MA4198 Report

Project name: Quaternion algebras

Student: Yuan Ze

Supervisor: Prof. Loke Hung Yean

Department of Mathematics
National University of Singapore
Semester 2, AY2023/2024

Contents

Chapter 1

Hyper Complex Systems

A hypercomplex system is a finite-dimensional unital algebra over the field $\mathbb{R}$ of real numbers. A hypercomplex number is then an element of one of these algebras. In this chapter, we shall explain what does it mean and how does this work. We will also discuss some of the typical examples and constrcutions down to detail.

1.1 Algebra over a Field

An algebra (over a field $\mathbb{F}$) A is a vector space over $\mathbb{F}$ together with a binary operation $(x, y) \mapsto xy$ that is compatibly defined to be satisfying the following properties

- **Left distribution:** $z(x + y) = zx + zy$ for all $x, y, z \in A$.
- **Right distribution:** $(x + y)z = xz + yz$ for all $x, y, z \in A$.
- **Compatibility with scalars:** $(ax)(by) = (ab)(xy)$ for all $x, y \in A$ and $a, b \in \mathbb{F}$.

In other words, $(x, y) \mapsto xy$ is a bilinear map. It is often referred to be the multiplication in A, and A is unital if there exists a multiplicative identity. In particular, A is called an associative algebra if $(xy)z = x(yz)$ for all $x, y, z \in A$.

An algebra A is alternative if $(xx)y = x(xy)$ and $x(yy) = (xy)y$ for all $x, y, z \in A$. Alternative algebras are so named because they are the algebras for which the associator $[x, y, z] = (xy)z-x(yz)$ is an alternating multilinear map, in the sense that it vanishes whenever two of its arguments are equal. Indeed by the alternativity we have

$$\begin{aligned} 0 &= [x, x + y, x + y] \\ &= [x, x, x] + [x, y, x] + [x, x, y] + [x, y, y] \\ &= [x, y, x] \end{aligned}$$

and consequently the associator is alternating. Conversely an algebra with alternating associator is clearly alternative. In particular, associative algebras are obviously alternative as the associator vanishes for all. Note that if there is an algebra for which each subalgebra generated by two elements is associative, the algebra must be alternative. The converse is a theorem due to Emil Artin [Sch17].

Theorem 1.1. *Every Subalgebra generated by two elements of an alternative algebra is associative.*

For subalgebras of a unital algebra, it makes no difference to the associativity by adjoining the multiplicative identity, we then have the following corollary.

Corollary 1.2. *Every Subalgebra generated by two elements together with the multiplicative identity of a unital alternative algebra is associative.*

In particular, the Artin's theorem implies that alternative algebras are power-associative, in the sense that any subalgebra generated by a single element is associative. We shall see later that the following inclusions are strict over $\mathbb{R}$.

$$\{\text{associative algebras}\} \subset \{\text{alternative algebras}\} \subset \{\text{power-associative algebras}\} \tag{1.1}$$

1.2 Division Algebra

Among all of the hypercomplex systems, those in which we can talk about division are of one of the most interesting kind.

A division algebra D is a non-zero algebra over a field, in which division is always possible except by zero. To be precise, for any $u, v \in D$ with $u \neq 0$, the equation $xu = v$ has always a unique solution in D, and same for the equation $uy = v$. In particular, the solutions are called the corresponding left quotient and right quotient respectively.

For associative algebras, being a division algebra is equivalent to being unital with all non-zero elements invertible. To see this, let D be an associative division algebra, choose non-zero $u \in D$ arbitrarily, there is $x \in D$ such that $xu = u$, hence we have $(zx)u = z(xu) = zu$ for all $z \in D$. By the uniqueness of left quotient, we must have $zx = z$ for all $z \in D$, i.e. x is a right multiplicative identity. Similarly y is a left multiplicative identity where y is the solution of $uy = u$, in which case we have $y = yx = x$, hence $x = y$ is in fact the unique multiplicative identity, which will be denoted by 1_D. The multiplicative inverse of u also exists, since now we have $x', y' \in D$ such that $x'u = 1_D$ and $uy' = 1_D$, which implies $x' = x'uy' = y'$, i.e. $x' = y'$ is the unique multiplicative inverse of u, and will be denoted by u^{-1}. Conversely a unital associative algebra in which all non-zero elements are invertible is clearly a division algebra, as we can solve the equations for quotients explicitly by multiplying the inverse of u on both sides, and obtain unique solutions.

Let D be a finite-dimensional algebra, then D is a division algebra if and only if D does not contain any zero-divisors, i.e. $D \setminus \{0\}$ is closed under multiplication. To see this, suppose D contains no zero-divisors, let $u \in D \setminus \{0\}$, consider the linear transformations

$$R_u : x \mapsto xu, \quad L_u : y \mapsto uy$$

defined by the right and left multiplication by u. D is a division algebra if and only if for any $u \in D \setminus \{0\}$, the two transformations are bijective, i.e. the kernels of the two transformations are zero, this is equivalent to the non-existence of zero-divisors.

Example 1.3. The complex numbers $\mathbb{C}$ naturally form a 2-dimensional division algebra over $\mathbb{R}$.

In particular, $\mathbb{C}$ is a division hypercomplex system. Here we consider $\mathbb{C}$ as spanned by $\{\mathbf{1}, \mathbf{i}\}$ over $\mathbb{R}$, where the multiplication is defined as

$$(a\mathbf{1} + b\mathbf{i})(c\mathbf{1} + d\mathbf{i}) = (ac - db)\mathbf{1} + (da + bc)\mathbf{i}$$

In the next section we will imitate this construction to obtain a family of hypercomplex systems.

1.3 Cayley-Dickson Construction

A *-algebra is a unital algebra over $\mathbb{R}$ together with a notion of conjugation, which is an anti-automorphism $x \mapsto \overline{x}$ that is also an involution, i.e. a linear automorphism $x \mapsto \overline{x}$ satisfying both $\overline{xy} = \overline{y} \cdot \overline{x}$ and $\overline{\overline{x}} = x$. Given a *-algebra S, we can define a new *-algebra structure on the vector space $S \oplus S$ as follow

- $(a, b)(c, d) := (ac - \overline{d}b, da + b\overline{c})$
- $\overline{(a, b)} := (\overline{a}, -b)$

The above multiplication on $S \oplus S$ has identity $(1, 0)$, and is bilinear provided by the bilinearity of the multiplication and the linearity of the conjugation in S. The new conjugation is also valid as it is clearly linear and satisfying the defining properties as follow

$$\begin{aligned}
\overline{(a, b)(c, d)} &= \overline{(ac - \overline{d}b, da + b\overline{c})} = (\overline{ac - \overline{d}b}, -da - b\overline{c}) \\
&= (\overline{c} \cdot \overline{a} - \overline{b}d, -b\overline{c} - da) = (\overline{c}, -d)(\overline{a}, -b) = \overline{(c, d)} \cdot \overline{(a, b)} \\
\overline{\overline{(a, b)}} &= \overline{(\overline{a}, -b)} = (a, b)
\end{aligned}$$

The new *-algebra we obtained will be denoted by $S^{(2)}$. In particular S is identified as a subalgebra of $S^{(2)}$ by the canonical embedding $a \mapsto (a, 0)$. The game continues, eventually we get an infinite sequence of *-algebras by repeatedly applying the construction we described above, know as the **Cayley-Dickson construction**.

We say a *-algebra S is nicely normed if it satisfies the following conditions

- $x + \overline{x} \in \mathbb{R} \cdot 1_S$ for all $x \in S$.
- $x\overline{x} \in \mathbb{R}^+ \cdot 1_S$ for all $x \in S \setminus \{0\}$.

