Quaternions and Algebras

Soham Dam

July 15, 2023

Abstract

Quaternions are numbers composed of a real part an an imaginary part, similar to complex numbers. The imaginary part uses 3 imaginary units, i, j, and k, which satisfy the equation $i^2 = j^2 = k^2 = ijk = -1$. First discovered by William Rowan Hamilton in 1843, quaternions have formed the foundation of a new algebra over the real numbers. This paper will discuss quaternion algebras, quadratic forms, and applications of quaternions in the real world.

Contents

1 Introduction

Section 2 will introduce quaternion algebras and connections to ring theory. Section 3 will discuss involutions and opposite algebras. The next two sections will discuss quadratic forms, with section 5 diving more into ternary quadratic forms. This section will also discuss a special case of the characteristic of the field where the algebra is based. Section 6 will discuss simple algebras and simple modules. Finally, in section 7, we prove that the only composition algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, the set of all quaternions, and $\mathbb{O}$, the set of all octonions. In section 9, we uncover a couple of applications of quaternions in the real world.

2 Quaternion algebras

First, we must define the new number systems of quaternions and octonions:

Definition 2.1. A quaternion is a number of the form $t + xi + yj + zk$ for $t, x, y, z \in \mathbb{R}$, with $i^2 = j^2 = k^2 = ijk = -1$. An octonion is a number of the form $x_\infty + \sum_{n=0}^{6} x_n i_n$ where the i_n satisfy $i_n^2 = -1$ and

$$i_{n+1}i_{n+2} = i_{n+4} = -i_{n+2}i_{n+1}$$
$$i_{n+2}i_{n+4} = i_{n+1} = -i_{n+4}i_{n+2}$$
$$i_{n+4}i_{n+1} = i_{n+2} = -i_{n+1}i_{n+4}$$

with indices taken modulo 7.

2.1 Groups, rings, and fields

An introduction to ring theory is required for the study of quaternion algebras and quadratic forms. The basic definitions are given below:

Definition 2.2. A group G is a set with binary operation $*$ with the following properties:

- For $a, b, c \in G$, $*$ satisfies $(a * b) \in G$ and $(a * b) * c = a * (b * c)$;
- There exists $e \in G$ such that $a * e = e * a = a$ for all $a \in G$; and
- For all $a \in G$, there exists $b \in G$ such that $a * b = b * a = e$.

A group is abelian if $a * b = b * a$ for all $a, b \in G$.

Definition 2.3. A ring R is a set with two binary operations, $+$ and $\times$ such that:

- R is an abelian group under $+$ with additive identity 0;
- For $a, b, c \in R$, $\times$ satisfies $(a \times b) \in R$ and $(a \times b) \times c = a \times (b \times c)$; and
- For all $a, b, c \in R$, we have $a \times (b + c) = a \times b + a \times c$ and $(a + b) \times c = a \times c + b \times c$.

Definition 2.4. A field F is a set with operations $+$ and $\times$ such that F is abelian under $+$, and $F \setminus \{0\}$ is abelian under $\times$. The characteristic of F is how many times the multiplicative identity must be added to get the additive identity. If this cannot be done, then $\operatorname{char} F = 0$.

Throughout the paper, let F be a commutative field with an algebraic closure. Assume that all rings are associative, not necessarily commutative, and have multiplicative identity 1.

2.2 Quaternion algebras

Below is the definition of an algebra provided on page 21 of [Voi21]:

Definition 2.5. An algebra over a field F is a ring B with homomorphism $F \to B$ such that the image of F lies in the center $Z(B)$ of B, defined by

$$Z(B) := \{\alpha \in B : \alpha\beta = \beta\alpha \quad \forall \beta \in B\}.$$

Definition 2.6. A homomorphism of F-algebras is a ring homomorphism which restricts to the identity on F. Such homomorphism is necessarily F-linear. An endomorphism is a homomorphism of $B \to B$. An isomorphism is an invertible F-algebra homomorphism, and an invertible endomorphism is an automorphism.

Definition 2.7. An algebra B is a quaternion algebra if there exist $i, j \in B$ such that $1, i, j, ij$ is an F-basis for B and

$$i^2 = a, \ j^2 = b, \text{ and } ji = -ij.$$

A division ring is a ring D such that $D \setminus \{0\}$ is a group under multiplication. A division algebra is an algebra that is also a division ring.

Lemma 2.8. *An F-algebra B is a quaternion algebra if and only if there exist nonzero $i, j \in B$ that generate B as an F-algebra and satisfy*

$$i^2 = a,\ j^2 = b,\ \text{and } ji = -ij.$$

with $a, b \in F^\times$.

Proof. It is necessary and sufficient to show that $1, i, j, ij$ are linearly independent. Suppose that $\alpha = t + xi + yj + zij = 0$ with $t, x, y, z \in F$. Using the relations given, we find that

$$0 = i(\alpha i + i\alpha) = 2a(t + xi).$$

Since $\operatorname{char} F \neq 2$ and $a \neq 0$, we conclude that $t + xi = 0$. Similarly, we have $t + yj = t + zij = 0$. Thus,

$$\alpha - (t + xi) - (t + yj) - (t - zij) = -2t = 0.$$

Since i, j are nonzero and B is not the zero ring, we have $1 \neq 0$, so $t = 0$. If $x \neq 0$, we have $i^2 = 0^2 = a = 0$, which is impossible, so x is zero. Similarly, y and z are zero. □

We define conjugation as we do for complex numbers: just negate the imaginary part. We can also represent Hamilton's quaternions in matrix form. We have

$$t + xi + yj + zk = u + j\overline{v} \mapsto \begin{bmatrix} u & -v \\ \overline{v} & \overline{u} \end{bmatrix}$$

with $u := w + xi$ and $v := y + zi$.

