the symbols \(i, j, k\) has certain hidden dangers, for example, if we try to deal with quaternions with complex instead of real numbers as coefficients.

\(\mathbb{R}e\) is an \(\mathbb{R}\)-subalgebra of \(\mathbb{H}\). In contrast to our practice with \(\mathbb{C}\), we do not however identify \(\mathbb{R}e\) with \(\mathbb{R}\), and therefore we consistently write \(e\) and not 1 for the unit element of \(\mathbb{H}\).

**2. The Matrix Algebra \( \mathcal{H} \) and the Isomorphism \( F: \mathbb{H} \to \mathcal{H} \).** The set \( \mathcal{C} \) of all real \(2 \times 2\) matrices \(\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}\), \(\alpha, \beta \in \mathbb{R}\), is an \(\mathbb{R}\)-subalgebra of \(\text{Mat}(2, \mathbb{R})\), and the mapping \(\alpha + \beta i \mapsto \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}\) is an \(\mathbb{R}\)-algebra isomorphism \(\mathbb{C} \to \mathcal{C}\) (see 3.2.5). In analogy with this, we have the following.

**Theorem.** The set \( \mathcal{H} := \left\{ \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix} : w, z \in \mathbb{C}\right\} \) is an \(\mathbb{R}\)-subalgebra of \(\text{Mat}(2, \mathbb{C})\), with unit element \(E := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\). Every matrix \(A = \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix}\) \(\in \mathcal{H}\) satisfies, over \(\mathbb{R}\), the quadratic equation

\[
(1) \quad A^2 - (\text{trace } A)A + (\det A)E = 0
\]

where \(\text{trace } A = 2 \text{ Re } w, \ \det A = |w|^2 + |z|^2\)

\(\mathcal{H}\) is a 4-dimensional, associative division algebra.

**Proof.** 1) It is easily verified, by direct calculation, that \(\mathcal{H}\) is a four-dimensional \(\mathbb{R}\)-vector subspace of \(\text{Mat}(2, \mathbb{C})\) which is closed under matrix multiplication. The matrix equation \(A^2 - (\text{trace } A)A + (\det A)E = 0\) can be checked in the same way.
Hamilton's Quaternions

2) The algebra $\mathcal{H}$ is associative because $\text{Mat}(2, \mathbb{C})$ is. To see that $\mathcal{H}$ is a division algebra we need to use the criterion R.5. Accordingly suppose $A, B \in \mathcal{H}$ and $AB = 0$. It then follows that $\det A \cdot \det B = 0$, and hence $\det A = 0$ or $\det B = 0$. As $\det \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix} = |w|^2 + |z|^2$ vanishes only for $w = z = 0$, the required statement follows. $\square$

The equation (1) is the statement of the so-called Cayley theorem (or of the Cayley–Hamilton theorem for the special case of $2 \times 2$ matrices) (See S. Lang, An Introduction to Linear Algebra, 2nd ed., Springer-Verlag.)

Lemma. The mapping

$$F: \mathbb{H} \rightarrow \mathcal{H}, \quad (\alpha, \beta, \gamma, \delta) \mapsto \begin{pmatrix} \alpha + \beta i & -\gamma - \delta i \\ \gamma - \delta i & \alpha - \beta i \end{pmatrix},$$

is an $\mathbb{R}$-algebra isomorphism, and

$$F(e_1) = E, \quad F(e_2) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} =: I,$$

$$F(e_3) = \begin{pmatrix} 0 & -1 \\ +1 & 0 \end{pmatrix} =: J, \quad F(e_4) = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix} =: K.$$

Proof. The mapping $F$ is obviously $\mathbb{R}$-linear and bijective. It remains to be shown (see R.6) that $F(e_\mu)F(e_\nu) = F(e_\mu e_\nu)$ for $\mu, \nu = 1, 2, 3, 4$. This however is clear, because the matrices $E, I, J, K$ are the images under $F$ of $e_1, e_2, e_3, e_4$ and satisfy the same laws of multiplication as $e_1, e_2, e_3, e_4$. The relations $I^2 = J^2 = -E, IJ = -JI = K$ are easily checked and the remaining relations can be derived from the associative law, for example, $K^2 = (IJ)(-JI) = -IJ^2I = I^2 = -E.$ $\square$

Corollary. The Hamiltonian algebra $\mathbb{H}$ is an associative division algebra.

By Lemma R.5, $\mathcal{H} \setminus \{0\}$ is a group with respect to multiplication. One can immediately verify that:

The set $\{E, -E, I, -I, J, -J, K, -K\}$ is a noncommutative subgroup of $\mathcal{H} \setminus \{0\}$, each of whose elements other than $\pm E$ is of order 4.

This group, and any group isomorphic to it, is known, in the literature, as the (finite) quaternion group.

The representation of quaternions by complex $2 \times 2$ matrices which we have used here was already familiar to Cayley in 1858. In his famous "Memoir on the theory of matrices" (Math. Papers 2, p. 491) he writes: "It may be noticed in passing, that if $L, M$ are skew convertible matrices
§1. The Quaternion Algebra \( \mathbb{H} \)

of the order 2, and if these matrices are also such that \( L^2 = -1, M^2 = -1 \), then putting \( N = LM = -ML \), we obtain

\[
L^2 = -1, \quad M^2 = -1, \quad N^2 = -1,
\]

\[
L = MN = -NM, \quad M = NL = -NL[sic], \quad N = LM = -ML,
\]

which is a system of relations precisely similar to that in the theory of quaternions.” CAYLEY does not however give explicit examples for \( L, M \).

As calculations with complex matrices can be performed more elegantly than with quaternions, it is often preferable—as above—to prove theorems about \( \mathbb{H} \), by first proving them for the algebra \( \mathcal{H} \), and then using the isomorphism \( F: \mathbb{H} \to \mathcal{H} \) to “lift” them to \( \mathbb{H} \). We shall use this principle again later.

As with complex numbers earlier, there are many possible ways of representing the quaternion algebra \( \mathbb{H} \) as an \( \mathbb{R} \)-subalgebra of \( \text{Mat}(2, \mathbb{C}) \). One can choose three matrices \( I_2, I_3, I_4 \in \text{Mat}(2, \mathbb{C}) \), in any way one likes as long as the nine Hamiltonian conditions are satisfied. The mapping

\[
\mathbb{H} \to \text{Mat}(2, \mathbb{C}), \quad (\alpha, \beta, \gamma, \delta) \mapsto \alpha E + \beta I_2 + \gamma I_3 + \delta I_4
\]

is then an \( \mathbb{R} \)-algebra monomorphism. It can be shown, as a generalization of the theorem in 3.2.5, that:

If \( g: \mathbb{H} \to \text{Mat}(2, \mathbb{C}) \) is an \( \mathbb{R} \)-algebra monomorphism, then there is an invertible matrix \( W \in \text{Mat}(2, \mathbb{C}) \), such that the associated “inner automorphism” \( \iota_W: \text{Mat}(2, \mathbb{C}) \to \text{Mat}(2, \mathbb{C}), A \mapsto W^{-1}AW \) has the property \( g = \iota_W \circ F \).

3. The Imaginary Space of \( \mathbb{H} \). We use the standard basis \( e, i, j, k \). The three-dimensional vector subspace

\[
(1) \quad \text{Im} \mathbb{H} := \mathbb{R}i + \mathbb{R}j + \mathbb{R}k
\]

of \( \mathbb{H} \) is called—in analogy to the complex numbers—the imaginary space of \( \mathbb{H} \). Its elements are called “purely imaginary.” \( \mathbb{H} \) is a direct sum of the vector spaces \( \mathbb{R}e \) and \( \text{Im} \mathbb{H} \)

\[
(2) \quad \mathbb{H} = \mathbb{R}e \oplus \text{Im} \mathbb{H}.
\]

The line \( \mathbb{R}e \) is defined invariantly by the unit element \( e \). The definition of \( \text{Im} \mathbb{H} \) is initially dependent on the basis. In order to characterize \( \text{Im} \mathbb{H} \) invariantly, we note that the quaternion \( x = \alpha e + \beta i + \gamma j + \delta k \) satisfies, by Theorem 2, the quadratic equation

\[
(3) \quad x^2 = 2\alpha x - (\alpha^2 + \beta^2 + \gamma^2 + \delta^2)e.
\]
As \( x \in \text{Im} \mathbb{H} \) if and only if \( \alpha = 0 \), we obtain the basis-free representation

\[
(4) \quad \text{Im} \mathbb{H} = \{ x \in \mathbb{H} : x^2 \in \mathbb{R} \text{ and } x \notin \mathbb{R} \setminus \{0\} \}.
\]

\( \text{Im} \mathbb{H} \) is not an \( \mathbb{R} \)-subalgebra of \( \mathbb{H} \). We note that:

For purely imaginary quaternions \( u, v \), the following relations hold: \( u^2 = -\omega e \) with \( \omega \geq 0 \) and \( uv + vu \in \mathbb{R} \).

**Proof.** If \( u = \beta i + \gamma j + \delta k \) then \( u^2 = -(\beta^2 + \gamma^2 + \delta^2)e \) with \( \beta^2 + \gamma^2 + \delta^2 \geq 0 \). Since \( u, v, u + v \) all belong to \( \text{Im} \mathbb{H} \), it follows that \( uv + vu = (u + v)^2 - u^2 - v^2 \in \mathbb{R} \). \( \square \)

In particular for every \( u \in \text{Im} \mathbb{H}, u \neq 0 \), a scalar \( \rho \) (namely \( \rho := \sqrt{\omega^{-1}} \)) can be found such that \( (\rho u)^2 = -e \) (normalization).

The imaginary space \( \text{Im} \mathbb{H} \) plays a dominant role in the theory of quaternions. Its elements are also called vectorial (or pure) quaternions. The expression “vector” first appears in HAMILTON’s writings in 1845, (Q. Jl. Math. 1, p. 56). In the long drawn-out war of resistance against the vector calculus, Lord KELVIN was even in 1896 expressing the opinion that: “Vector is a useless survival, or offshoot from quaternions, and has never been of the slightest use to any creature.”

