석 사 학 위 논 문

# 사원수 선형 대수에 관한 연구

제주대학교 대학원

수학과

양 수 연

2020년 6월 8일

# 사원수 선형 대수에 관한 연구

지도교수 김성운

양 수 연

이 논문을 이학 석사학위논문으로 제출함.

2020년 6월 8일

양 수 연의 이학 석사학위 논문을 인준함.

심사위원장 유 상 욱 (인)

심 사 위 원 김 성 운 (인)

심 사 위 원 이 종 락 (인)

제주대학교 대학원

2020년 6월 8일

# A note on quaternion linear algebra

Sooyeon Yang
(Supervised by professor Sungwoon Kim)

A thesis submitted in partial fulfillment of the requirement for the degree of Master of Science

2020. 6. 8.

This thesis has been examined and approved.

---

Thesis director, Sungwoon Kim, Professor of Mathematics

Department of Mathematics
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY

**초 록**

이 논문에서는 사원수의 대수적인 구조들 및 사원수 쌍곡 공간들의 기하적인 특성들에 대해 연구하였다.

**핵 심 낱 말** 사원수, 나눗셈환, 가군, 사원수 선형대수, 사원수 쌍곡 공간.

**Abstract**

We study the algebraic structures of quaternions and geometric properties of quaternionic hyperbolic spaces.

**Keywords** Quaternion, division ring, module, quaternionic linear algebra and quaternionic hyperbolic space.

# Contents

# Chapter 1. Introduction

From the fact that the complex numbers are presented as points on a 2-dimensional plane, Hamilton tried to find a way to interpret points on a 3-dimensional space in a similar way as complex numbers. First, he thought a number system that consists of triples of numbers but he failed to solve the problem of multiplication and division. In the end, he realized that there is a number system consisting of quadruples of numbers with a notion of division.

After Hamilton's death, Peter Guthrie Tait who is Hamilton's student continued studying quaternions [4]. But from the mid-1880's, people were not interested in quaternions. But quaternions were again in the spotlight in the late 20th century because of their utility in describing spatial rotations. Describing spatial rotations by quaternions is more efficient and quicker to compute than describing spatial rotations by matrices. For this reason, quaternions are used in computer graphics [5], control theory, signal processing, attitude control, physics, bioinformatics [6, 7], molecular dynamics, computer simulations and orbital mechanics.

We will start with the definition of quaternions first and then study their basic properties concerning quaternion linear algebra. The main difference of quaternions from complex numbers is non-commutativity, which makes all difficult. For example, some quadratic equation has infinitely many solution in quaternions. It turns out that the number system of quaternions is not a field but fortunately a division ring. From this, it follows that modules over quaternions are similar to vector spaces over a field. We will give self-contained proofs for well-known results in quaternion linear algebra.

# Chapter 2. Quaternions

In this chapter. we will study quaternion's definition and basic theorems. We start with the definition of quaternions.

**Definition 2.1** (Quaternions)**.** A *quaternion* is an expression of the form

$$a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k},$$

where $a, b, c, d$ are real numbers, and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are just symbols. If one of $a, b, c, d$ is 0, the corresponding term is omitted if $a, b, c, d$ are all zero, the quaternion is the zero quaternion, denoted 0; if one of $b, c, d$ equals 1, the corresponding term is written simply $\mathbf{i}, \mathbf{j}$, or $\mathbf{k}$. The quaternion $b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ is called the *imaginary part* of $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$, and $a$ is the *real part* of $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$. Let $\mathbb{H}$ be the set of quaternions. Define two binary operations $+$ and $\cdot$ on $\mathbb{H}$ as follows. First, the addition $+$ on $\mathbb{H}$ is defined by

$$(a_1+b_1\mathbf{i}+c_1\mathbf{j}+d_1\mathbf{k})+(a_2+b_2\mathbf{i}+c_2\mathbf{j}+d_2\mathbf{k}) = (a_1+a_2)+(b_1+b_2)\mathbf{j}+(c_1+c_2)\mathbf{j}+(d_1+d_2)\mathbf{k}.$$

The multiplication $\cdot$, called the Hamilton product, is defined by

$$\begin{aligned}
&(a_1 + b_1\mathbf{i} + c_1\mathbf{j} + d_1\mathbf{k}) \cdot (a_2 + b_2\mathbf{i} + c_2\mathbf{j} + d_2\mathbf{k}) \\
&\quad = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)\mathbf{i} \\
&\qquad + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)\mathbf{j} + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)\mathbf{k}.
\end{aligned}$$

From the definition of quaternions with the addition and multiplication, it immediately follows that

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i} \cdot \mathbf{j} \cdot \mathbf{k} = -1$$

and

$$\mathbf{i} \cdot \mathbf{j} = -\mathbf{j} \cdot \mathbf{i} = \mathbf{k},\ \mathbf{j} \cdot \mathbf{k} = -\mathbf{k} \cdot \mathbf{j} = \mathbf{i},\ \mathbf{k} \cdot \mathbf{i} = -\mathbf{i} \cdot \mathbf{k} = \mathbf{j}.$$

**Lemma 2.2.** *Let $x$ be a quaternion. Then $x \cdot y = y \cdot x$ for all $y \in \mathbb{H}$ if and only if $x$ is a real number.*

*Proof.* Suppose that $xy = yx$ for all $y \in \mathbb{H}$. Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$. Then

$$\begin{aligned}
x \cdot y &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \cdot (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\
&\quad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k}
\end{aligned}$$

and

$$
\begin{aligned}
y \cdot x &= (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \cdot (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \\
&= (y_0x_0 - y_1x_1 - y_2x_2 - y_3x_3) + (y_0x_1 + y_1x_0 + y_2x_3 - y_3x_2)\mathbf{i} \\
&\quad + (y_0x_2 - y_1x_3 + y_2x_0 + y_3x_1)\mathbf{j} + (y_0x_3 + y_1x_2 - y_2x_1 + y_3x_0)\mathbf{k}.
\end{aligned}
$$

To be $xy = yx$, we have that for all $y_1, y_2, y_3 \in \mathbb{R}$,

$$
\begin{aligned}
x_2y_3 - x_3y_2 &= 0, \\
x_1y_3 - x_3y_1 &= 0, \\
x_1y_2 - x_2y_1 &= 0.
\end{aligned}
$$

Taking $y_1 = y_2 = 1$ and $y_3 = 0$, the equations above reduce to

$$
-x_3 = 0, \ \ x_1 = x_2. \tag{2.1}
$$

Furthermore, taking $y_1 = y_3 = 1$ and $y_2 = 0$, the equations reduce to

$$
x_2 = 0, \ \ x_1 - x_3 = 0. \tag{2.2}
$$

From (2.1) and (2.2), it follows that $x_1 = x_2 = x_3 = 0$, which implies that $x$ is a real number. The converse is obvious. □

Lemma 2.2 means that the center of $\mathbb{H}$ for the multiplication on $\mathbb{H}$ is the set of real numbers.

**Proposition 2.3.** *The addition* $+$ *and multiplication* $\cdot$ *turn* $\mathbb{H}$ *into a real vector space of dimension* 4.

*Proof.* Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ and $z = z_0 + z_1\mathbf{i} + z_2\mathbf{j} + z_3\mathbf{k}$ and $a, b \in \mathbb{R}$. Then,

$$
\begin{aligned}
&(x + y) + z \\
&\quad = ((x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k})) + (z_0 + z_1\mathbf{i} + z_2\mathbf{j} + z_3\mathbf{k}) \\
&\quad = ((x_0 + y_0) + (x_1 + y_1)\mathbf{i} + (x_2 + y_2)\mathbf{j} + (x_3 + y_3)\mathbf{k}) + (z_0 + z_1\mathbf{i} + z_2\mathbf{j} + z_3\mathbf{k}) \\
&\quad = ((x_0 + y_0) + z_0) + ((x_1 + y_1) + z_1)\mathbf{i} + ((x_2 + y_2) + z_2)\mathbf{j} + ((x_3 + y_3) + z_3)\mathbf{k} \\
&\quad = (x_0 + (y_0 + z_0)) + (x_1 + (y_1 + z_1))\mathbf{i} + (x_2 + (y_2 + z_2))\mathbf{j} + (x_3 + (y_3 + z_3))\mathbf{k} \\
&\quad = (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + ((y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) + (z_0 + z_1\mathbf{i} + z_2\mathbf{j} + z_3\mathbf{k})) \\
&\quad = x + (y + z)
\end{aligned}
$$

This implies the associativity of the quaternion addition. The fourth equality follows from the associativity of real number under addition. The commutativity of real numbers under addition gives the commutativity of quaternion addition as follows.

$$\begin{aligned} x+y &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \\ &= (x_0 + y_0) + (x_1 + y_1)\mathbf{i} + (x_2 + y_2)\mathbf{j} + (x_3 + y_3)\mathbf{k} \\ &= (y_0 + x_0) + (y_1 + x_1)\mathbf{i} + (y_2 + x_2)\mathbf{j} + (y_3 + x_3)\mathbf{k} \\ &= y + x. \end{aligned}$$

One can easily check that 0 is the identity for quaternion addition. For all $x \in \mathbb{H}$,

$$\begin{aligned} x+0 &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + (0 + 0\mathbf{i} + 0\mathbf{j} + 0\mathbf{k}) \\ &= (x_0 + 0) + (x_1 + 0)\mathbf{i} + (x_2 + 0)\mathbf{j} + (x_3 + 0)\mathbf{k} \\ &= (0 + x_0) + (0 + x_1)\mathbf{i} + (0 + x_2)\mathbf{j} + (0 + x_3)\mathbf{k} \\ &= x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k} \\ &= 0 + x \\ &= x. \end{aligned}$$

This implies that $x + 0 = 0 + x = x$. For each $x \in \mathbb{H}$, the additive inverse of $x$ is $-x$, that is,

$$\begin{aligned} x+(-x) &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + (-x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}) \\ &= (x_0 - x_0) + (x_1 - x_1)\mathbf{i} + (x_2 - x_2)\mathbf{j} + (x_3 - x_3)\mathbf{k} \\ &= (-x_0 + x_0) + (-x_1 + x_1)\mathbf{i} + (-x_2 + x_2)\mathbf{j} + (-x_3 + x_3)\mathbf{k} \\ &= (-x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}) + (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \\ &= (-x) + x \\ &= 0. \end{aligned}$$

The associativity of scalar multiplication immediately follows from the associativity of real numbers under multiplication:

$$\begin{aligned} a \cdot (b \cdot x) &= a(bx_0 + bx_1\mathbf{i} + bx_2\mathbf{j} + bx_3\mathbf{k}) \\ &= a(bx_0) + a(bx_1)\mathbf{i} + a(bx_2)\mathbf{j} + a(bx_3)\mathbf{k} \\ &= (ab)x_0 + (ab)x_1\mathbf{i} + (ab)x_2\mathbf{j} + (ab)x_3\mathbf{k} \\ &= (ab) \cdot x. \end{aligned}$$

The distributivities of scalar sums and quaternion sums follow from the distributivity of real numbers as follows.

$$\begin{aligned}(a+b)\cdot x &= (a+b)x_0 + (a+b)x_1\mathbf{i} + (a+b)x_2\mathbf{j} + (a+b)x_3\mathbf{k}\\ &= (ax_0+bx_0) + (ax_1+bx_1)\mathbf{i} + (ax_2+bx_2)\mathbf{j} + (ax_3+bx_3)\mathbf{k}\\ &= (ax_0+ax_1\mathbf{i}+ax_2\mathbf{j}+ax_3\mathbf{k}) + (bx_0+bx_1\mathbf{i}+bx_2\mathbf{j}+bx_3\mathbf{k})\\ &= a\cdot x + b\cdot x,\end{aligned}$$

and

$$\begin{aligned}a\cdot(x+y) &= a(x_0+y_0) + a(x_1+y_1)\mathbf{i} + a(x_2+y_2)\mathbf{j} + a(x_3+y_3)\mathbf{k}\\ &= (ax_0+ay_0) + (ax_1+ay_1)\mathbf{i} + (ax_2+ay_2)\mathbf{j} + (ax_3+ay_3)\mathbf{k}\\ &= a\cdot x + a\cdot y.\end{aligned}$$

Finally, it easily follows that the number 1 is the scalar multiplication identity:

$$\begin{aligned}1\cdot x &= 1\cdot(x_0+x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k})\\ &= x_0+x_1\mathbf{i}+x_2\mathbf{j}+x_3\mathbf{k}\\ &= x.\end{aligned}$$

Therefore $(\mathbb{H},+,\cdot)$ satisfies all conditions in the definition of real vector space. Hence $(\mathbb{H},+,\cdot)$ is a real vector space. $\square$

**Definition 2.4** (Ring and Division Ring)**.** A *ring* is a set $R$ equipped with two binary operations $+$ and $\cdot$ satisfying the following three sets of axioms, called the ring axioms

1. $R$ is an abelian group under addition, meaning that:

$$(a+b)+c = a+(b+c)$$

for all $a,b,c \in R$ (that is, $+$ is associative).

$$a+b = b+a$$

for all $a,b \in R$ (that is, $+$ is commutative). There is an element 0 in $R$ such that

$$a+0 = a$$

for all $a$ in $R$ (that is, 0 is the additive identity). For each $a$ in $R$ there exists $-a$ in $R$ such that

$$a+(-a) = 0$$

(that is, $-a$ is the additive inverse of $a$).