Denote by $\mathbb{H}^0$ the set of quaternions with imaginary part 0. Then $vw = -v \cdot w + v \times w$ where $\cdot$ is the dot product and $\times$ is the cross product. Two vectors are orthogonal when their dot product is 0.

Lemma 2.9. *For all $v, w \in \mathbb{H}^0$, we have the following:*

1. *$vw \in \mathbb{H}^0$ if and only if v, w are orthogonal.*
2. *$v^2 = -||v||^2 \in \mathbb{R}$*
3. *$wv = -vw$ if and only if v, w are orthogonal.*

Proof. Direct application of the equation $vw = -v \cdot w + v \times w$. □

3 Involutions

3.1 Definition of an involution

Definition 3.1. An involution $\_\_ : B \to B$ is an F-linear map that satisfies:

1. $\overline{1} = 1$, where 1 is the multiplicative identity;
2. $\overline{\overline{\alpha}} = \alpha$ for all $\alpha \in B$; and
3. $\overline{\alpha\beta} = \overline{\beta}\,\overline{\alpha}$ for all $\alpha, \beta \in B$.

3.2 Properties of involutions

Definition 3.2. An involution is standard if $\alpha\overline{\alpha} \in F$ for all $\alpha \in B$.

Examples of standard involutions include the identity map on F as an F-algebra, and the $\mathbb{R}$-algebra $\mathbb{C}$.

Lemma 3.3. *Suppose that* $\operatorname{char} F \neq 2$*, and let* $B = \left(\frac{a,b}{F}\right)$*. The map*

$$\alpha = t + xi + yj + zij \mapsto \overline{\alpha} = t - xi - yj - zij$$

is a standard involution on B*.*

Proof. The first two properties are satisfied as $\overline{1} = 1$ and $\overline{\overline{\alpha}} = \alpha$. For the third property, we only need to verify for a basis due to F-linearity:

$$\overline{ij} = -ij = ji = (-j)(-i) = \overline{j}\overline{i}.$$

The other multiplications can be proven similarly. Finally, we have

$$\alpha\bar{\alpha} = (t + xi + yj + zij)(t - xi - yj - zij) = t^2 - ax^2 - by^2 + abz^2 \in F.$$

□

We now move to reduced trace and reduced norm. We define the reduce trace as

$$\operatorname{trd} : B \to F,\ \alpha \mapsto \alpha + \overline{\alpha} \tag{3.1}$$

and similarly the reduced norm

$$\operatorname{nrd} : B \to F,\ \alpha \mapsto \alpha\overline{\alpha} \tag{3.2}$$

3.3 Quadratic fields

We next explore the degree of an algebra. The degree is the smallest nonnegative integer m such that every element $\alpha \in B$ satisfies a monic polynomial of degree m, if such an integer exists; otherwise, the degree is ∞.

We now state the following theorem:

Theorem 3.4. *Suppose* $\operatorname{char} F \neq 2$ *and let* B *be a division* F*-algebra. Then* B *has degree at most 2 if and only if one of the following holds:*

1. $B = F$*;*
2. $B = K$ *is a quadratic field extension of* F*; or*
3. B *is a division quaternion algebra over* F*.*

The proof is given on page 41 on [Voi21].

Corollary 3.5 (Frobenius)**.** *Let* B *be a division algebra of finite degree over* $\mathbb{R}$*. Then either* $B = \mathbb{R}$ *or* $B \simeq \mathbb{C}$ *or* $B \simeq \mathbb{H}$ *as* $\mathbb{R}$*-algebras.*

Proof. Suppose $\alpha \in B \setminus \mathbb{R}$. Then we know that $\mathbb{R}(\alpha) \simeq \mathbb{C}$, so α satisfies a polynomial of degree 2. If $B \neq \mathbb{R}$, then B has degree 2, so $B \simeq \mathbb{C}$ or B is a division quaternion algebra over $\mathbb{R}$, meaning $B \simeq \mathbb{H}$. □

4 Quadratic forms

From Equation 3.2, we know that a quaternion algebra B has a reduced norm map, which defines a quadratic form, a homogeneous polynomial of degree 2 in $F[t, x, y, z]$.

Let $Q : V \to F$ be a quadratic form. Then Q can be diagonalized by a change of variables: there is a basis $e_1, \ldots, e_n$ of V such that

$$Q(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = Q(x_1, x_2, \ldots, x_n) = a_1x_1^2 + a_2x_2^2 + \cdots + a_nx_n^2 \tag{4.1}$$

with $a_i \in F$. The discriminant is defined by $\mathrm{disc}(Q) := a_1a_2 \cdots a_n/2^n \in F/F^{\times 2}$. A quadratic form is nondegenerate if the discriminant is nonzero. Now we explore the orthogonal group of a quadratic form:

Definition 4.1. A similarity from Q to another quadratic form $Q' : V' \to F$ is a pair (f, u) where $f : V \tilde{\to} V'$ is an F-linear isomorphism and $u \in F^\times$ satisfy $Q'(f(x)) = uQ(x)$ for all $x \in V$. An isometry is a similarity with $u = 1$.

Definition 4.2. The orthogonal group of Q is the group of self-isometries of Q. In other words,

$$O(Q)(F) := \{f \in \mathrm{Aut}_F(V) : Q(f(x)) = Q(x)\ \forall x \in V\}.$$

[Voi21] states the following main result about quadratic forms and quaternion algebras:

Theorem 4.3. *Let B be an F-algebra. Then B has a nondegenerate standard involution if and only if one of the following holds:*

1. *$B = F$;*
2. *$B = K$ has $\dim_F K = 2$ and either $K \simeq F \times F$ or K is a field; or*
3. *B is a quaternion algebra over F.*

Now, we are ready to define a quadratic form:

Definition 4.4. A quadratic form is a map $Q : V \to F$ on an F-vector space V satisfying:

1. $Q(ax) = a^2Q(x)$ for all $a \in F$ and $x \in V$; and
2. The map $T : V \times V \to F$ defined by

$$T(x, y) = Q(x + y) - Q(x) - Q(y)$$

is F-bilinear.