By (2) every quaternion \( x \) can be expressed uniquely in the form

\[
(5) \quad x = \alpha e + u \quad \text{with} \quad \alpha \in \mathbb{R} \text{ and } u \in \text{Im} \mathbb{H}.
\]

In this expression \( \alpha e \) is sometimes called the **scalar part** (or **real part**) and \( u \) the **vector(ial) part** (or **imaginary part**) of \( x \).

Every plane in \( \mathbb{H} \), containing the straight line \( \mathbb{R}e \), is a subalgebra of \( \mathbb{H} \), isomorphic to \( \mathbb{C} \). It is however fundamentally impossible to make \( \mathbb{H} \) “somehow or other” into a \( \mathbb{C} \)-algebra.\(^3\)

4. **Quaternion Product, Vector Product and Scalar Product.** For vectorial quaternions \( u = \beta i + \gamma j + \delta k, v = \rho i + \sigma j + \tau k \) we have

\[
(1) \quad uv = -(\beta \rho + \gamma \sigma + \delta \tau)e + (\gamma \tau - \delta \rho)i + (\delta \rho - \beta \tau)j + (\beta \sigma - \gamma \rho)k.
\]

Here the “scalar part” is, apart from sign, the canonical Euclidean scalar product \( \langle u, v \rangle \) of the vectors \( u = (\beta, \gamma, \delta), v = (\rho, \sigma, \tau) \in \mathbb{R}^3 \); the “vectorial

\[^{3}\text{In fact we can improve on this with the following.}

**Theorem.** Every finite dimensional complex division algebra with unit element is isomorphic to \( \mathbb{C} \).

This is proved in the same way as the theorem in the footnote on page 190: the left-multiplication \( L_a \) now has a complex eigenvalue \( \lambda \) (Fundamental theorem of algebra).
part" of $uv$ is the vector product (cross product) of these two vectors. We thus obtain the aesthetically pleasing formula

\[(2) \quad uv = -(u, v)e + u \times v, \quad u, v, u \times v \in \text{Im} \mathbb{H},\]

The mapping $(u, v) \mapsto u \times v$ is by definition bilinear and anticommutative:

\[(3) \quad u \times v = -v \times u, \quad u, v \in \text{Im} \mathbb{H}.\]

From (3) we get immediately using (2)

\[(4) \quad u \times v = \frac{1}{2}(uv - vu), \quad (u, v)e = -\frac{1}{2}(uv + vu) \text{ for all } u, v \in \text{Im} \mathbb{H}.\]

The vector product is not associative. We note that

\[(5) \quad u \times (v \times w) = \frac{1}{2}(uvw - vw)u, \quad u, v, w \in \text{Im} \mathbb{H}.\]

Proof. Since $uvw = -(v, w)u + u(v \times w)$ and $vwu = -(v, w)u + (v \times w)u$ by (2), the identity (5) follows from $u(v \times w) - (v \times w)u = 2u \times (v \times w)$. □

Exercise. Show that every quaternion $a \in \mathbb{H}$ can be represented (in infinitely many different ways) as the product $a = bc$ of two purely imaginary quaternions $b, c$.

As a substitute to some extent for the associative law we have the

\textbf{GRASSMANN Identity:} $u \times (v \times w) = (u, w)v - (u, v)w.$

Proof. This follows from (5) with the help of (3), since

\[uvw - vw = (uv + vu)w - v(uw + wu) = -2(u, v)w + 2(u, w)v. \quad \Box\]

If we introduce $u, v, w$ cyclically in (5) or in the GRASSMAN identity, and add, we obtain the

\textbf{JACOBI Identity:} $u \times (v \times w) + v \times (w \times u) + w \times (u \times v) = 0.$

This identity and (4) assert that the span $\mathbb{R}^3 \cong \text{Im} \mathbb{H}$, with the vector product, is a Lie algebra (see R.2.3).

It follows directly from (1), or from (2) with the product rule 2.2(4), that:

\[(6) \quad (u, v)^2 + |u \times v|^2 = |u|^2|v|^2.\]

This is a strengthened version of the Cauchy–Schwarz inequality. If we write $(u, v) = |u||v|\cos \varphi$ with $\varphi \in [0, \pi]$, we obtain

\[|u \times v| = |u||v|\sin \varphi.]
Thus $|u \times v|$ is the area of the parallelogram spanned by the vectors $u, v$. The equation (6) plays a central role in the theory of vector product algebras (see 10.3.1).

The triple (scalar) product of three vectors $u, v, w \in \text{Im}\mathbb{H}$ is the real number $(u \times v, w)$. Since $\mathbb{R}e$ and $\text{Im}\mathbb{H}$ are orthogonal, it follows from (2) that

$$\langle u \times v, w \rangle = \langle uv, w \rangle, \quad u, v, w \in \text{Im}\mathbb{H}.$$ 

We can immediately deduce from (1) that $\langle u \times v, u \rangle = 0$. After replacing $u$ by $u + v$, and taking account of (4), we get the

**Interchange Rule:** $\langle u \times v, w \rangle = \langle u, v \times w \rangle$, $u, v, w \in \text{Im}\mathbb{H}$.

It is at once clear from this that the mapping $(u, v, w) \mapsto \langle u \times v, w \rangle$ is a determinant function of the vector space $\text{Im}\mathbb{H}$.

**Exercise.** Show that, for $2 \times 2$ matrices $A, B, C$

$$[A, [B, C]] = \{2\sigma(AB) - \sigma(A)\sigma(B))C - \{2\sigma(AC) - \sigma(A)\sigma(C))B \]

$$

$$+ \{\sigma(B)\sigma(AC) - \sigma(C)\sigma(AB)\}E,$$

where $[A, B] := AB - BA$, $\sigma(A) := \text{trace } A$, $E := \text{unit matrix}$. Show also that, in the case where $A, B, C \in F(\text{Im}\mathbb{H})$, this is the GRASSMANN identity.

**Historical Remarks.** Vector multiplication was discovered by H. GRASSMANN in 1844 (one year after HAMILTON's discovery of quaternions) as a special case of a much more general so-called "exterior product." The algebra of vectors in $\mathbb{R}^3$ first became popular however in the eighties of the last century through the works of the American physicist and mathematician Josiah Willard GIBBS (1839-1903) who was a professor at Yale University. GIBBS maintained amongst other arguments—what seems to us nowadays almost self-evident—that the scalar product $(u, v)$ and the vector product $u \times v$ have their own independent meaning and that the quaternion product $uv$ in which these two products are combined with one another has no essential significance in many problems. GIBBS was an opponent of the quaternionists, and it was because of this controversy that a colleague of GIBBS founded in 1895 at Yale the association for the worldwide promotion of quaternions, mentioned earlier in the introduction.

5. **Noncommutativity of $\mathbb{H}$. The Center.** The fact that $\mathbb{H}$ is not commutative leads to many unusual consequences. Thus polynomials can have more zeros than is indicated by their degree. For example, the quadratic polynomial $X^2 + e$ has, as zeros, all purely imaginary quaternions $u = \beta i + \gamma j + \delta k$, whose "length" $\beta^2 + \gamma^2 + \delta^2$ equals one. These quaternions
represent the surface of the unit sphere in the three-dimensional space $\mathbb{R}^3$ or the real number triples $(\beta, \gamma, \delta)$.

Another statement we can make is that:

*There are cubic polynomials over $\mathbb{H}$, for example, $X^2iXi + iX^2iX - iXiX^2 - XixX^2i$, which assume the value zero for all quaternions.*

As every quaternion satisfies an equation $X^2 = \alpha X + \beta e$ the truth of this statement can be proved by substituting $\alpha X + \beta e$ for $X^2$ in the above polynomial. \(\square\)

Since $\mathbb{H}$ is not commutative, the naive definition of determinants fails. For example neither

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc \quad \text{nor} \quad \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - cb$$

would be a suitable definition. In the first case, we would have

$$\det \begin{pmatrix} i & j \\ i & j \end{pmatrix} = ij - ji = 2k \neq 0,$$

and in the second case we would have

$$\det \begin{pmatrix} i & i \\ j & j \end{pmatrix} = ij - ji \neq 0,$$

so that neither determinant would vanish even though the first has equal columns and the second equal rows.

To measure the departure from commutativity of an algebra $\mathcal{A}$, we consider its center

$$Z(\mathcal{A}) := \{z \in \mathcal{A} : zx = xz \text{ for all } x \in \mathcal{A}\}.$$ 

If $\mathcal{A}$ is associative, $Z(\mathcal{A})$ is a subalgebra of $\mathcal{A}$, and $Z(\mathcal{A}) = \mathcal{A}$ if and only if $\mathcal{A}$ is commutative. For algebras with a unit element $e$, $\Re e \in Z(\mathcal{A})$. The extreme case $Z(\mathcal{A}) = \Re e$ can occur.

*For the algebra $\mathbb{H}$ we have $Z(\mathbb{H}) = \Re e = \{x \in \mathbb{H} : xu = ux \text{ for all } u \in \text{Im} \mathbb{H}\}$. *

This is included in the following statement:

*For all $u \in \mathbb{H} \setminus \Re e$, $\{x \in \mathbb{H} : xu = ux\} = \Re e + \Re u$. *

**Proof.** Since $\{x \in \mathbb{H} : xu = ux\} = \{x \in \mathbb{H} : xv = vx\}$ for $v := u - (\Re u)e$ one can assume that $u \in \text{Im} \mathbb{H}$, $u \neq 0$. One can even assume $u^2 = -e$ and

---

4 The reason for this phenomenon is that polynomials over $\mathbb{H}$ no longer factor in the usual way. For example $(X - x)(X - y) = X^2 - xX - Xy + xy$, and the linear terms cannot be combined into $-(x + y)X$. 

---
$x^2 = -e$ (we pass from $x$ to $x - \alpha e$ and normalize). It then follows that $(x-u)(x+u) = x^2 - ux + xu - u^2 = 0$, so that $x = \pm u$. 