2. $R$ is a monoid under multiplication, meaning that:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c)$$

for all $a, b, c \in R$ (that is, $\cdot$ is associative). There is an element 1 in $R$ such that

$$a \cdot 1 = a$$

and

$$1 \cdot a = a$$

for all $a$ in $R$ (that is, 1 is the multiplicative identity).

3. Multiplication is distributive with respect to addition, meaning that:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

for all $a, b, c \in R$ (left distributivity).

$$(b + c) \cdot a = (b \cdot a) + (c \cdot a)$$

for all $a, b, c \in R$ (right distributivity).

A ring $R$ is called a *division ring* if every nonzero element of $R$ has a multiplicative inverse.

**Example 2.5.** The set $M_2(\mathbb{C})$ of $2 \times 2$ complex matrices is a ring.

*Proof.* Let $+$ and $\cdot$ be the addition and multiplication on $M_2(\mathbb{C})$. Let $X, Y$ and $Z$ be $2 \times 2$ complex matrices. Write

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}, \ Y = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}, \ Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22}. \end{bmatrix}.$$

Then from the associativity of the addition of complex numbers, we have that

$$\begin{aligned} (X+Y)+Z &= \begin{bmatrix} (x_{11}+y_{11})+z_{11} & (x_{12}+y_{12})+z_{12} \\ (x_{21}+y_{21})+z_{21} & (x_{22}+y_{22})+z_{22} \end{bmatrix} \\ &= \begin{bmatrix} x_{11}+y_{11}+z_{11} & x_{12}+y_{12}+z_{12} \\ x_{21}+y_{21}+z_{21} & x_{22}+y_{22}+z_{22} \end{bmatrix} \\ &= \begin{bmatrix} x_{11}+(y_{11}+z_{11}) & x_{12}+(y_{12}+z_{12}) \\ x_{21}+(y_{21}+z_{21}) & x_{22}+(y_{22}+z_{22}) \end{bmatrix} = X+(Y+Z). \end{aligned}$$

Thus the addition $+$ on $M_2(\mathbb{C})$ is associative. Moreover, from the commutativity of the addition of complex numbers,

$$\begin{aligned} X+Y &= \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \\ &= \begin{bmatrix} x_{11}+y_{11} & x_{12}+y_{12} \\ x_{21}+y_{21} & x_{22}+y_{22} \end{bmatrix} \\ &= \begin{bmatrix} y_{11}+x_{11} & y_{12}+x_{12} \\ y_{21}+x_{21} & y_{22}+x_{22} \end{bmatrix} \\ &= \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} + \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = Y+X. \end{aligned}$$

Thus the addition $+$ on $M_2(\mathbb{C})$ is commutative. Let $\mathbf{0}$ denote the zero $2\times 2$ matrix. For any $2\times 2$ complex matrix $X$,

$$X+\mathbf{0} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} x_{11}+0 & x_{12}+0 \\ x_{21}+0 & x_{22}+0 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = X.$$

This implies that the zero matrix $\mathbf{0}$ is the additive identity in $M_2(\mathbb{C})$. To complete the proof that $M_2(\mathbb{C})$ is an abelian group under addition, we only need to show the existence of the additive inverse. For each $X \in M_2(\mathbb{C})$, we set

$$-X = \begin{bmatrix} -x_{11} & -x_{12} \\ -x_{21} & -x_{22} \end{bmatrix}.$$

Then it follows that

$$\begin{aligned} X+(-X) &= \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} + \begin{bmatrix} -x_{11} & -x_{12} \\ -x_{21} & -x_{22} \end{bmatrix} \\ &= \begin{bmatrix} x_{11}-x_{11} & x_{12}-x_{12} \\ x_{21}-x_{21} & x_{22}-x_{22} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathbf{0}. \end{aligned}$$

Hence $(-X)$ is the additive inverse of $X$. All results above imply that $M_2(\mathbb{C})$ is an abelian group under addition.

Next we will prove that $M_2(\mathbb{C})$ is a monoid under multiplication. First, it can be easily checked that the multiplication on $M_2(\mathbb{C})$ is associative as follows. A simple computation gives that the $(i,j)$-entry of $(X\cdot Y)\cdot Z$ is

$$(x_{i1}y_{11}+x_{i2}y_{21})z_{1j} + (x_{i1}y_{12}+x_{i2}y_{22})z_{2j}$$

and $(i,j)$-entry of $X\cdot(Y\cdot Z)$ is

$$(x_{i1}(y_{11}z_{1j}+y_{12}z_{2j}) + x_{i2}(y_{21}z_{1j}+y_{22}z_{2j}),$$

which deduces that they equal to each other. Thus it follows that $(X \cdot Y) \cdot Z = X \cdot (Y \cdot Z)$.

Let $I$ denote the identity matrix in $M_2(\mathbb{C})$ that is the $2 \times 2$ matrix with ones on the main diagonal and zeros elsewhere. For each $X \in M_2(\mathbb{C})$,

$$X \cdot I = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = X,$$

and

$$I \cdot X = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = X.$$

This implies that the identity matrix $I$ is the multiplicative identity in $M_2(\mathbb{C})$. From the results above, we conclude that $M_2(\mathbb{C})$ is a monoid under multiplication.

Lastly, it remains to verifty that the multiplication on $M_2(\mathbb{C})$ is distributive with respect to the addition on $M_2(\mathbb{C})$. For any $X, Y, Z \in M_2(\mathbb{C})$,

$$\begin{aligned}
X \cdot (Y + Z) &= \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \cdot \begin{bmatrix} y_{11} + z_{11} & y_{12} + z_{12} \\ y_{21} + z_{21} & y_{22} + z_{22} \end{bmatrix} \\
&= \begin{bmatrix} x_{11}(y_{11} + z_{11}) + x_{12}(y_{21} + z_{21}) & x_{11}(y_{12} + z_{12}) + x_{12}(y_{22} + z_{22}) \\ x_{21}(y_{11} + z_{11}) + x_{22}(y_{21} + z_{21}) & x_{21}(y_{12} + z_{12}) + x_{22}(y_{22} + z_{22}) \end{bmatrix} \\
&= \begin{bmatrix} (x_{11}y_{11} + x_{12}y_{21}) + (x_{11}z_{11} + x_{12}z_{21}) & (x_{11}y_{12} + x_{12}y_{22}) + (x_{11}z_{12} + x_{12}z_{22}) \\ (x_{21}y_{11} + x_{22}y_{21}) + (x_{21}z_{11} + x_{22}z_{21}) & (x_{21}y_{12} + x_{22}y_{22}) + (x_{21}z_{12} + x_{22}z_{22}) \end{bmatrix} \\
&= \begin{bmatrix} x_{11}y_{11} + x_{12}y_{21} & x_{11}y_{12} + x_{12}y_{22} \\ x_{21}y_{11} + x_{22}y_{21} & x_{21}y_{12} + x_{22}y_{22} \end{bmatrix} + \begin{bmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} \end{bmatrix} \\
&= X \cdot Y + X \cdot Z.
\end{aligned}$$

Thus $X \cdot (Y + Z) = X \cdot Y + X \cdot Z$. Furthermore,

$$\begin{aligned}
(X + Y) \cdot Z &= \begin{bmatrix} x_{11} + y_{11} & x_{12} + y_{12} \\ x_{21} + y_{21} & x_{22} + y_{22} \end{bmatrix} \cdot \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \\
&= \begin{bmatrix} (x_{11} + y_{11})z_{11} + (x_{12} + y_{12})z_{21} & (x_{11} + y_{11})z_{12} + (x_{12} + y_{12})z_{22} \\ (x_{21} + y_{21})z_{11} + (x_{22} + y_{22})z_{21} & (x_{21} + y_{21})z_{12} + (x_{22} + y_{22})z_{22} \end{bmatrix} \\
&= \begin{bmatrix} x_{11}z_{11} + x_{12}z_{21} + y_{11}z_{11} + y_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} + y_{11}z_{12} + y_{12}z_{22} \\ x_{21}z_{11} + x_{22}z_{21} + y_{21}z_{11} + y_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} + y_{21}z_{12} + y_{22}z_{22} \end{bmatrix} \\
&= \begin{bmatrix} x_{11}z_{11} + x_{12}z_{21} & x_{11}z_{12} + x_{12}z_{22} \\ x_{21}z_{11} + x_{22}z_{21} & x_{21}z_{12} + x_{22}z_{22} \end{bmatrix} + \begin{bmatrix} y_{11}z_{11} + y_{12}z_{21} & y_{11}z_{12} + y_{12}z_{22} \\ y_{21}z_{11} + y_{22}z_{21} & y_{21}z_{12} + y_{22}z_{22} \end{bmatrix} \\
&= \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \cdot \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} + \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \cdot \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \\
&= X \cdot Z + Y \cdot Z.
\end{aligned}$$

We complete a proof for the left and right distributivity. Therefore $M_2(\mathbb{C})$ equipped with the addition $+$ and multiplication $\cdot$ satisfies all the ring axioms. In other words $(M_2(\mathbb{C}), +, \cdot)$ is a ring. $\square$

**Proposition 2.6.** *The set* $\mathbb{H}$ *of quaternions is a non-commutative division ring.*

*Proof.* First note that we have shown that $\mathbb{H}$ is an abelian group under addition. Hence it remains to prove that $\mathbb{H}$ is a monoid under multiplication and the multiplication on $\mathbb{H}$ is distributive with respect to addition. For $a, b$ and $c$ in $\mathbb{H}$,