Definition 4.4 can help us verify the statement in Equation 4.1. Let $e_1, e_2, \ldots, e_n$ be a basis for V with finite dimension. Then we have

$$Q(x_1e_1 + x_2e_2 + \cdots + x_ne_n) = \sum_i Q(e_i)x_i^2 + \sum_{i<j} T(e_i, e_j)x_ix_j \in F[x_1, x_2, \ldots, x_n].$$

The diagonalization makes the second sum above equal to 0, thus making the above have the same form as Equation 4.1. In matrix form, we have

$$[T] := (T(e_i, e_j))_{i,j} \in M_n(F)$$

where $M_n(F)$ is the algebra of $n \times n$ matrices with entries in F. $[T]$ is known as the Gram matrix of Q. This allows for a new way of viewing orthogonal vectors in a vector space: x and y in V are orthogonal if $T(x, y) = 0$.

Lemma 4.5. *Let B be an F-algebra with a standard involution. Then* $\operatorname{nrd} : B \to F$ *is a quadratic form on B. Indeed,* $\operatorname{nrd}(a\alpha) = a^2\alpha$ *for all $\alpha \in B$, and we have*

$$T(\alpha, \beta) = (\alpha + \beta)\overline{(\alpha + \beta)} - \alpha\overline{\alpha} - \beta\overline{\beta} = \alpha\overline{\beta} + \beta\overline{\alpha} = \alpha\overline{\beta} + \overline{\alpha\overline{\beta}} = \operatorname{trd}(\alpha\overline{\beta}).$$

We then have

$$\operatorname{trd}(\alpha\overline{\beta} = \operatorname{trd}(\alpha(\operatorname{trd}(\beta) - \beta)) = \operatorname{trd}(\alpha)\operatorname{trd}(\beta) - \operatorname{trd}(\alpha\beta).$$

This means that α and β are orthogonal if and only if $\operatorname{trd}(\alpha\overline{\beta}) = \alpha\overline{\beta} + \beta\overline{\alpha} = 0$ *if and only if* $\operatorname{trd}(\alpha\beta) = \operatorname{trd}(\alpha)\operatorname{trd}(\beta)$.

Definition 4.6 (Discriminant in matrix form)**.** The discriminant of a quadratic form Q is $\operatorname{disc}(Q) := 2^{-n} \det T$ where T is the Gram matrix. The signed discriminant is $(-1)^{n(n-1)/2} \operatorname{disc}(Q)$.

We are now ready to prove Theorem 4.3:

Proof. If $B = F$, then the standard involution is the identity and nrd is nondegenerate. If $\dim_F K = 2$, then after completing the square, we have $K \simeq F[x]/(x^2 - a)$ and in the basis $1, x$ we find $\operatorname{nrd} \simeq \langle 1, a\rangle$. Thus, nrd is nondegenerate if and only if $a \in F^\times$ if and only if $K \simeq F \times F$.

Now suppose $\dim_F B > 2$. Let $1, i, j$ be part of a normalized basis for B with respect to the quadratic form nrd. Then $T(1, i) = 0$, so $i^2 = a \in F^\times$ since nrd is nondegenerate. Similarly, $j^2 = b \in F^\times$, so we have $ij + ji = 0$. We have $T(1, ij) = \operatorname{trd}(ij) = 0$ and $T(ij, i) = \operatorname{trd}(\bar{i}(ij)) = -a\operatorname{trd}(j) = 0$ and similarly $T(ij, j) = 0$. Thus, ij is orthogonal to 1, i, and j. If $ij = 0$, then $i(ij) = aj = 0$, meaning $j = 0$, contradiction. Since nrd is nondegenerate, it follows that the set $\{1, i, j, ij\}$ is linearly independent.

Therefore, the subalgebra A of B generated by i and j satisfies $A \simeq \left(\dfrac{a, b}{F}\right)$. If $\dim_F B = 4$, we are done. So let $k \in A^\perp$; then $\operatorname{trd}(k) = 0$ and $k^2 = c \in F^\times$. Thus, $k \in B^\times$, with $k^{-1} = c^{-1}k$. By Lemma 4.5, we have $k\alpha = \overline{\alpha}k$ for any $\alpha \in A$ since $\bar{k} = -k$. But then

$$k(ij) = (\overline{ij})k = \bar{j}\bar{i}k = \bar{j}ki = k(ji),$$

meaning $ij = ji = -ij$ since $k \in B^\times$, which is a contradiction. □

Theorem 4.7 (Cartan-Dieudonné)**.** *Let (V,Q) be a nondegenerate quadratic space with* $\dim_F V = n$. *Then every isometry $f \in O(Q)(F)$ is a product of at most n reflections.*

Proof. See the references given in [Voi21]. The proof uses induction on n. □

5 Ternary quadratic forms

5.1 More quadratic forms

We begin the section with a proposition relating quadratic spaces and quaternion algebras:

Proposition 5.1. *Let B and B' be quaternion algebras over F. Then the following are equivalent:*

1. *$B \simeq B'$ are isomorphic as F-algebras;*
2. *$B \simeq (B')^{op}$ are isomorphic as F-algebras;*
3. *$B \simeq B'$ are isometric as quadratic spaces; and*
4. *$B^0 \simeq (B')^0$ are isometric as quadratic spaces.*

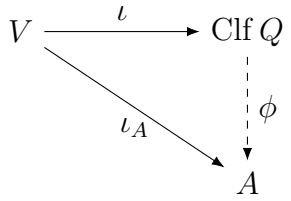
An isometry $f : B^0 \tilde{\to} (B')^0$ extends uniquely to an isomorphism for $B \tilde{\to} B'$ or $B \tilde{\to} (B')^{op}$.