The noncommutativity of $\mathbb{H}$ is also the reason why $\mathbb{H}$ has many $\mathbb{R}$-algebra automorphisms: every $a \in \mathbb{H}, a \neq 0$ induces a so-called inner automorphism $h_a: \mathbb{H} \to \mathbb{H}, x \mapsto axa^{-1}$. Since $Z(A) = \mathbb{R}e$, we have $h_a = h_b$, if and only if $b^{-1}a \in \mathbb{R}e$. We shall show in 3.2 that the $\mathbb{R}$-algebra $\mathbb{H}$ has no other automorphisms.

**Exercise.** Show that, for any two elements $a, b \in \mathbb{H}$, the following statements are equivalent:

i) $ab = ba$.

ii) $e, a, b$ are linearly independent.

iii) there is a subalgebra of $\mathbb{H}$, isomorphic to $\mathbb{C}$, and containing $a$ and $b$.

6. The Endomorphisms of the $\mathbb{R}$-Vector Space $\mathbb{H}$. For any two quaternions $a, b$ the mapping $\mathbb{H} \to \mathbb{H}, x \mapsto axb$ is an $\mathbb{R}$-linear mapping of $\mathbb{H}$ into itself (an endomorphism). We denote by $\text{End} \mathbb{H}$ the $\mathbb{R}$-vector space of all endomorphisms of $\mathbb{H}$.

**Theorem.** If $a_1, \ldots, a_4$ is a basis of $\mathbb{H}$, the mapping $\mathbb{H}^4 \to \text{End} \mathbb{H}$,

$$(b_1, b_2, b_3, b_4) \mapsto f \in \text{End} \mathbb{H} \text{ with } f(x) := \sum_{\nu=1}^{4} a_{\nu}xb_{\nu}$$

is $\mathbb{R}$-linear and bijective.

**Proof.** The $\mathbb{R}$-linearity is obvious. Since $\dim \mathbb{H}^4 = \dim (\text{End} \mathbb{H}) = 16$, as $\dim \mathbb{H} = 4$, it only remains to prove the injectivity of the mapping in question. This is the case $n = 4$ of the following auxiliary proposition:

Let $n = 1, 2, 3$ or $4$ and suppose $\sum_{1}^{n} a_{\nu}xb_{\nu} = 0$ for all $x \in \mathbb{H}$ then $b_1 = \cdots = b_n = 0$.

We argue by induction, the case $n = 1$ being clear. Suppose $n > 1$, then if $b_1$ were not zero, we should have

$$a_1x + \sum_{2}^{n} a_{\nu}xq_{\nu} = 0 \text{ with } q_{\nu} := b_{\nu}b_1^{-1}.$$  

If we now multiply this equation on the right by $y$, and subtract from it the equation obtained by replacing $x$ by $xy$ (in the original equation) we obtain

$$\sum_{2}^{n} a_{\nu}x(q_{\nu}y - yq_{\nu}) = 0 \text{ and hence } q_{\nu}y = yq_{\nu} \text{ for all } y \in \mathbb{H}$$
by the inductive hypothesis. Since \( Z(\mathbb{H}) = \mathbb{Re} \) it follows that \( q_\nu = \alpha_\nu e, \alpha_\nu \in \mathbb{R} \). We now deduce from (*) that

\[
\left( a_1 + \sum_{2}^{n} \alpha_\nu a_\nu \right) x = 0, \quad \text{that is} \quad a_1 + \sum_{2}^{n} \alpha_\nu a_\nu = 0,
\]

or in other words \( a_1, \ldots, a_4 \) would be linearly dependent. It follows that \( b_1 = 0 \), and similarly \( b_2 = b_3 = b_4 = 0 \). \( \Box \)

**Example.** The conjugation \( x \mapsto \bar{x} \) (see §2.1) belongs to \( \text{End} \mathbb{H} \), and with respect to the basis \( 1, i, j, k \), we have:

\[
\bar{x} = \frac{1}{2}(x + ixi + jxj + kxk).
\]

The theorem proved here is to be found in HAMILTON's work *Elements of quaternions*, which was published by his son in 1866. The analogue of this theorem does not hold for the field \( \mathbb{C} \), where (see 3.3.1) the \( \mathbb{R} \)-linear mappings \( \mathbb{C} \to \mathbb{C} \) have the form \( z \mapsto az + \bar{b}z \). The fact that one can work without conjugate quaternions is really tied in with the fact that \( \mathbb{H} \) has a one-dimensional center, whereas this is not true of \( \mathbb{C} \).

## 7. Quaternion Multiplication and Vector Analysis

HAMILTON applied quaternion multiplication to derive important formulae in vector analysis in an elegant fashion. He introduced the “Nabla” operator

\[
\nabla := \frac{\partial}{\partial x} i + \frac{\partial}{\partial y} j + \frac{\partial}{\partial z} k
\]

(he chose the word nabla because of the similarity of the shape of the symbol to that of an Hebrew musical instrument of that name). The application of \( \nabla \) to a differentiable function \( f(x, y, z) \) of three real variables, gives the gradient of \( f \)

\[
\nabla f := \frac{\partial f}{\partial x} i + \frac{\partial f}{\partial y} j + \frac{\partial f}{\partial z} k = \text{grad} f.
\]

Application of \( \nabla \) to a “differentiable quaternion field” \( F(x, y, z) = u(x, y, z)i + v(x, y, z)j + w(x, y, z)k \) gives, when formally expanded:

\[
\nabla F = - \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) e + \left( \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \right) i
\]

\[
+ \left( \frac{\partial u}{\partial z} - \frac{\partial w}{\partial x} \right) j + \left( \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \right) k.
\]

The real part, up to sign, is the divergence of \( F \), and the imaginary part is the curl \( F \), of the vector field \( F \):

\[
\nabla F = -\text{div} F + \text{curl} F.
\]
Applying the operator $\nabla$ twice to a function $f$ leads to the well-known LAPLACE operator of potential theory, more precisely:

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

All this works amazingly well. Felix KLEIN, in the first volume of his Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert (p. 188) writes: “Die Leichtigkeit und Eleganz ist in der Tat überraschend, und es läß sich wohl von hier aus die alles andere ablehnende Begeisterung der Quaternionisten für ihr System begreifen, die bald über vernünftige Grenzen hinauswuchs, in einer weder der Mathematik als Ganzem noch der Quaternionentheorie selbst förderlichen Weise.” [The ease and elegance is indeed astonishing, and may well account for the enthusiasm of the quaternionists for their system and their rejection of all others; an enthusiasm which soon outgrew all reasonable bounds and advanced neither the theory of quaternions itself nor mathematics as a whole.]

8. The Fundamental Theorem of Algebra for Quaternions. It is easily seen that

Every polynomial $X^n - a$, $a \in \mathbb{H}$, of degree $n > 0$ has zeros in every plane in $\mathbb{H}$ containing 0, e, and $a$.

**Proof.** Every such plane $E$ is a subalgebra of $\mathbb{H}$, isomorphic to $\mathbb{C}$, and so $X^n - a$, by the fundamental theorem of algebra for complex numbers, always has zeros in $E$. $\square$

The number of zeros of $X^n - a$ can be infinite, for example, for $X^2 + e$.

The reader may care to show that:

If $\text{Im} a \neq 0$, then $X^n - a$ has exactly $n$ zeros in $\mathbb{H}$. $\square$

The exponential series

$$\exp x := \sum_{n=0}^{\infty} \frac{x^n}{n!}, \quad x \in \mathbb{H}$$

converges absolutely, and uniformly on compact subsets (that is, w.r.t. the norm) of $\mathbb{H}$. By multiplication of series, or by reduction to the complex case (see the proof above and Exercise 7.1.5), one obtains the

**Addition Theorem.** $(\exp x)(\exp y) = \exp(x + y)$, if $xy = yx$.

Quaternions have a “representation in polar coordinates”:

$$a = |a|(\exp u) \quad \text{with} \quad u \in \text{Im} \mathbb{H}.$$
§1. The Quaternion Algebra \( \mathbb{H} \)

To see this, we again consider a plane in \( \mathbb{H} \), containing 0, \( e \), and \( a \), and transfer to it what we know about \( \mathbb{C} \). With this representation, the roots of \( X^n - a \) can be given an explicit solution:

\[
\text{if } a = |a| (\exp \alpha), \text{ then } b := \sqrt[2]{|a|} (\exp \frac{\alpha}{n}) \text{ is a zero of } X^n - a. \quad \Box
\]

It is not immediately apparent that quadratic polynomials \( X^2 + aX + b \), \( a, b \in \mathbb{H} \), always have zeros in \( \mathbb{H} \) (the reduction to pure polynomials by "completing the square" only works if \( ab = ba \)). Nevertheless as we shall see a fundamental theorem of algebra does in fact hold for \( \mathbb{H} \) as well. We first define, inductively, the concept of a monomial (over \( \mathbb{H} \)). Any constant \( a \neq 0 \) is a monomial of degree 0. The "indeterminate" \( X \) is a monomial of degree 1. If \( m_1 \) and \( m_2 \) are monomials of degree \( k_1 \) and \( k_2 \) respectively, their product \( m_1 m_2 \) is a monomial of degree \( k_1 + k_2 \). The general monomial of the \( n \)th degree accordingly has the form

\[
a_1 X^{n_1} a_2 X^{n_2} a_3 \cdots a_r X^{n_r} a_{r+1}, a_p \in \mathbb{H} \setminus \{0\}, \quad n = n_1 + n_2 + \cdots + n_r.
\]

Any finite sum of monomials is said to be a polynomial (over \( \mathbb{H} \)). Every polynomial \( p \) defines a continuous mapping \( p: \mathbb{H} \rightarrow \mathbb{H}, x \mapsto p(x) \).

**Fundamental Theorem of Algebra for Quaternions.** Let \( p \) be a polynomial over \( \mathbb{H} \) of degree \( n > 0 \) of the form \( m + q \) where \( m \) is a monomial of degree \( n \) and \( q \) a polynomial of degree \( < n \). Then the mapping \( p: \mathbb{H} \rightarrow \mathbb{H} \) is surjective, and in particular \( p \) has zeros in \( \mathbb{H} \).

