Quaternion Algebras and the Algebraic Legacy of Hamilton’s Quaternions

DAVID W. LEWIS

Abstract. We describe the basic definitions and fundamental properties of quaternion algebras over fields and proceed to give an account of how Hamilton’s 1843 discovery of the quaternions was a major turning point in the subject of algebra. Noncommutative algebra started here! We will emphasize especially the theory of division algebras and other kinds of algebras which emanated from this discovery.

The 16th of October 1843 was a momentous day in the history of mathematics and in particular a major turning point in the subject of algebra. On this day William Rowan Hamilton had a brain wave and came up with the idea of the quaternions. He carved the multiplication formulae with his knife on the stone of Brougham Bridge (nowadays known as Broomebridge) in Dublin,

\[ i^2 = j^2 = -1, \quad ij = -ji. \]

The story of this discovery has been told many times, and translated into many different languages. See [20] for this story and for a full biography of Hamilton. Broomebridge may be regarded as the birthplace of noncommutative algebra and it has been visited over the years by many mathematicians from many different lands, especially by algebraists to some of whom it is regarded as a kind of pilgrimage. Hamilton’s discovery was the precursor to a vast range of new kinds of algebraic structures in which the basic assumptions of properties such as commutativity and associativity were cast aside. Something similar happened in the same century in the subject of geometry where the old Euclidean geometry axioms were modified leading to various kinds of noneuclidean geometry.

This article is an expanded version of a talk given at the Irish Mathematical Society September Meeting at Dublin City University,
1–2 September 2005. The plan of the talk and of this article is as follows:
1. Basic results on quaternion algebras over fields.
2. The algebraic legacy.

Thanks are due to the referee who made some useful suggestions for improving the first version of this article.

1. Quaternion Algebras

The theory of quaternion algebras is intimately related to the algebraic theory of quadratic forms. The following books are a good source for the basic definitions and results about quaternion algebras over fields, [29], [30], [45], the last one also containing much material about the associative algebras discussed later in the article. See also [31].

We will write $H$ for the real quaternions. The elements of $H$ are expressions of the form

$$\alpha + \beta i + \gamma j + \delta k,$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, the field of real numbers.

We can regard $H$ as a 4-dimensional vector space over $\mathbb{R}$ with the natural definitions of addition and scalar multiplication, i.e.,

$$(\alpha + \beta i + \gamma j + \delta k) + (\alpha' + \beta' i + \gamma' j + \delta' k) = (\alpha + \alpha') + (\beta + \beta') i + (\gamma + \gamma') j + (\delta + \delta') k,$$

$$\lambda(\alpha + \beta i + \gamma j + \delta k) = \lambda\alpha + \lambda\beta i + \lambda\gamma j + \lambda\delta k$$

for all $\lambda, \alpha, \beta, \gamma, \delta, \alpha', \beta', \gamma', \delta' \in \mathbb{R}$. The set $\{1, i, j, k\}$ is a natural basis for this vector space.

The set $H$ is made into a ring by defining multiplication via the rules

$$i^2 = j^2 = -1, \quad ij = -ji = k$$

together with the usual distributive law.

We identify $\mathbb{R}$ with the set of quaternions with $\beta = \gamma = \delta = 0$ and in this way regard $\mathbb{R}$ as a subset of $H$.

In order to define the more general notion of a quaternion algebra we will replace $\mathbb{R}$ by any field $F$. (We will assume for simplicity that $1+1 \neq 0$ in $F$, i.e. $F$ does not have characteristic 2. A modification of the definition is required in characteristic 2.) We choose two nonzero elements $a, b$ in $F$. A choice with $a = b$ is not excluded. We define
the quaternion algebra \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) to be the set of all expressions of the form
\[
\alpha + \beta i + \gamma j + \delta k,
\]
where \( \alpha, \beta, \gamma, \delta \in F \).

We can regard \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) as a 4-dimensional vector space over \( F \) with the same definitions of addition and scalar multiplication as in \( \mathbb{H} \).

The multiplication rules are now
\[
i^2 = a, \quad j^2 = b, \quad ij = -ji = k.
\]
It easily follows that \( k^2 = -ab \) and that \( i, j, k \) anti-commute just like in \( \mathbb{H} \). We have then that \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) is a 4-dimensional algebra over \( F \).

In this notation we can write \( \mathbb{H} = \left( \begin{array}{c} -1 \\ -1 \\ \frac{\text{over } \mathbb{R}} \end{array} \right) \).

Now \( \mathbb{H} \) is a division algebra, i.e. an algebra which is also a division ring, i.e., each nonzero element of \( \mathbb{H} \) is invertible. In general \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) need not be a division algebra. It depends on the choice of \( F, a, \) and \( b \).

In fact there are only two possibilities:

1. \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) is a division algebra,
2. \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) is isomorphic to \( M_2 F \), the algebra of all \( 2 \times 2 \) matrices with entries from \( F \).

(If an \( F \)-algebra is isomorphic to a full matrix algebra over \( F \) we say that the algebra is split, so (2) is the split case.)

The quick way to see this is to note that \( \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) \) is a central simple algebra, i.e., it has centre \( F \) and is simple in the sense that it has no nontrivial two-sided ideals, and then appeal to a famous 1907 theorem of J. H. M. Wedderburn. This theorem says that any central simple algebra, finite-dimensional over its centre, is isomorphic to an algebra \( M_n D \) for some integer \( n \) and some division algebra \( D \) over \( F \). See [53], p. 282 for a proof of Wedderburn’s theorem. Considering dimensions over \( F \) we have \( \dim_F \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) = 4 \) and \( \dim_F M_n D = n^2 \dim_F D \). Hence the only possibilities are \( n = 1, \left( \begin{array}{c} a \\ b \\ \frac{\text{over } F} \end{array} \right) = D \) or else \( n = 2, D = F \).

This can also be proved, by longer arguments, without appealing to Wedderburn’s theorem. See [29], theorem 2.7.
How can we tell which of the two possibilities we have?

The answer is given by the norm form of a quaternion algebra, which we will now define:

Let \( q = \alpha + \beta i + \gamma j + \delta k \).