$$
\begin{aligned}
(a \cdot b) \cdot c &= [(a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})] \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= [a_0b_0c_0 - a_1b_1c_0 - a_2b_2c_0 - a_3b_3c_0 - a_0b_1c_1 - a_1b_0c_1 - a_2b_3c_1 + a_3b_2c_1 \\
&\quad - a_0b_2c_2 + a_1b_3c_2 - a_2b_0c_2 - a_3b_1c_2 - a_0b_3c_3 - a_1b_2c_3 + a_2b_1c_3 - a_3b_0c_3] \\
&\quad + [a_0b_0c_1 - a_1b_1c_1 - a_2b_2c_1 - a_3b_3c_1 + a_0b_1c_0 + a_1b_0c_0 + a_2b_3c_0 - a_3b_2c_0 \\
&\quad + a_0b_2c_3 - a_1b_3c_3 + a_2b_0c_3 + a_3b_1c_3 - a_0b_3c_2 - a_1b_2c_2 + a_2b_1c_2 - a_3b_0c_2]\mathbf{i} \\
&\quad + [a_0b_0c_2 - a_1b_1c_2 - a_2b_2c_2 - a_3b_3c_2 - a_0b_1c_3 - a_1b_0c_3 - a_2b_3c_3 + a_3b_2c_3 \\
&\quad + a_0b_2c_0 - a_1b_3c_0 + a_2b_0c_0 + a_3b_1c_0 + a_0b_3c_1 + a_1b_2c_1 - a_2b_1c_1 + a_3b_0c_1]\mathbf{j} \\
&\quad + [a_0b_0c_3 - a_1b_1c_3 - a_2b_2c_3 - a_3b_3c_3 + a_0b_1c_2 + a_1b_0c_2 + a_2b_3c_2 - a_3b_2c_2 \\
&\quad - a_0b_2c_1 + a_1b_3c_1 - a_2b_0c_1 - a_3b_1c_1 + a_0b_3c_0 + a_1b_2c_0 - a_2b_1c_0 + a_3b_0c_0]\mathbf{k} \\
&= a_0[(b_0c_0 - b_1c_1 - b_2c_2 - b_3c_3) + (b_0c_1 + b_1c_0 + b_2c_3 - b_3c_2)\mathbf{i} \\
&\quad + (b_0c_2 - b_1c_3 + b_2c_0 + b_3c_1)\mathbf{j} + (b_0c_3 + b_1c_2 - b_2c_1b_3c_0)\mathbf{k}] \\
&\quad + a_1[(-b_1c_0 - b_0c_1 + b_3c_2 - b_2c_3) + (-b_1c_1 + b_0c_0 - b_3c_3 - b_2c_2)\mathbf{i} \\
&\quad + (-b_1c_2 - b_0c_3 - b_3c_0 + b_2c_1)\mathbf{j} + (-b_1c_3 + b_0c_2 + b_3c_1 + b_2c_0)\mathbf{k}] \\
&\quad + a_2[(-b_2c_0 - b_3c_1 - b_0c_2 + b_1c_3) + (-b_2c_1 + b_3c_0 + b_0c_3 + b_2c_2)\mathbf{i} \\
&\quad + (-b_2c_2 - b_3c_3 + b_0c_0 - b_1c_1)\mathbf{j} + (-b_2c_3 + b_3c_2 - b_0c_1 - b_1c_0)\mathbf{k} \\
&\quad + a_3[(-b_3c_0 + b_2c_1 - b_1c_2 + b_2c_1) + (-b_3c_1 - b_2c_0 + b_1c_3 - b_2c_0)\mathbf{i} \\
&\quad + (-b_3c_2 + b_2c_3 + b_1c_0 + b_0c_1)\mathbf{j} + (-b_3c_3 - b_2c_2 - b_1c_1 + b_0c_0)\mathbf{k}] \\
&= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot [(b_0c_0 - b_1c_1 - b_2c_2 - b_3c_3) + (b_0c_1 + b_1c_0 + b_2c_3 - b_3c_2)\mathbf{i} \\
&\quad + (b_0c_2 - b_1c_3 + b_2c_0 + b_3c_1)\mathbf{j} + (b_0c_3 + b_1c_2 - b_2c_1 + b_3c_0)\mathbf{k}] \\
&= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot [(b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k})] \\
&= a \cdot (b \cdot c).
\end{aligned}
$$

Thus $(a \cdot b) \cdot c = a \cdot (b \cdot c)$. This implies the associativity of the multiplication on $\mathbb{H}$.

For each $a \in \mathbb{H}$,

$$a \cdot 1 = (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot 1 = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = a = 1 \cdot a.$$

This implies that the identity 1 is the multiplicative identity in $\mathbb{H}$. From the results above, we conclude that $\mathbb{H}$ is a monoid under multiplication.

Next we will verify that the multiplication on $\mathbb{H}$ is distributive with respect to the addition on $\mathbb{H}$. For any $a, b, c \in \mathbb{H}$,

$$\begin{aligned}
a \cdot (b+c) &= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot [(b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) + (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k})] \\
&= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot [(b_0 + c_0) + (b_1 + c_1)\mathbf{i} + (b_2 + c_2)\mathbf{j} + (b_3 + c_3)\mathbf{k}] \\
&= [a_0(b_0 + c_0) - a_1(b_1 + c_1) - a_2(b_2 + c_2) - a_3(b_3 + c_3)] \\
&\quad\; + [a_0(b_1 + c_1) + a_1(b_0 + c_0) + a_2(b_3 + c_3) - a_3(b_2 + c_2)]\mathbf{i} \\
&\quad\; + [a_0(b_2 + c_2) - a_1(b_3 + c_3) + a_2(b_0 + c_0) + a_3(b_1 + c_1)]\mathbf{j} \\
&\quad\; + [a_0(b_3 + c_3) + a_1(b_2 + c_2) - a_2(b_1 + c_1) + a_3(b_0 + c_0)]\mathbf{k} \\
&= (a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) + (a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2)\mathbf{i} \\
&\quad\; + (a_0b_2 - a_1b_3 + a_2b_0 + a_3b_1)\mathbf{j} + (a_0b_3 + a_1b_2 - a_2b_1 + a_3b_0)\mathbf{k} \\
&\quad\; + (a_0c_0 - a_1c_1 - a_2c_2 - a_3c_3) + (a_0c_1 + a_1c_0 + a_2c_3 - a_3c_2)\mathbf{i} \\
&\quad\; + (a_0c_2 - a_1c_3 + a_2c_0 + a_3c_1)\mathbf{j} + (a_0c_3 + a_1c_2 - a_2c_1 + a_3c_0)\mathbf{k} \\
&= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \\
&\quad\; + (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= a \cdot b + a \cdot c.
\end{aligned}$$

In addition, we have

$$\begin{aligned}
(a+b) \cdot c &= [(a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) + (b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k})] \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= [(a_0 + b_0) + (a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}] \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= [(a_0 + b_0)c_0 - (a_1 + b_1)c_1 - (a_2 + b_2)c_2 - (a_3 + b_3)c_3] \\
&\quad + [(a_0 + b_0)c_1 + (a_1 + b_1)c_0 + (a_2 + b_2)c_3 - (a_3 + b_3)c_2]\mathbf{i} \\
&\quad + [(a_0 + b_0)c_2 - (a_1 + b_1)c_3 + (a_2 + b_2)c_0 + (a_3 + b_3)c_1]\mathbf{j} \\
&\quad + [(a_0 + b_0)c_3 + (a_1 + b_1)c_2 - (a_2 + b_2)c_1 + (a_3 + b_3)c_0]\mathbf{k} \\
&= (a_0c_0 - a_1c_1 - a_2c_2 - a_3c_3) + (a_0c_1 + a_1c_0 + a_2c_3 - a_3c_2)\mathbf{i} \\
&\quad\; + (a_0c_2 - a_1c_3 + a_2c_0 + a_3c_1)\mathbf{j} + (a_0c_3 + a_1c_2 - a_2c_1 + a_3c_0)\mathbf{k} \\
&\quad\; + (b_0c_0 - b_1c_1 - b_2c_2 - b_3c_3) + (b_0c_1 + b_1c_0 + b_2c_3 - b_3c_2)\mathbf{i} \\
&\quad\; + (b_0c_2 - b_1c_3 + b_2c_0 + b_3c_1)\mathbf{j} + (b_0c_3 + b_1c_2 - b_2c_1 + b_3c_0)\mathbf{k} \\
&= (a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&\quad\; + (b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \cdot (c_0 + c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}) \\
&= a \cdot c + b \cdot c.
\end{aligned}$$

Two identities above imply the left and right distributivity. Finally, we conclude that $\mathbb{H}$ equipped with the addition $+$ and multiplication $\cdot$ satisfies all the ring axioms. In other words, $(\mathbb{H}, +, \cdot)$ is a ring. To verify that $\mathbb{H}$ is a division ring, let $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$

be an arbitrary non-zero quaternion. Then we can give an inverse of $a$ explicitly as follows. Define the inverse $a^{-1}$ of $a$ by

$$a^{-1} = \frac{1}{a_0^2 + a_1^2 + a_2^2 + a_3^2}(a_0 - a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}).$$

Then a trivial verification shows that $a \cdot a^{-1} = a^{-1} \cdot a = 1$, which means that $a^{-1}$ is the multiplicative inverse of $a$. Lastly, one can easily see that $\mathbb{H}$ is non-commutative by $\mathbf{k} = \mathbf{i} \cdot \mathbf{j} \neq \mathbf{j} \cdot \mathbf{i} = -\mathbf{k}$. Thus it is derived that $\mathbb{H}$ is a non-commutative division ring. $\square$

**Definition 2.7** (Norm on a division ring)**.** Let $(R, +, \cdot)$ be a division ring whose zero is denoted by 0. A norm on $R$ is a mapping from $R$ to the non-negative reals $\mathbb{R}_{\geq 0}$:

$$|\cdot| : R \to \mathbb{R}_{\geq 0}$$

satisfying the norm axioms: For all $x$ and $y$ in $R$,

1. (Positive definiteness) If $|x| = 0$, then $x = 0$.
2. (Multiplicativity) $|x \cdot y| = |x||y|$.
3. (Triangle inequality) $|x + y| \leq |x| + |y|$.

Define a norm $|\cdot|$ on $\mathbb{H}$ by $|a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2}$.

**Proposition 2.8.** *The norm $|\cdot|$ on $\mathbb{H}$ is a norm on the division ring $\mathbb{H}$.*

*Proof.* Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$. Suppose that $|x| = 0$. Then $|x|^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0$, which implies that $x_0 = x_1 = x_2 = x_3 = 0$ and hence $x = 0$. The norm $|\cdot|$ is positive definite. An easy computation gives that

$$\begin{aligned}
|x \cdot y|^2 &= |(x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\
&\qquad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k}|^2 \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3)^2 + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)^2 \\
&\qquad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)^2 + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)^2 \\
&= (x_0^2 + x_1^2 + x_2^2 + x_3^2)(y_0^2 + y_1^2 + y_2^2 + y_3^2) = |x|^2|y|^2.
\end{aligned}$$

From this, the multiplicativity of the norm immediately follows. It only remains to check the triangle inequality. By a direct computation,

$$\begin{aligned}
(|x| + |y|)^2 - |x + y|^2 = (x_0y_1 - x_1y_0)^2 + (x_0y_2 - x_2y_0)^2 + (x_0y_3 - x_3y_0)^2 \\
+ (x_1y_2 - x_2y_1)^2 + (x_1y_2 - x_3y_1)^2 + (x_2y_3 - x_3y_2)^2.
\end{aligned}$$

This gives rise to the triangle inequality $|x + y| \leq |x| + |y|$ and the equality holds if and only if $y = r \cdot x$ for some real number $r$. Therefore we complete the proof. $\square$

Define the conjugate $\bar{x}$ of a quaternion $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ by $\bar{x} = x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}$.

**Theorem 2.9.** *Let $x$ and $y$ be quaternions. Then the following holds.*

(1) $x\bar{x} = \bar{x}x = |x|^2$.

(2) *If $x \neq 0$, the inverse of $x$ is $\bar{x}/|x|^2$.*

(3) $\overline{xy} = \bar{y}\bar{x}$.

*Proof.* Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ be a quaternion. Then it can be easily seen that

$$x\bar{x} = \bar{x}x = |x|^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2.$$

Furthermore, if $x \neq 0$, it immediately follows that

$$x \cdot \frac{\bar{x}}{|x|^2} = \frac{\bar{x}}{|x|^2} \cdot x = 1.$$

Thus the inverse of $x$ is $\bar{x}/|x|^2$. For a quaternion $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$,

$$\begin{aligned}
\overline{xy} &= (x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k})(y_0 - y_1\mathbf{i} - y_2\mathbf{j} - y_3\mathbf{k}) \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) - (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\
&\qquad - (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} - (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k}
\end{aligned}$$

and

$$\begin{aligned}
\bar{y}\bar{x} &= (y_0 - y_1\mathbf{i} - y_2\mathbf{j} - y_3\mathbf{k})(x_0 - x_1\mathbf{i} - x_2\mathbf{j} - x_3\mathbf{k}) \\
&= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) - (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\
&\qquad - (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} - (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k}.
\end{aligned}$$

Thus $\overline{xy} = \bar{y}\bar{x}$. We finish the proof. □

**Definition 2.10** (Ring homomorphism)**.** Let $(R, +_R, \cdot_R)$ and $(S, +_S, \cdot_S)$ be rings. Then a function $f : R \to S$ is said to be a *ring homomorphism* if for any two elements $a, b \in R$ the following conditions are satisfied:

$$\begin{array}{ll}
f(a +_R b) = f(a) +_S f(b) & \text{addition preserving,} \\
f(a \cdot_R b) = f(a) \cdot_S f(b) & \text{multiplication preserving,} \\
f(1_R) = 1_S & \text{unit preserving,}
\end{array}$$

where $1_R$ and $1_S$ are the identities of $R$ and $S$ respectively.

Define a map $\phi : \mathbb{H} \to M_2(\mathbb{C})$ by

$$\phi(x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) = \begin{bmatrix} x_0 + x_1\mathbf{i} & x_2 + x_3\mathbf{i} \\ -x_2 + x_3\mathbf{i} & x_0 - x_1\mathbf{i} \end{bmatrix}.$$

Jacobson [1] suggested to define quaternions as the subset of the ring $M_2(\mathbb{C})$ of $2 \times 2$ matrices with complex number entries as follows.