**Remark.** The hypothesis that only one monomial of highest degree is present in \( p \), is essential to the validity of this theorem. Thus, for example, the linear polynomial \( iX - Xi + 1 \) has no zero in \( \mathbb{H} \) (because for any \( a \in \mathbb{H} \), the polynomial \( aX - Xa \) assumes values only in \( \text{Im} \mathbb{H} \), since for all \( a = \alpha e + u, x = \xi e + v, u, v \in \text{Im} \mathbb{H} \), we have \( ax - xa = 2u \times v \)). \( \Box \)

The usual proofs of the fundamental theorem for \( \mathbb{C} \) do not carry over to \( \mathbb{H} \). With the help of the more powerful methods of topology a proof can be given as follows. Since in \( p \), the monomial of the \( n \)th degree dominates the remaining terms for large \( x \in \mathbb{H} \), we have the growth equation:

\[
\lim_{x \to \infty} |p(x)| = \infty.
\]

Thus \( p \) can be extended to a continuous mapping \( \hat{p}: S^4 \to S^4 \) of the four-dimensional sphere into itself with \( \hat{p}(\infty) := \infty \). \( S^4 \) is regarded as the compactification of \( \mathbb{H} \cong \mathbb{R}^4 \) by the addition of a point \( \infty \). It can now be shown that the mapping \( \hat{p} \) has degree \( n \) (in the sense of topology). Since \( n \neq 0 \), it follows from a general theorem of topology that \( \hat{p} \) is surjective, and this means that \( p(\mathbb{H}) = (\mathbb{H}) \).

**Historical Note.** The fundamental theorem was proved in 1944 by Eilenberg and Niven, using the notion of the degree of a mapping, in the paper: The "fundamental theorem of algebra for quaternions," in *Bull. AMS* 50, 246–248, after Niven had already resolved the special case in which all
terms in \( p \) have the form \( aX^k \) and the indeterminate \( X \) commutes with \( a \) (see I. Niven: Equations in quaternions, in *Am. Math. Monthly* 48, 654–661). The topological proof is also given in the textbook of S. Eilenberg and N. Steenrod: *Foundations of algebraic topology*, Princeton University Press, 1952, 306–311. It is rather surprising that, until the year 1941, the subject of the fundamental theorem for quaternions was never treated in the literature.

§2. **The Algebra \( \mathbb{H} \) as a Euclidean Vector Space**

If \( V \) is a real vector space, then a *bilinear form* \( V \times V \to \mathbb{R}, (x, y) \mapsto \langle x, y \rangle \), is said to be a *scalar product*, if it is *symmetric* and *positive definite*, that is

\[
\langle x, y \rangle = \langle y, x \rangle \text{ and } \langle x, x \rangle > 0 \text{ for } x \neq 0.
\]

\( V \) together with a scalar product is called a *Euclidean* vector space. The number \( |x| := +\sqrt{\langle x, x \rangle} \geq 0 \) is called the (*Euclidean*) *length*, or the *norm*, of the vector \( x \in V \). Two vectors \( x, y \in V \) are said to be *orthogonal* (or to be *perpendicular to each other*), if \( \langle x, y \rangle = 0 \).

The object of this section is to introduce a scalar product in the quaternion algebra \( \mathbb{H} \), which fits in well with the multiplication in \( \mathbb{H} \). In \( \mathbb{C} = \mathbb{R}^2 \), the scalar product \( \langle w, z \rangle = \text{Re}(w\bar{z}) \) is an optimal choice which is compatible with multiplication in \( \mathbb{C} \), as the product rule \( |wz| = |w||z| \) shows (see 3.3.4). We shall see that an analogous situation applies to \( \mathbb{H} = \mathbb{R}^4 \), if one defines, for any two quaternions \( x = ae + bi + cj + dk, x' = a'e + b'i + c'j + d'k \in \mathbb{H} \), the *canonical scalar product*

\[
\langle x, x' \rangle := \alpha \alpha' + \beta \beta' + \gamma \gamma' + \delta \delta' \in \mathbb{R}.
\]

Then it is clear that \( e, i, j, k \) constitute an *orthonormal basis* of \( \mathbb{H} \). By (1) the length \( |x| \) of \( x \) is given by

\[
|x|^2 := \langle x, x \rangle = \alpha^2 + \beta^2 + \gamma^2 + \delta^2.
\]

1. **Conjugation and the Linear Form** \( \mathbb{R}e \). By 1.3(5) every quaternion \( x \) has the basis-independent representation \( x = ae + u, u \in \text{Im}\mathbb{H} \). We shall discuss the \( \mathbb{R} \)-linear *conjugation* (mapping) defined (by analogy with conjugation in \( \mathbb{C} \)) by

\[
\mathbb{H} \to \mathbb{H}, \quad x \mapsto \bar{x} := ae - u.
\]

We then have

\[
\bar{x} = x, \quad \text{Im}\mathbb{H} = \{ x \in \mathbb{H} : \bar{x} = -x \},
\]

and the *fixed point set* is the *straight line* \( \mathbb{R}e \). It is also clear that

\[
|x| = |\bar{x}|, \quad x \in \mathbb{H} \quad (\text{preservation of length}).
\]
We shall continually make use of the multiplication rule

\[(4) \quad xy = \bar{y}\bar{x};\]

which follows, for example, from the product formula 1.1, though one need verify it only for the basis quaternions \(e, i, j, k\), because its general validity is then a consequence of the fact that the mapping \((x, y) \mapsto \bar{xy} - \bar{y}\bar{x}\) is bilinear. In view of the identities (2) and (4) the mapping \(x \mapsto \bar{x}\) is called an involution of the quaternion algebra \(\mathbb{H}\).

We also simulate the real part mapping in \(\mathbb{C}\), and introduce the \(\mathbb{R}\)-linear form

\[(5) \quad \text{Re}: \mathbb{H} \rightarrow \mathbb{R}, \quad x \mapsto \text{Re}(x) := \alpha, \quad \text{where} \quad x = \alpha e + u, \quad u \in \text{Im}\mathbb{H}.\]

Clearly \(\text{Re}\) is characterized by the properties

\[
\text{Re}(e) = 1 \quad \text{and} \quad \text{kernel } \text{Re} = \text{Im}\mathbb{H}.
\]

It is also clear from the definition that (analogously to 3.3.1)

\[(6) \quad x + \bar{x} = 2 \text{Re}(x)e \quad \text{and} \quad \text{Re}(\bar{x}) = \text{Re}(x).\]

The important quadratic equation (3) in 1.3 can now be written as

\[(7) \quad x^2 = 2 \text{Re}(x)x - |x|^2e.\]

Since \((x, y) \mapsto \text{Re}(xy) - \text{Re}(yx)\) is bilinear,

\[(8) \quad \text{Re}(xy) = \text{Re}(yx),\]

holds generally, because it obviously holds for \(e, i, j, k\). Incidentally it may be mentioned that \(\text{Re}(xy)\) is the bilinear form of the Lorentz metric in \(\mathbb{R}^4\), because (by the product formula 1.1)

\[
\text{Re}(xy) = \alpha \alpha' - \beta \beta' - \gamma \gamma' - \delta \delta'
\]

for

\[
x = \alpha e + \beta i + \gamma j + \delta k, \quad x' = \alpha' e + \beta' i + \gamma' j + \delta' k \in \mathbb{H}.
\]

**Remark.** The proofs of the rules (4) and (8) become more readily understandable if one makes use of the algebra isomorphism introduced in 1.2, namely

\[
F: \mathbb{H} \rightarrow \mathcal{H}, \quad x = (\alpha, \beta, \gamma, \delta) \mapsto F(x) = \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix},
\]

\[
w := \alpha + \beta \bar{i}, \quad z := \gamma + \delta \bar{i} \in \mathbb{C},
\]
and works in the matrix algebra $\mathcal{H}$. Thus, writing $A^t$ for the transpose of a matrix $A$, so that
\[ F(\bar{x}) = \overline{F(x)^t}, \quad \text{Re } x = \frac{1}{2} \text{trace } F(x), \]
the familiar rules of the matrix calculus give us
\[
 F(\bar{xy}) = \overline{F(xy)^t} = (F(x)F(y))^t = (F(x)F(y))^t = F(y)^t F(x)^t = F(y) F(\bar{x}) = F(y \bar{x}),
\]
and therefore $\bar{xy} = \bar{yx}$ since $F$ is injective. Since $\text{Re}(xy) = \frac{1}{2} \text{trace}(F(x)F(y))$ (8) follows immediately from the commutativity of the trace: $\text{trace}(AB) = \text{trace}(BA)$.

2. Properties of the Scalar Product. In the introduction, the scalar product $\langle x, x' \rangle$ was defined by (1) in terms of the basis $e, i, j, k$ of $\mathbb{H}$. It is easy to describe it in terms independent of the basis by means of conjugation. We first verify

(1) $x \bar{x} = \bar{x} x = \langle x, x \rangle e$, and in particular $x^{-1} = |x|^{-2} \bar{x}$ for $x \neq 0$.

Writing $x + y$ in place of $x$, we have

(2) $\langle x, y \rangle e = \frac{1}{2} (x \bar{y} + y \bar{x})$

in view of the bilinearity of $\langle x, y \rangle$. We deduce at once from (2) the

**Orthogonality Criterion:** $\langle x, y \rangle = 0 \iff x \bar{y} = -y \bar{x} \iff x \bar{y} \in \text{Im } \mathbb{H}$.

The scalar product in $\mathbb{C}$ is given by $\text{Re}(w \bar{z})$. The same formula applies for $\mathbb{H}$:

(3) $\langle x, y \rangle = \text{Re}(xy) = \text{Re}(\bar{xy})$, in particular $\langle x, e \rangle = \text{Re}(x)$.

If one wishes to avoid the straightforward but tedious deduction of (3) from the product formula, one can argue as follows. Since $x \bar{y} + \bar{x} y = 2 \text{Re}(xy)e$ by 1(6), and since $\bar{xy} = y \bar{x}$ by 1(4), the relation (3) follows from (2). A second proof of (3) is contained in the remark that the mapping $\mathbb{H} \times \mathbb{H} \rightarrow \mathbb{R}$, $(x, y) \mapsto \text{Re}(xy)$, is easily seen to be bilinear, symmetric and positive definite and that $e, i, j, k$ form an orthonormal basis.