Define the conjugate of \( q \), denoted \( \bar{q} \), by \( \bar{q} = \alpha - \beta i - \gamma j - \delta k \).

Define the norm map \( N : \left( \frac{a,b}{F} \right) \to F \) by \( N(q) = \bar{q}q = \alpha^2 - a\beta^2 - b\gamma^2 + ab\delta^2 \)

so that \( N \) may be regarded as a quadratic form in four variables, \( \alpha, \beta, \gamma, \delta \).

(We recall that a quadratic form is a homogeneous polynomial of degree two, or more conceptually it is a map \( Q : V \to F \), \( V \) being an \( F \)-vector space with the following two properties:

(i) \( Q(\alpha v) = \alpha^2 Q(v) \) for all \( v \in V \), and all \( \alpha \in F \),

(ii) the map \( B : V \times V \to F, B(v,w) = Q(v+w) - Q(v) - Q(w), \) is bilinear.)

The quadratic form \( N \) is known as the norm form of the quaternion algebra. In the standard notation of quadratic form theory, cf. [29], this form is denoted \( <1,-a,-b,ab> \). This notation corresponds to representing the quadratic form by a diagonal matrix with \( 1, -a, -b, ab \) as the diagonal entries.

(In the special case of Hamilton’s quaternions \( \mathbb{H} \) the norm form is \( \alpha^2 + \beta^2 + \gamma^2 + \delta^2 \), a sum of four squares, i.e. \( <1,1,1,1> \).)

It is clear that whenever \( N(q) \neq 0 \) the element \( q \) is invertible, its inverse being \( \frac{1}{N(q)}q \). Indeed \( q \) is invertible if and only if \( N(q) \neq 0 \) because \( N(q) = 0 \) implies that \( q \) is a zero divisor. Thus we have the following:

**Theorem 1.** The quaternion algebra \( \left( \frac{a,b}{F} \right) \) is a division algebra if and only if its norm form does not represent zero nontrivially, (i.e. \( N(q) = 0 \Rightarrow q = 0 \)).

A note on terminology: A quadratic form \( Q : V \to F \) on an \( F \)-vector space \( V \) is said to be anisotropic if \( Q(v) = 0 \Rightarrow v = 0 \).

Hence, in the language of quadratic form theory, the above theorem says that \( \left( \frac{a,b}{F} \right) \) is a division algebra if and only if its norm form is anisotropic.
How can we tell when two quaternion algebras \((\frac{a,b}{F})\) and \((\frac{a',b'}{F'})\) are isomorphic as algebras?

The answer is again provided by the norm form. We first need some more terminology from quadratic form theory.

Two quadratic forms \(Q_1 : V_1 \to F\) and \(Q_2 : V_2 \to F\) are said to be isometric if there exists a vector space isomorphism \(\gamma : V_1 \to V_2\) such that \(Q_2(\gamma(x)) = Q_1(x)\) for all \(x \in V_1\). The map \(\gamma\) is called an isometry.

**Theorem 2.** The quaternion algebras \((\frac{a,b}{F})\) and \((\frac{a',b'}{F'})\) are isomorphic as algebras if and only if their norm forms are isometric as quadratic forms.

One way around this implication is clear since any isomorphism of the algebras will give an isometry of the norm forms. For the converse first observe that an isometry of norm forms is not in general an isomorphism of the quaternion algebras, but it is an isomorphism of the three-dimensional subalgebras of pure quaternions. (A pure quaternion \(q\) is one satisfying \(\bar{q} = -q\), i.e., \(q = \beta i + \gamma j + \delta k\) for some \(\beta, \gamma, \delta \in F\).) From this one can deduce that quaternion algebras are isomorphic. See [29], theorem 2.5.

**Examples 3.**

(i) \((\frac{a,b}{F})\) is isomorphic to \((\frac{b,a}{F})\) because their norm forms, \(<1, -a, -b, ab>\) and \(<1, -b, -a, ab>\), are isometric.

(ii) \((\frac{2}{F})\) is isomorphic to \(M_2F\) for any \(a \in F\) because the norm form \(<1, -a, -1, a>\) is isotropic.

(We write \(\hat{F}\) for the set of nonzero elements of \(F\).)

When \(a = -1\) we can write down an explicit isomorphism from \((\frac{-1}{F})\) to \(M_2F\) by mapping \(i\) to \(
\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}
\)
and \(j\) to \(
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\)
so that \(\alpha + \beta i + \gamma j + \delta k\) is mapped to the matrix \(
\begin{pmatrix} \alpha + \delta & \beta + \gamma \\ -\beta + \gamma & \alpha - \delta \end{pmatrix}
\).

(iii) \((\frac{2^1}{F})\) is isomorphic to \(M_2F\) for any \(a \in F\), \(a \neq 1\), because the norm form \(<1, -a, -1, a(a - 1)>\) is isotropic.

For some familiar fields we can easily describe what happens. We should first remark that the isometry class of a quadratic form is unchanged when any of the entries of a standard diagonal representation of the form are multiplied by a square. Hence \((\frac{a,b}{F})\) is isomorphic to \((\frac{c^2 ad}{F})\) for any \(c, d \in F\).
(a) The complex numbers \( \mathbb{C} \):
\((\frac{a+b}{\sqrt{2}})\) is isomorphic to \( M_2 \mathbb{C} \) for any nonzero \( a, b \in \mathbb{C} \) because the norm form \( \langle 1, -a, -b, ab \rangle \) is always isotropic. (Each element of \( \mathbb{C} \) is a square.) So \((\frac{a+b}{\sqrt{2}})\) is always split.

(b) The real numbers \( \mathbb{R} \):
\((\frac{a+b}{\sqrt{2}})\) is isomorphic to \( \mathbb{H} \) whenever \( a < 0 \) and \( b < 0 \). Otherwise \((\frac{a+b}{\sqrt{2}})\) is split. (Each positive element of \( \mathbb{R} \) is a square.)

(c) The finite fields \( \mathbb{F}_n \) with \( n \) elements:
\((\frac{a+b}{\sqrt{2}})\) is always split because of the 1905 Theorem of Wedderburn [61] that any finite division ring is commutative.