Proof. See page 67 of [Voi21]. □

The ternary quadratic forms are those of quaternions whose real part is 0. This quadratic forms carries the same information as a quadratic form for all quaternions. Next, we explore the Clifford algebra of a quadratic form Q:

Proposition 5.2. *Let $Q : V \to F$ be a quadratic form where F has finite dimension and arbitrary characteristic. Then there exists an F-algebra $\operatorname{Clf} Q$ such that:*

1. *There is an F-linear map $\iota : V \to \operatorname{Clf} Q$ such that $\iota(x)^2 = Q(x)$ for all $x \in V$; and*
2. *$\operatorname{Clf} Q$ has the following universal property: if A is an F-algebra and $\iota_A : V \to A$ is a map such that $\iota_A(x)^2 = Q(x)$ for all $x \in V$, then there exists a unique F-algebra homomorphism $\phi : \operatorname{Clf} Q \to A$ such that the diagram*



commutes.

Proof. See page 69 of [Voi21]. □

5.2 What if $\operatorname{char} F = 2$?

The theorems in the previous sections assumed that $\operatorname{char} F \neq 2$. Here, we explore this special case. First, we define the basis of a quaternion algebra B over F with $\operatorname{char} F = 2$:

Definition 5.3. A quaternion algebra B over F with $\operatorname{char} F = 2$ has a basis $1, i, j, k$ such that

$$i^2 + i = a,\ j^2 = b,\ \textit{and}\ k = ij = j(i+1).$$

We now prove the following theorem:

Theorem 5.4. *Let B be a division F-algebra with a standard involution that is not the identity. Then either B is a separable quadratic field extension of F or B is a quaternion algebra over F.*

To prove this theorem, we first prove the following:

Theorem 5.5. *Let B be an F-algebra with $\operatorname{char} F = 2$. Then B has a nondegenerate standard involution if and only if one of the following holds:*

1. *$B = F$;*
2. *$B = K$ is a separable quadratic F-algebra; or*
3. *B is a quaternion algebra over F.*

Proof of Theorem 5.5. If $B = F$, then the standard involution is the identity.

If $\dim_F B = 2$, then $B = K$ has a unique standard involution. This involution is nondegenerate if and only if K is separable.

So suppose $\dim_F B = 2$. Since B has a nondegenerate standard involution, there exists an element $i \in B$ such that $T(i, 1) = \operatorname{trd}(i) \neq 0$. i is not in F since the reduced trace of elements in F is 0. Without loss of generality, suppose that $\operatorname{trd}(i) = 1$, whence $i^2 = i + a$ for $a \in F$, and $\operatorname{nrd}\|_{F+Fi} = [1, a]$.

By nondegeneracy, there exists j orthogonal to 1 and i such that $\operatorname{nrd} j = b \neq 0$. Thus, $\operatorname{trd}(j) = 0$ so $\bar{j} = j$ and $j^2 = b \in F^\times$. Furthermore,

$$0 = \operatorname{trd}(ij) = ij + j\bar{i} = ij + j(i+1)$$

meaning $ij = j(i+1)$. The remainder of the proof finishes as in the proof of Theorem 4.3. □

Now, Theorem 5.4 can be proven using the above proof.

6 Simple algebras

6.1 Motives and introduction

Among the set of all algebras are the simple algebras, ones that cannot be broken down further. A ring A is simple if it has no nontrivial two-sided ideals, meaning that the only two-sided ideals are $\{0\}$ and A. If $\phi : A \to A'$ is a ring homomorphism and A is simple, then ϕ is either injective or the zero map. The main result we wish to prove is:

Theorem 6.1 (Wedderburn-Artin)**.** *Let F be a field and B be a finite-dimensional F-algebra. Then B is simple if and only if $B \simeq M_n(D)$ where n is a positive integer and D is finite-dimensional division F-algebra.*

It is also convenient to work with semisimple algebras, which are finite direct products of simple algebras. The second main result concerns with simple subalgebras of simple algebras:

Theorem 6.2 (Skolem-Noether)**.** *Let A, B be simple F-algebras and suppose that B is central. Suppose that $f, g : A \to B$ are homomorphisms. Then there exists $\beta \in B$ such that $f(\alpha) = \beta^{-1} g(\alpha)\beta$ for all $\alpha \in A$.*

6.2 Simple modules

Throughout this section, let B be a finite-dimensional F-algebra.

A representation of B over F is a vector space V over F together with an F-algebra homomorphism $B \to \mathrm{End}_F(V)$. For example, the space of column vectors F^n is a left module of $M_n(F)$ and the space of row vectors is a right $M_n(F)$-module.

Definition 6.3. Let V be a left B-module. V is simple if $V \neq \{0\}$ and the only B-submodules of V are $\{0\}$ and V.

Lemma 6.4. *A finite-dimensional left B module V admits a filtration*

$$V = V_0 \supsetneq V_1 \supsetneq \cdots \supsetneq V_r = \{0\}$$

such that V_i/V_{i+1} is simple for each i.

Lemma 6.5 (Schur)**.** *Let B be an F-algebra, and V_1, V_2 be simple B-modules. Then any homomorphism $\phi : V_1 \to V_2$ is either zero or an isomorphism.*

Proof. We have that $\ker \phi$ and $\mathrm{img}\, \phi$ are B-submodules of V_1 and V_2, respectively, so either $\phi = 0$ or $\ker \phi = \{0\}$ and $\mathrm{img}\, \phi = V_2$, hence, $V_1 \simeq V_2$. □

One point to take away is that endomorphisms of a left module act on the right, but the more common convention is that endomorphisms act on the left. This is where opposite algebras are involved.