**Proposition 2.11.** *The map $\phi : \mathbb{H} \to M_2(\mathbb{C})$ is an injective ring homomorphism. Furthermore,*

(1) $|x|^2 = \det \phi(x)$.

(2) *the eigenvalues of* $\phi(x)$ *are* $\mathrm{Re}(x) \pm |\mathrm{Im}(x)|\mathbf{i}$.

*Proof.* First, we will show the map $\phi$ is a ring homomorphism. For $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$ in $\mathbb{H}$,

$$\begin{aligned} f(x+y) &= f((x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) + (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k})) \\ &= f((x_0+y_0) + (x_1+y_1)\mathbf{i} + (x_2+y_2)\mathbf{j} + (x_3+y_3)\mathbf{k}) \\ &= \begin{bmatrix} (x_0+y_0) + (x_1+y_1)\mathbf{i} & (x_2+y_2) + (x_3+y_3)\mathbf{i} \\ -(x_2+y_2) + (x_3+y_3)\mathbf{i} & (x_0+y_0) - (x_1+y_1)\mathbf{i} \end{bmatrix} \\ &= \begin{bmatrix} x_0 + x_1\mathbf{i} & x_2 + x_3\mathbf{i} \\ -x_2 + x_3\mathbf{i} & x_0 - x_1\mathbf{i} \end{bmatrix} + \begin{bmatrix} y_0 + y_1\mathbf{i} & y_2 + y_3\mathbf{i} \\ -y_2 + y_3\mathbf{i} & y_0 - y_1\mathbf{i} \end{bmatrix} \\ &= f(x) + f(y). \end{aligned}$$

Hence the map $\phi$ is an addition preserving map. Furthermore,

$$\begin{aligned} f(x \cdot y) &= f((x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \cdot_R (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k})) \\ &= f((x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\ &\qquad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k}), \end{aligned}$$

A straightforward computation gives that the $(1,1)$-entry of $f(x \cdot y)$ is

$$(x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i}$$

and the $(1,1)$-entry of $f(x) \cdot f(y)$ is

$$(x_0 + x_1\mathbf{i})(y_0 + y_1\mathbf{i}) + (x_2 + x_3\mathbf{i})(-y_2 + y_3\mathbf{i})$$

and they equal to each other. Similarly, it can be easily seen that $(1,2)$-entry of $f(x \cdot y)$ equals to the $(1,2)$-entry of $f(x) \cdot f(y)$. Therefore, it is derived that $f(x \cdot y) = f(x) \cdot f(y)$. In other words, $\phi$ is a multiplication preserving map.

For $1 \in \mathbb{H}$, it is easy to check that

$$f(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

1 is the identity in $\mathbb{H}$ and $I$ is the identity matrix in $M_2(\mathbb{C})$. Thus the map $\phi$ is a unit preserving map.

Second, we will show that the map $\phi$ is injective. Showing that $\phi$ is injective is equivalent to show that if $\phi(x) = \phi(y)$, then $x = y$. Suppose that $\phi(x) = \phi(y)$. Then

$$\begin{bmatrix} x_0 + x_1\mathbf{i} & x_2 + x_3\mathbf{i} \\ -x_2 + x_3\mathbf{i} & x_0 - x_1\mathbf{i} \end{bmatrix} = \begin{bmatrix} y_0 + y_1\mathbf{i} & y_2 + y_3\mathbf{i} \\ -y_2 + y_3\mathbf{i} & y_0 - y_1\mathbf{i} \end{bmatrix}.$$

This immediately implies that $x_0 = y_0$, $x_1 = y_1$, $x_2 = y_2$ and $x_3 = y_3$ and hence $x = y$. Thus $\phi$ is an injective map.

Given a quaternion $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$, let $z_1 = x_0 + x_1\mathbf{i}$ and $z_2 = x_2 + x_3\mathbf{i}$. Then $\phi(x)$ is written by

$$\phi(x) = \begin{bmatrix} z_1 & z_2 \\ -\bar{z}_2 & \bar{z}_1 \end{bmatrix}.$$

A direct computation gives

$$\det \phi(x) = |z_1|^2 + |z_2|^2 = x_0^2 + x_1^2 + x_2^2 + x_3^2 = |x|^2.$$

Next to find the eigenvalues of $\phi(x)$, we compute the characteristic polynomial of $\phi(x)$ as follows.

$$\begin{aligned} \det(\lambda I - \phi(x)) &= \begin{bmatrix} \lambda - z_1 & -z_2 \\ \bar{z}_2 & \lambda - \bar{z}_1 \end{bmatrix} \\ &= \lambda^2 - 2\lambda \mathrm{Re}(z_1) + |z_1|^2 + |z_2|^2 \\ &= \lambda^2 - 2\lambda \mathrm{Re}(x) + |x|^2. \end{aligned}$$

Here note that $\mathrm{Re}(z_1) = \mathrm{Re}(x) = x_0$. By a well-known quadratic formula,

$$\mathrm{Re}(x) \pm \sqrt{\mathrm{Re}(x)^2 - |x|^2} = \mathrm{Re}(x) \pm |\mathrm{Im}(x)|\mathbf{i}$$

are eigenvalues of $\phi(x)$. We complete the proof. □

**Lemma 2.12.** *For any quaternions $x, y \in \mathbb{H}$,*

$$\mathrm{Re}(xy) = \mathrm{Re}(yx).$$

*Proof.* Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ and $y = y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}$. Then

$$\begin{aligned} x \cdot y &= (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \cdot (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \\ &= (x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3) + (x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2)\mathbf{i} \\ &\quad + (x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1)\mathbf{j} + (x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0)\mathbf{k} \end{aligned}$$

and

$$\begin{aligned} y \cdot x &= (y_0 + y_1\mathbf{i} + y_2\mathbf{j} + y_3\mathbf{k}) \cdot (x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) \\ &= (y_0x_0 - y_1x_1 - y_2x_2 - y_3x_3) + (y_0x_1 + y_1x_0 + y_2x_3 - y_3x_2)\mathbf{i} \\ &\quad + (y_0x_2 - y_1x_3 + y_2x_0 + y_3x_1)\mathbf{j} + (y_0x_3 + y_1x_2 - y_2x_1 + y_3x_0)\mathbf{k}. \end{aligned}$$

Therefore $\mathrm{Re}(xy) = \mathrm{Re}(yx) = x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3$. □

**Definition 2.13** (Similarity)**.** Two quaternions $x$ and $y$ are said to be *similar* if there exists a nonzero quaternion $u$ such that $u^{-1}xu = y$. This is written as $x \sim y$.

Obviously, $x$ and $y$ are similar if and only if there is a unit quaternion $v$ such that $v^{-1}xv = y$, and two similar quaternions have the same norm. It is routine to check that $\sim$ is an equivalence relation on the quaternions. We denote by $[x]$ the equivalence class containing $x$.

**Proposition 2.14** ([8, 9]). *Let $x$ and $y$ be quaternions. Then $x$ is similar to $y$ if and only if* $\mathrm{Re}(x) = \mathrm{Re}(y)$ *and* $|\mathrm{Im}(x)| = |\mathrm{Im}(y)|$.

*Proof.* Suppose that $x$ and $y$ are similar quaternions. Then there is a non-zero quaternion $q$ such that $x = qyq^{-1}$. Then by Lemma 2.12,

$$\mathrm{Re}(x) = \mathrm{Re}(qyq^{-1}) = \mathrm{Re}(q^{-1}qy) = \mathrm{Re}(y).$$

Furthermore, from the following equation

$$\mathrm{Re}(x)^2 + |\mathrm{Im}(x)|^2 = |x|^2 = |qyq^{-1}|^2 = |y|^2 = \mathrm{Re}(y)^2 + |\mathrm{Im}(y)|^2,$$

it is derived that $|\mathrm{Im}(x)| = |\mathrm{Im}(y)|$.

For the converse, suppose that $\mathrm{Re}(x) = \mathrm{Re}(y)$ and $|\mathrm{Im}(x)| = |\mathrm{Im}(y)|$. Let

$$p = \sqrt{x_1^2 + x_2^2 + x_3^2} + x_1 - x_3\mathbf{j} + x_2\mathbf{k} \text{ and } q = \sqrt{y_1^2 + y_2^2 + y_3^2} + y_1 - y_3\mathbf{j} + y_2\mathbf{k}.$$

Then it is not difficult to see that

$$pxp^{-1} = \mathrm{Re}(x) + |\mathrm{Im}(x)|\mathbf{i} = \mathrm{Re}(y) + |\mathrm{Im}(y)|\mathbf{i} = qyq^{-1}$$

which implies that $x$ and $y$ are similar. $\square$

By the proof of Proposition 2.14, we immediately get the following corollary.

**Corollary 2.15.** *Every quaternionic number is similar to a complex number.*

# Chapter 3. Algebra of Quaternions

To study the algebra of quaternions, we start with a simple linear equation over quaternions. Let $a, b$ and $c$ be quaternions. Consider the equation $xa - bx = c$ over $\mathbb{H}$. We will consider two cases:

**Case 1** ($a$ is similar to $b$)**.** There is a quaternion $u$ such that $b = u^{-1}au$ and thus the equation $xa - bx = c$ is written by $xa - u^{-1}aux = c$. Left-multiplying both sides by $u$, we get $uxa - aux = uc$. Replacing $ux$ and $uc$ by $x$ and $c$ respectively, the equation $xa - bx = c$ is converted into an equation of the form $xa - ax = c$. Let $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $c = 2c_0 + 2c_1\mathbf{i} + 2c_2\mathbf{j} + 2c_3\mathbf{k}$ and $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. Then by a direct computation,

$$\begin{aligned} x \cdot a - a \cdot x = [&(x_0a_0 - x_1a_1 - x_2a_2 - x_3a_3) + (x_0a_1 + x_1a_0 + x_2a_3 - x_3a_2)\mathbf{i} \\ &+ (x_0a_2 - x_1a_3 + x_2a_0 + x_3a_1)\mathbf{j} + (x_0a_3 + x_1a_2 - x_2a_1 + x_3a_0)\mathbf{k}] \\ &- [(a_0x_0 - a_1x_1 - a_2x_2 - a_3x_3) + (a_0x_1 + a_1x_0 + a_2x_3 - a_3x_2)\mathbf{i} \\ &+ (a_0x_2 - a_1x_3 + a_2x_0 + a_3x_1)\mathbf{j} + (a_0x_3 + a_1x_2 - a_2x_1 + a_3x_0)\mathbf{k}] \\ = &\ 2(x_0a_1 + x_1a_0)\mathbf{i} + 2(x_0a_2 + x_2a_0)\mathbf{j} + 2(x_0a_3 + x_3a_0)\mathbf{k} \\ = &\ 2c_0 + 2c_1\mathbf{i} + 2c_2\mathbf{j} + 2c_3\mathbf{k} = c. \end{aligned}$$

Hence the equation $xa - ax = c$ is equivalent to the following system of equations.

$$\begin{aligned} 0 &= 2c_0 \\ 2(a_1x_0 + a_0x_1) &= 2c_1 \\ 2(a_2x_0 + a_0x_2) &= 2c_2 \\ 2(a_3x_0 + a_0x_3) &= 2c_3 \end{aligned}$$

which is written by

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ a_1 & a_0 & 0 & 0 \\ a_2 & 0 & a_0 & 0 \\ a_3 & 0 & 0 & a_0 \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

The determinant of the matrix above is 0. If $c_0 \neq 0$, this equation has no solution. If $c_0 = 0$ and $a_0 \neq 0$, the equation has infinite solutions. If $c_0 = 0$ and $a_0 = 0$, then

$$a_1x_0 = c_1, \quad a_2x_0 = c_2, \quad a_3x_0 = c_3.$$

If $c_0 = 0$, $a_0 = 0$ and $a_1 : a_2 : a_3 = c_1 : c_2 : c_3$, the equation has infinite solutions. If $c_0 = 0$, $a_0 = 0$ and $a_1 : a_2 : a_3 \neq c_1 : c_2 : c_3$, there is no solution to the equation.