(d) The rational numbers \( \mathbb{Q} \):
There exist infinitely many non-isomorphic quaternion algebras over \( \mathbb{Q} \).

(e) Algebraic number fields, i.e. finite extension fields of \( \mathbb{Q} \):
In general there can exist infinitely many non-isomorphic quaternion algebras over such a field.

(f) The \( p \)-adic fields \( \mathbb{Q}_p \) where \( p \) is a prime, i.e. the completion of the field \( \mathbb{Q} \) with respect to the \( p \)-adic absolute value on \( \mathbb{Q} \):
(Recall that the field of real numbers can be viewed as the completion of \( \mathbb{Q} \) with respect to the usual absolute value on \( \mathbb{Q} \).)
For each prime \( p \) there is a unique quaternion division algebra over \( \mathbb{Q}_p \). This follows from the theory of quadratic forms over \( \mathbb{Q}_p \). Up to isometry there exists a unique anisotropic quadratic form of dimension four and this is the norm form of a quaternion division algebra. See chapter 6 of [29] for further details.

1.1. **Characterization of quaternion algebras.** We give two possible characterizations, one via central simple algebras, the other via involutions. For a proof see p. 74 of [29].

**Theorem 4.** If \( A \) is a central simple algebra of dimension four over a field \( F \) then \( A \) is a quaternion algebra.

For the second characterization we need the notion of *involition* on an algebra, i.e. an anti-automorphism of period two, i.e. a map \( \sigma : A \rightarrow A \) which is \( F \)-linear and satisfies

(i) \( \sigma(x + y) = \sigma(x) + \sigma(y) \) for all \( x, y \in A \),
(ii) \( \sigma(xy) = \sigma(y)\sigma(x) \) for all \( x, y \in A \),
(iii) \( \sigma(\sigma(x)) = x \) for all \( x \in A \).
Conjugation on a quaternion algebra is the standard example of an involution. Involutions as defined above are known as involutions of the first kind. (Involutions of the second kind are maps which satisfy (i), (ii), (iii) above but which are not $F$-linear.)

**Theorem 5** (E. H. Moore [37], Dieudonné [14]). Let $A$ be a non-commutative associative finite-dimensional division algebra over a field $F$. If there exists an involution $\sigma : A \rightarrow A$ for which $\sigma(x) + x$ and $\sigma(x)x$ belong to $F$ for all $x \in F$, then $A$ is a quaternion algebra.

For further discussion of this theorem see [56], also [29], p. 74–75.

1.2. **Biquaternion algebras.** Let $A_1$ and $A_2$ be quaternion algebras over a field $F$. The tensor product $A_1 \otimes_F A_2$ is called a biquaternion algebra. It is a 16-dimensional algebra. Write $A$ for this algebra. Then, as a consequence of Wedderburn’s theorem on central simple algebras there are exactly three possibilities:

1. $A$ is a division algebra.
2. $A$ is split, i.e. $A$ is isomorphic to $M_4 F$.
3. $A$ is isomorphic to $M_2 D$ for some quaternion division algebra $D$.

Some more quadratic form theory enables us to distinguish between the different possibilities.

The *Albert quadratic form* of the biquaternion algebra

$$A = \left( \frac{a_1, b_1}{F} \right) \otimes_F \left( \frac{a_2, b_2}{F} \right)$$

is the six-dimensional form $< a_1, b_1, -a_1 b_1, -a_2, -b_2, a_2 b_2 >$. We will write $\phi_A$ for this form. The way we have defined $\phi_A$ seems to depend on the particular formulation of $A_1 = \left( \frac{a_1, b_1}{F} \right)$ and $A_2 = \left( \frac{a_2, b_2}{F} \right)$. However it can be shown that $\phi_A$ is uniquely determined up to a multiple by a scalar from $F$.

**Theorem 6.** Let $A$ be a biquaternion algebra. Then

1. $A$ is a division algebra if and only if $\phi_A$ is anisotropic.
2. $A$ is split, i.e. $A$ is isomorphic to $M_4 F$, if and only if $\phi_A$ is hyperbolic, i.e $\phi_A$ is isometric to $< 1, -1, 1, -1, 1, -1 >$.

Otherwise $A$ is isomorphic to $M_2 D$ for some quaternion division algebra $D$.

For a proof of this theorem of Albert, see [29], theorem 4.8.
Remark 7. Possibility (1) does not occur for any of the fields of types (a)–(f) listed earlier, i.e. biquaternion algebras are never division algebras over these fields. For other kinds of fields, biquaternion division algebras may certainly exist, e.g. \( \left( \frac{x_1}{x_2} \right) \otimes \left( \frac{x_3}{x_4} \right) \) where \( F = \mathbb{R}(x_1, x_2, x_3, x_4) \), the field of rational functions in four indeterminates \( x_1, x_2, x_3, x_4 \) over \( \mathbb{R} \). Its Albert form is anisotropic.

1.3. A remark on involutions of the first kind. Much of the pioneering work on the structure of algebras and of involutions on algebras was done by Albert [1].

Another theorem of Albert says that a central simple algebra \( A \) over \( F \) admits an \( F \)-linear involution, as defined above, if and only if \( A \) is isomorphic as an \( F \)-algebra to its opposite algebra \( A^{op} \). (We recall that \( A^{op} \) is identical to \( A \) except that the multiplication in \( A^{op} \) is the reverse of that in \( A \).) See chapter 8 of [53] for a proof of this theorem.

This property clearly holds for all quaternion algebras, and also is easily seen to hold for all tensor products of quaternion algebras, i.e. any tensor product of quaternion algebras always admits an \( F \)-linear involution. Given quaternion algebras \( A_i, i = 1, 2, \ldots, n \) equipped with involutions \( \sigma_i, i = 1, 2, \ldots, n \), we have an involution of the first kind \( \sigma_1 \otimes \sigma_2 \otimes \ldots \otimes \sigma_n \) defined on the tensor product \( A_1 \otimes A_2 \otimes \ldots \otimes A_n \).