Lemma 6.6. *Let B be a finite-dimensional simple F-algebra. Then there exists a simple left B-module that is unique up to isomorphism.*

Proof. Since B is finite-dimensional over F, there exists a nonzero left ideal I of B of minimal dimension, and such an ideal I is necessarily simple. Moreover, if $v \in I$ is nonzero then $Bv = I$, since $Bv \subseteq I$ is nonzero and I is simple. Let $I = Bv$ with $v \in I$.

Now let V be any simple B-module; we will show $I \simeq V$ as B-modules. Since B is simple, the natural map $B \to \mathrm{End}_F(V)$ is injective (since it is nonzero). Therefore, there exists $x \in V$ such that $vx \neq 0$, so $Ix \neq \{0\}$. Thus, the map $I \to V$ by $\beta \mapsto \beta x$ is a nonzero B-module homomorphism, so it is an isomorphism by Lemma 6.5. □

6.3 Semisimple Modules

Assume that B is an F-algebra of finite dimension, and V is a B-module of finite dimension.

Definition 6.7. A B-module V is semisimple if V is isomorphic to a finite direct sum of simple B-modules $V \simeq \bigoplus_i V_i$. B is a semisimple F-algebra if B is semisimple as a left B-module.

Lemma 6.8. *The following statements are always true:*

1. *A B-module V is semisimple if and only if it is the sum of simple B-modules.*
2. *A submodule or quotient module of a semisimple B-module is semisimple.*
3. *If B is a semisimple F-algebra, then every B-module is semisimple.*

Proof. For item 1, let $V = \sum_i V_i$ be the sum of simple B-modules. Since V is finite-dimensional, we can rewrite it as an irredundant finite sum; and then since each V_i is simple, the intersection of any two distinct summands is $\{0\}$, so the sum is direct.

For item 2, Let W be a submodule of semisimple B-module V. Every $x \in W$ with $x \neq 0$ is contained in a simple B-module of W by minimality, so $W = \sum_i W_i$ is a sum of B-modules. The result now follows from item 1 for submodules. For quotient modules, let $\phi : V \to Z$ be a surjective B-module homomorphism, and $\phi^{-1}(Z) = \sum_i W_i$ is a sum of simple B-modules. The proof finishes by Schur's Lemma.

For item 3, let V be a B-module. Since V is finitely generated, there is a surjective homomorphism $B^r \to V$. for some $r \geq 1$. Since B^r is semisimple, so is V. □

7 Hurwitz integral quaternions

We must define the set of quaternions that are integral. One obvious choice is to let $t + xi + yj + zk$ be integral when $t, x, y, z \in \mathbb{Z}$. These numbers are known as the Lipschitz integers (L). However, Hurwitz found an alternate definition with nicer properties:

Definition 7.1. A quaternion $t + xi + yj + zk$ is integral when $t, x, y, z \in \frac{1}{2}\mathbb{Z}$, where $\frac{1}{2}\mathbb{Z}$ denotes the set of all integers and half-integers. Denote the set of Hurwitz integers by H.

The reason that this definition is more suitable is because it satisfies the "division with small remainder" property.

7.1 Hurwitz units and primes

In integers, $p = 1 \times p = p \times 1$. Similarly, in Hurwitz integers, we have

$$P = P' \times U = V \times P''$$

We must now find the Hurwitz units, namely the Hurwitz integers of norm 1:

Theorem 7.2. *There are precisely* 24 *Hurwitz units. They are* $\pm 1, \pm i, \pm j, \pm k$*, and the* 16 *others* $\pm\frac{1}{2} \pm \frac{1}{2}i \pm \frac{1}{2}j \pm \frac{1}{2}k$*.*

Proof. For a Lipschitz unit, we must have $|t|, |x|, |y|, |z| \geq 1$ and $t^2 + x^2 + y^2 + z^2 = 1$, so we get that one of the variables is ± 1 and the rest are 0. For the other units, we have $|t|, |x|, |y|, |z| \geq 1/2$, so all the variables are $1/2$. □

This means that the only factorizations of a prime P into two Hurwitz integers are $P = PU^{-1} \times U$ and $P = V \times V^{-1}P$, as U and V run across the 24 Hurwitz units.

Theorem 7.3. *Let* Q *be a primitive Hurwitz integer with norm* q*. Suppose* $q = p_0 p_1 \cdots p_k$ *is a factorization of* q *into rational primes. Let the factorization* $Q = P_0 P_1 \cdots P_k$ *be modelled on* $q = p_0 p_1 \cdots p_k$ *if the norm of* P_i *is* p_i*. The other factorizations of* Q *modelled over the factorization of* q *are of the form*

$$Q = P_0 U_1 \cdot U_1^{-1} P_1 U_2 \cdot \cdots \cdot U_k^{-1} P_k.$$

Proof. The ideal $p_0 H + QH$ must be principal, so that we have $p_0 H + QH = P_0 H$ for some P_0. Here, $[P_0]$ must divide $[p_0] = p_0^2$. so it is one of 1, p_0, or p_0^2.

However, if $[P_0] = 1$, then $p_0 H + QH$ would be all of H, since its typical element $p_0 a + Qb$ has norm

$$[p_0 a] + 2[p_0 a, Qb] + [Qb] = p_0^2[a] + 2p_0[a, Qb] + p_0 \cdots p_k[b],$$

which is divisible by p_0. We also cannot have $[P_0] = p_0^2$; this would mean that $p_0 = P_0 U$ for a Hurwitz unit U. The only possibility is that $[P_0] = p_0$, showing that P_0 is a Hurwitz prime dividing Q.