**Case 2** ($a$ is not similar to $b$)**.** Let $a = a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $b = b_0 + b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, $c = 2c_0 + 2c_1\mathbf{i} + 2c_2\mathbf{j} + 2c_3\mathbf{k}$ and $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$. Then $xa - bx = c$ is equivalent to the following system of linear equations.

$$\begin{bmatrix} a_0 - b_0 & -a_1 + b_1 & -a_2 + b_2 & -a_3 + b_3 \\ a_1 - b_1 & a_0 - b_0 & a_3 + b_3 & -a_2 - b_2 \\ a_2 - b_2 & -a_3 - b_3 & a_0 - b_0 & a_1 + b_1 \\ a_3 - b_3 & a_2 + b_2 & -a_1 - b_1 & a_0 - b_0 \end{bmatrix} \cdot \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2c_0 \\ 2c_1 \\ 2c_2 \\ 2c_3 \end{bmatrix}. \tag{3.1}$$

Denote by $B$ the $4 \times 4$ matrix in (3.1). Then by a straightforward computation,

$$\begin{aligned} \det B = (|\mathrm{Im}(a)|^2 - |\mathrm{Im}(b)|^2)^2 + (a_0 - b_0)^2[(a_0 - b_0)^2 + (a_1 - b_1)^2 + (a_1 + b_1)^2 \\ + (a_2 - b_2)^2 + (a_2 + b_2)^2 + (a_3 - b_3)^2 + (a_3 + b_3)^2]. \end{aligned}$$

If $a$ is not similar to $b$, either $\mathrm{Re}(a) \neq \mathrm{Re}(b)$ or $|\mathrm{Im}(a)| \neq |\mathrm{Im}(b)|$, which implies that $a_0 - b_0 \neq 0$ or $|\mathrm{Im}(a)|^2 \neq |\mathrm{Im}(b)|^2$. Then $\det B \neq 0$ and thus there is a unique solution to the equation $xa - bx = c$.

We recover the following classical theorem for quaternion linear equation.

**Theorem 3.1** (Johnson [3])**.** *If $a, b, c \in \mathbb{H}$, and $a$ and $b$ are not similar, then $xa - bx = c$ has a unique solution.*

We now study quaternion quadratic equations. As well known, any complex quadratic equation has at most two solutions. However this does not work for quaternion quadratic equations as follows.

**Lemma 3.2.** *There are infinitely many solutions in $\mathbb{H}$ to $x^2 + 1 = 0$.*

*Proof.* Let $x = x_0 + x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}$ be a solution of the equation $x^2 + 1 = 0$. Then

$$x^2 = x_0^2 - x_1^2 - x_2^2 - x_3^2 + 2x_0(x_1\mathbf{i} + x_2\mathbf{j} + x_3\mathbf{k}) = -1$$

which is equivalent to

$$(\mathrm{Re}x)^2 - |\mathrm{Im}x|^2 = -1 \text{ and } \mathrm{Re}x \cdot \mathrm{Im}x = 0.$$

If $\mathrm{Re}x = 0$, then $|\mathrm{Im}x|^2 = 1$ and thus $|\mathrm{Im}x| = 1$. If $\mathrm{Im}x = 0$, then $(\mathrm{Re}x)^2 = -1$. There are no solutions in this case. Therefore, all quaternions $x$ with $|\mathrm{Im}x| = 1$ are solutions of $x^2 + 1 = 0$. Hence there are infinitely many solutions to $x^2 + 1 = 0$. $\square$

More generally, Eilenberg and Niven proved a result for quaternion polynomial equations as follows.

**Theorem 3.3** (Eilenberg and Niven [2]). *Let*

$$f(x) = a_0 x a_1 x \cdots x a_n + \phi(x),$$

*where* $a_0, a_1, \cdots, a_n$ *are nonzero quaternions* ($a_i \neq 0$ *for* $i = 0, \ldots, n$)*,* $x$ *is a quaternion indeterminant, and* $\phi(x)$ *is a sum of a finite number of similar monomials* $b_0 x b_1 x \cdots x b_k$*,* $k < n$*. Then* $f(x) = 0$ *has at least one quaternion solution.*

# Chapter 4. Quaternionic vector space

We recall definitions concerning modules and some theories on the set of quaternion matrices.

**Definition 4.1** (Right or left modules over a ring)**.** Suppose that $(R, +_R, \cdot_R)$ is a ring and $1_R$ is its multiplicative identity. A *left $R$-module $M$* consists of an abelian group $(M, +_M)$ and an operation $\cdot : R \times M \to M$ such that for all $r, s$ in $R$ and $x, y$ in $M$, we have:

$$\begin{aligned} r \cdot (x +_M y) &= r \cdot x +_M r \cdot y, \\ (r +_R s) \cdot x &= r \cdot x +_M s \cdot x, \\ (r \cdot_R s) \cdot x &= r \cdot (s \cdot x), \\ 1_R \cdot x &= x. \end{aligned}$$

The operation of the ring on $M$ is called scalar multiplication, and is usually written by juxtaposition, i.e. as $rx$ for $r$ in $R$ and $x$ in $M$, though here it is denoted as $r \cdot x$ to distinguish it from the ring multiplication operation, denoted here by juxtaposition. A *right $R$-module $M$* consists of an abelian group $(M, +_M)$ and an operation $\cdot : M \times R \to M$ such that for all $r, s$ in $R$ and $x, y$ in $M$, we have:

$$\begin{aligned} (x +_M y) \cdot r &= x \cdot r +_M y \cdot r, \\ x \cdot (r +_R s) &= x \cdot r +_M x \cdot s, \\ x \cdot (r \cdot_R s) &= (x \cdot r) \cdot s, \\ x \cdot 1_R &= x. \end{aligned}$$

Let $\mathbb{H}^n$ be the set of $n$-tuples of quaternions. Define $+$ on $\mathbb{H}^n$ and an operation $\cdot : \mathbb{H} \times \mathbb{H}^n \to \mathbb{H}^n$ by

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

$$(x_1, \ldots, x_n) \cdot r = (x_1 r, \ldots, x_n r) \text{ or } r \cdot (x_1, \ldots, x_n) = (r x_1, \ldots, r x_n).$$

**Lemma 4.2.** *The addition $+$ on $\mathbb{H}^n$ and the operation $\cdot : \mathbb{H} \times \mathbb{H}^n \to \mathbb{H}^n$ turn $\mathbb{H}^n$ into a right or left $\mathbb{H}$-module.*

*Proof.* For $r, s \in \mathbb{H}$ and $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n) \in \mathbb{H}^n$,

$$\begin{aligned}
r \cdot (x + y) &= r \cdot [(x_1, \ldots, x_n) + (y_1, \ldots, y_n)] \\
&= r \cdot (x_1 + y_1, \ldots, x_n + y_n) \\
&= (r \cdot x_1, \ldots, r \cdot x_n) + (r \cdot y_1, \ldots, r \cdot y_n) \\
&= r \cdot x + r \cdot y, \\
(r + s) \cdot x &= (r + s) \cdot (x_1, \ldots, x_n) \\
&= r \cdot (x_1, \ldots, x_n) + s \cdot (x_1, \ldots, x_n) \\
&= r \cdot x + s \cdot x, \\
(r \cdot s) \cdot x &= (r \cdot s) \cdot (x_1, \ldots, x_n) \\
&= r \cdot (s \cdot (x_1, \ldots, x_n)) \\
&= r \cdot (s \cdot x), \\
1 \cdot x &= 1 \cdot (x_1, \ldots, x_n) = x.
\end{aligned}$$

This leads to a conclusion that $\mathbb{H}^n$ is a left $\mathbb{H}$-module.

$$\begin{aligned}
(x + y) \cdot r &= [(x_1, \ldots, x_n) + (y_1, \ldots, y_n)] \cdot r \\
&= (x_1 + y_1, \ldots, x_n + y_n) \cdot r \\
&= (x_1 \cdot r, \ldots, x_n \cdot r) + (y_1 \cdot r, \ldots, y_n \cdot r) \\
&= x \cdot r + y \cdot r, \\
x \cdot (r + s) &= (x_1, \ldots, x_n) \cdot (r + s) \\
&= (x_1, \ldots, x_n) \cdot r + (x_1, \ldots, x_n) \cdot s \\
&= x \cdot r + x \cdot s, \\
x \cdot (r \cdot s) &= (x_1, \ldots, x_n) \cdot (r \cdot s) \\
&= ((x_1, \ldots, x_n) \cdot r) \cdot s \\
&= (x \cdot r) \cdot s, \\
x \cdot 1 &= (x_1, \ldots, x_n) \cdot 1 = x.
\end{aligned}$$

This implies that $\mathbb{H}^n$ is a right $\mathbb{H}$-module. □

**Definition 4.3** (Basis of right(left) $\mathbb{H}$-modules)**.** A subset $X$ of an $R$-module $A$ is said to be *linearly independent* provided that for distinct $x_1, \ldots, x_n \in X$ and $r_i \in R$.

$$r_1 x_1 + r_2 x_2 + \cdots + r_n x_n = 0 \Rightarrow r_i = 0 \ \text{ for every } i.$$

A set that is not linearly independent is said to be *linearly dependent.* If $A$ is generated as an $R$-module by a set $Y$, then we say that $Y$ spans $A$. If $R$ has an identity and $A$ is unitary, then $Y$ spans $A$ if and only if every element of $A$ may be written as a linear combination of elements of $Y$. A linearly independent subset of $A$ that spans $A$ is called a basis of $A$. Observe that the empty set is (vacuously) linearly independent and is a basis of the zero module.

**Lemma 4.4.** *Let $V$ be a right or left $\mathbb{H}$-module. All bases of $V$ have the same number of elements.*

*Proof.* Let $V$ be a right $\mathbb{H}$-module and, $\{u_1, \dots, u_m\}$ and $\{v_1, \dots, v_k\}$ be bases of $V$. Let $u_i = v_1 a_{1i} + v_2 a_{2i} + \dots + v_k a_{ki}$ for $i = 1, \dots, m$. Suppose that $u_1 x_1 + \dots + u_m x_m = 0$. Then

$$u_1 x_1 + \dots + u_m x_m = \sum_{j=1}^{m} \left( \sum_{i=1}^{k} v_i a_{ij} \right) x_j = \sum_{i=1}^{k} v_i \left( \sum_{j=1}^{m} a_{ij} x_j \right) = 0.$$

Since $\{v_1, \dots, v_k\}$ is linearly independent, $\sum_{j=1}^{m} a_{ij} x_j = 0$ for any $i = 1, \dots, k$. If $k < m$, the number of variables is greater than the number of equations and hence there are infinite solutions. However, it contradicts the fact that If $u_1 x_1 + \dots + u_m x_m = 0$, then $x_1 = \dots = x_m = 0$, which follows from the linearly independence of $\{u_1, \dots, u_m\}$. Thus $k \geq m$. Switching $\{u_1, \dots, u_m\}$ and $\{v_1, \dots, v_k\}$, it follows that $m \geq k$. Therefore we get $m = k$. This proof works for left $\mathbb{H}$-modules in the same way. $\square$

Due to Lemma 4.4, we can define the dimension of a right or left $\mathbb{H}$-module $V$, denoted by $\dim_{\mathbb{H}} V$, by the number of elements of a basis of $V$.

**Definition 4.5** (Right or left $\mathbb{H}$-module homomorphism)**.** Let $(M, +_M, \cdot_M)$ and $(N, +_N, \cdot_N)$ be left $R$-modules with the scalar multiplication $\cdot$ on them. A map $f : M \to N$ is said to be a *homomorphism of left $R$-modules* if for any $m, n$ in $M$ and $r, s$ in $R$,

$$f(r \cdot m +_M s \cdot n) = r \cdot f(m) +_N s \cdot f(n).$$

In other words, $f$ preserves the structure of left $R$-modules. Another name for a homomorphism of left $R$-modules is a left $R$-linear map.

If $(M, +_M, \cdot_M)$ and $(N, +_N, \cdot_N)$ are right $R$-modules, a map $f : M \to N$ is said to be a *homomorphism of right $R$-modules* if for any $m, n$ in $M$ and $r, s$ in $R$,

$$f(m \cdot r +_M n \cdot s) = f(m) \cdot r +_N f(n) \cdot s.$$

If a homomorphism $f : M \to N$ of $R$-modules is bijective, then $f$ is called an *isomorphism*. We say that $M$ is *isomorphic* to $N$ if there is an isomorphism between $M$ and $N$. The kernel of a module homomorphism $f : M \to N$ is the submodule of $M$ consisting of all elements that are sent to zero by $f$ and the image of $f$ is also the submodule of $N$.