For algebras of dimension 16 these are the only possible ones because of the following theorem of Albert [1]:

**Theorem 8.** If \( A \) is a central simple algebra of dimension sixteen over a field \( F \) and if \( A \) admits an involution of the first kind then \( A \) is a biquaternion algebra.

For three different proofs of this theorem see p. 233–234 of [27], this book also containing an abundance of information on involutions.

For a long time tensor products of quaternion algebras were the only known examples of finite-dimensional division algebras with an involution of the first kind. In 1978 Amitsur, Rowen, Tignol [4] produced the first example of a division algebra admitting an \( F \)-linear involution but which is not a tensor product of quaternion algebras.

However, as a consequence of a celebrated theorem of Merkurjev in 1981, it turns out that any division algebra \( D \) admitting an \( F \)-linear involution is similar to a tensor product of quaternion algebras in
the sense that there exists an integer $n$ such that the matrix algebra $M_n D$ is isomorphic to a tensor product of quaternion algebras. We can give a brief outline of Merkurjev’s Theorem, but a complete proof is highly nontrivial and way beyond the scope of this article. The theorem says that the quaternion symbol map $\alpha : k_2 F \to Br_2 F$ is an isomorphism where $k_2 F, Br_2 F$, and $\alpha$ are defined as follows:

$k_2 F$ is the reduced Milnor K-theory group of the field $F$, a multiplicative group generated by the bimultiplicative symbols $[a, b]$ with $a, b \in \hat{F}$ satisfying the set of relations

\[
\begin{align*}
[a, 1 - a] &= 1 \quad (a \in \hat{F}, a \neq 1) \\
[a, b] &= [b, a] \quad (a, b \in \hat{F}) \\
[a, a] &= [a, -1] \quad (a \in \hat{F}).
\end{align*}
\]

$Br_2 F$ is the subgroup of elements of order $\leq 2$ in the Brauer group $Br F$, where $Br F$ is the group of similarity classes of central simple $F$-algebras. The relation of similarity $\sim$ is defined by $A \sim B$ if and only if there are integers $r, s$ such that $M_r A$ and $M_s B$ are isomorphic $F$-algebras. The group operation is tensor product, the identity element is the class of $M_n F$, the inverse of the class of $A$ is the class of $A^{op}$, the opposite algebra. Note that $Br_2 F$ consists precisely of the classes of central simple $F$-algebras which admit an $F$-linear involution.

The map $\alpha : k_2 F \to Br_2 F$ sends the symbol $[a, b]$ to the class of the quaternion algebra $\left( \frac{a, b}{F} \right)$. This map is well-defined because we have seen earlier that quaternion algebras satisfy the same kind of relations as the defining relations for the symbols in $k_2 F$.

The surjectivity of the map $\alpha$ implies that any central simple $F$-algebra $A$ admitting an $F$-linear involution is similar to a tensor product of quaternion algebras.

See [29], chapter 5, section 6, for some more details and see [58] for a full proof of Merkurjev’s theorem.

We finish this section with some remarks about solving equations in $\mathbb{H}$, and about quaternionic eigenvalues.

1.4. Equations in $\mathbb{H}$. The subject of algebra originated in attempts to solve equations, e.g. linear, quadratic, cubic, and higher order polynomial equations, and it then naturally developed into the study of more abstract algebraic structures.
Any polynomial of degree \( n \) with coefficients from the field \( \mathbb{C} \) of complex numbers has at least one root in \( \mathbb{C} \). (The Fundamental Theorem of Algebra, first proved by Gauss in 1799.) Moreover any polynomial with coefficients in any field \( F \) can have at most \( n \) roots in \( F \). For polynomials with coefficients from \( \mathbb{H} \) the situation is somewhat different.

**Example 9.** The equation \( x^2 = -1 \) has infinitely many different solutions in \( \mathbb{H} \). Any pure quaternion of norm one, i.e. \( \beta i + \gamma j + \delta k \) with \( \beta^2 + \gamma^2 + \delta^2 = 1 \), will satisfy this equation. There are infinitely many such pure quaternions because they correspond to points on the unit sphere in \( \mathbb{R}^3 \).

**Example 10.** The equation \( x^2 i - ix^2 = 1 \) has no solutions in \( \mathbb{H} \). To see this note that \( i(x^2 i - ix^2)i = x^2 i - ix^2 \) but \( i1i = -1 \). (Alternatively observe that the left hand side has trace zero but right hand side has trace two, where trace \( q = q + \bar{q} \).

The notion of polynomial in \( \mathbb{H} \) becomes more complicated due to lack of commutativity. We may have several terms of a given degree. In degree two we may have terms like \( ax^2, xax, x^2a, axbx \), etc. The above example has two terms of degree two.

Nevertheless there is a Fundamental Theorem of Algebra for \( \mathbb{H} \), see [19], also [38], which says that if the polynomial has only one term of highest degree then there exists a root in \( \mathbb{H} \). The proof is topological, using a little homotopy theory.

There is also a Fundamental Theorem of Algebra for the octonions, [25], proved in a similar manner.

In the context of solving polynomial equations we should mention the Wedderburn Factorization Theorem for polynomials over division algebras.

**Theorem 11** (Wedderburn Factorization Theorem). Let \( D \) be a division ring with centre \( F \) and let \( p(t) \) be an irreducible monic polynomial of degree \( n \) with coefficients from the field \( F \). If there exists \( d \in D \) such that \( p(d) = 0 \) then we can write

\[
p(t) = (t - d)(t - d_2)(t - d_3) \cdots (t - d_n)
\]

and each \( d_i \) is conjugate to \( d \), i.e. there exist nonzero elements \( s_i \in D, i = 2, 3, \ldots, n \), such that \( d_i = s_i ds_i^{-1} \) for each \( i = 2, 3, \ldots, n \).

This theorem says that if the polynomial has one root in \( D \) then it factorizes completely as a product of linear factors over \( D \). This
is in stark contrast to the situation in field theory where there may well exist just one root of a polynomial in a given field. See [48], chapter 7, for a proof.