Thus, we have $Q = P_0 Q_1$ with $[Q_1] = p_1 \cdots p_k$ and P_0 is unique up to right-multiplication by a unit. Similarly, we have

$$Q_1 = P_1 Q_2, \quad Q_2 = P_2 Q_3, \quad \ldots$$

with P_i a Hurwitz prime of norm p_i. This gives $Q = P_0 P_1 \cdots P_k Q'$ for a unit Q', but this can be absorbed with P_k. Thus, prime factorization is unique up to unit-migration. □

7.2 Factoring a rational prime over quaternions

Next, we dive into the quaternionic factorizations of rational integers, especially rational primes.

Definition 7.4. A quadratic residue r satisfies $r \equiv a^2 \pmod{p}$ for some $a \neq 0$. The quadratic non-residues are the numbers that cannot be represented in this form.

Lemma 7.5. *Each quadratic non-residue* n *satisfies* $n \equiv a^2 + b^2 \pmod{p}$ *for* $a, b \neq 0$*, and*

Lemma 7.6. $0 \equiv a^2 + b^2 + c^2 \pmod{p}$ *for* $a, b, c \neq 0$*.*

Proof of Lemmas 7.5 and 7.6. Note that the non-residues are quadratic residue multiples of one of the non-residues. We also know that there exists a non-residue that is one more than a residue, so for this choice of non-residue n, we have $n = a^2 + 1^2$. Therefore, $-1 \equiv x^2 + y^2$ for $x, y \neq 0$, so $0 \equiv x^2 + y^2 + 1^2 \pmod{p}$. □

Theorem 7.7. *Each rational prime p can be factored into $p = P_0\overline{P_0}$.*

Proof. Since $x^2 \equiv (p-x)^2$, we can assume in Lemma 7.6 that $0 \leq a, b, c \leq p/2$. Thus, $a^2 + b^2 + c^2 = mp$ for $0 < m < p$, so we have a quaternion $Q = a + bi + cj$ with norm mp. Therefore, if $p\mathbb{H} + Q\mathbb{H} = P_0\mathbb{H}$, then $[P_0] = p$. □

In fact, p has as many factorizations as a quaternion P of norm p. It can be shown that the number of factorizations is 24 when $p = 2$ and $24(p+1)$ otherwise.

7.3 Factoring the Lipschitz integers

Geometrically, the Lipschitz integers L form a 4-dimensional hypercube I_4. The Hurwitz integers are formed by three copies of this hypercube; they are $L = I_4$, ωL, and $\overline{\omega} L$ where $\omega = \frac{-1+i+j+k}{2}$. We attempt to count Lipschitzian factorizations.

Lemma 7.8. *Any factorization of a Lipschitzian into Hurwitzians is equivalent by unit-migration to one into Lipschitzians.*

Proof. We study the effect of ω migration:

$$\alpha = \beta\gamma \to \alpha = \beta\omega \cdot \overline{\omega}\gamma.$$

Since α is in [0] of [1], the diagram on page 63 of [CS03] includes all possibilities for the cosets of β and γ, and there is a factorization in each triple where both factors are Lipschitzian. □

In fact, the number of Lipschitzian factorizations equivalent to $P_1P_2\cdots P_k$ is $8^{k-1}3^{l-1}$ where l is the number of factors of even norm.

Between p_i and $p_{i+1}\cdots p_k$, we can always transfer the eight Lipschitzian units and the three powers of ω when the factors have even norm.

8 Proof of the main theorem

We are now ready to prove the main theorem explained in the introduction:

Theorem 8.1 (Hurwitz)**.** *The only composition algebras with identity on a real Euclidean space are $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and $\mathbb{O}$.*

The proof will follow the outline in [CS03]. Each subsection will prove an important lemma or set of laws that will contribute to the final proof of the theorem. Let $[a, b]$ denote the inner product of a and b.

8.1 Multiplication laws

We first deduce some consequences of

Law 8.2 (Composition Law)**.** $[xy] = [x][y]$.

Law 8.3 (Scaling Laws)**.** $[xy, xz] = [x][y, z]$ (*and* $[xz, yz] = [x, y][z]$).

Proof. Replace y with $y + z$ in Law 8.2 to get

$$[xy] + [xz] + 2[xy, xz] = [x]([y] + 2[y, z] + [z]),$$

from which we cancel some terms and divide by 2. □

Law 8.4 (Exchange Law)**.** $[xy, uz] = 2[x, u][y, z] - [xz, uy]$.

Proof. Replace x with $x + u$ in Law 8.3 to get

$$[xy, xz] + [xy, uz] + [uy, xz] + [uy, uz] = ([x] + 2[x, u] + [u])[y, z],$$

from which we cancel some terms and rearrange. □

8.2 Conjugation laws

Now, we prove three laws involving the conjugation $\bar{x} = 2[x, 1] - x$.

Law 8.5 (Braid Laws)**.** $[xy, z] = [y, \bar{x}z]$ *(and* $[xy, z] = [x, z\bar{y}]$*).*

Proof. Substitute $u = 1$ in Law 8.4 to get

$$2[x, 1][y, z] - [xz, y] = [y, (2[x, 1] - x)z].$$

□

Remark 8.6. *Six inner products can be equated in a cycle with the Braid Laws:*

$$[xy, z] = [y, \bar{x}z] = [y\bar{z}, \bar{x}] = [\bar{z}, \bar{y}\bar{x}] = [\bar{z}x, \bar{y}] = [x, z\bar{y}] = [xy, z].$$

Law 8.7 (Biconjugation)**.** $\overline{\bar{x}} = x$.