Right or left $\mathbb{H}$-module homomorphisms between finite-dimensional right or left $\mathbb{H}$-modules can be described by quaternion matrices as follows. Let $f : \mathbb{H}^n \to \mathbb{H}^m$ be a map such that for each $i = 1, \ldots, n$,

$$f(e_j) = (a_{1j}, a_{2j}, \ldots, a_{mj}) \in \mathbb{H}^m.$$

Let $A$ be an $m \times n$ quaternion matrix whose $(i, j)$-entry is $a_{ij}$ and $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathbb{H}^n$. If $f$ is a right $\mathbb{H}$-module homomorphism, then

$$\begin{aligned} f(\mathbf{x}) = f(x_1, \ldots, x_n) = f\left(\sum_{i=1}^{n} e_i x_i\right) &= \sum_{i=1}^{n} f(e_i) x_i \\ &= \sum_{i=1}^{n} \left(\sum_{k=1}^{m} e_k a_{ki}\right) x_i \\ &= \sum_{k=1}^{m} e_k \left(\sum_{i=1}^{n} a_{ki} x_i\right) = A \cdot \mathbf{x}. \end{aligned}$$

On the other hand, if $f$ is a left $\mathbb{H}$-module homomorphism,

$$\begin{aligned} f(\mathbf{x}) = f(x_1, \ldots, x_n) = f\left(\sum_{i=1}^{n} x_i e_i\right) &= \sum_{i=1}^{n} x_i f(e_i) \\ &= \sum_{i=1}^{n} x_i \left(\sum_{k=1}^{m} a_{ki} e_k\right) \\ &= \sum_{k=1}^{m} \left(\sum_{i=1}^{n} x_i a_{ki}\right) e_k = \mathbf{x}^T \cdot A^T. \end{aligned}$$

**Lemma 4.6.** *Let $(a_1, \ldots, a_n) \in \mathbb{H}^n$ be a non-zero quaternion vector. Then the set of all solutions $(x_1, \ldots, x_n) \in \mathbb{H}^n$ of $a_1 x_1 + \cdots + a_n x_n = 0$ (resp. $x_1 a_1 + \cdots + x_n a_n = 0$) is a right (resp. left) $\mathbb{H}$-module of dimension $n - 1$.*

*Proof.* Let $M = \{(x_1, \ldots, x_n) \in \mathbb{H}^n \mid a_1 x_1 + \cdots + a_n x_n = 0\}$. To verify that $M$ is a right $\mathbb{H}$-module of dimension $n - 1$, we first prove that the right multiplication of $\mathbb{H}$ on $M$ is well defined. For any $q \in \mathbb{H}$ and any $x = (x_1, \ldots, x_n) \in M$,

$$a_1(x_1 q) + \cdots + a_n(x_n q) = (a_1 x_1 + \cdots + a_n x_n) q = 0 \cdot q = 0$$

which implies $xq \in M$. In other words, the right multiplication of $\mathbb{H}$ on $M$ is well defined. Thus $M$ is a right $\mathbb{H}$-module.

Let $a_1x_1 + \cdots + a_nx_n = 0$. We may assume that $a_n \neq 0$. Then

$$x_n = -a_n^{-1}a_1x_1 - \cdots - a_n^{-1}a_{n-1}x_{n-1}.$$

Any $x \in V$ is written by

$$\begin{aligned} x &= (x_1, x_2, \ldots, x_{n-1}, -a_n^{-1}a_1x_1 - \cdots - a_n^{-1}a_{n-1}x_{n-1}) \\ &= (e_1 - a_n^{-1}a_1e_n)x_1 + \cdots + (e_{n-1} - a_n^{-1}a_{n-1}e_n)x_{n-1} \end{aligned}$$

which implies that $\{e_1 - a_n^{-1}a_1e_n, \ldots, e_{n-1} - a_n^{-1}a_{n-1}e_n\}$ is a basis of $M$. Thus the dimension of a right $\mathbb{H}$-module $M$ is $n-1$. □

Let $M_{m\times n}(\mathbb{H})$ denote the set of $m \times n$ quaternion matrices. For $A \in M_{m\times n}(\mathbb{H})$, define $A^* = \bar{A}^T$.

**Theorem 4.7.** *Let $A \in M_{m\times n}(\mathbb{H})$ and $B \in M_{n\times l}(\mathbb{H})$. Then*
(1) $\overline{A}^T = \overline{A^T}$.
(2) $(AB)^* = B^*A^*$.
(3) $(A^*)^{-1} = (A^{-1})^*$ *if $A$ is invertible.*
(4) $\overline{AB} \neq \bar{A}\bar{B}$ *in general.*
(5) $(AB)^T \neq B^TA^T$ *in general.*
(6) $\bar{A}^{-1} \neq \overline{A^{-1}}$ *in general.*
(7) $(A^T)^{-1} \neq (A^{-1})^T$ *in general.*

*Proof.* By a straightforward computation, (1), (2) and (3) can be easily seen. For (4)-(7), we give counterexamples. For (4) and (5), let $a = 1 + \mathbf{i}$ and $b = 2 - \mathbf{j}$. Then

$$\overline{ab} = \overline{(1+\mathbf{i})\cdot(2-\mathbf{j})} = \overline{2 + 2\mathbf{i} - \mathbf{j} - \mathbf{k}} = 2 - 2\mathbf{i} + \mathbf{j} + \mathbf{k}.$$

and

$$\bar{a}\bar{b} = (\overline{1+\mathbf{i}})(\overline{2-\mathbf{j}}) = (1-\mathbf{i})\cdot(2+\mathbf{j}) = 2 - 2\mathbf{i} + \mathbf{j} - \mathbf{k}.$$

Furthermore,

$$(ab)^T = ((1+\mathbf{i})\cdot(2-\mathbf{j}))^T = (2 + 2\mathbf{i} - \mathbf{j} - \mathbf{k})^T = 2 + 2\mathbf{i} - \mathbf{j} - \mathbf{k}$$

and

$$b^Ta^T = ((2-\mathbf{j})\cdot(1+\mathbf{i})) = 2 + 2\mathbf{i} - \mathbf{j} + \mathbf{k}.$$

Thus $\overline{ab} \neq \bar{a}\bar{b}$ and $(ab)^T \neq b^Ta^T$.

For (6) and (7), let $A = \begin{bmatrix} \mathbf{i} & -\mathbf{j} \\ 0 & \mathbf{k} \end{bmatrix}$. Then

$$\overline{A}^{-1} = \begin{bmatrix} \mathbf{i} & 1 \\ 0 & \mathbf{k} \end{bmatrix} \neq \begin{bmatrix} \mathbf{i} & -1 \\ 0 & \mathbf{k} \end{bmatrix} = \overline{A^{-1}}.$$

and

$$(A^T)^{-1} = \begin{bmatrix} -\mathbf{i} & 0 \\ 1 & -\mathbf{k} \end{bmatrix} \neq \begin{bmatrix} -\mathbf{i} & 0 \\ -1 & -\mathbf{k} \end{bmatrix} = (A^{-1})^T.$$

Thus $\bar{A}^{-1} \neq \overline{A^{-1}}$ and $(A^T)^{-1} \neq (A^{-1})^T$. □

# Chapter 5. Jordan canonical form

In this chapter, we reprove the well-known Jordan canonical forms of $2 \times 2$ quaternion matrices. Brenner [8] proved the following theorem. We here give a self-contained proof.

**Theorem 5.1.** *Let $A$ be a $2 \times 2$ quaternion matrix. Then there exists an invertible quaternion matrix $Q$ such that $Q^{-1}AQ$ is of the form*

$$\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} \ \textit{or} \ \begin{bmatrix} a & 1 \\ 0 & d \end{bmatrix}.$$

*Proof.* We first prove that there is a right eigenvector of $A$. In other words, there is a non-zero vector $v \in \mathbb{H}^2$ such that $Av = v\lambda$ for some $\lambda \in \mathbb{H}$. Write

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } v = \begin{bmatrix} x \\ y \end{bmatrix}.$$

If $v$ is a right eigenvector of $A$, then for some $\lambda \in \mathbb{H}$, the following equations hold.

$$ax + by = x\lambda, \tag{5.1}$$

$$cx + dy = y\lambda. \tag{5.2}$$

If $c = 0$, then $(x, y, \lambda) = (1, 0, a)$ is a solution to (5.1) and (5.2). Otherwise we have $x = -c^{-1}dy + c^{-1}y\lambda$ from (5.2). Then (5.1) is written by

$$(b - ac^{-1}d)y + (ac^{-1} + c^{-1}d)y\lambda - c^{-1}y\lambda^2 = 0. \tag{5.3}$$

Let $y = 1$. Then by Theorem 3.3, there is a solution $\lambda_0$ to (5.3). Hence $(x_0, y_0, \lambda_0)$ is a solution to (5.1) and (5.2) where $x_0 = -c^{-1}d + c^{-1}\lambda_0$. In other words, there is a non-zero vector $v_0 = (x_0, y_0) \in \mathbb{H}^2$ such that $v_0$ is a right eigenvector of $A$.

If $v_0$ is a right eigenvector with eigenvalue $\lambda_0$, then it can be easily seen that $v_0 q$ is also right eigenvector with eigenvalue $q^{-1}\lambda_0 q$ since

$$A(v_0 q) = (Av_0)q = (v_0\lambda)q = (v_0 q)(q^{-1}\lambda_0 q).$$

Hence we can assume that $\|v_0\| = 1$. Choose a non-zero vector $v_1 \in \mathbb{H}^2$ such that $\langle v_0, v_1 \rangle = 0$ and $\|v_1\| = 1$. Let $v_1 = (x_1, y_1)$. Set $Q = \begin{bmatrix} x_0 & x_1 \\ y_0 & y_1 \end{bmatrix}$, $e_1 = (1, 0)$. Then

$$(Q^{-1}AQ)e_1 = Q^{-1}Av_0 = Q^{-1}(v_0\lambda_0) = (Q^{-1}v_0)\lambda_0 = e_1\lambda_0,$$

which implies that $Q^{-1}AQ$ is of the form $\begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$. If $b = 0$, then $Q^{-1}AQ$ is of the form $\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$. From now on we suppose that $b \neq 0$. For any $x \in \mathbb{H}$, we have

$$\begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & ax + b - xd \\ 0 & d \end{bmatrix}.$$

If there is a solution $x_0$ to $ax + b - xd = 0$, then

$$\begin{bmatrix} 1 & x_0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} 1 & x_0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}.$$

Otherwise,

$$\begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \begin{bmatrix} b & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} b^{-1}ab & 1 \\ 0 & d \end{bmatrix}$$

which is of the form $\begin{bmatrix} a & 1 \\ 0 & d \end{bmatrix}$. We complete the proof. $\square$

# Chapter 6. Quaternionic Lorentzian space

Quaternionic Lorentzian $(n+1)$-space is the inner product space consisting of the right $\mathbb{H}$-module $\mathbb{H}^{n+1}$ together with the $(n+1)$-dimensional Lorentzian inner product. The Lorentzian inner product is defined by

$$\langle x, y\rangle = \bar{x}_1 y_{n+1} + \bar{x}_{n+1} y_1 + \bar{x}_2 y_2 + \cdots + \bar{x}_n y_n.$$

where $x = (x_1, \ldots, x_{n+1})$ and $y = (y_1, \ldots, y_{n+1})$. We denote by $\mathbb{H}^{n,1}$ the quaternionic Lorentzian $(n+1)$-space.

**Definition 6.1.** A vector $x$ is said to be *spacelike* (resp. *timelike* and *lightlike*) if $\langle x, x\rangle > 0$ (resp. $\langle x, x\rangle < 0$ and $\langle x, x\rangle = 0$).

**Example 6.2** (Spacelike, timelike and lightlike vectors)**.** We give examples of spacelike, timelike and lightlike vectors. First, $(0, 1, 0)$ is a spacelike vector with positive norm. An example of timelike vector is $(\mathbf{i}\sqrt{2}/2, 0, -\mathbf{i}\sqrt{2}/2)$ whose norm is $-1 < 0$. Lastly, $(1, 0, 0)$ is lightlike vector whose norm is 0.

Let $\mathbf{0}$ denote the origin vector $(0, 0, \ldots, 0)$ in $\mathbb{H}^{n+1}$.