Wedderburn’s Theorem has proved to be a surprisingly powerful tool, yielding neat proofs of certain results on division algebras. Here are a few examples:

(i) This author’s criterion for zero being a nontrivial sum of squares in a division algebra [35].

(ii) Wadsworth’s criterion for extending valuations to division algebras [59].

(iii) Hazrat’s proof of Platonov’s congruence theorem in the reduced K-theory of division algebras [21].

(iv) The cyclicity of division algebras of degree three. (See the section on cyclic algebras later in this paper.)

1.5. Quaternionic eigenvalues. Let $A$ be an $n \times n$ matrix with entries from $\mathbb{H}$. Lack of commutativity means that we have two possible notions of eigenvalue—left eigenvalue and right eigenvalue.

A left eigenvalue of $A$ is an element $\lambda \in \mathbb{H}$ such that there exists a column vector $x \in \mathbb{H}^n$ such that $Ax = \lambda x$.

A right eigenvalue of $A$ is an element $\lambda \in \mathbb{H}$ such that there exists a column vector $x \in \mathbb{H}^n$ such that $Ax = x\lambda$.

Do left or right eigenvalues necessarily exist for a given matrix?

The existence of right eigenvalues has been shown via algebraic methods in [7] and [34], and recently by a topological argument using the Lefschetz fixed point theorem by Baker [6]. The algebraic methods rely on writing a quaternionic matrix $A$ in the form $A = B + Cj$ where $B$ and $C$ are $n \times n$ matrices with complex entries. (Any element $\alpha + \beta i + \gamma j + \delta k \in \mathbb{H}$ may be written as $\alpha + \beta i + (\gamma + \delta i)j$. This viewpoint enables us to view quaternions as ordered pairs of complex numbers in the same way that complex numbers are viewed as ordered pairs of real numbers.)

The existence of left eigenvalues was first proved topologically by R. Wood, [63], using homotopy theory, specifically using $\pi_3GL_n\mathbb{H}$, the third homotopy group of the group of invertible quaternionic $n \times n$ matrices. The existence of a left eigenvalue $\lambda$ for $A$ amounts to showing that the matrix $\lambda I_n - A$ is not invertible. The assumption that $\lambda I_n - A$ is invertible for every $\lambda$ leads to the construction of two different elements of $\pi_3GL_n\mathbb{H}$ which can be shown to be homotopic to each other, a contradiction.
2. The Algebraic Legacy

The quaternions were the progenitor to a host of different algebraic structures in which the traditional restrictive assumptions such as commutativity and associativity were abandoned. In December 1843 John Graves discovered the octonions \( \mathbb{O} \), an 8-dimensional algebra over \( \mathbb{R} \) which is nonassociative as well as being noncommutative. They have a basis \( \{1, e_0, e_1, \ldots, e_6\} \) with the multiplication rules as follows:

\[
\begin{align*}
 e_i^2 &= -1 \text{ for each } i \\
 e_{n+1}e_{n+2} &= e_{n+4} = -e_n+2e_{n+1} \\
 e_{n+2}e_{n+4} &= e_{n+1} = -e_n+4e_{n+2} \\
 e_{n+4}e_{n+1} &= e_{n+2} = -e_{n+1}e_{n+4},
\end{align*}
\]

where the subscripts run modulo 2; see [10]. These were rediscovered by Cayley in 1845 and are also known sometimes as the Cayley numbers. Each nonzero element of \( \mathbb{O} \) still has an inverse so that \( \mathbb{O} \) is a division ring, albeit a nonassociative one. The octonions do retain some vestige of associativity in that they are an \textit{alternative algebra}, i.e. they have the property \( x(xy) = (xx)y \) and \( (yx)x = y(xx) \) for all \( x, y \in \mathbb{O} \). For a most comprehensive account of the octonions see [5].

The process of going from \( \mathbb{R} \) to \( \mathbb{C} \), from \( \mathbb{C} \) to \( \mathbb{H} \), and from \( \mathbb{H} \) to \( \mathbb{O} \), is in each case a kind of “doubling process”. At each stage something is lost—from \( \mathbb{R} \) to \( \mathbb{C} \) we lose the property that \( \mathbb{R} \) is ordered, from \( \mathbb{C} \) to \( \mathbb{H} \) we lose commutativity and from \( \mathbb{H} \) to \( \mathbb{O} \) we lose associativity. This process has been generalized to algebras over fields and indeed over rings. It is called “Dickson doubling” or “Cayley–Dickson Doubling”. See [10], [42]. If we apply the Cayley–Dickson doubling process to the octonions we obtain a structure called the sedenions, which is a 16-dimensional nonassociative algebra. The sedenions are not a division ring and are not even an alternative algebra, so that they seem to be of very limited use.

In 1844 Grassmann discovered exterior algebras, which subsequently became important in differential geometry. They have generators \( e_1, e_2, e_3, \ldots, e_n \) which anti-commute and satisfy \( e_i^2 = 0 \) for all \( i \) and are associative algebras.
In 1878 W. K. Clifford discovered Clifford algebras, defined to have generators $e_1, e_2, e_3, \ldots, e_n$ which anti-commute and satisfy $e_i^2 = a_i$ for chosen elements $a_1, a_2, a_3, \ldots, a_n$ of $\mathbb{R}$. Choosing $a_i = 0$ for all $i$ yields the exterior algebras as a special case, while choosing $n = 2$ and nonzero $a_i$ yields the quaternion algebras. In 1884 R. Lipschitz showed how Clifford algebras are related to quadratic forms, the orthogonal group and the spin group. The physicist P. A. M. Dirac rediscovered Clifford algebras in the 1920s and used them in connection with electron spin. Clifford algebras may be defined over fields in general and are useful in various parts of algebra, analysis and physics.

In 1878 Frobenius proved that the only associative finite-dimensional real division algebras are $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{H}$. See [24], p. 430–431, for a proof.

In 1881 there appeared a very long article [41] entitled “Linear Associative Algebras” by the Harvard mathematician Benjamin Peirce, the first such treatise on algebras per se.