The fundamental property is

(4) $|xy| = |x||y|$ (product rule).

**Proof.** Using (1), 1(4), the associative law, and then (1) a second time, we have

\[
|xy|^2 e = (xy, xy)e = (\bar{xy})(xy) = y(\bar{x}x)y = \langle x, x \rangle \bar{y}y
= \langle x, x \rangle \langle y, y \rangle e = |x|^2 |y|^2 e.
\]
Finally we prove yet another formula, which will prove useful in 3.2, and which expresses in a surprising way a triple product of the form $yxy$ as a linear combination of $y$ and $\bar{x}$:

$$yxy = 2(\bar{x}, y)y - \langle y, y \rangle \bar{x}, \quad x, y \in \mathbb{H}$$ (triple product identity).

**Proof.** The identity (2) is equivalent to $2(\bar{x}, y)e = \bar{x}y + yx$. Right-multiplication by $y$ now gives the required result, since $\bar{y}y = (yy)e$.

**Remark.** If we consider, in the algebra $\mathcal{H}$, the mapping

$$\mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R}, \quad (A, B) \mapsto (A, B) := \frac{1}{2} \text{trace}(AB^t),$$

it is clear, from the remark in §2.1, that the algebra isomorphism $F: \mathbb{H} \rightarrow \mathcal{H}$ has the property $\langle F(x), F(y) \rangle = \text{Re}(x\bar{y})$. Formula (3) says therefore that $\langle F(x), F(y) \rangle = \langle x, y \rangle$. This means that $(A, B)$ is a scalar product in $\mathcal{H}$ (which could of course be verified directly) and that $F: \mathbb{H} \rightarrow \mathcal{H}$ is an orthogonal mapping (see 3.1 for this concept). Since (4) $\text{trace}(A^tA) = 2 \det A$, it follows that $\det F(x) = |x|^2$, so that the product rule (4) translates into the product rule for determinants.

3. The “Four Squares Theorem.” In 3.3.4 we deduced the “two-squares” theorem from the product rule for $\mathbb{C}$. In the same way we deduce, from the product rule for $\mathbb{H}$, the famous

**Four Squares Theorem.** For all $\alpha, \beta, \gamma, \delta, \alpha', \beta', \gamma', \delta' \in \mathbb{R}$ we have:

$$(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)(\alpha'^2 + \beta'^2 + \gamma'^2 + \delta'^2)$$

$$= (\alpha \alpha' - \beta \beta' - \gamma \gamma' - \delta \delta')^2 + (\alpha \beta' + \beta \alpha' + \gamma \delta' - \delta \gamma')^2$$

$$+ (\alpha \gamma' + \gamma \alpha' + \delta \beta' - \beta \delta')^2 + (\alpha \delta' + \delta \alpha' + \beta \gamma' - \gamma \beta')^2.$$

**Proof.** The identity follows from the product rule 2(4) and the product formula in 1.1.

The “four squares theorem” was discovered by EULER in 1748 (letter to Goldbach of the 4th May; see “Correspondance entre Leonhard Euler et Chr. Goldbach 1729–1763,” in Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIème siècle, ed. P.-H. Fuss, St. Petersbourg 1843, vol. 1, p. 452). Euler was trying to prove the theorem, which had already been stated by Fermat in 1659, that every natural number is the sum of four squares of natural numbers; by means of his identity he was able to reduce this theorem to the corresponding assertion for primes. The first complete proof of the theorem stated by Fermat was given in 1770 by Lagrange (further information on this will be found...

Gauss remarked (in an unpublished manuscript found after his death, Werke 3, 383-4) that, if complex numbers are used, the “four squares theorem” is contained in the identity

\[(|u|^2 + |v|^2)(|w|^2 + |z|^2) = |uw + vz|^2 + |u\bar{z} - v\bar{w}|^2, \quad u, v, w, z \in \mathbb{C},\]

which is nothing else than the theorem on the product of determinants applied to the matrices \(\begin{pmatrix} u & v \\ \bar{v} & \bar{u} \end{pmatrix}\) and \(\begin{pmatrix} w & \bar{z} \\ z & \bar{w} \end{pmatrix}\) in \(\mathcal{H}\).

Hamilton, as we have already pointed out in the introduction to this chapter, elevated the “four squares theorem” into a “touchstone” to test the value of his quaternions. Once the four squares formula has been found, it is obvious (as in the case of two squares) that it must be true in any commutative ring.

4. Preservation of Length, and of the Conjugacy Relation Under Automorphisms. The excellent interplay within \(\mathbb{H}\) between the operations of conjugation, multiplication, and the formation of the scalar product is again underlined by the following.

Theorem. Every \(\mathbb{R}\)-algebra automorphism \(h: \mathbb{H} \to \mathbb{H}\) has the following two properties:

\[h(\bar{x}) = \overline{h(x)}, \quad |h(x)| = |x|, \quad x \in \mathbb{H},\]

which assert that the mapping \(h\) preserves conjugacy and length respectively.

Proof. Since \(h(e) = e\) and \(\text{Im} \mathbb{H} = \{x \in \mathbb{H}: x^2 = -\omega e \text{ with } \omega \geq 0\}\) it follows that \(h(\text{Im} \mathbb{H}) \subset \text{Im} \mathbb{H}\). Hence, for \(x = \alpha e + u \in \mathbb{H}, \alpha \in \mathbb{R}, u \in \text{Im} \mathbb{H}\), it also follows that \(h(x) = \alpha e + h(u)\) with \(h(u) \in \text{Im} \mathbb{H}\). This implies \(h(x) = \alpha e - h(u) = h(\alpha e - u) = h(\bar{x})\). Moreover \(|h(x)|^2 e = h(x)\overline{h(x)} = h(x\bar{x}) = |x|^2 e\), that is \(|h(x)| = |x|\). \(\square\)

In the theorem we have just proved, the bijectivity of \(h\) is used nowhere. In fact the statement holds good for all \(\mathbb{R}\)-algebra endomorphisms \(h \neq 0\) of \(\mathbb{H}\), because we always have kernel \(h = 0\), as \(\mathbb{H}\) is a division algebra, and \(h(e) = e\). The above theorem was used in 3.2 to prove that all automorphisms of \(\mathbb{H}\) are of the form \(x \mapsto axa^{-1}, a \neq 0\). For the \(\mathbb{R}\)-algebra \(\mathbb{C}\) the corresponding statement is trivial because \(\mathbb{C}\) has only two \(\mathbb{R}\)-automorphisms, namely, the identity mapping and the conjugation mapping (see 3.3.2).

5. The Group \(S^3\) of Quaternions of Length 1. As with complex numbers (see 3.3.4) the product rule gives us immediately
The set $S^3 := \{ x \in \mathbb{H} : |x| = 1 \}$ of all quaternions of length 1 constitutes a group with respect to multiplication in $\mathbb{H}$, which is a subgroup of the multiplicative group $\mathbb{H}^\times := (\mathbb{H} \setminus \{0\}, \cdot)$.

As $e, i, j, k \in S^3$ it is clear that the group $S^3$ is not abelian. In $\mathbb{R}^4 \simeq \mathbb{H}$, the set $S^3$ is the “surface of the unit (hyper)sphere” whose center is at the origin. $S^3$ is compact, it is also called the three-dimensional sphere; topologically $S^3$ can be obtained from $\mathbb{R}^3$, the familiar space of our physical intuition, by compactification through the addition of a point at infinity. The group $S^3$ will play a central role in the next section when we come to study the orthogonal mappings of $\mathbb{H}$ and of $\text{Im} \mathbb{H}$.

The group $S^3$ is its own commutator subgroup. In particular, for every $x \in S^3$ there are elements $u, v \in S^3 \cap \text{Im} \mathbb{H}$ with $x = uvu^{-1}v^{-1}$.

**Proof.** For any such $x$ there is a $y \in S^3$ with $y^2 = x$. From Exercise 7.1.4 there are elements $u, v \in \text{Im} \mathbb{H}$ with $y = uv$. We may assume that $u, v \in S^3$. Then $u^{-1} = -u$ and $v^{-1} = -v$. Therefore $x = (uv)^2 = uvu^{-1}v^{-1}$. \hfill \Box

In the space $\mathbb{R}^{n+1}$ of $(n+1)$-tuples $x = (\xi_0, \ldots, \xi_n)$, $y = (\eta_0, \ldots, \eta_n)$ with the scalar product $(x, y) = \sum_0^n \xi_v \eta_v$ we define the “$n$-dimensional sphere,” by $S^n := \{ x \in \mathbb{R}^{n+1} : |x| = 1 \}$. A nontrivial theorem states that $S^1$ and $S^3$ are the only spheres with a “continuous” group structure.

The following relationship exists between the multiplicative groups $\mathbb{H}^\times$, $S^3$ and $\mathbb{R}_+^\times := \{ x \in \mathbb{R}, x > 0 \}$:

The mapping $\mathbb{H}^\times \rightarrow \mathbb{R}_+^\times \times S^3$, $x \mapsto (|x|, x/|x|)$ is a (topological) isomorphism of the (topological) group $\mathbb{H}^\times$ onto the product of the (topological) groups $\mathbb{R}_+^\times$ and $S^3$.