We define the real part of x as $\mathrm{Re}(x) := (x + \overline{x})/2$, the imaginary part of x as $\mathrm{Im}(x) := (x - \overline{x})/2$, and the norm of x as $\|x\| := \sqrt{x\overline{x}}$. It turns out that the Cayley-Dickson construction preserves the property of being nicely normed. Indeed, let S be a nicely normed *-algebra and $(a, b) \in S^{(2)}$, suppose $a + \overline{a} = r_1 \cdot 1_S$ and $\|a\|^2 + \|b\|^2 = r_2 \cdot 1_S$ where $r_1 \in \mathbb{R}$ always and $r_2 \in \mathbb{R}^+$ whenever $(a, b) \neq 0$, we have

$$\begin{aligned}
(a, b) + \overline{(a, b)} &= (a, b) + (\overline{a}, -b) = (a + \overline{a}, 0) \\
&= r_1(1_S, 0) = r_1 \cdot 1_{S^{(2)}} \in \mathbb{R} \cdot 1_{S^{(2)}} \\
(a, b)\overline{(a, b)} &= (a, b)(\overline{a}, -b) = (a\overline{a} + \overline{b}b, -ab + ab) = (\|a\|^2 + \|b\|^2, 0) \\
&= r_2(1_S, 0) = r_2 \cdot 1_{S^{(2)}} \in \mathbb{R}^+ \cdot 1_{S^{(2)}} \quad \text{whenever } (a, b) \neq 0
\end{aligned}$$

where $\overline{b}b = \|b\|^2$ is due to the fact that $b\overline{b} \in \mathbb{R} \cdot 1_S$ and $\overline{1_S} = 1_S$ (an anti-automorphism always sends the multiplicative identity to itself).

Example 1.4. The set of real numbers $\mathbb{R}$ has a natural structure of a 1-dimensional nicely normed *-algebra, where the multiplication is the usual one and the conjugation is the identity map. By applying the Cayley-Dickson construction based on $\mathbb{R}$ we obtain an infinite sequence of nicely normed *-algebras with dimension 2^n, where $n = 0, 1, 2, \ldots$. In fact, $\mathbb{R}^{(2)}$ coincides with $\mathbb{C}$ equipped with the complex conjugation via the canonical decomposition $\mathbb{C} = \mathbb{R} \oplus i\mathbb{R}$. Moreover, $\mathbb{H} := \mathbb{C}^{(2)}$ is called the algebra of quaternions, and $\mathbb{O} := \mathbb{H}^{(2)}$ is called the algebra of octonions. We shall further mention that $\mathbb{S} := \mathbb{O}^{(2)}$ is called the algebra of sedenions.

In a nicely nomred *-algebra, every element $x \in S \setminus \{0\}$ has at least one multiplicative inverse given by $\overline{x}/\|x\|^2$, but it may not be unique without further restriction. That is to say, the Cayley-Dickson construction may not preserve the possibility of doing division. Indeed, each time we apply the construction, the algebra gets a bit worse. Let S be a nicely normed *-algebra, the following propositions illustrate such phenomenon.

Proposition 1.5. *The conjugation in S is the identity map if and only if $S^{(2)}$ is commutative.*

Proof. If the conjugation in S is the identity map, then S is commutative since the identity map is now an anti-automorphism. Let $(a, b), (c, d) \in S^{(2)}$ we have

$$\begin{aligned}
(a, b)(c, d) &= (ac - db, da + bc) \\
&= (ca - bd, bc + da) = (c, d)(a, b)
\end{aligned}$$

i.e. $S^{(2)}$ is commutative. Conversely if $S^{(2)}$ is commutative, for each $a \in S$, we have

$$(a, 0) = (0, -a)(0, 1_S) = (0, 1_S)(0, -a) = (\overline{a}, 0)$$

i.e. $\overline{a} = a$, the conjugation in S is the identity map. □

Proposition 1.6. *S is commutative and associative if and only if $S^{(2)}$ is associative.*

Proof. If S is commutative and associative, let $(a,b),(c,d),(e,f) \in S^{(2)}$ we have

$$\begin{aligned}
((a,b)(c,d))(e,f) &= (ac - \overline{d}b, da + b\overline{c})(e,f) \\
&= (ace - \overline{d}be - \overline{f}da - \overline{f}b\overline{c}, fac - f\overline{d}b + da\overline{e} + b\overline{c} \cdot \overline{e}) \\
(a,b)((c,d)(e,f)) &= (a,b)(ce - \overline{f}d, fc + d\overline{e}) \\
&= (ace - a\overline{f}d - \overline{c} \cdot \overline{f}b - e\overline{d}b, fca + d\overline{e}a + b\overline{e} \cdot \overline{c} - b\overline{d}f)
\end{aligned}$$

where the two products coincide, hence $S^{(2)}$ is associative. Conversely if $S^{(2)}$ is associative, then S is of course associative as being a subalgebra of $S^{(2)}$. On the other hand, for any $a, b \in S$, we have $((b,0)(0,-a))(0,1_S) = (b,0)((0,-a)(0,1_S))$, but we also have

$$\begin{aligned}
((b,0)(0,-a))(0,1_S) &= (0,-ab)(0,1_S) = (ab,0) \\
(b,0)((0,-a)(0,1_S)) &= (b,0)(a,0) = (ba,0)
\end{aligned}$$

thus $ab = ba$, S is commuatative. □

Proposition 1.7. *S is associative if and only if $S^{(2)}$ is alternative.*

Proof. If S is associative, let $(a,b),(c,d) \in S^{(2)}$, we have

$$\begin{aligned}
((a,b)(a,b))(c,d) &= (a^2 - \overline{b}b, ba + b\overline{a})(c,d) \\
&= (a^2c - \overline{b}bc - \overline{d}ba - \overline{d}b\overline{a}, da^2 - d\overline{b}b + ba\overline{c} + b\overline{a} \cdot \overline{c}) \\
&= (a^2c - \|b\|^2c - 2\mathrm{Re}(a)\overline{d}b, da^2 - \|b\|^2d + 2\mathrm{Re}(a)b\overline{c}) \\
(a,b)((a,b)(c,d)) &= (a,b)(ac - \overline{d}b, da + b\overline{c}) \\
&= (a^2c - a\overline{d}b - \overline{a} \cdot \overline{d}b - c\overline{b}b, da^2 + b\overline{c}a + b\overline{c} \cdot \overline{a} - b\overline{b}d) \\
&= (a^2c - 2\mathrm{Re}(a)\overline{d}b - \|b\|^2c, da^2 + 2\mathrm{Re}(a)b\overline{c} - \|b\|^2d)
\end{aligned}$$

where the two products coincide. It is by symmetry to see that $(a,b)((c,d)(c,d)) = ((a,b)(c,d))(c,d)$ is also true, hence $S^{(2)}$ is alternative. Conversely if $S^{(2)}$ is alternative, it is true for any $a,b,c \in S$ that $((\overline{a},c)(\overline{a},c))(0,\overline{b}) = (\overline{a},c)((\overline{a},c)(0,\overline{b}))$, but we also have

$$\begin{aligned}
((\overline{a},c)(\overline{a},c))(0,\overline{b}) &= (\overline{a^2} - \overline{c}c, c\overline{a} + ca)(0,\overline{b}) = (\overline{a^2} - \|c\|^2, 2\mathrm{Re}(a)c)(0,\overline{b}) \\
&= (-2\mathrm{Re}(a)(bc), \overline{b} \cdot \overline{a^2} - \|c\|^2\overline{b}) \\
(\overline{a},c)((\overline{a},c)(0,\overline{b})) &= (\overline{a},c)(-bc, \overline{b} \cdot \overline{a}) = (-\overline{a}(bc) - (ab)c, (\overline{b} \cdot \overline{a})\overline{a} - c(\overline{c} \cdot \overline{b}))
\end{aligned}$$

By comparing the first argument we deduce that

$$\overline{a}(bc) + (ab)c = 2\mathrm{Re}(a)(bc) = \overline{a}(bc) + a(bc)$$

which implies $(ab)c = a(bc)$, and S is then associative since a,b,c are arbitrary. □

According to the fact that the complex numbers $\mathbb{C}$ is commutative and associative but with conjugation not being the identity map, combining Proposition 1.5, Proposition 1.6 and Proposition 1.7 we see that $\mathbb{H}$ is associative but not commutative, and $\mathbb{O}$ is alternative but not associative. In fact, both $\mathbb{H}$ and $\mathbb{O}$ are division algebras, this can be shown by the following proposition.