Proof. Substitute $y = 1$ and $z = t$ and use Law 8.5 twice to get

$$[x, t] = [1, \bar{x}t] = [\overline{\bar{x}}, t],$$

which holds for all t. □

Law 8.8 (Product Conjugation)**.** $\overline{xy} = \bar{y}\bar{x}$.

Proof. Repeated use of Law 8.5 gives

$$[\bar{y}\bar{x}, t] = [\bar{x}, yt] = [\bar{x}\bar{t}, y] = [\bar{t}, xy] = [\bar{t}\overline{xy}, 1] = [\overline{xy}, t].$$

This conforms to the properties of an involution. □

8.3 Doubling laws

Before introducing the three laws that follow, we must first know how to construct a Dickson double. Let H be an n-dimensional subalgebra containing 1, let i be a unit vector orthogonal to H, and let $a, b, c, \ldots$ denote elements of H. Then $\bar{i} = -i$ and $[i, a] = 0$. This will cause the $2[x, u][y, z]$ term to vanish when applying the Exchange Law. The next three laws investigate inner product, conjugation, and product on the Dickson double algebra $H + iH$ in terms of those on H.

Law 8.9 (Inner Product Doubling)**.** $[a + ib, c + id] = [a, c] + [b, d]$.

Proof. To prove this law, we use these three equations:

$$[a, id] = [a\bar{d}, i] = 0,\ [ib, c] = [i, c\bar{b}] = 0,\ [ib, id] = [i][b, d] = [b, d].$$

The inner product is distributive over addition, so the proof is complete. □

Law 8.10 (Conjugation Doubling)**.** $\overline{a + ib} = \bar{a} - ib$.

Proof. $\overline{ib} = 2[ib, 1] - ib = -ib$. Add $\bar{a}$ to both sides. Note that $ib = -\overline{ib} = -\bar{b}\bar{i} = \bar{b}i$. □

Law 8.11 (Composition Doubling)**.** $(a + ib)(c + id) = (ac - d\bar{b}) + i(cb + \bar{a}d)$.

Proof. We use the following three equations:

$$[a(id), t] = [id, \bar{a}t] = 0 - [it, \bar{a}d] = [t, i(\bar{a}d)], \tag{8.1}$$

$$[(ib)c, t] = [ib, t\bar{c}] = [\bar{b}i, t\bar{c}] = 0 - [\bar{b}\bar{c}, ti] = [(\overline{cb})i, t] = [i(cb), t], \tag{8.2}$$

$$[(ib)(id), t] = -[ib, t(id)] = 0 + [i(id), tb] = -[id, i(tb)] = -[i][d, tb] = [-d\bar{b}, t]. \tag{8.3}$$

Equation 8.1 results from using Laws 8.5, 8.4, and 8.5 in that order. Equation 8.2 results from using Laws 8.5, 8.10, 8.4, 8.5, and 8.10 in that order. Equation 8.3 results from using Laws 8.5, 8.4, 8.5, 8.3, and 8.5 in that order. We finish using distributive properties. □

8.4 Completing Hurwitz's Theorem

We have shown that a composition algebra Z that has a proper subalgebra contains its Dickson double. Therefore, if Z is finite-dimensional, then it must be the result of repeated doubling of its smallest subalgebra $\mathbb{R}$. We now show that the composition property can last for only three doubling operations.

Lemma 8.12. *$Z = Y + i_Z Y$ is a composition algebra when Y is an associative composition algebra.*

Proof. For $a, b, c, d \in Y$, we have

$$\begin{aligned}[a + i_Z b, c + i_Z d] &= [(ac - d\bar{b}) + i_Z(cb + \bar{a}d)] \\ &= [a][c] + [a][d] + [b][c] + [b][d] \\ &= [ac] - 2[ac, d\bar{b}] + [d\bar{b}] + [cb] + 2[cb, \bar{a}d] + [\bar{a}d]\end{aligned}$$

Therefore, $[ac, d\bar{b}] = [cb, \bar{a}d]$, so $(ac)b = a(cb)$. □

Lemma 8.13. *$Y = X + i_Y X$ is an associative composition algebra when X is a commutative associative composition algebra.*

Proof. Since $(i_Y b)c = i_Y(cb)$, we must have $bc = cb$ due to the associative property that X has due to Y's associative property. □

Lemma 8.14. *$X = W + i_X W$ is a commutative associative composition algebra when W is a commutative associative composition algebra with trivial conjugation.*

Proof. We have $ix = \bar{x}i$ for all $x \in W$. Since X has commutative properties, W must have commutative properties as well, so $x = \bar{x}$. □

We have proven Theorem 8.1. Now we extend up to isotopy for algebras without identities.

Let u and v be elements in an arbitrary composition algebra of norm 1. Then $x \to xv$ and $y \to uy$ are orthogonal maps, so they have inverses α and β, respectively. Now, define a new multiplication by $x \star y = x^\alpha y^\beta$. Then

$$[x \star y] = [x^\alpha y^\beta] = [x^\alpha][y^\beta] = [x][y],$$

showing that $\star$ still gives a composition algebra; and

$$uv \star uy = (uv)^\alpha (uy)^b eta = uy$$

$$xv \star uv = (xv)^\alpha (uv)^\beta = xv,$$

showing that uv is a two-sided identity for the new multiplication.

Remark 8.15. *Any norm 1 element can be converted into an identity element by applying an isotopy, so that the monotopies are transitive on such elements.*

8.5 Other properties of the algebras

We define $x^{-1} = \bar{x}/[x]$ for $x \neq 0$.

Law 8.16 (Inverse Laws)**.** $\bar{x}(xy) = [x]y = yx(\bar{x})$

Proof. $[\bar{x}(xy), t] = [xy, xt] = [x][y, t] = [[x]y, t]$. □

Law 8.17 (Alternative Laws)**.** $x(xy) = x^2 y$ *and* $(yx)x = yx^2$

Proof. Substitute $\bar{x} = 2[x, 1] - x$ in $\bar{x}(xy) = (\bar{x}x)y$. □

Law 8.18 (Moufang Laws)**.** $(xy)(zx) = (x(yz))x = x((yz)x)$.