In 1898 A. Hurwitz proved that the only real composition algebras are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$. (A real composition algebra is an algebra $A$ over $\mathbb{R}$, not necessarily associative or finite-dimensional, equipped with a nonsingular quadratic form $Q : A \to F$ such that $Q(ab) = Q(a)Q(b)$ for all $a, b \in A$. The form $Q$ is given by the norm. The multiplication in $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$ also yield the 1, 2, 4, and 8-square identities. See [10], [52] for more on all this.)

In 1899 D. Hilbert gave the first example of an associative division algebra which is infinite-dimensional over its centre. It is an algebra of twisted Laurent series. See [16] for details.

Also the late 19th century heralded the start of the study of Lie algebras by Killing [26], and Cartan [8]. At that time they were called “infinitesimal groups” and later were christened “Lie algebras” by Hermann Weyl in 1930. These are nonassociative algebras that have become of vital importance in mathematics and physics. (The 1888 paper of Killing has been described as the “greatest mathematical paper of all time” [9]!)

Another important class of nonassociative algebras was introduced in 1930 by the physicist P. Jordan. They were christened “Jordan algebras” by Albert in the 1940s and were studied by Albert and some other pure mathematicians. This was unbeknown to Jordan who was surprised, on visiting the USA, to discover that
these algebras had been named after him. See [40] for the history of Jordan algebras.

For more on the basic theory of nonassociative algebras see [51].

2.1. Division algebras. We will now concentrate on associative noncommutative finite-dimensional algebras, especially division algebras, and the developments in the 20th century on this topic.

One natural generalization of quaternion algebras is the notion of a symbol algebra, also sometimes known as a power norm residue algebra. This is defined as follows:

Let $F$ be a field of characteristic not two, and choose nonzero elements $a, b \in F$. (A choice with $a = b$ is not excluded.) Fix a positive integer $n \geq 2$ and assume that $F$ contains an element $\omega$ which is a primitive $n$-th root of unity. (Primitive means that $\{1, \omega, \omega^2, \ldots, \omega^{n-1}\}$ is a full set of all $n$-th roots of unity.) We introduce symbols $i$ and $j$ and consider the set of all expressions of the form

$$\sum_{r=0}^{n-1} \sum_{s=0}^{n-1} a_{rs} i^r j^s$$

where each $a_{rs} \in F$.

We make this set into an $F$-vector space of dimension $n^2$ by defining addition in the obvious way (add corresponding coefficients) and scalar multiplication in the obvious way (multiply each coefficient by the scalar). A natural basis is $\{i^r j^s\}$ where $r = 0, 1, \ldots, n - 1$ and $s = 0, 1, \ldots, n - 1$. We make this set into a ring via the usual distributive laws and the multiplication rules

$$i^n = a, \quad j^n = b, \quad ij = \omega ji.$$

The elements of $F$ are to commute with everything and the resulting $F$-algebra is called a symbol algebra and we label it $\mathbb{[a, b, F, n, \omega]}$.

It can be shown to be a central simple $F$-algebra and hence, by Wedderburn’s Theorem, is isomorphic to $M_k D$ for some integer $k$ and division algebra $D$ over $F$. For appropriate choices of $a, b, F, n, \omega$ it may be a division algebra.

Note that $[−1, −1, \mathbb{R}, 2, −1] = \mathbb{H}$ and $[a, b, F, 2, −1] = \left(\frac{a b}{F}\right)$ so that quaternion algebras are special cases of symbol algebras.

These symbol algebras are very natural generalizations of quaternion algebras. In [36] Milnor calls them symbol algebras because of their connection with Steinberg symbols and the K-theory of a field.
Earlier authors, interested in algebras over number fields, had used such algebras and referred to them as power norm residue algebras.

Another related kind of algebra is the cyclic algebra which was introduced by L. E. Dickson in 1914.

Let $F$ be a field and let $L$ be a cyclic field extension of degree $n$ of $F$, i.e., the Galois group of $L$ over $F$ is a cyclic group of order $n$. Choose a generator $\sigma$ of this Galois group. Adjoin an element $z$ to the field $L$ and consider the set of all expressions of the form

$$\sum_{i=0}^{n-1} \lambda_i z^i$$

where each $\lambda_i \in L$. Make this set into an $F$-vector space by defining addition in the obvious way (add corresponding coefficients) and multiplication by scalars from $F$ in the natural way, i.e. $\mu \sum_{i=0}^{n-1} \lambda_i z^i = \sum_{i=0}^{n-1} \mu \lambda_i z^i$ for $\mu \in F$. Define multiplication by the usual distributive law and the formulae

$$z^n = \delta, \quad z\lambda = \sigma(\lambda)z$$

for $\lambda \in L$,

where $\delta$ is some chosen nonzero element of $F$. The elements of $F$ commute with everything.

The resulting algebra will be labelled $(L/F, \sigma, \delta)$ and is called a cyclic algebra. It can be shown to be a central simple $F$-algebra of dimension $n^2$. For appropriate choices of $L, F, \sigma, \delta$ it may be a division algebra. Dickson in 1914 gave a specific example of a cyclic division algebra of degree three.

**Example 12.** Let $\theta$ be a primitive 7-th root of unity in $\mathbb{C}$. Let $\omega = \theta + \theta^2 + \theta^3, \alpha = \theta + \theta^5, \beta = \theta^2 + \theta^3, \gamma = \theta^3 + \theta^4$. Let $F = \mathbb{Q}(\omega)$ and $L = F(\alpha)$ which is a cyclic extension of degree three of $F$. The Galois group of $L$ over $F$ is generated by the automorphism $\sigma$ which cyclically permutes $\alpha, \beta, \gamma$. Then the cyclic algebra $(L/F, \sigma, \delta)$ is a division algebras for the choice of $\delta = \bar{\omega}/\omega$. (Here $\bar{\omega}$ is the complex conjugate of $\omega$.)

If we assume that the base field $F$ contains a primitive $n$-th root of unity then cyclic algebras and symbol algebras are more or less the same thing. (In the symbol algebra as defined above we may take $L = F(a^{1/n})$ which will be a cyclic extension provided that $L$ is an extension of degree $n$.)
Note that any quaternion division algebra \((\frac{a,b}{F})\) is necessarily a cyclic algebra since we may take \(L = F(i)\), a quadratic extension, i.e. a cyclic degree two extension field, and take \(z = j, \delta = b, \sigma(i) = -i\), i.e. \(\sigma\) is the nontrivial automorphism of \(L\).