Proposition 1.8. *An alternative nicely normed *-algebra of finite dimension is a division algebra.*

Proof. Let S be a finite-dimensional alternative nicely normed *-algebra and $a, b \in S$. Note that $\overline{a} = a - 2\mathrm{Re}(a)$ and $\overline{b} = b - 2\mathrm{Re}(b)$, then $\overline{a}$ and $\overline{b}$ are contained inside the subalgebra of S generated by a, b and 1_S which is associative by Corollary 1.2, hence

$$\begin{aligned} \|ab\|^2 &= (ab)\overline{(ab)} = (ab)(\overline{b} \cdot \overline{a}) \\ &= a(b\overline{b})\overline{a} = \|b\|^2 a\overline{a} = \|a\|^2\|b\|^2 \end{aligned}$$

i.e. the norm in S is multiplicative. This implies that $S \setminus \{0\}$ is closed under multiplication, hence S is a division algebra. □

Similar argument will not be working for the sedenions $\mathbb{S}$, as by Proposition 1.7 we see that $\mathbb{S}$ can not be alternative. In fact, an explicit calculation shows that there are tons of zero-divisors inside $\mathbb{S}$ [Caw04], hence $\mathbb{S}$ can not be a division algebra, and so are the rest algebras we constructed later on which contain $\mathbb{S}$ as a subalgebra. More generally, a famous classification result due to Adolf Hurwitz is as follow.

Theorem 1.9 (Hurwitz)**.** *A finite-dimensional unital real division algebra endowed with a positive-definite quadratic form that is also multiplicative, should always be isomorphic to one of the real numbers, the complex numbers, the quaternions, and the octonions.*

For associative algebras, there is also a theorem due to Ferdinand Georg Frobenius.

Theorem 1.10 (Frobenius)**.** *A finite-dimensional real division algebra should be always isomorphic to one of the real numbers, the complex numbers, and the quaternions.*

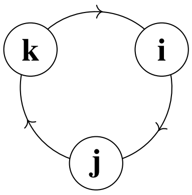
One property that remains to be true for the sedenions is the power-associativity. Indeed, if S is a nicely normed *-algebra that is also alternative, then $S^{(2)}$ is always power-associative. The verification is routine. It is noted that the inclusions in Equation (1.1) are indeed strict.

1.4 Quaternions

In this section we illustrate the quaternions explicitly over $\mathbb{R}$. Recall $\mathbb{H} := \mathbb{C}^{(2)}$, each of its elements is written as (z, w) for some $z, w \in \mathbb{C}$, hence $\mathbb{H} = \mathrm{Span}_{\mathbb{R}}\{(\mathbf{1}, 0), (\mathbf{i}, 0), (0, \mathbf{1}), (0, \mathbf{i})\}$ as a 4-dimensional algebra over $\mathbb{R}$. Denote $(\mathbf{1}, 0)$ by $\mathbf{1}$, $(\mathbf{i}, 0)$ by $\mathbf{i}$ (abuse of notation) and $(0, \mathbf{1})$ by $\mathbf{j}$, $(0, \mathbf{i})$ by $\mathbf{k}$. For $a \in \mathbb{R}$, $a\mathbf{1}$ is always simplified as a, then each element of $\mathbb{H}$ is of the form $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ for some $a, b, c, d \in \mathbb{R}$. To work out the explicit multiplication law in $\mathbb{H}$ over basis $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$, it suffices to specify how $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ are multiplied with each other, since the general law is then clear by extending it linearly. The calculation goes as follow

$$\begin{aligned} \mathbf{ij} &= (\mathbf{i}, 0)(0, 1) = (0, \mathbf{i}) = \mathbf{k}, \quad & \mathbf{ji} &= (0, 1)(\mathbf{i}, 0) = (0, -\mathbf{i}) = -\mathbf{k} \\ \mathbf{jk} &= (0, 1)(0, \mathbf{i}) = (\mathbf{i}, 0) = \mathbf{i}, \quad & \mathbf{kj} &= (0, \mathbf{i})(0, 1) = (-\mathbf{i}, 0) = -\mathbf{i} \\ \mathbf{ki} &= (0, \mathbf{i})(\mathbf{i}, 0) = (0, 1) = \mathbf{j}, \quad & \mathbf{ik} &= (\mathbf{i}, 0)(0, \mathbf{i}) = (0, -1) = -\mathbf{j} \end{aligned}$$

and it is immediate that $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -\mathbf{1}$. A mnemonic diagram is given by



in the sense that the product of any two of them is always the third or its negative, depending on whether the direction of the shortest arc joining the first factor and the second is from the first factor to the second or not. The explicit conjugation in $\mathbb{H}$ over $\mathbb{R}$ is as follow

$$\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}} = \overline{(a + b\mathbf{i}, c + d\mathbf{i})} = (a - b\mathbf{i}, -c - d\mathbf{i}) = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$$

We then have $\mathrm{Re}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = a$ and $\mathrm{Im}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Let $z = a + b\mathbf{i}$ and $w = c + d\mathbf{i}$, the explicit norm formula is given as follow

$$\begin{aligned} \|a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\|^2 &= (z, w)\overline{(z, w)} \\ &= (\|z\|^2 + \|w\|^2, 0) = a^2 + b^2 + c^2 + d^2 \end{aligned}$$

Since $\mathbb{H}$ is a nicely normed division *-algebra, each non-zero element is invertible, and doing division in $\mathbb{H}$ is equivalent as finding multiplicative inverses, which is done explicitly by the norm formula above. Let $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \neq 0$, the inverse is given by

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})^{-1} = \frac{a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}}{a^2 + b^2 + c^2 + d^2}$$

Recall that the norm in $\mathbb{H}$ is multiplicative, the explicit formula then provide a proof for the following result known as the 4-square identity.

Theorem 1.11. *The product of sums of four squares is again a sum of four squares.*

Moreover, let $a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 \in \mathbb{R}$, by writing donw explicitly

$$\begin{aligned} &(a_1^2 + a_2^2 + a_3^2 + a_4^2)(b_1^2 + b_2^2 + b_3^2 + b_4^2) \\ &\quad = (a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4)^2 + (a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)^2 \\ &\quad\quad + (a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2)^2 + (a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)^2 \end{aligned}$$

we see that each term on the right hand side is a square of a bilinear form on $\mathbb{H}$ over $\mathbb{R}$. Such an identity is called a **bilinear identity**.