**Lemma 6.3.** *Let $x$ and $y$ be vectors in $\mathbb{H}^{n,1}$. Then $\langle x, y\rangle = 0$ for all $y \in \mathbb{H}^{n,1}$ if and only if $x = \mathbf{0}$.*

*Proof.* Let $x = (x_1, \ldots, x_{n+1})$. Suppose that $\langle x, y\rangle = 0$ for any $y \in \mathbb{H}^{n,1}$. Let $\{e_1, \ldots, e_{n+1}\}$ be the canonical basis of $\mathbb{H}^{n+1}$. If $y = e_1$, then $\langle x, e_1\rangle = \bar{x}_{n+1} = 0$ and thus $x_{n+1} = 0$. If $y = e_2$, then $\langle x, e_2\rangle = \bar{x}_2 = 0$ and thus $x_2 = 0$. In this way, we show that $x_i = 0$ for all $i = 1, \ldots, n+1$. This implies that $x = \mathbf{0}$. The converse is obvious. Thus we complete the proof. □

**Lemma 6.4.** *For $x, y \in \mathbb{H}^{n,1}$, $\langle x, y\rangle = \overline{\langle y, x\rangle}$.*

*Proof.* Let $x = (x_1, \ldots, x_{n+1})$ and $y = (y_1, \ldots, y_{n+1})$. By definition,

$$\langle x, y\rangle = \bar{x}_1 y_{n+1} + \bar{x}_{n+1} y_1 + \sum_{i=2}^{n} \bar{x}_i y_i \text{ and } \langle y, x\rangle = \bar{y}_1 x_{n+1} + \bar{y}_{n+1} x_1 + \sum_{i=2}^{n} \bar{y}_i x_i.$$

Thus

$$\overline{\langle y, x\rangle} = \overline{\bar{y}_1 x_{n+1}} + \overline{\bar{y}_{n+1} x_1} + \sum_{i=2}^{n} \overline{\bar{y}_i x_i} = \bar{x}_{n+1} y_1 + \bar{x}_1 y_{n+1} + \sum_{i=2}^{n} \bar{x}_i y_i = \langle x, y\rangle,$$

which implies the Lemma. □

If $\langle x,y\rangle \in \mathbb{R}$, then $\overline{\langle x,y\rangle} = \langle x,y\rangle$ and thus $\langle y,x\rangle = \langle x,y\rangle$ by Lemma 6.4. We have the following corollary.

**Corollary 6.5.** *If $\langle x,y\rangle \in \mathbb{R}$, then $\langle y,x\rangle = \langle x,y\rangle$.*

Let $v$ be a nonzero vector in $\mathbb{H}^{n,1}$. Denote by $v^\perp$ the set of all vectors $x \in \mathbb{H}^{n,1}$ with $\langle v,x\rangle = 0$ i.e.,

$$v^\perp = \{(x_1,\dots,x_n) \in \mathbb{H}^{n+1} \mid \bar{v}_1 x_{n+1} + \bar{v}_{n+1}x_1 + \bar{v}_2 x_2 + \cdots \bar{v}_n x_n = 0\}.$$

By Lemma 4.6, $v^\perp$ is a right $\mathbb{H}$-submodule of $\mathbb{H}^{n+1}$ of dimension $n$. From now on, we will focus on the case of $n = 2$.

**Lemma 6.6.** *Let $x$ be a nonzero timelike vector in $\mathbb{H}^{2,1}$. Then $x \notin x^\perp$ and every nonzero vector of $x^\perp$ is spacelike.*

*Proof.* First of all, since $x$ is timelike, it follows that $\langle x,x\rangle < 0$. Thus obviously $x \notin x^\perp$. Suppose that $y = (y_1,y_2,y_3) \in x^\perp$. Then $yq \in x^\perp$ for all quaternions $q \in \mathbb{H}$. This follows from $\langle x,yq\rangle = \langle x,y\rangle q = 0$. Furthermore by observing

$$\langle yq,yq\rangle = \bar{q}\langle y,y\rangle q = |q|^2\langle y,y\rangle$$

in order to prove that $y$ is spacelike, it is sufficient to prove that $yq$ is spacelike for some non-zero quaternion $q$.

Let $x = (x_1,x_2,x_3)$. From the assumption that $x$ is timelike, it follows that $x_1 \neq 0$. By scaling $x$, we may assume that $x_1 = 1$. The conditions of $\langle x,x\rangle < 0$ and $\langle x,y\rangle = 0$ give

$$2\mathrm{Re}(x_3) + |x_2|^2 < 0, \tag{6.1}$$
$$y_3 + \bar{x}_2 y_2 + \bar{x}_3 y_1 = 0. \tag{6.2}$$

If $y_1 = 0$, then $\langle y,y\rangle = |y_2|^2$. If $y_1 = 0$ and $y_2 = 0$, (6.2) forces $y_3 = 0$, contrary to $y \neq (0,0,0)$. Hence if $y_1 = 0$, then $y_2 \neq 0$, which implies that $\langle y,y\rangle = |y_2|^2$ is positive. In other words, $y$ is spacelike.

From now on, we suppose that $y_1 \neq 0$. Then by scaling $y$, we can assume that $y_1 = 1$. Then (6.2) is written by $y_3 + \bar{x}_2 y_2 + \bar{x}_3 = 0$ and

$$\begin{aligned}\langle y,y\rangle &= 2\mathrm{Re}(y_3) + |y_2|^2\\ &= -2\mathrm{Re}\bar{x}_2 y_2 - 2\mathrm{Re}\bar{x}_3 + |y_2|^2\\ &> -2\mathrm{Re}\bar{x}_2 y_2 + |x_2|^2 + |y_2|^2 = |x_2 - y_2|^2 \geq 0.\end{aligned}$$

This leads to a conclusion that every non-zero vector in $x^\perp$ is spacelike. $\square$

**Lemma 6.7.** *Let $x$ be a nonzero spacelike vector in $\mathbb{H}^{2,1}$. Then $x \notin x^{\perp}$ and $x^{\perp}$ contains timelike, spacelike and lightlike vectors. Furthermore, there are linearly independent two lightlike vectors in $x^{\perp}$.*

*Proof.* First of all, since $\langle x, x\rangle > 0$, it immediately follows that $x \notin x^{\perp}$. We claim that there is a timelike vector in $x^{\perp}$. Suppose, contrary to our claim, there is no timelike vector in $x^{\perp}$. Then every vector in $x^{\perp}$ is either spacelike or lightlike. If there is a spacelike vector $y \in x^{\perp}$, then we can find an orthogonal basis $\{x, y, z\}$ where $z \in x^{\perp} \cap y^{\perp}$. By hypothesis, $\langle z, z\rangle \geq 0$. Then every vector $v \in \mathbb{H}^{2,1}$ is written by $v = xa + yb + zc$ for some $a, b, c \in \mathbb{H}$ and

$$\langle v, v\rangle = |a|^2\langle x, x\rangle + |b|^2\langle y, y\rangle + |c|^2\langle z, z\rangle > 0$$

which contradicts the existence of a lightlike vector in $\mathbb{H}^{2,1}$.

Now we can assume that every vector in $x^{\perp}$ is lightlike. Then there are linear independent vectors $y$ and $z$ in $x^{\perp}$. Since $\dim_{\mathbb{H}} x^{\perp} = 2$, it follows that $\mathrm{span}_{\mathbb{H}}\{y, z\} = x^{\perp}$. From the assumption that every vector in $x^{\perp}$ is lightlike, for all $a, b \in \mathbb{H}$,

$$0 = \langle ya + zb, ya + zb\rangle = |a|^2\langle y, y\rangle + |b|^2\langle z, z\rangle + 2\mathrm{Re}(\bar{a}\langle y, z\rangle b) = 2\mathrm{Re}(\bar{a}\langle y, z\rangle b)$$

which leads to a conclusion that $\langle y, z\rangle = 0$. Then $\{x, y, z\}$ is an orthogonal basis of $\mathbb{H}^{2,1}$. Since $x$ is spacelike and, $y$ and $z$ are lightlike, it is derived that every vector has non-negative norm. This also contradicts the existence of a timelike vector in $\mathbb{H}^{2,1}$. Therefore, the claim holds. In other words, there is a timelike vector $y \in x^{\perp}$.

Choosing a vector $z \in x^{\perp} \cap y^{\perp}$, we have an orthogonal basis $\{y, z\}$ of $x^{\perp}$. By scaling $y$ and $z$, we may assume that $\langle y, y\rangle = -1$ and $\langle z, z\rangle = 1$. Since $x^{\perp} = \mathrm{span}_{\mathbb{H}}\{y, z\}$, it can be easily seen that $y + z$ and $y - z$ are linearly independent lightlike vectors in $x^{\perp}$. Summarizing, there are a timelike vector $y$, a spacelike vector $z$ and a lightlike vector $y + z$ in $x^{\perp}$. Furthermore, there are linearly independent two lightlike vectors $y + z$ and $y - z$ in $x^{\perp}$. □

**Lemma 6.8.** *Let $x$ be a nonzero lightlike vector in $\mathbb{H}^{2,1}$. Then $x \in x^{\perp}$ and every vector of $x^{\perp}$ is either spacelike or lightlike. Furthermore, $y$ is a lightlike vector in $x^{\perp}$ if and only if $y = xq$ for some nonzero $q \in \mathbb{H}$.*

*Proof.* It follows from the assumption of $\langle x, x\rangle = 0$ that $x \in x^{\perp}$. First observe that there are no timelike vectors in $x^{\perp}$. If there is a timelike vector perpendicular to $x$, then $x$ must be spacelike, contrary to the hypothesis that $x$ is lightlike. Thus every vector of $x^{\perp}$ is either spacelike or lightlike. It remains to prove the second statement of the Lemma.

Suppose that there is a lightlike vector $y$ such that $\{x, y\}$ is linearly independent. Then $x^{\perp} = \mathrm{span}_{\mathbb{H}}\{x, y\}$. For any $a, b \in \mathbb{H}$,

$$\langle xa + yb, xa + yb\rangle = |a|^2\langle x, x\rangle + |b|^2\langle y, y\rangle + 2\mathrm{Re}(\bar{a}\langle x, y\rangle b) = 2\mathrm{Re}(\bar{a}\langle x, y\rangle b).$$

If $\langle x, y\rangle = 0$, then

$$2\mathrm{Re}(\bar{x}_1 x_3) + |x_2|^2 = 0, \tag{6.3}$$

$$2\mathrm{Re}(\bar{y}_1 y_3) + |y_2|^2 = 0, \tag{6.4}$$

$$\bar{x}_1 y_3 + \bar{x}_2 y_2 + \bar{x}_3 y_1 = 0, \tag{6.5}$$

where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$. If $x_1 = 0$, then $x_2 = 0$ by (6.3) and hence $x_3 \neq 0$. Then $y_1 = 0$ by (6.5) and $y_2 = 0$ by (6.4). This implies that $x, y \in \mathrm{span}_{\mathbb{H}}\{e_3\}$ and hence, $x$ and $y$ are linearly dependent, contrary to the hypothesis. Thus $x_1 \neq 0$ and $y_1 \neq 0$. By scaling $x$ and $y$, we may assume that $x_1 = y_1 = 1$. Then

$$2\mathrm{Re}(x_3) + |x_2|^2 = 0, \tag{6.6}$$

$$2\mathrm{Re}(y_3) + |y_2|^2 = 0, \tag{6.7}$$

$$y_3 + \bar{x}_2 y_2 + \bar{x}_3 = 0. \tag{6.8}$$

If we take the real part on both sides of (6.8), we have

$$0 = 2\mathrm{Re}(y_3) + 2\mathrm{Re}(\bar{x}_2 y_2) + 2\mathrm{Re}(\bar{x}_3) = -|y_2|^2 + 2\mathrm{Re}(\bar{x}_2 y_2) - |x_2|^2 = -|x_2 - y_2|^2.$$

Therefore $x_2 = y_2$. By (6.6) and (6.7), we have $\mathrm{Re}(x_3) = \mathrm{Re}(y_3)$. Taking the imaginary part on both sides of (6.8),

$$0 = \mathrm{Im}(y_3) + \mathrm{Im}(\bar{x}_2 y_2) + \mathrm{Im}(\bar{x}_3) = \mathrm{Im}(y_3) + \mathrm{Im}(|x_2|^2) - \mathrm{Im}(x_3) = \mathrm{Im}(y_3) - \mathrm{Im}(x_3).$$

Since $\mathrm{Re}(x_3) = \mathrm{Re}(y_3)$ and $\mathrm{Im}(x_3) = \mathrm{Im}(y_3)$, we conclude that $x_3 = y_3$. Therefore $x = y$, contrary to the assumption that $x$ and $y$ are linearly independent. Finally, there are no lightlike vectors in $x^{\perp}$ which are linearly independent with $x$. □

**Lemma 6.9.** *Let $V$ be a 2-dimensional right $\mathbb{H}$-submodule of $\mathbb{H}^3$. Then there is a nonzero vector $x \in \mathbb{H}^3$ such that $V = x^{\perp}$.*

*Proof.* Since $V$ is a 2 dimensional right $\mathbb{H}$ submodule of $\mathbb{H}^3$, there exists a basis $\{u, v\}$.

$$V = \{uq_1 + vq_2 \ | \ q_1, q_2 \in \mathbb{H}\} = span_{\mathbb{H}}\{u, v\}.$$

Find an element $x \in \mathbb{H}^{2,1}$ such that $x^{\perp} \in V$.