For every quaternion $x \neq 0$, $x \bar{x}^{-1} \in S^3$. One can verify directly that:

The mapping $h: \mathbb{H} \setminus \{0\} \rightarrow S^3$, $x \mapsto x \bar{x}^{-1} = |x|^{-1}x^2$, is surjective:

(1) $h \left( e + \frac{b}{1 + \alpha} \right) = \alpha + b$, if $\alpha e + b \in S^3 \setminus \{-e\}$, $\alpha \in \mathbb{R}$, $b \in \text{Im} \mathbb{H}$;

\[ h(i) = -e. \]

If one puts $x = \kappa e + b$, $b \in \text{Im} \mathbb{H}$, then $b^2 = -|b|^2 e$ and hence

(2) $h(x) = \frac{\kappa^2 - |b|^2}{\kappa^2 + |b|^2} e + \frac{2\kappa}{\kappa^2 + |b|^2} b$;

we have thus obtained the following parametric representation for the group $S^3$

\[ S^3 = \left\{ \frac{1}{\kappa^2 + |b|^2} [(\kappa^2 - |b|^2)e + 2\kappa b] : (\kappa, b) \in (\mathbb{R} \times \text{Im} \mathbb{H}) \setminus \{0\} \right\} \]

as a generalization of the parametric representation 3.5.4(2') of the circle group $S^1$. The equations (1) and (2) also yield the result (whose proof is left as an exercise):
Every "rational" quaternion \( a e + \beta_1 i + \beta_2 j + \beta_3 k \in S^3 \setminus \{e\}, \alpha, \beta, \in \mathbb{Q} \), has the form

\[
(4) \quad \alpha = \frac{1 - q^2}{1 + q^2}, \quad \beta_\nu = \frac{2q_\nu}{1 + q^2} \quad \text{with} \quad q_\nu := \frac{\beta_\nu}{1 + \alpha} \in \mathbb{Q}, \quad 1 \leq \nu \leq 3,
\]

\[
q^2 := q_1^2 + q_2^2 + q_3^2 \in \mathbb{Q}.
\]

This representation can be utilized (by analogy with 3.5.4) to parametrize Pythagorean quintuplets, that is to say 5-tuples \((k, l, m, n, p)\) of nonzero natural numbers satisfying the equation \(k^2 + l^2 + m^2 + n^2 = p^2\). The reader interested in this may care to work through the simple calculations.

6. The Special Unitary Group \( SU(2) \) and the Isomorphism \( S^3 \to SU(2) \). The set

\[
(1) \quad U(2) := \{ U \in GL(2, \mathbb{C}) : U \bar{U}^t = E \}
\]
of all unitary \(2 \times 2\) matrices is an important subgroup of the group \( GL(2, \mathbb{C}) \) of all complex, nonsingular \(2 \times 2\) matrices. Since \( \det \bar{A}^t = \det A \) we have \( |\det U| = 1 \) for all \( U \in U(2) \). The special unitary group \( SU(2) \) is the normal subgroup of the group \( U(2) \) defined by

\[
SU(2) := \{ U \in U(2) : \det U = 1 \}.
\]

In terms of the subalgebra \( \mathcal{H} \) of \( \text{Mat}(2, \mathbb{C}) \) defined as in 1.2, we now have the

**Theorem.** \( SU(2) = \{ A \in \mathcal{H} : \det A = 1 \} \), and in particular \( SU(2) \subset \mathcal{H} \).

**Proof.** The equation \( A \bar{A}^t = (\det A) \cdot E \) can be immediately verified for

\[
A = \begin{pmatrix} w & -z \\ \bar{z} & \bar{w} \end{pmatrix} \in \mathcal{H}
\]

and from this follows the inclusion relation \( \{ A \in \mathcal{H} : \det A = 1 \} \subset SU(2) \).

For \( U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SU(2) \) we have \( U^{-1} = \bar{U}^t = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix} \) by (1).

Since however \( U^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \), it follows that \( d = \bar{a}, \ c = -\bar{b} \), that is \( U \in \mathcal{H} \).

This immediately yields the

**Isomorphism Theorem.** The algebra isomorphism \( F : \mathbb{H} \to \mathcal{H} \) maps the group \( S^3 \) of all quaternions of length 1 isomorphically onto the special unitary group \( SU(2) \).

**Proof.** Since \( |x|^2 = \det F(x) \) we have \( F(S^3) = \{ A \in \mathcal{H} : \det A = 1 \} = SU(2) \).
§3. The Orthogonal Groups $O(3)$, $O(4)$ and Quaternions

An "Eulerian parametrization" of the group $SU(2)$ can now be obtained from 5(3)

$$SU(2) = \left\{ \frac{1}{\kappa^2 + \lambda^2 + \mu^2 + \nu^2} \begin{pmatrix} \kappa^2 - \lambda^2 - \mu^2 - \nu^2 + 2k\lambda i & 2k\mu + 2k\nu i \\ -2k\mu + 2k\nu i & \kappa^2 - \lambda^2 - \mu^2 - \nu^2 - 2k\lambda i \end{pmatrix} \right\},$$

where $\kappa, \lambda, \mu, \nu$ run through all real quadruples $\neq 0$.

The reader should compare the results discussed in this section with those considered in 5.3 and 5.4 of Chapter 3.

§3. THE ORTHOGONAL GROUPS $O(3)$, $O(4)$ AND QUATERNIONS

HAMILTON tried for many years to find an algebraic structure in the space of our physical world, with whose help the Euclidean geometry of the $\mathbb{R}^3$ would be more easily understood. We have seen that the structure of a division algebra cannot be realized until we have embedded the $\mathbb{R}^3$ in an $\mathbb{R}^4$, and that there are interesting connections between quaternion multiplication and the natural scalar product in $\mathbb{R}^4$. It now turns out that with the "purely imaginary quaternions" one can also give a very elegant interpretation of rotations in $\mathbb{R}^3$ in terms of quaternion multiplication.

Already in 1844, that is within a year after the discovery of quaternions, HAMILTON and CAYLEY were aware that every properly orthogonal mapping of $\mathbb{R}^3$ has the form

$$\text{Im } \mathbb{H} \to \text{Im } \mathbb{H}, \quad u \mapsto au a^{-1},$$

where $a$ runs through all quaternions $\neq 0$. (See HAMILTON, Quaternions: applications in geometry, in Math. Papers 3, 353–362, in particular formula (i') in the footnote on page 361; and CAYLEY: On certain results relating to quaternions, in Math. Papers 1, 123–126). CAYLEY himself assigns the priority to HAMILTON: "the discovery of the formula $q(ix + jy + kz)q^{-1} = ix' + jy' + kz'$, as expressing a rotation, was made by Sir W.R. HAMILTON some months previous to the date of this paper" (Math. Papers 1, p. 586).

In 1855 CAYLEY remarked in a paper which appeared in Vol. 50 of Crelle's Journal (p. 312; Math. Papers 2, p. 214), that every properly orthogonal mapping of $\mathbb{R}^4 = \mathbb{H}$ has the form

$$\mathbb{H} \to \mathbb{H}, \quad x \mapsto \frac{axb}{|a||b|},$$

where $a, b$ independently of each other run through all quaternions $\neq 0$. In the paragraphs which follow these theorems of HAMILTON and CAYLEY will be discussed in some detail. We shall, departing from the usual procedure, first deal with the situation in $\mathbb{R}^4 = \mathbb{H}$, and then obtain the perhaps more interesting case of $\mathbb{R}^3$ as a "gift" from the natural embedding of $\mathbb{R}^3 = \text{Im } \mathbb{H}$ in $\mathbb{H}$. 
1. **Orthogonal Groups.** Let $V$ denote a finite dimensional inner product space. A linear mapping $f: V \to V$ is said to be orthogonal, if

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \text{for} \quad x, y \in V;$$

this holds if and only if $f$ is length-preserving: $|f(x)| = |x|$ for all $x \in V$. Every orthogonal mapping is bijective, and its inverse mapping is likewise orthogonal. The orthogonal mappings of $V$ form a group $O(V)$ under composition; $O(V)$ is called the orthogonal group of the inner product space $V$.

Every endomorphism $f: V \to V$ has a determinant. The determinant has the value

$$\det f = \pm 1 \quad \text{when} \quad f \in O(V).$$

The subgroup $SO(V)$ of the properly orthogonal mappings is defined by

$$SO(V) = O^+(V) := \{ f \in O(V) : \det f = 1 \};$$

the coset of reflections is given by

$$O^-(V) := \{ f \in O(V) : \det f = -1 \},$$

and thus $O(V) = O^+(V) \cup O^-(V)$.

The groups $O(\mathbb{R}^n)$ and $SO(\mathbb{R}^n)$ of the Euclidean number space $\mathbb{R}^n$ are traditionally denoted by $O(n)$ and $SO(n)$ and are often identified with the matrix groups $\{ A \in GL(n, \mathbb{R}) : A^tA = E \}$ and $\{ A \in GL(n, \mathbb{R}) : A^tA = E \text{ and } \det A = 1 \}$ respectively.

The mappings $s_a: V \to V, \quad x \mapsto x - 2\langle a, x \rangle a, \quad a \in V, \quad |a| = 1,$

play a particularly important role. $s_a$ is always orthogonal, and represents a reflection in the hyperplane $\{ x \in V : \langle a, x \rangle = 0 \}$ orthogonal to the line $\mathbb{R}a$. We have

1) $s_a \in O^-(V), \quad s_a^2 = \text{id}, \quad f \circ s_a \in O^+(V) \quad \text{for} \quad f \in O^-(V).$

2) $f \circ s_a = s_{f(a)} \circ f \quad \text{for} \quad f \in O(V).$

We state the following theorem taken from S. Lang, *An Introduction to Linear Algebra*, 2nd ed., Springer-Verlag.

**Generation Theorem for the Orthogonal Group.** The group $O(V)$ is generated by its reflections. The mappings $f \in SO(V)$ are (just) the products of an even number $k$ of reflections, where $k \leq \dim V$.

2. **The Group $O(\mathbb{H})$. CAYLEY's Theorem.** Every mapping

$$\mathbb{H} \to \mathbb{H}, \quad x \mapsto axb, \quad \mathbb{H} \to \mathbb{H}, \quad x \mapsto axb, \quad a, b \in S^3,$$

is orthogonal by virtue of the product rule 2.2(4). To show that these exhaust the orthogonal mappings of $\mathbb{H}$, we invoke the mappings $s_a: \mathbb{H} \to \mathbb{H}.$
§3. The Orthogonal Groups $O(3)$, $O(4)$ and Quaternions

It follows directly from the triple product identity 2.2(5) that

\[(1) \quad s_a(x) = -a\bar{a}x \text { for all } a \in S^3, x \in \mathbb{H};\]

in particular: $s_e(x) = -\bar{x}$. We denote by $p_a$ the mapping $x \mapsto axa$, $a \in S^3$. It follows from (1) that

\[(2) \quad s_a \circ s_b = p_a \circ p_b \text { for all } a, b \in S^3,\]

and in particular $p_a = s_a \circ s_e$.