1.5 Octonions

In this section we study the octonions explicitly. Recall $\mathbb{O} := \mathbb{H}^{(2)}$, each of its elements is written as (p, q) for some $p, q \in \mathbb{H}$, hence

$$\mathbb{O} = \mathrm{Span}_{\mathbb{R}}\{(\mathbf{1}, 0), (\mathbf{i}, 0), (\mathbf{j}, 0), (\mathbf{k}, 0), (0, \mathbf{1}), (0, \mathbf{i}), (0, \mathbf{j}), (0, \mathbf{k})\}$$

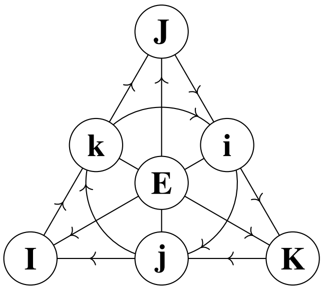
as a 8-dimensional algebra over $\mathbb{R}$. Denote $(\mathbf{1}, 0)$ by $\mathbf{1}$, $(\mathbf{i}, 0)$ by $\mathbf{i}$, $(\mathbf{j}, 0)$ by $\mathbf{j}$, $(\mathbf{k}, 0)$ by $\mathbf{k}$ (abuse of notation), and $(0, \mathbf{1})$ by $\mathbf{E}$, $(0, \mathbf{i})$ by $\mathbf{I}$, $(0, \mathbf{j})$ by $\mathbf{J}$, $(0, \mathbf{k})$ by $\mathbf{K}$, each element of $\mathbb{O}$ is then of the following form

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}$$

for some $a, b, c, d, A, B, C, D \in \mathbb{R}$. It is also easy to calculate the multiplication table by playing with the Cayley-Dickson construction

$$\begin{gathered}\mathbf{Ik} = (0, \mathbf{i})(\mathbf{k}, 0) = \mathbf{J}, \quad \mathbf{kJ} = (\mathbf{k}, 0)(0, \mathbf{j}) = \mathbf{I}, \quad \mathbf{JI} = (0, \mathbf{j})(0, \mathbf{i}) = \mathbf{k} \\ \mathbf{Ji} = (0, \mathbf{j})(\mathbf{i}, 0) = \mathbf{K}, \quad \mathbf{iK} = (\mathbf{i}, 0)(0, \mathbf{k}) = \mathbf{J}, \quad \mathbf{KJ} = (0, \mathbf{k})(0, \mathbf{j}) = \mathbf{i} \\ \cdots\end{gathered}$$

and so on. Instead of listing the whole multiplication table which turns out to be totally unenlightening, we have the Fano plane below serving as a mnemonic diagram to memorize the multiplication law of octonions



in the sense that each three elements lying on a straight line or a circle form a multiplicative cycle that is similar with the imaginary basis of quaternions in the indicated orientation.

The explicit conjugation in $\mathbb{O}$ over $\mathbb{R}$ is as follow

$$\begin{aligned}\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}} &= (\overline{a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}}, -A - B\mathbf{i} - C\mathbf{j} - D\mathbf{k}) \\ &= a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} - A\mathbf{E} - B\mathbf{I} - C\mathbf{J} - D\mathbf{K}\end{aligned}$$

and we then have $\mathrm{Re}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}) = a$ and

$$\mathrm{Im}(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}) = b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}$$

Let $p = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ and $q = A + B\mathbf{i} + C\mathbf{j} + D\mathbf{k}$, the explicit norm formula is given as

$$\begin{aligned}\|a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K}\|^2 &= (p, q)\overline{(p, q)} = (\|p\|^2 + \|q\|^2, 0) \\ &= a^2 + b^2 + c^2 + d^2 + A^2 + B^2 + C^2 + D^2\end{aligned}$$

Similar as $\mathbb{H}$, the octonions $\mathbb{O}$ is a nicely normed division *-algebra, where each non-zero element is invertible, hence doing division in $\mathbb{O}$ is equivalent as finding multiplicative inverses. This is also done explicitly by the norm formula. Let $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K} \neq 0$, the inverse is given by

$$(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} + A\mathbf{E} + B\mathbf{I} + C\mathbf{J} + D\mathbf{K})^{-1} = \frac{a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} - A\mathbf{E} - B\mathbf{I} - C\mathbf{J} - D\mathbf{K}}{a^2 + b^2 + c^2 + d^2 + A^2 + B^2 + C^2 + D^2}$$

Recall that the norm in $\mathbb{O}$ is multiplicative, the explicit norm formula then provide a proof for the following result known as the 8-square identity.

Theorem 1.12. *The product of sums of eight squares is again a sum of eight squares.*

It is routine to check that the identity defined by the multiplicativity of the norm in $\mathbb{O}$ is a bilinear identity. Similar method does not work for the sedenions, as our assumption for multiplicativity of the norm is based on the alternativity. In fact, it is proved by Adolf Hurwitz that there is no similar bilinear identity for 16 squares or any other number of squares except for $1, 2, 4$ and 8. This is a consequence of Theorem 1.9 and is called the Hurwitz's $1, 2, 4, 8$ theorem. If the requirement of being bilinear is weakened in the sense that we are suppose to find solutions for z's such that

$$(x_1^2 + x_2^2 + \cdots + x_n^2)(y_1^2 + y_2^2 + \cdots + y_n^2) = z_1^2 + z_2^2 + \cdots + z_n^2$$

where z's can be any rational functions of x's and y's instead of only bilinear functions, it can always be done when n is any power of 2, known as the Pfister's theorem [Con16].

1.6 Clifford Algebra

As the octonions $\mathbb{O}$ is a generalization of the quaternions $\mathbb{H}$ by means of the Cayley-Dickson construction dragging out the information of the conjugation, there is another approach that is invented by William Clifford concerning with quadratic forms.

In general, given a field $\mathbb{F}$ together with a non-degenerate quadratic form Q, let V be a vector space over $\mathbb{F}$, the Clifford algebra $\text{Cliff}(V, Q)$ is defined as

$$\text{Cliff}(V, Q) := \left(\bigoplus_{k=0}^{\infty} T^k V \right) \Big/ v \otimes v \sim Q(v)$$

where $T^k V$ is the tensor product of k copies of V for $k > 0$ and $T^0 V = \mathbb{F}$. The Clifford algebra $\text{Cliff}(V, Q)$ is the freest unital associative algebra generated by V subject to the equivalence relation, in the sense that given any unital associative algebra A over $\mathbb{F}$ with any linear map $j : V \to A$ satisfying $j(v)^2 = Q(v) \cdot 1_A$ for all $v \in V$, there is a unique algebra homomorphism $f : \text{Cliff}(V, Q) \to A$ fitting into the following commutative diagram

$$\begin{tikzcd}
V \arrow[r, hook, "i"] \arrow[dr, hook, "j"'] & \text{Cliff}(V, Q) \arrow[d, dotted, "f"] \\
& A
\end{tikzcd}$$

where i is the canonical embedding.

When fixing $\mathbb{F} = \mathbb{R}$, we denote $\mathrm{Cliff}(n) := \mathrm{Cliff}(\mathbb{R}^n, -\|\bullet\|_n)$ where $n \geq 1$ is an integer and $\|\bullet\|_n$ is the Euclidean norm in $\mathbb{R}^n$. We also identify $\mathbb{R}^n$ and $\mathbb{R}$ with canonical embeddings into $\mathrm{Cliff}(n)$, and denote $\cdot$ for the multiplication in $\mathrm{Cliff}(n)$ instead of $\otimes$ for simplicity. Explicitly, $\mathrm{Cliff}(n)$ is an associative algebra generated by n anti-commuting square roots of -1. Indeed, let $\{e_i\}_{i=1}^n$ be the standard orthonormal basis of $\mathbb{R}^n$, we have $(e_i)^2 = -\|e_i\|_n^2 = -1$ for all i, and

$$\begin{aligned} -2 = -\|e_i + e_j\|_n^2 = (e_i + e_j)^2 &= (e_i)^2 + e_ie_j + e_je_i + (e_j)^2 \\ &= -\|e_i\|^2 - \|e_j\|^2 + e_ie_j + e_je_i \\ &= -2 + e_ie_j + e_je_i \implies e_ie_j + e_je_i = 0 \end{aligned}$$