2.2. The degree of a central simple algebra. It can be shown that the dimension of any central simple \(F\)-algebra \(A\), in particular any \(F\)-division algebra with centre \(F\), is a square. Thus \(\dim_F A = n^2\) for some integer \(n\). We define the degree of \(A\) to equal this integer \(n\), i.e. degree \(A = \sqrt{\dim_F A}\). See for example chapter 13 of [45].

Recall that the degree of a field extension \(E\) of a field \(F\) is defined as the dimension of \(E\) as a vector space over \(F\) and is denoted \([E : F]\).

It can be shown that a central simple \(F\)-algebra \(A\) is a cyclic algebra if and only if \(A\) contains a strictly maximal subfield \(E\) which is a cyclic extension of \(F\). Strictly maximal means that \([E : F] = n\), where \(n\) is the degree of \(A\).

2.3. Division algebras of prime degree. Any division algebra of degree two over a field \(F\) must be a quaternion division algebra, see section 1 of this article, and thus is a cyclic algebra.

In 1921 Wedderburn proved that any division algebra of degree three over a field \(F\) is necessarily a cyclic algebra [62]. See also the elegant proof of Haile appearing in [27], p. 303–304. The Wedderburn Factorization Theorem is used in these proofs.

Open Question: Let \(p\) be a prime. Is every division algebra of degree \(p\) a cyclic algebra?

This is still open for all primes \(p \geq 5\).

For some familiar fields a lot more is known. For the real field \(\mathbb{R}\) we know by Frobenius’ Theorem that the only finite-dimensional noncommutative division algebra is \(\mathbb{H}\), and for \(p\)-adic fields it is not too hard to show that the only noncommutative finite-dimensional division algebras are cyclic ones. A major accomplishment of the early 1930s is the following result, a combination of the work of four mathematicians:

**Theorem 13** (The Albert–Brauer–Hasse–Noether Theorem). Every finite-dimensional noncommutative division algebra over an algebraic number field is a cyclic algebra.

See [47]. The proof uses localization and a local-global principle.
The following is a natural question to ask.

Question: Is every finite-dimensional division algebra a cyclic algebra?

The answer is no. In 1932 Albert gave the first example of a noncyclic division algebra. His example is a biquaternion algebra. Specifically we take \( F = \mathbb{R}(x, y) \) and \( A = \left( \frac{-1}{F} \right) \otimes \left( \frac{-x}{F} \right) \). This is a division algebra because the Albert form can be shown to be anisotropic. On the other hand it can be shown that \( A \) does not contain any cyclic subfield of degree 4 over \( F \) and hence is not a cyclic algebra. See chapter 15 of [45].

We should remark that a biquaternion algebra may sometimes have the structure of a cyclic algebra. For a discussion of when this can happen see [32].

2.4. Crossed product algebras. Following ideas of Dickson the notion of crossed product algebra was introduced in the 1920s by Noether and Brauer. It is a generalization of the idea of a cyclic algebra.

Let \( L \) be a Galois extension field of \( F \) and let \( G \) be the Galois group of all \( F \)-automorphisms of \( L \). We introduce a set of symbols \( u_\sigma \) indexed by the elements \( \sigma \in G \), and we consider the set of all expressions of the form

\[
\sum_{\sigma \in G} a_\sigma u_\sigma
\]

where each \( a_\sigma \in L \).

We make this set into an \( F \)-algebra in a similar way to cyclic algebras, but now using the multiplication formulae

\[
\begin{align*}
    u_\sigma a &= \sigma(a)u_\sigma \quad \text{for } a \in L \\
    u_\sigma u_\tau &= k_{\sigma \tau} u_{\sigma \tau}
\end{align*}
\]

where \( k_{\sigma \tau} \in L \) is a 2-cocycle. The set \( \{k_{\sigma \tau}, \sigma \in G, \tau \in G\} \) must satisfy the cocycle condition

\[
\rho(k_{\sigma \tau})k_{\rho(\sigma \tau)} = k_{\rho \sigma}k_{(\rho \sigma)\tau}
\]

for all \( \rho, \sigma, \tau \in G \). This condition is needed in order that associativity of multiplication is satisfied.

A set \( \{k_{\sigma \tau}, \sigma \in G, \tau \in G\} \) satisfying this cocycle condition is called a factor set. (These are known as Noether factor sets. A variation of the definition yields Brauer factor sets. See [23].) We will denote this factor set by \( k \) for short and the resulting algebra
will be labelled \((L/F, k)\) and called a crossed product algebra. It can be shown that \((L/F, k)\) is a central simple \(F\)-algebra.

Cyclic algebras are crossed products, the factor set for the cyclic algebra \((L/F, \sigma, \delta)\) being given by

\[
\begin{align*}
k_{\sigma^i \sigma^j} &= 1 \text{ if } i + j < n \\
k_{\sigma^i \sigma^j} &= \delta \text{ if } i + j \geq n
\end{align*}
\]

where \([L : F] = n\).

It can be shown that a central simple \(F\)-algebra \(A\) is a crossed product algebra if and only if \(A\) contains a strictly maximal subfield \(E\) which is a Galois extension of \(F\).

Albert’s example of a noncyclic division algebra can be shown to be a crossed product. Indeed any biquaternion division algebra is a crossed product [32] with group \(G\) being the Klein 4-group.

It is now natural to ask:

**Question:** Is every finite-dimensional division algebra a crossed product algebra?

The question remained open until 1972 when Amitsur [3] produced examples of non-crossed product division algebras of certain degrees \(n\). These are the so-called universal division algebras of degree \(n\) which he showed are not crossed products at least in the case when \(n\) is either divisible by 8 or by \(p^2\) where \(p\) is an odd prime. See [45].

This result says nothing about algebras of prime degree and so does not answer our earlier question about whether or not they are cyclic.