We can now immediately deduce from (1), (2) and the generation theorem in 1 above the

**Generation Theorem for $O(\mathbb{H})$.** Every orthogonal mapping $f \in SO(\mathbb{H})$ is a product of at most four mappings $p_a$, $a \in S^3$.

The group $O(\mathbb{H})$ is generated by the two mappings $x \mapsto axa$, $a \in S^3$, and $x \mapsto -\bar{x}$.

**Example.** For the mapping $g: \mathbb{H} \to \mathbb{H}$, $x \mapsto -x$, we have

$$g = p_i \circ p_j \circ p_k.$$  

An immediate deduction from the generation theorem for $O(\mathbb{H})$ is the following result:

**Theorem (CAYLEY).** To every orthogonal mapping $f: \mathbb{H} \to \mathbb{H}$ correspond two quaternions $a, b \in S^3$ with the following properties:

a) $f(x) = axb$, if $f \in O^+(\mathbb{H})$.

b) $f(x) = a\bar{x}b$, if $f \in O^-(\mathbb{H})$.

**Proof.** a) When $f \in O^+(\mathbb{H})$ we have $f = p_{a_1} \circ \cdots \circ p_{a_4}$ with $a_1, \ldots, a_4 \in S^3$. If we put $a := a_1a_2a_3a_4$, $b := a_4a_3a_2a_1$, then $a, b \in S^3$ and $f(x) = axb$.

b) When $f \in O^- (\mathbb{H})$ we have $f \circ s_e \in O^+(\mathbb{H})$, hence $f(-\bar{x}) = f \circ s_e (x) = cxb$ with $b, c \in S^3$ by a). We thus see that $f(x) = a\bar{x}b$ with $a := -c$.

From this theorem of CAYLEY can be obtained the result already announced in 1.5.

**Theorem.** Every $\mathbb{R}$-algebra automorphism $h: \mathbb{H} \to \mathbb{H}$ has the form $h(x) = axa^{-1}$, $a \in S^3$.

**Proof.** By Theorem 2.4, $h \in O(\mathbb{H})$. As $h(e) = e$ it follows that

$$h(x) = axa^{-1} \text { or } h(x) = a\bar{x}a^{-1} \text { with } a \in S^3.$$  

The second case is impossible since we should then have

$$h(xy) = a\bar{y}a^{-1} = a\bar{y}a^{-1}a\bar{x}a^{-1} = h(y)h(x).$$
3. The Group $O(\text{Im } \mathbb{H})$. HAMILTON's Theorem. Every orthogonal mapping $\mathbb{H} \rightarrow \mathbb{H}$, $x \mapsto \pm ax\bar{a}$, $a \in S^3$, maps the subspace $\text{Im } \mathbb{H} = \{u \in \mathbb{H} : u = -u\}$ of all purely imaginary quaternions onto itself, since $a\bar{u}\bar{a} = a\bar{a} = -au\bar{a}$, and thus induces an orthogonal mapping $\text{Im } \mathbb{H} \rightarrow \text{Im } \mathbb{H}$, $u \mapsto \pm au\bar{a}$. We assert that all orthogonal mappings of $\text{Im } \mathbb{H}$ can be obtained in this way, for, since the space $\text{Im } \mathbb{H}$ is orthogonal to the line $\mathbb{R}e$, every orthogonal mapping $f$ of $\text{Im } \mathbb{H}$ can be extended uniquely to an orthogonal mapping $\hat{f}: \mathbb{H} \rightarrow \mathbb{H}$ by defining

$$\hat{f} := \text{id on } \mathbb{R}e, \quad \hat{f} := f \text{ on } \text{Im } \mathbb{H}.$$ 

In matrix notation the matrix associated with $\hat{f}$ is $\begin{pmatrix} 1 & 0 \\ 0 & B \end{pmatrix}$, where $B$ is the $3 \times 3$ matrix associated with $f$. It is therefore clear that

$$\det \hat{f} = \det f,$$

so that in particular $f \in O^+(\text{Im } \mathbb{H}) \Leftrightarrow \hat{f} \in O^+(\mathbb{H})$.

We can now easily derive

HAMILTON's Theorem. To every orthogonal mapping $f: \text{Im } \mathbb{H} \rightarrow \text{Im } \mathbb{H}$ there corresponds a quaternion $a \in S^3$ with the following property

a) $f(u) = au\bar{a}$, if $f \in O^+(\text{Im } \mathbb{H})$.

b) $f(u) = -au\bar{a}$, if $f \in O^-(\text{Im } \mathbb{H})$.

Proof. a) Suppose $f \in O^+(\text{Im } \mathbb{H})$. Then $\hat{f} \in O^+(\mathbb{H})$, so that, by a) of Theorem 2, $f(x) = axb$ with $a,b \in S^3$. From $f(e) = e$ it follows that $ab = e$, or in other words $b = a^{-1} = \bar{a}$.

b) This clearly follows from a) since $f \in O^-(\text{Im } \mathbb{H})$ implies $-f \in O^+(\text{Im } \mathbb{H})$, as $\text{Im } \mathbb{H}$ is of dimension 3. \hfill $\Box$

4. The Epimorphisms $S^3 \rightarrow SO(3)$ and $S^3 \times S^3 \rightarrow SO(4)$. The theorems of HAMILTON and CAYLEY provide some important information about the classical groups $SO(3)$ and $SO(4)$. With every $a \in S^3$, and with every pair $(a,b) \in S^3 \times S^3$ we may associate the orthogonal mappings

$$\varphi(a): \text{Im } \mathbb{H} \rightarrow \text{Im } \mathbb{H}, \quad u \mapsto au\bar{a}, \quad \text{and} \quad \psi(a,b): \mathbb{H} \rightarrow \mathbb{H}, \quad x \mapsto axb\bar{b}.$$ 

We consider the mappings $\varphi: S^3 \rightarrow O(\text{Im } \mathbb{H})$, $\psi: S^3 \times S^3 \rightarrow O(\mathbb{H})$. Just as $S^3$ forms a compact non-abelian multiplicative group, so the Cartesian product $S^3 \times S^3$ forms a compact non-abelian group with respect to the composition $(a,b) \cdot (c,d) = (ac, bd)$.

Theorem. The mappings $\varphi: S^3 \rightarrow O(\text{Im } \mathbb{H})$ and $\psi: S^3 \times S^3 \rightarrow O(\mathbb{H})$
are group homomorphisms. The kernel groups each have two elements: kernel $\varphi = \{ \pm e \}$, kernel $\psi = \{ \pm (e, e) \}$. The image groups satisfy $\varphi(S^3) = SO(\Im \mathbb{H})$, $\psi(S^3 \times S^3) = SO(\mathbb{H})$.

**Proof.** That each is a homomorphism is a direct consequence of the definitions of the mappings concerned. For example $\psi((a, b) \cdot (c, d)) = \psi(a, b) \circ \psi(c, d)$ because

$$[\psi(a, b) \circ \psi(c, d)](x) = \psi(a, b)(cxd) = acxd = (ac)x(bd) = \psi(ac, bd)(x), \ x \in \mathbb{H}.$$ 

Suppose that $a \in \text{kernel } \varphi$, so that $u = au$ for all $u \in \Im \mathbb{H}$. By 1.5 such is the case if and only if $a \in \mathbb{R}$. Since $|a| = 1$, it follows that $a = \pm e$, and hence kernel $\varphi = \{ \pm e \}$. Suppose furthermore that $(a, b) \in \text{kernel } \psi$, and thus $axb = x$ for all $x \in \mathbb{H}$. If $x := e$ then $a = b$ and thus $a \in \text{kernel } \varphi$, that is $a = \pm e$ whence kernel $\psi = \{ \pm (e, e) \}$.

Theorems 2 and 3 yield the non-trivial inclusion relations $\varphi(S^3) \supset SO(\Im \mathbb{H})$, $\psi(S^3 \times S^3) \supset SO(\mathbb{H})$. In both cases we in fact have equality: this follows immediately on continuity grounds (by the usual argument based on determinants) or directly, as follows. If there were, for example, a $\psi(a, b) \in O^-(\mathbb{H})$, then by b) of Theorem 2 there would be elements $c, d \in S^3$ such that $axb = x$ for all $x \in \mathbb{H}$. Thus we should always have $\tilde{x} = pxq^{-1}$ with $p := c^{-1}a$, $q := db$. For $x := e$ it would follow that $p = q$ and for $x := p$, we should therefore have $\tilde{p} = p$, and hence $p \in \mathbb{R}$. This however leads to the absurdity $\tilde{x} = x$.

We see from the foregoing that there are natural group epimorphisms $S^3 \rightarrow SO(3)$, $S^3 \times S^3 \rightarrow SO(4)$, whose kernels each have 2 elements. As $S^3$, by 2.6 is isomorphic to $SU(2)$ there are also correspondingly epimorphisms $SU(2) \rightarrow SO(3)$, $SU(2) \times SU(2) \rightarrow SO(4)$, with kernels of 2 elements.

As $S^3$ is of dimension 3, and $S^3 \times S^3$ of dimension 6, the following consequences among others may be noted:

The group $SO(3)$ is 3-dimensional, the group $SO(4)$ 6-dimensional, (and generally dim $SO(n) = \frac{1}{2}n(n - 1)$).