for all $i \neq j$. It turns out that $\mathrm{Cliff}(n)$ is in fact the universal associative algebra over such condition, we see that $\mathrm{Cliff}(1) \simeq \mathbb{C}$ and $\mathrm{Cliff}(2) \simeq \mathbb{H}$ as so do $\mathbb{C}$ and $\mathbb{H}$. In particular, $\mathbb{O}$ is not a Clifford algebra as it is not associative. Further calculations provide us the following table

n	Cliff(n)
1	$\mathbb{C}$
2	$\mathbb{H}$
3	$\mathbb{H} \oplus \mathbb{H}$
4	$\mathbb{H}[2]$

n	Cliff(n)
5	$\mathbb{C}[4]$
6	$\mathbb{R}[8]$
7	$\mathbb{R}[8] \oplus \mathbb{R}[8]$
8	$\mathbb{R}[16]$

where $\mathbb{F}[n]$ denotes the algebra consists of $n \times n$ matrices with entries in $\mathbb{F}$. The table continues in the sense of the periodicity $\mathrm{Cliff}(n+8) \simeq \mathrm{Cliff}(n) \otimes \mathbb{R}[16]$. In other words, $\mathrm{Cliff}(n+8)$ consists of 16×16 matrices with entries in $\mathrm{Cliff}(n)$ [ABS64].

Chapter 2

Geometry of Quaternions

In this chapter, we return to Hamilton's original design, say, using quaternions model rotations in Euclidean space. Recall there is a canonical norm defined on quaternions, which induces a metric topology and makes $\mathbb{H}$ homeomorphic to $\mathbb{R}^4$ by the canonical map $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \mapsto (a, b, c, d)$.

2.1 Quaternions in Vector Algebra

Consider $\mathbb{H}$ as $\mathbb{C}^{(2)}$, we have an $\mathbb{R}$-algebra embedding $\mathbb{H} \hookrightarrow M(2, \mathbb{C})$ given by

$$\lambda :\ (z, w) \mapsto \begin{pmatrix} z & -w \\ \overline{w} & \overline{z} \end{pmatrix}$$

This is of course a linear map with zero kernel, and the homomorphism property is checked directly

$$\begin{aligned} \lambda((z_1, w_1)(z_2, w_2)) = \lambda(z_1 z_2 - \overline{w_2} w_1, w_2 z_1 + w_1 \overline{z_2}) &= \begin{pmatrix} z_1 z_2 - \overline{w_2} w_1 & -w_2 z_1 - w_1 \overline{z_2} \\ \overline{z_1} \cdot \overline{w_2} + z_2 \overline{w_1} & \overline{z_2} \cdot \overline{z_1} - \overline{w_1} w_2 \end{pmatrix} \\ &= \begin{pmatrix} z_1 & -w_1 \\ \overline{w_1} & \overline{z_1} \end{pmatrix} \begin{pmatrix} z_2 & -w_2 \\ \overline{w_2} & \overline{z_2} \end{pmatrix} = \lambda(z_1, w_1) \lambda(z_2, w_2) \end{aligned}$$

Note that $\det(\lambda(z, w)) = \|z\|^2 + \|w\|^2 = \|(z, w)\|^2$, it provides another proof showing that the norm in $\mathbb{H}$ is multiplicative. We define the group of unit quaternions as $\mathrm{Sp}(1) := \{q \in \mathbb{H} :\ \|q\| = 1\}$ which is canonically homeomorphic to the 3-dimensional sphere $\mathbb{S}^3$. The embedding λ restricts to an isomorphism $\mathrm{Sp}(1) \simeq SU(2)$ where $SU(2)$ is the special unitary group $\{A \in M(2, \mathbb{C}) :\ A^* A = I \text{ and } \det(A) = 1\}$ of rank 2, and A^* is the conjugate transpose of A. Indeed for $(u, v) \in \mathrm{Sp}(1)$ we have

$$\lambda(u, v)^* \lambda(u, v) = \begin{pmatrix} \overline{u} & v \\ -\overline{v} & u \end{pmatrix} \begin{pmatrix} u & -v \\ \overline{v} & \overline{u} \end{pmatrix} = \begin{pmatrix} \|u\|^2 + \|v\|^2 & 0 \\ 0 & \|u\|^2 + \|v\|^2 \end{pmatrix} = I$$

and we also have $\det(\lambda(u, v)) = \|u\|^2 + \|v\|^2 = 1$.

On the other hand, consider $\mathbb{H}$ as being spanned by $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ over $\mathbb{R}$, define $\mathbb{H}^0 := \mathrm{Span}_{\mathbb{R}}\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ which is the subspace of all pure quaternions with no real part. We shall identify the $\mathbb{H}^0$ with $\mathbb{R}^3$

as vector spaces via the canonical map without further mention. It is noted that $\mathbb{H}^0$ is not a subalgebra of $\mathbb{H}$ as it is not closed under multiplication, the following proposition describes how the multiplication behaves on $\mathbb{H}^0$ in detail.

Proposition 2.1. *Let $u, v \in \mathbb{H}^0$, we have $uv = -\langle u, v\rangle + [u, v]$, where $\langle , \rangle$ and $[,]$ are carried from the standard dot product and cross product in $\mathbb{R}^3$ respectively.*

Proof. To be precise, the operation $\langle , \rangle$ is defined as $\langle (u_1, u_2, u_3), (v_1, v_2, v_3)\rangle = u_1v_1 + u_2v_2 + u_3v_3$, and the operation $[,]$ is defined as

$$[(u_1, u_2, u_3), (v_1, v_2, v_3)] = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix} = (u_2v_3 - u_3v_2, u_3v_1 - u_1v_3, u_1v_2 - u_2v_1)$$

It suffices to check that $(u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k})(v_1\mathbf{i} + v_2\mathbf{j} + v_3\mathbf{k}) = -(u_1v_1 + u_2v_2 + u_3v_3) + (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k}$, which is totally routine. □

Let $u, v \in \mathbb{H}^0$, then u and v are said to be orthogonal to each other if it happens so in $\mathbb{R}^3$, i.e. $\langle u, v\rangle = 0$. Note that $\langle , \rangle$ is symmetric and $[,]$ is anti-symmetric, the proposition then derives the following results.

Corollary 2.2. *Let $u, v \in \mathbb{H}^0$, then $uv \in \mathbb{H}^0$ if and only if u and v are orthogonal to each other.*

Corollary 2.3. *Let $u, v \in \mathbb{H}^0$, then $uv = -vu$ if and only if u and v are orthogonal to each other.*

Corollary 2.4. *Let $u \in \mathbb{H}^0$, then $u^2 = -\|u\|^2$ is always real (since $[u, u] = 0$).*

2.2 Quaternions and Rotation

Let $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathrm{Sp}(1)$, then $a^2 + \|b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\|^2 = 1$, there exists a unique angle $0 \leq \theta \leq \pi$, such that $a = \cos\theta$ and $\|b\mathbf{i} + c\mathbf{j} + d\mathbf{k}\| = \sin\theta$. Define $I(q) := (b\mathbf{i} + c\mathbf{j} + d\mathbf{k})/\sin\theta$, then $\|I(q)\| = 1$, and we have

$$q = \cos\theta + I(q)\sin\theta$$

Now take $v \in \mathbb{H}^0$ that is orthogonal to $I(q)$, the multiplication of q by v on the right has a clear geometric meaning. To see this, we note that $qv = v\cos\theta + I(q)v\sin\theta$, by $\langle I(q), v\rangle = 0$ we must have $I(q)v = [I(q), v]$ which is orthogonal to both $I(q)$ and v, with norm $\|I(q)\|\|v\|\sin\pi/2 = \|v\|$. Denote $[I(q), v]$ by $\widetilde{v}$, then $\{v, \widetilde{v}, I(q)\}$ forms a right-hand system, and we have

$$q = v\cos\theta + \widetilde{v}\sin\theta$$

which is obtained from v by a rotation through angel θ about the axis $I(q)$. It can also be thought as the effect of q acting on v through the left multiplication, but with restriction that v is orthogonal to $I(q)$. In fact, it turns out that we can represent all kinds of rotations of the whole space by a more complicated action of $\mathrm{Sp}(1)$ on $\mathbb{H}^0$.