Proof.

$$\begin{aligned}[(xy)(zx), t] = [xy, t(\bar{x}\bar{z})] &= 2[x,t][y,\bar{x}\bar{z}] - [x(\bar{x}\bar{z}), ty] \\ &= 2[x,t][yz,\bar{x}] - [\bar{x}\bar{z}, \bar{x}(ty)] \\ &= 2[yz,\bar{x}][x,t] - [x][\bar{z}\bar{y}, t] \\ &= 2[x,\overline{yz}][x,t] - [x][\overline{yz}, t].\end{aligned}$$

Thus, $(xy)(zx)$ is a function in x and yz only. We can replace y and z with any two elements with the same product, and so deduce that $(xy)(zx) = (x(yz))(1x) = (x(yz))x$ and similarly on the other side. The third alternative law results from plugging in $z = 1$. □

8.6 Left-sided, right-sided, and both-sided multiplication

The breakdown of the associative law motivates us to study multiplicative operators. We define left-multiplication, right-multiplication, and bi-multiplication as follows:

Definition 8.19. $L_x : y \to xy$, $R_x : y \to yx$, $B_x : y \to xyx$.

The third alternative law shows that B_x can be obtained by multiplying L_x and R_x in either order. We will create a geometrical link between B_x and reflections:

Definition 8.20. $\mathrm{ref}(x) : t \to t - \frac{2[x,t]}{[x]}x$ represents a reflection in a vector x.

We see that $(xy)(zx) = -[x]\overline{yz}^{\mathrm{ref}(x)} = [x](yz)^{\mathrm{ref}(1)\cdot\mathrm{ref}(x)}$.

8.7 Coordinates of quaternions and octonions

We now recover our original definitions of quaternions and octonions in Definition 2.1. Let i be the unit that extends $\mathbb{R}$ to $\mathbb{C}$; this is the well-known square root of -1. If we let j be the unit that extends $\mathbb{C}$ to $\mathbb{H}$, we have Hamilton's equations for quaternions: $i^2 = j^2 = k^2 = -1$ and $ij = k$, etc.

Finally, we extend $\mathbb{H}$ to $\mathbb{O}$. We do this by letting $i = i_1$, $j = i_2$, and $k = i_4$. The unit that extends the quaternions to the octonions is i_0, and the seven imaginary units $i_0, \ldots, i_6$ satisfy $i_0 i_n = i_{3n}$, with indices taken modulo 7.

8.8 N-square identities

By Theorem 8.1, the identity

$$(x_1^2 + x_2^2 + \cdots x_N^2)(y_1^2 + y_2^2 + \cdots y_N^2) = z_1^2 + z_2^2 + \cdots z_N^2$$

holds when $N = 1, 2, 4, 8$. However, A. Pfister proved that in 1967 that the N-square identity holds when $N = 2^n$ for a nonnegative integer n. Please see reference 34 in [CS03] for more details.

9 Applications of quaternions

Several applications of quaternions exist in the real world. Hamilton created his quaternions to model three-dimensional rotations, which are used in computer graphics, robotics, navigation systems, and many areas of physics.

9.1 Hamilton's quaternions

Hamilton created his quaternions in order to mathematically describe three-dimensional rotation about an axis through some angle.

Theorem 9.1. *The map* $[q] : v \to q^{-1}vq$ *is a congruence of Euclidean space.*

This map is a simple rotation, which is a rotation in n dimensions that fixes an $(n-2)$-dimensional subspace. Additionally, the product of simple rotations $x \to q_1^{-1}xq_1$ and $x \to q_2^{-1}xq_2$ is $x \to (q_1q_2)^{-1}x(q_1q_2)$ due to the fact that $q_2^{-1}q_1^{-1} = (q_1q_2)^{-1}$.

However, this method of modeling rotations has been phased out due to other methods using linear algebra being used instead of quaternions. However, quaternions are used in computer graphics due to the smaller amount of memory required to store quaternions versus a rotation matrix.

9.2 Computer Graphics

Two rotations in 3D space can be modeled using linear interpolation between the Euler angles. However, this method can lead to gimbal lock, which can severely affect the smoothness of an animation. Using quaternions leads to a more realistic animation. A technique that is gaining popularity is spherical linear interpolation, which relies on the fact that the set of all unit quaternions form a unit sphere in 4D hyperspace.

9.3 Aerodynamics

Gimbal lock is caused when two of the rotation axes in a set of three rotations overlap. This phenomenon was encountered during the Apollo 11 mission on the inertial measurement unit of the lunar module.

Quaternions provide an axis of rotation and a size (angle) of rotation, which makes them more efficient to use than using a set of three rotations along the three coordinate axes.

9.4 Other areas in physics

Rotation with quaternionic models spinors very well due to the property that you only need to rotate a point two full revolutions to get to the original, which is what happens with particles of spin 1/2.

Additionally, quaternions can be used to express the Lorentz transform, which makes it useful for Special and General Theories of Relativity. They can also be used in scattering experiments such as crystallography.

Acknowledgements

The author would like to thank Euler Circle director Simon Rubinstein-Salzedo and teaching assistant Pariya Akhiani for extensive help with the project.

References

[CS03] John H. Conway and Derek A. Smith. *On quaternions and octonions: their geometry, arithmetic, and symmetry*. Natick, MA: A K Peters, 2003.

[Voi21] John Voight. *Quaternion algebras*, volume 288 of *Grad. Texts Math.* Cham: Springer, 2021.