The sets $G := \psi(S^3 \times e)$, $G' := \psi(e \times S^3)$ are normal subgroups of $SO(4)$, which are isomorphic to the group $S^3$ under the isomorphisms $a \mapsto \psi(a, e)$, $b \mapsto \psi(e, b)$, respectively. While $G \cdot G' = SO(4)$, we have $G \cap G' = \pm \text{id}$ as is readily proved. We see in particular that:

The group $SO(4)$ contains normal subgroups isomorphic to the group $S^3$ and is therefore not a "simple" Lie group. On the other hand all groups

---

5 As multiplication in $S^3$ and $S^3 \times S^3$ is non-abelian, one would no longer have a homomorphism if one associated with every $a \in S^3$, and with every $(a, b) \in S^3 \times S^3$, the isometric mappings $u \mapsto au$ and $x \mapsto axb$, respectively. Note that $\bar{a} = a^{-1}$ for $a \in S^3$. 
SO(n), n > 4 are simple, that is to say they contain no nontrivial connected normal subgroups. The groups SO(2n + 1) in fact have no proper normal subgroups \( \neq \{e\} \) at all; the groups SO(2n) have just the one nontrivial normal subgroup \( \{(\pm e)\} \).

5. Axis of Rotation and Angle of Rotation. For \( a \in S^3 \) let \( f_a : \text{Im}_3 \to \text{Im}_3 \) be defined by \( f_a(u) := au \bar{a} \). Clearly \( f_a \in O(\text{Im}_3) \), and \( f_a \) is the identity if and only if \( a = \pm e \) (note that \( f_a = \varphi(a) \) by 4).

If \( f_a \neq \text{id} \), it follows that \( 0 \neq a - \bar{a} \in \text{Im}_3 \) and \( f_a(a - \bar{a}) = a - \bar{a} \); each point of the line generated by \( a - \bar{a} \) is thus invariant under the mapping \( f_a \).

**Proof.** Since \( a \neq \pm e \), it follows that \( 0 \neq a - \bar{a} \in \text{Im}_3 \) and also that \( f_a(a - \bar{a}) = a(a - \bar{a}) \bar{a} = a(a \bar{a}) - (a \bar{a}) \bar{a} = a - \bar{a} \). \( \square \)

To describe the mapping \( f_a \) in a different way we use the following Lemma.

**Lemma.** Every quaternion \( a \in S^3 \setminus \{\pm e\} \) has a unique representation in the form

\[
(1) \quad a = \cos \frac{1}{2} \omega \cdot e + \sin \frac{1}{2} \omega \cdot q \quad \text{with} \quad q \in \text{Im}_3, \quad |q| = 1, \quad \text{and} \quad 0 < \omega < 2\pi.
\]

**Proof.** We write \( a = \alpha e + \beta q \) with \( q \in \text{Im}_3, \quad |q| = 1 \) and \( \beta > 0 \). Since \( \alpha^2 + \beta^2 = 1 \) there is just one \( \omega \in (0, 2\pi) \) such that \( \alpha = \cos \frac{1}{2} \omega, \quad \beta = \sin \frac{1}{2} \omega \). \( \square \)

We shall now show that \( f_a \) is a rotation about the axis \( \mathbb{R}q \) through the angle \( \omega \), or in other words that the plane in \( \text{Im}_3 \) perpendicular to the line \( \mathbb{R}q \) is rotated through the angle \( \omega \). This and more is implicit in the following

**Theorem.** If for any \( a \in S^3 \setminus \{\pm e\} \) the quantities \( \omega \) and \( q \) are chosen to satisfy the equations (1) then

\[
f_a(u) = \cos \omega \cdot u + \sin \omega \cdot q \times u + (1 - \cos \omega) (q, u) q \quad \text{for all} \quad u \in \text{Im}_3.
\]

**Proof.** Using the abbreviations \( \alpha := \cos \frac{1}{2} \omega, \quad \beta := \sin \frac{1}{2} \omega \) we have

\[
au \bar{a} = (\alpha e + \beta q)(\alpha e - \beta q) = \alpha^2 u + \beta aq - \alpha \beta uq - \beta^2 quq.
\]

From the definition of the vector product (see 1.4) we have \( 2q \times u = qu - uq \).

As \( \bar{u} = -u \) and \( \langle q, q \rangle = 1 \), it follows that \( quq = u - 2(q, u)q \) by the triple product identity (2.2(5)), and consequently

\[
f_a(u) = (\alpha^2 - \beta^2) u + 2\alpha \beta q \times u + 2\beta^2 (q, u) q, \quad u \in \text{Im}_3.
\]
§3. The Orthogonal Groups $O(3)$, $O(4)$ and Quaternions

From the definitions of $\alpha$, $\beta$ and the elementary formulae of trigonometry, it follows that $\alpha^2 - \beta^2 = \cos \omega$, $2\alpha\beta = \sin \omega$, $2\beta^2 = 1 - (\alpha^2 - \beta^2) = 1 - \cos \omega$.

Corollary. $f_a(q) = q$ and $\langle f_a(u), u \rangle = \cos \omega$ for all $u \in \text{Im} \mathbb{H}$ with $|u| = 1$ and $\langle u, q \rangle = 0$.

We deduce from the foregoing results that $f_a$ is a rotation about the axis $\mathbb{R}q$ through the angle $\omega$. Incidentally it is easily shown that $\cos \omega = \text{Re}(a^2)$. If $a$ is purely imaginary then $\omega = \pi$ and $f_a = -s_a$ is a rotation of $180^\circ$ about the axis $\mathbb{R}a$.

**Remark.** As is well known every properly orthogonal mapping $\neq \text{id}$ of $\mathbb{R}^3$ is a rotation about a uniquely defined axis. Every $f \in SO(\text{Im} \mathbb{H}) \setminus \{\text{id}\}$ is therefore a rotation about an axis $\mathbb{R}q$, $q \in \text{Im} \mathbb{H}$, $|q| = 1$, through an angle $\omega$, $0 < \omega < 2\pi$. If we now define $a \in S^3$ by (1), then $a = \pm e$ and $f_a \in SO(\text{Im} \mathbb{H})$ is by the theorem a rotation through $\omega$ about the axis $\mathbb{R}q$. We have thus proved afresh statement a) of HAMILTON's theorem in 3, namely every $f \in SO(\text{Im} \mathbb{H})$ has the form $f_a$ with $a \in S^3$.

6. **Euler's Parametric Representation of $SO(3)$**. The mapping

$$
\mathbb{H} \setminus \{0\} \to SO(\text{Im} \mathbb{H}), \quad a \mapsto h_a \quad \text{with} \quad h_a : \text{Im} \mathbb{H} \to \text{Im} \mathbb{H},
$$

$$
u \mapsto \frac{1}{|a|^2} au\bar{a} = aua^{-1},
$$

is by Theorem 4, an epimorphism of the multiplicative group $\mathbb{H} \setminus \{0\}$ with $\Re \setminus \{0\}$ as kernel. If one sets $a := \kappa e + \lambda i + \mu j + \nu k$ and writes $u := xi + yj + zk$ as a column vector, we have

$$
h_a(u) = A \begin{pmatrix} x \\ y \\ z \end{pmatrix}
$$

where $A$ is a properly orthogonal matrix $A \in SO(3)$. This matrix is found by expressing $|a|^{-2} au\bar{a}$ in terms of the basis $i, j, k$ of $\text{Im} \mathbb{H}$. One obtains in this way the result discovered by EULER in 1770 (Opera omnia 6, Ser. 1, 287–315), the well-known

Rational Parametric Representation of Orthogonal $3 \times 3$ Matrices.

For every quadruple $(\kappa, \lambda, \mu, \nu) \in \mathbb{R}^4 \setminus \{0\}$ the $3 \times 3$ matrix

$$
\begin{pmatrix}
\kappa^2 + \lambda^2 - \mu^2 - \nu^2 & -2\kappa\nu + 2\lambda\mu & 2\kappa\mu + 2\lambda\nu \\
2\kappa\nu + 2\lambda\mu & \kappa^2 - \lambda^2 + \mu^2 - \nu^2 & -2\kappa\lambda + 2\mu\nu \\
-2\kappa\mu + 2\lambda\nu & 2\kappa\lambda + 2\mu\nu & \kappa^2 - \lambda^2 - \mu^2 + \nu^2
\end{pmatrix}
$$

(1)
is properly orthogonal, and all properly orthogonal $3 \times 3$ matrices can be expressed in this form.

**Proof.** If we set $a := ke + b \in (\mathbb{R} \times \text{Im} \mathbb{H}) \setminus \{0\}$, then $aua = (ke+b)u(ke-b) = \kappa^2 u + 2\kappa b \times u - bub$. As $bub = |b|^2 u - 2(b, u)b$ by the triple product identity since $\bar{u} = -u$, it is clear that $aua = (\kappa^2 - |b|^2)u + 2\kappa b \times u + 2(b, u)b$. The representation (1) follows at once from this if $b := \lambda i + \mu j + \nu k$. □

The parametric representation for properly orthogonal $2 \times 2$ matrices given in 3.5.4 follows from (1), if we put $\mu = 0, \nu = 0$ (and write $-\lambda$ for $\lambda$) in the leading minor.

As the epimorphism $\mathbb{H} \setminus \{0\} \rightarrow SO(\text{Im} \mathbb{H})$, $a \mapsto h_a$, has the group $\mathbb{R}e \setminus \{0\}$ as its kernel, the same matrix $A$ defined by (1) appertains to the two distinct quadruples $a, a' \in \mathbb{R}^4 \setminus \{0\}$ if and only if $a' = aa$ with $a \neq 0$, or in other words if and only if $a$ and $a'$ define the same point in the real projective 3-dimensional space $\mathbb{P}^3(\mathbb{R})$ with the homogeneous coordinates $\kappa, \lambda, \mu, \nu$. 

**Euler's theorem** can therefore also be expressed as follows:

The mapping $\mathbb{P}^3(\mathbb{R}) \rightarrow SO(3)$ defined by (1) is bijective, and in particular $SO(3)$ is a rational manifold.

This statement was generalized by **Cayley** in 1846 (*Math. Papers* 1, 332–336): 

The group $SO(n)$ is an $\frac{1}{2}n(n-1)$-dimensional rational manifold. The $\frac{1}{2}n(n-1)$-dimensional real projective space is mapped birationally into $SO(n)$ by the **Cayley** mapping

$$X \mapsto (\kappa E - X)^{-1}(\kappa E + X), \quad X \in \text{Mat}(n, \mathbb{R}),$$

where $X$ is skew-symmetric.

The case $n = 3$ of this **Cayley** representation is none other than the **Euler** parametric representation.

**Reference**