The map $\varphi\colon \mathrm{Sp}(1) \to GL_{\mathbb{R}}(\mathbb{H}^0)$ given by $\varphi(q)v = qvq^{-1}$ defines a linear representation of Sp(1), which is called the **adjoint representation**. Indeed, for $q \in \mathrm{Sp}(1)$ we have

$$\overline{\varphi(q)v} = \overline{qvq^{-1}} = \overline{q^{-1}} \cdot \overline{v} \cdot \overline{q} = qvq^{-1} = \varphi(q)v \quad \text{for all } v \in \mathbb{H}^0$$

i.e. $\varphi(q) \in \mathbb{H}^0$, and for $p, q \in \mathbb{H}$ we have

$$\varphi(p)\varphi(q)v = p(qvq^{-1})p^{-1} = (pq)v(pq)^{-1} = \varphi(pq)v \quad \text{for all } v \in \mathbb{H}^0$$

which makes φ be a group homomorphism from Sp(1) to $GL_{\mathbb{R}}(\mathbb{H}^0)$ as $\varphi(q)$ is clearly an invertible linear map for each $q \in \mathbb{H}$.

Proposition 2.5. *Let $q \in Sp(1)$, then the action of q on $\mathbb{H}^0$ through the adjoint representation is given by a rotation through angel 2θ about the axis $I(q)$.*

Proof. Let $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathrm{Sp}(1)$, any element $v \in \mathbb{H}^0$ can be decomposed into $v_1 + v_2$ where v_1 is orthogonal to $I(q)$ and v_2 is a multiple of $I(q)$. It is clear that $qv_2 = v_2 q$, hence $qv_2q^{-1} = v_2qq^{-1} = v_2$. For the part of v_1, note that $q^{-1} = \cos\theta - I(q)\sin\theta$, we have

$$qv_1q^{-1} = qv_1\cos\theta - (qv_1)I(q)\sin\theta$$

We already see that qv_1 is obtained by rotating v_1 through angel θ about the axis $I(q)$, hence is orthogonal to $I(q)$, by Corollary 2.3 we have $(qv_1)I(q) = -I(q)(qv_1)$. We also saw earlier that $\widetilde{qv_1} := I(q)(qv_1)$ is obtained by rotating qv_1 through $\pi/2$ about the axis $I(q)$, and we have

$$qv_1q^{-1} = qv_1\cos\theta + \widetilde{qv_1}\sin\theta$$

hence qv_1q^{-1} is obtained by rotating qv_1 through angel θ about the axis $I(q)$, that is to say, by rotating v_1 through angel 2θ about the axis $I(q)$. In conclusion, qvq^{-1} is obtained by rotating v through angel 2θ about the axis $I(q)$. □

It is well know that each element of $SO(3)$ is a rotation about some axis, hence given any element in $SO(3)$ we can explicitly cook up a quaternion such that the adjoint action provides the exactly same rotation. It turns out that φ is in fact a surjective group homomorphism from Sp(1) to $SO(3)$.

Example 2.6 (The composition of rotations)**.** Consider a rotation through an angle θ_1 about an axis determined by a unit vector p_1, followed by another rotation through an angle θ_2 about an axis determined by a unit vector p_2. We can easily find the angle and axis of the resultant rotation by the help of quaternions. Indeed, the first rotation will be characterized by $\varphi(q_1)$ where $q_1 = \cos\theta_1/2 + p_1\sin\theta_1/2$, and the resultant rotation will be $\varphi(q_2)\circ\varphi(q_1) = \varphi(q_2q_1)$ where $q_2 = \cos\theta_2/2 + p_2\sin\theta_2/2$. Note that $q_2q_1 \in \mathrm{Sp}(1)$ and we know how to calculate explicitly, we then obtain $2\arccos\mathrm{Re}(q_2q_1)$ and $\mathrm{Im}(q_2q_1)/\|\mathrm{Im}(q_2q_1)\|$ which are precisely the angel and axis of the resultant rotation.

Example 2.7 (Hopf fibration)**.** Note that the 3-dimensional rotation group $SO(3)$ acts transitively on $\mathbb{S}^2$, hence SO(3) acting on any point of $\mathbb{S}^2$ gives the whole sphere. Let the point be $(1, 0, 0)$ for simplicity, and recall that $\varphi :$ Sp(1) $\to SO(3)$ is surjective, we obtained a a surjective map from Sp(1) to $\mathbb{S}^2$ given by $h : q \mapsto \varphi(q)(1, 0, 0) = q\mathbf{i}q^{-1}$. In real coordinates it is given as

$$\begin{aligned} h(a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}) &= (a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k})\mathbf{i}(a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}) \\ &= (a^2 + b^2 - c^2 - d^2)\mathbf{i} + 2(ad + bc)\mathbf{j} + 2(bd - ac)\mathbf{k} \end{aligned}$$

which defines a continuous map from $\mathbb{S}^3$ to $\mathbb{S}^2$ as

$$h(a, b, c, d) = (a^2 + b^2 - c^2 - d^2, 2(ad + bc), 2(bd - ac))$$

Suppose $H(p) = H(q)$, then $p^{-1}q$ commutes with $\mathbf{i}$, hence is a complex number with unit norm and can be written as $\cos\alpha + \mathbf{i}\sin\alpha$ for some α. Conversely if $q = p(\cos\alpha + \mathbf{i}\sin\alpha)$ for some α, then of course $p^{-1}q$ commutes with $\mathbf{i}$, and we have $H(p) = H(q)$. As a consequence, the fiber of each point on $\mathbb{S}^2$ is parameterized as

$$q(\cos\alpha + \mathbf{i}\sin\alpha), \quad 0 \le \alpha < 2\pi$$

for some q inside the fiber, and is in fact an inclusion of $\mathbb{S}^1$ into $\mathbb{S}^3$. The map H is an example of the so called Hopf fibration. In topology, the Hopf fibration is an influential early example of a fiber bundle, where the structure is denoted as $\mathbb{S}^1 \hookrightarrow \mathbb{S}^3 \to \mathbb{S}^2$. It also provides a non-trivial element in the third homotopy group $\pi_3(\mathbb{S}^2)$ of $\mathbb{S}^2$, which turns out to be in fact generating $\pi_3(\mathbb{S}^2)$.

Corollary 2.8. *The adjoint representation defines a surjective group homomorphism $\varphi : Sp(1) \to SO(3)$ with $\{\pm\mathbf{1}\}$ being the kernel.*

Proof. If $q \in \mathrm{Ker}(\varphi) \subset$ Sp(1), then q commutes with all $v \in \mathbb{H}^0$, hence can only be real, in which case $\|q\| = 1$ enforces q to be $\pm\mathbf{1}$. □

Corollary 2.9. *There exists a two-to-one group homomorphism from $SU(2)$ onto $SO(3)$.*

Proof. It is clear by Sp(1) $\simeq SU(2)$. □

Instead of working in dimension 3, we now identify $\mathbb{H}$ with $\mathbb{R}^4$ as vector spaces in the canonical way, then the map $\psi :$ Sp(1) $\times$ Sp(1) $\to GL_{\mathbb{R}}(\mathbb{H})$ given by $\psi(p, q)u = puq^{-1}$ defines a linear representation of Sp(1) $\times$ Sp(1). Indeed, for $(p_1, q_1), (p_2, q_2) \in$ Sp(1) we have

$$\psi((p_1, q_1)(p_2, q_2))u = p_1(p_2uq_2^{-1})q_1^{-1} = (p_1p_2)u(q_1q_2)^{-1} \quad \text{for all } u \in \mathbb{H}$$

which makes ψ be a group homomorphism from Sp(1) $\times$ Sp(1) to $GL_{\mathbb{R}}(\mathbb{H})$ as $\psi(p, q)$ is clearly an invertible linear map for each $(p, q) \in$ Sp(1) $\times$ Sp(1).