2.5. The Tannaka–Artin problem. In the context of division algebras we should also mention the Tannaka–Artin problem.

Let \(D\) be a finite-dimensional division algebra over a field \(F\). There is a natural reduced norm map \(N: D \to F\).

Write \(\text{Ker}N\) for the kernel of this map, i.e. \(\text{Ker}N = \{x \in D : N(x) = 1\}\).

Write \([\hat{D}, \hat{D}]\) for the subgroup of \(\hat{D}\) generated by all of the multiplicative commutators \(xyz^{-1}y^{-1}\) where \(x, y \in \hat{D}\). Then, since the norm is multiplicative, the group \([\hat{D}, \hat{D}]\) is contained in \(\text{Ker}N\).

The reduced Whitehead group \(SK_1(D)\) is defined to be the quotient group \(\text{Ker}N/[\hat{D}, \hat{D}]\).

The problem of determining whether or not \(SK_1(D) = 1\) is known as the Tannaka–Artin problem.
It was believed for a long time that $SK_1(D) = 1$ always. It is true that $SK_1(D) = 1$ for all local and global fields [60], but Platonov [44] in 1975 showed that in general it can be nontrivial.

Some easy examples using biquaternion algebras are produced in [16]. Specifically we take any field $K$ containing a primitive 4-th root of unity $\omega$ and consider the iterated Laurent series field $F = K((t_1))((t_2))((t_3))((t_4))$.

Let $D = (\mathbb{Q}(\sqrt{2}) \otimes \mathbb{Q}(\sqrt{3}))$, a biquaternion division algebra. Then $SK_1(D) \neq 1$ because the element $\omega$ has norm 1 but can be shown not to be a commutator.

2.6. Group theory. The quaternions have had an impact on group theory and group representation theory.

The quaternion group of order 8 is the finite multiplicative subgroup $\{\pm 1, \pm i, \pm j, \pm k\}$ of the multiplicative group $\mathbb{H}$, and is the first in the set of generalized quaternion groups of order $2^n$, $n \geq 3$.

Quaternionic representations, that is, the representation of finite groups by matrices of quaternions, play a role in group representation theory, with matrices over $\mathbb{H}$ and other quaternion algebras appearing as components of the group algebra of certain groups. See [12], also [47], p. 383.

In 1940 H. S. M. Coxeter, [11], classified the binary polyhedral groups and other generalizations of the quaternion group, it being related to his work on polytopes.

A recent breakthrough linking division algebras and group theory is the following result of Rapinchuk, Segev, and Seitz, [46], that finite quotients of division algebras are solvable. More precisely it says that if $D$ is a division ring, finite-dimensional over its centre, and if $N$ is a normal subgroup of the multiplicative group $\hat{D}$ such that the quotient group $\hat{D}/N$ is a finite group then this group $\hat{D}/N$ is a solvable group.

2.7. Other parts of the legacy. We summarize quickly some other areas of mathematics in which the quaternions have had an influence. A detailed study is outside the scope of this article (and this author!)

2.7.1. Number theory. Hurwitz introduced what he called the ring of integral quaternions. This is the subring of $\mathbb{H}$ consisting of all quaternions of the form $\alpha + \beta i + \gamma j + \delta k$ where either each of $\alpha, \beta, \gamma, \delta \in \mathbb{Z}$, or else each of $\alpha, \beta, \gamma, \delta$ is congruent to 1/2 modulo $\mathbb{Z}$. This turns out to be a ring with a good division algorithm. If we just allowed
quaternions with integer coefficients then the division property will not work. See [10] for a discussion of factorization questions, and [22], chapter 7, where the integral quaternions are used to prove the famous Lagrange theorem that every positive integer is a sum of at most four squares.

There has been a lot of work on orders (i.e. full lattices) in quaternion algebras, mainly over the local and global fields of number theory. See [18], [57].

A discussion of “integral octonions” can be found in [10] and they are used in constructing the $E_8$ lattice, related to sphere packing problems.

2.7.2. Analysis. Complex analysis, i.e. the theory of functions of a complex variable, has been extended to quaternionic analysis. Indeed there is a more general Clifford analysis, working with functions taking values in a Clifford algebra, see for example [13]. Many of the notions of complex analysis, e.g. contour integration, have been extended to this context. The subject has some applications in mathematical physics.

2.7.3. Geometry. The geometric legacy is at least as great as the algebraic legacy. There is a wealth of interpretations, consequences and applications of quaternions in a geometric context.

Quaternions can be used to describe rotations in 3-dimensional and in 4-dimensional space. For example see [17], [10], [2].

Quaternions have been useful in topology. Topologists have utilized quaternionic projective space (and also octonion projective space), defined in similar fashion to real and complex projective space. See [43]. In homotopy theory there are the Hopf maps between spheres, $S^3 \to S^2$, $S^7 \to S^4$, and $S^{15} \to S^8$. They are defined using, respectively, complex numbers, quaternions, and octonions. These were the first examples of homotopically nontrivial maps from a sphere to a lower dimensional sphere and were discovered by H. Hopf in the early 1930s. The quaternions and octonions also have links with Bott periodicity, i.e. the periodic nature of the homotopy groups of the classical groups. See [5] for more on this.

Geometers and topologists have studied quaternionic structures on manifolds, these also being of considerable interest to physicists. See [33].
2.7.4. **Physics.** Quaternions have had applications in physics, although they have never achieved what Hamilton had hoped for when he made his discovery in 1843. In recent years the octonions have turned out to be particularly important in string theory, the so-called “theory of everything” which attempts to unify general relativity and quantum theory. Many of the kinds of algebras mentioned earlier, as well as some we did not mention, have been used in physics. See Okubo [39], Dixon [15].

2.7.5. **Other Applications.** Quaternions have been used in inertial navigation systems [28], in computer graphics for computer animation [55], in image processing [49], [50], and in coding theory [54].

**References**


David W. Lewis,
School of Mathematical Sciences,
University College Dublin,
Belfield,
Dublin 4, Ireland
david.lewis@ucd.ie

Received on 10 October 2005 and in revised form on 3 November 2005.