Proposition 2.10. *ψ defines a surjective group homomorphism $\psi : Sp(1) \times Sp(1) \to SO(4)$ with $\{\pm(\mathbf{1}, \mathbf{1})\}$ being the kernel.*

Proof. For simplicity we prove the proposition by the help of Lie theory. Let $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$ and $SO(4)$ be equipped with the usual Lie group structure, it is clear that $\mathrm{Im}(\psi) \subset O(4)$ as ψ preserves the norm. Note that $\mathrm{Sp}(1) \times \mathrm{Sp}(1)$ now is homeomorphic to $\mathbb{S}^3 \times \mathbb{S}^3$, hence is connected, whose image under ψ should be inside the connected component $SO(4)$ of $O(4)$ containing the identity, i.e. ψ is into $SO(4)$. To determine the kernel, we choose $\mathbf{1}, \mathbf{i}, \mathbf{j}$ as testing elements. Suppose $(p, q) \in \mathrm{Ker}(\psi)$, then $p\mathbf{1}q^{-1} = \mathbf{1}$ implies $p = q$. Let $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, we further have $q\mathbf{i} = \mathbf{i}q$ implies $c = d = 0$ and then $q\mathbf{j} = \mathbf{j}q$ implies $b = 0$. Now $\|q\| = 1$ enforces q to be $\pm\mathbf{1}$, hence we have $(p, q) = \pm(\mathbf{1}, \mathbf{1})$. Since $\mathrm{Ker}(\psi)$ is discrete, the induced map on Lie algebra is injective, hence is also surjective as they are of the same dimension by $\dim(\mathfrak{su}(2) \oplus \mathfrak{su}(2)) = 6 = \dim(\mathfrak{so}(4))$ (we have Lie group isomorphism $\mathrm{Sp}(1) \simeq SU(2)$), then $\mathrm{Im}(\psi)$ contains a neighborhood of the identity. Since a connected Lie group is generated by any neighborhood of the identity, we conclude that ψ is onto $SO(4)$. □

Proposition 2.10 could also be proved by using purely matrix algebra, decomposing every 4-dimensional rotation into a left–isoclinic and a right–isoclinic rotation, see for instance [PGT17].

Chapter 3

Generalized Quaternion Algebras

3.1 Definition

In this chapter we always assume $\mathbb{F}$ to be a field with characteristic not equal to 2.

Definition 3.1. A unital associative $\mathbb{F}$-algebra A is a **(generalized) quaternion algebra** if there exist non-zero $\mathbf{i}, \mathbf{j} \in A$ generate A, satisfying $\mathbf{i}^2 = a$, $\mathbf{j}^2 = b$ and $\mathbf{ij} = -\mathbf{ji}$ with $a, b \in \mathbb{F} \setminus \{0\}$.

In such case, let $\mathbf{k} = \mathbf{ij}$, then $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$ is automatically an $\mathbb{F}$-basis of A. To see this, suppose $\beta = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = 0$, then

$$0 = \mathbf{i}(\mathbf{i}\beta + \beta\mathbf{i}) = 2a(w + x\mathbf{i})$$

by the fact that $a \in \mathbb{F} \setminus \{0\}$ and $\operatorname{char}(\mathbb{F}) \neq 2$, we have $w + x\mathbf{i} = 0$. By symmetry we also have $w + y\mathbf{j} = w + z\mathbf{k} = 0$, hence

$$2w = (w + x\mathbf{i}) + (w + y\mathbf{j}) + (w + z\mathbf{k}) - \beta = 0$$

which implies $w = 0$. Now by $x\mathbf{i} = y\mathbf{j} = z\mathbf{k} = 0$ we conclude that $x = y = z = 0$, hence $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ are linearly independent, and the claim is true since $\{\mathbf{i}, \mathbf{j}\}$ generates A, by which A is clearly spanned by $\{\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$.

We define $\left(\dfrac{a, b}{\mathbb{F}}\right)$ to be the quaternion algebra over $\mathbb{F}$ with basis $\mathbf{1}, \mathbf{i}, \mathbf{j}, \mathbf{k}$ satisfying $\mathbf{i}^2 = a$, $\mathbf{j}^2 = b$ and $\mathbf{ij} = -\mathbf{ji}$. In particular, $\left(\dfrac{-1, -1}{\mathbb{R}}\right)$ is the $\mathbb{R}$-algebra of quaternions we discussed earlier. The matrix representation for a general quaternion algebras works like below

$$\begin{aligned} \lambda :\ A &\to \mathbb{F}\left(\sqrt{a}\right)[2] \\ t + x\mathbf{i} + y\mathbf{j} + \mathbf{k} &\mapsto \begin{pmatrix} t + x\sqrt{a} & b(y + z\sqrt{a}) \\ y - z\sqrt{a} & t - x\sqrt{a} \end{pmatrix} \end{aligned}$$

where $\mathbb{F}\left(\sqrt{a}\right)$ is the simple extension given by $\mathbb{F}[X]/(X^2 - a)$, and $\mathbb{F}[X]$ is the polynomial ring over $\mathbb{F}$. Direct calculation shows that λ is an injective $\mathbb{F}$-algebra homomorphism, hence gives an algebra isomorphism onto the image. The matrix representation conversely guarantees the existence of an arbitrary generalized quaternion algebra, since the subset

$$\left\{ \begin{pmatrix} t + x\sqrt{a} & b(y + z\sqrt{a}) \\ y - z\sqrt{a} & t - x\sqrt{a} \end{pmatrix} : t, x, y, z \in \mathbb{F}, a, b \in \mathbb{F} \setminus \{0\} \right\} \subset \mathbb{F}\left(\sqrt{a}\right)[2]$$

forms a subalgebra with $\begin{pmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{a} \end{pmatrix}, \begin{pmatrix} 0 & b \\ 1 & 0 \end{pmatrix}$ satisfying the defining relation in Definition 3.1.

Example 3.2. We have an isomorphism

$$\left(\frac{1, b}{\mathbb{F}}\right) \simeq \mathbb{F}[2]$$

$$t + x\mathbf{i} + y\mathbf{j} + \mathbf{k} \mapsto \begin{pmatrix} t + x & b(y + z) \\ y - z & t - x \end{pmatrix}$$

since the image is of dimension 4.

3.2 Division Condition

It is natural to introduce the notion of a conjugation in generalized quoternion algebras as we did before. Let $A = \left(\frac{a, b}{\mathbb{F}}\right)$, define the conjugation in A as follow

$$\overline{t + x\mathbf{i} + y\mathbf{j} + \mathbf{k}} = t - x\mathbf{i} - y\mathbf{j} - \mathbf{k}$$

which is clearly an anti-automorphism that is also an involution. Then for $q \in A$, we can also define $\mathrm{Re}(q) := (q + \overline{q})/2$, $\mathrm{Im}(q) := (q - \overline{q})/2$ and $N(q) := q\overline{q}$. Let $A^0 := \mathrm{Span}_{\mathbb{F}}\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be the subspace of elements with no real part, it is obvious that $\mathrm{Re}(q) \in \mathbb{F}$ and $\mathrm{Im}(q) \in A^0$. Suppose $q = t + x\mathbf{i} + y\mathbf{j} + \mathbf{k}$, we also have

$$\begin{aligned} N(q) = q\overline{q} &= (t + x\mathbf{i} + y\mathbf{j} + \mathbf{k})(t - x\mathbf{i} - y\mathbf{j} - \mathbf{k}) \\ &= t^2 - ax^2 - by^2 + abz^2 \end{aligned}$$

which belongs to $\mathbb{F}$. Since A is associative, N is immediately multiplicative

$$N(pq) = (pq)\overline{(pq)} = (pq)(\overline{q} \cdot \overline{p}) = p(q\overline{q})\overline{p} = (p\overline{p})N(q) = N(p)N(q)$$

Proposition 3.3. *A is a division algebra if and only if $N(q) \neq 0$ for all $q \neq 0$.*

Proof. Since A is associative and unital, being a division algebra is equivalent to the existence of a multiplicative inverse for each non-zero element. Suppose each $q \in \mathbb{F} \setminus \{0\}$ has inverse q^{-1}, then

$$1 = N(1) = N(qq^{-1}) = N(q)N(q^{-1})$$

implies $N(q) \neq 0$. Conversely if $N(q) \neq 0$ for all $q \in \mathbb{F} \setminus \{0\}$, then each $q \in \mathbb{F} \setminus \{0\}$ has inverse because $q\left(\overline{q}/N(q)\right) = 1$. □

Proposition 3.4. *A is a division algebra if and only if $t^2 = ax^2 + by^2$ has no nontrivial solution.*

Proof. Suppose $N(q) \neq 0$ for all $q \neq 0$, then

$$0 = t^2 - ax^2 - by^2 = N(t + x\mathbf{i} + y\mathbf{j})$$

implies that $t = x = y = 0$. Conversely if $t^2 = ax^2 + by^2$ has no nontrivial solution, suppose $v(t + x\mathbf{i} + y\mathbf{j} + \mathbf{k}) = 0$, then by $t^2 - by^2 = a(x^2 - bz^2)$ we have

$$a(x^2 - bz^2)^2 = (t^2 - by^2)(x^2 - bz^2) = (tx + byz)^2 - b(tz + xy)^2$$

i.e. $(tx + byz)^2 = a(x^2 - bz^2)^2 + b(tz + xy)^2$, hence we must have $x^2 - bz^2 = 0$, which further implies $t^2 - by^2 = 0$. Now the two equations again have no nontrivial solution, we deduce that $t = x = y = z = 0$. □

For instance let $A = \left(\frac{-1,-1}{\mathbb{R}}\right)$, note that $t^2 + x^2 + y^2 = 0$ clearly has no non trivial solution in $\mathbb{R}$, hence A is a division algebra by the proposition, as we saw earlier.

3.3 Isomorphism Classes

It is clear that $\left(\frac{a,b}{\mathbb{F}}\right) \simeq \left(\frac{b,a}{\mathbb{F}}\right)$ by interchanging $\mathbf{i}$ and $\mathbf{j}$, so Definition 3.1 is symmetric in a, b. Moreover, for $a, b, c, d \in \mathbb{F} \setminus \{0\}$, we have $\left(\frac{a,b}{\mathbb{F}}\right) \simeq \left(\frac{ac^2,bd^2}{\mathbb{F}}\right)$ by the canonical isomorphism

$$\theta : \ t + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \mapsto t + cx\mathbf{i} + dy\mathbf{j} + cdz\mathbf{k}$$

which is indeed into $\left(\frac{ac^2,bd^2}{\mathbb{F}}\right)$ (up to isomorphism) as now $\theta(\mathbf{i}), \theta(\mathbf{j})$ serve as defining generators. This implies that if the index $[\mathbb{F}^\times : (\mathbb{F}^\times)^2]$ is finite where $\mathbb{F}^\times$ is the group under multiplication with under lying set $\mathbb{F} \setminus \{0\}$, there can only be finitely many isomorphism classes. In particular if $\mathbb{F}^\times = (\mathbb{F}^\times)^2$, all quaternion algebras over $\mathbb{F}$ are isomorphic to $\left(\frac{1,1}{\mathbb{F}}\right)$, which is $\mathbb{F}[2]$ by Example 3.2.

Example 3.5. There are two isomorphism classes of quoternion algebras over $\mathbb{R}$, represented by $\left(\frac{-1,1}{\mathbb{F}}\right)$ and $\left(\frac{-1,-1}{\mathbb{F}}\right)$, which are $\mathbb{F}[2]$ and $\mathbb{H}$ respectively. Consequently, $\mathbb{H}$ is the only quaternion algebra over $\mathbb{R}$ that is also a division algebra (up to isomorphism).

Indeed, the elements a, b in $\left(\frac{a, b}{\mathbb{F}}\right)$ are far from unique in determining the isomorphism class. Let A and B be two quaternion algebras over $\mathbb{F}$, we have the following result

Proposition 3.6. *$A \simeq B$ as $\mathbb{F}$-algebras if and only if there is a linear isometry between A^0 and B^0.*

Proof. See [PP82]. □

Here the notion of distance is the one induced by the norm on each algebra. Equivalently, suppose $A = \left(\frac{a, b}{\mathbb{F}}\right)$ and $B = \left(\frac{a', b'}{\mathbb{F}}\right)$, we have

Proposition 3.7. *$A \simeq B$ if and only if $diag(-a, -b, ab)$ and $diag(-a', -b', a'b')$ are congruent.*

Proof. If there is a linear isometry $T : A^0 \to B^0$, by identifying elements in A^0 and B^0 with 3-dimensional column vectors, we have

$$v^{\top} \left(T^{\top} \begin{pmatrix} -a' & 0 & 0 \\ 0 & -b' & 0 \\ 0 & 0 & a'b' \end{pmatrix} T \right) v = \|Tv\|_B^2 = \|v\|_A^2 = v^{\top} \begin{pmatrix} -a & 0 & 0 \\ 0 & -b & 0 \\ 0 & 0 & ab \end{pmatrix} v$$

for all $v \in \mathbb{F}^3$, hence the two diagonal matrices are congruent. Conversely the congruence also defines a linear isometry. □

Bibliography

[ABS64] Michael F Atiyah, Raoul Bott, and Arnold Shapiro. Clifford modules. *Topology*, 3:3–38, 1964.

[Bae02] John Baez. The octonions. *Bulletin of the american mathematical society*, 39(2):145–205, 2002.

[Caw04] Raoul Cawagas. On the structure and zero divisors of the cayley-dickson sedenion algebra. *Discussiones Mathematicae-General Algebra and Applications*, 24(2):251–265, 2004.

[Con16] Keith Conrad. Pfister's theorem on sums of squares. *Web address accessed*, 27, 2016.

[Kan89] Isaiah Kantor. *Hypercomplex numbers: an elementary introduction to algebras*, volume 302. Springer, 1989.

[Lyo03] David W Lyons. An elementary introduction to the hopf fibration. *Mathematics magazine*, 76(2):87–98, 2003.

[PGT17] Alba Perez-Gracia and Federico Thomas. On cayley's factorization of 4d rotations and applications. *Advances in Applied Clifford Algebras*, 27:523–538, 2017.

[PP82] Richard S Pierce and Richard S Pierce. *The associative algebra*. Springer, 1982.

[Sch17] Richard D Schafer. *An introduction to nonassociative algebras*. Courier Dover Publications, 2017.

[Voi21] John Voight. *Quaternion algebras*. Springer Nature, 2021.