Since $\langle u, x\rangle = \langle v, x\rangle = 0$, we have $\langle uq_1 + vq_2, x\rangle = \overline{q_1}\langle u, x\rangle + \overline{q_2}\langle v, x\rangle = 0$ for any $q_1, q_2 \in \mathbb{H}$. Let $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3)$, then

$$\overline{u_3}x_1 + \overline{u_2}x_2 + \overline{u_1}x_3 = 0.$$

$$\overline{v_3}x_1 + \overline{v_2}x_2 + \overline{v_1}x_3 = 0.$$

We can find a soluton of these equations. Because the number of $x_i$ is larger than the number of equations. $\square$

**Corollary 6.10.** *Let $V$ be a 2-dimensional right $\mathbb{H}$-submodule of $\mathbb{H}^3$. Then*

(1) *if $V$ has linearly independent two lightlike vectors, then $V = x^{\perp}$ where $x$ is spacelike,*

(2) *if $V$ has only one lightlike vectors up to scaling by a nonzero quaternion, then $V = x^{\perp}$ where $x$ is lightlike.*

(3) *if $V$ has no lightlike vectors, then $V = x^{\perp}$ where $x$ is timelike.*

**Definition 6.11.** A basis $\{v_1, v_2, v_3\}$ of $\mathbb{H}^{2,1}$ is said to be *orthonormal* if $|\langle v_i, v_i\rangle| = 1$ for $i = 1, 2, 3$ and $\langle v_i, v_j\rangle = 0$ for all distinct $i, j$.

**Example 6.12.** An example of an orthonormal basis of $\mathbb{H}^{2,1}$ is $\left\{\frac{e_1+e_3}{\sqrt{2}}, e_2, \frac{e_1-e_3}{\sqrt{2}}\right\}$.

**Theorem 6.13.** *Let $\{v_1, v_2, v_3\}$ be an orthonormal basis of $\mathbb{H}^{2,1}$. Then it exactly consists of two spacelike vectors and one timelike vector.*

*Proof.* If everty vector $v_i$ is spacelike, for any $v = v_1a_1 + v_2a_2 + v_3a_3$,

$$\langle v, v\rangle = |a_1|^2\langle v_1, v_1\rangle + |a_2|^2\langle v_2, v_2\rangle + |a_3|^2\langle v_3, v_3\rangle \geq 0$$

which means that every vector in $\mathbb{H}^{2,1}$ is spacelike, contrary to the fact that there is a timelike vector in $\mathbb{H}^{2,1}$. If every vector $v_i$ is timelike, every vector in $\mathbb{H}^{2,1}$ is timelike in a similar way, contrary to the fact that there is a spacelike vector. Thus any orthonormal basis of $\mathbb{H}^{2,1}$ has at least one spacelike and one timelike vector.

An example of orthonormal basis of $\mathbb{H}^{2,1}$ is $\{\frac{e_1+e_3}{\sqrt{2}}, e_2, \frac{e_1-e_3}{\sqrt{2}}\}$. Here $\frac{e_1+e_3}{\sqrt{2}}$ and $e_2$ are spacelike vectors and $\frac{e_1-e_3}{\sqrt{2}}$ is a timelike vector. Let $v_1$ be a spacelike vector and $v_2, v_3$ timelike vectors. Then $\mathrm{span}_{\mathbb{H}}\{\frac{e_1+e_3}{\sqrt{2}}, e_2\} \cap \mathrm{span}_{\mathbb{H}}\{v_2, v_3\} = \{0\}$ and hence $\{\frac{e_1+e_3}{\sqrt{2}}, e_2, v_2, v_3\}$ is linearly independent. Thus the dimension of $\mathrm{span}_{\mathbb{H}}\{\frac{e_1+e_3}{\sqrt{2}}, e_2, v_2, v_3\}$ is at least 4. But this contradicts $\dim \mathbb{H}^{2,1} = 3$. Thus any orthonormal basis $\{v_1, v_2, v_3\}$ of $\mathbb{H}^{2,1}$ has two spacelike vectors and one timelike vector. $\square$

Let $f : \mathbb{H}^{2,1} \to \mathbb{H}^{2,1}$ be a right $\mathbb{H}$-module homomorphism preserving the Lorentzian inner product i.e.

$$\langle f(x), f(y)\rangle = \langle x, y\rangle.$$

Then the quaternion $3 \times 3$ matrix $A$ associated to $f$ satisfies $A^*JA = J$ where

$$J = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Let $\mathbf{Sp}(2,1)$ denote the set of all quaternion $3 \times 3$ matrices $A$ with $A^*JA = J$.

**Lemma 6.14.** *Every element of* $\mathbf{Sp}(2,1)$ *is invertible.*

*Proof.* For any $A \in \mathbf{Sp}(2,1)$, $A^*JA = J$ and hence $J \cdot A^*JA = J \cdot J = I$. Then it follows that $A^{-1} = JA^*J$. In other words every element of $\mathbf{Sp}(2,1)$ is invertible. □

**Proposition 6.15.** *The set* $\mathbf{Sp}(2,1)$ *is a group.*

*Proof.* Let $A$, $B$ and $C$ be matrices in $\mathbf{Sp}(2,1)$. The associativity of quaternions gives $A \cdot (B \cdot C) = (A \cdot B) \cdot C$. This implies that $\mathbf{Sp}(2,1)$ satisfies the associative law. Obviously, the identity matrix $I$ is in $\mathbf{Sp}(2,1)$ and thus it is the identity of $\mathbf{Sp}(2,1)$. By Lemma 6.14, each matrix $A$ in $\mathbf{Sp}(2,1)$ has an inverse $A^{-1} = JA^*J$. Moreover,

$$(JA^*J)^* \cdot J \cdot (JA^*J) = JA(JA^*J) = JAA^{-1} = J.$$

Thus $A^{-1} \in \mathbf{Sp}(2,1)$. Therefore $\mathbf{Sp}(2,1)$ is a group. □

Let $V_0$ be the set of lightlike vectors of $\mathbb{H}^{2,1}$. Let $V_+$ (resp. $V_-$) be the space of spacelike (resp. timelike) vectors whose norms are 1 (resp. $-1$).

There is a natural $\mathbf{Sp}(2,1)$-action the space of bases of $\mathbb{H}^{2,1}$ as follows: Let $A \in \mathbf{Sp}(2,1)$ and $\{v_1, v_2, v_3\}$ be an orthonormal basis of $\mathbb{H}^{2,1}$. Since $A$ preserves the Lorentzian inner product, it immediately follows that $\{Av_1, Av_2, Av_3\}$ is also an orthonormal basis of $\mathbb{H}^{2,1}$.

**Definition 6.16** (transitivity)**.** A group action $G \times X \to X$ is transitive if it possesses only a single group orbit, i.e., for every pair of elements $x$ and $y$, there is a group element $g$ such that $gx = y$. In this case, $X$ is isomorphic to the left cosets of the isotropy group, $X \sim G/G_x$. The space $X$, which has a transitive group action, is called a homogeneous space when the group is a Lie group.

If, for every two pairs of points $x_1, x_2$ and $y_1, y_2$, there is a group element $g$ such that $gx_i = y_i$, then the group action is called doubly transitive. Similarly, a group action can be triply transitive and, in general, a group action is $k$-transitive if every set $\{x_1, ..., y_k\}$ of $2k$ distinct elements has a group element $g$ such that $gx_i = y_i$.

**Theorem 6.17.** *The* $\mathbf{Sp}(2,1)$*-action on the set of orthonormal bases of* $\mathbb{H}^{2,1}$ *is transitive.*

*Proof.* Let $u_1 = (e_1 + e_3)/\sqrt{2}$, $u_2 = e_2$ and $u_3 = (e_1 - e_3)/\sqrt{2}$. Then we have shown that $\{u_1, u_2, u_3\}$ is an orthonormal basis of $\mathbb{H}^{2,1}$. To prove the theorem, it suffices to show that for a given orthonormal basis $\{v_1, v_2, v_3\} \in F_1$, there exists an element

$A \in \mathbf{Sp}(2,1)$ such that $A \cdot \{u_1, u_2, u_3\} = \{v_1, v_2, v_3\}$. Let $|v_1| = |v_2| = 1$ and $|v_3| = -1$. Define $A = \left[\frac{v_1+v_3}{\sqrt{2}}, v_2, \frac{v_1-v_3}{\sqrt{2}}\right]$. Then by a direct computation,

$$\begin{aligned}
Au_1 &= A\left(\frac{e_1+e_3}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\left(\frac{v_1+v_3}{\sqrt{2}} + \frac{v_1-v_3}{\sqrt{2}}\right) = v_1. \\
Au_2 &= Ae_2 = v_2. \\
Au_3 &= A\left(\frac{e_1-e_3}{\sqrt{2}}\right) = \frac{1}{\sqrt{2}}\left(\frac{v_1+v_3}{\sqrt{2}} - \frac{v_1-v_3}{\sqrt{2}}\right) = v_3.
\end{aligned}$$

To prove that $A \in \mathbf{Sp}(2,1)$, we verify that

$$A^*JA = \begin{bmatrix} \langle \frac{v_1+v_3}{\sqrt{2}}, \frac{v_1+v_3}{\sqrt{2}} \rangle & \langle \frac{v_1+v_3}{\sqrt{2}}, v_2 \rangle & \langle \frac{v_1+v_3}{\sqrt{2}}, \frac{v_1-v_3}{\sqrt{2}} \rangle \\ \langle v_2, \frac{v_1+v_3}{\sqrt{2}} \rangle & \langle v_2, v_2 \rangle & \langle v_2, \frac{v_1-v_3}{\sqrt{2}} \rangle \\ \langle \frac{v_1-v_3}{\sqrt{2}}, \frac{v_1+v_3}{\sqrt{2}} \rangle & \langle \frac{v_1-v_3}{\sqrt{2}}, v_2 \rangle & \langle \frac{v_1-v_3}{\sqrt{2}}, \frac{v_1-v_3}{\sqrt{2}} \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = J.$$

Therefore, the Lemma follows. □

**Lemma 6.18.** *The* $\mathbf{Sp}(2,1)$*-actions on* $V_0$*,* $V_+$ *and* $V_-$ *are all transitive.*

*Proof.* Let $x \in V_0$. It is sufficient to prove that there exists $A \in \mathbf{Sp}(2,1)$ such that $Ae_1 = x$. Choose $y \in V_0$ such that $\{x, y\} \subset V_0$ are linearly independent. Then $\langle x, y \rangle \neq 0$. By scaling $y$, we may assume that $\langle x, y \rangle = 1$. According to corollary 6.10, there is a spacelike vector $z$ such that $\langle z, z \rangle = 1$ and $z^\perp = \mathrm{span}_{\mathbb{H}}\{x, y\}$. Define $A = (x, z, y)$.

$$A^*JA = \begin{bmatrix} \langle x,x \rangle & \langle x,z \rangle & \langle x,y \rangle \\ \langle z,x \rangle & \langle z,z \rangle & \langle z,y \rangle \\ \langle y,x \rangle & \langle y,z \rangle & \langle y,y \rangle \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = J.$$

Thus $A \in \mathbf{Sp}(2,1)$ and $Ae_1 = x$.

Let $x \in V_+$. By corollary 6.10, there is a vector $z \in V_-$ such that $z \in x^\perp$. Choose $y \in V_+$ such that $y \in x^\perp \cap z^\perp$. Then $\{x, y, z\}$ is an orthonormal basis of $\mathbb{H}^{2,1}$. By Theorem 6.17, there exists an element $A \in \mathbf{Sp}(2,1)$ such that $Ae_2 = y$. Hence the $\mathbf{Sp}(2,1)$-action on $V_+$ is transitive. Similarly, one can prove that the $\mathbf{Sp}(2,1)$-action on $V_-$ is transitive. □

**Proposition 6.19.** *The* $\mathbf{Sp}(2,1)$*-action on* $V_0 \times V_0 \setminus \Delta$ *is transitive where* $\Delta = \{(x,x) \mid x \in V_0\}$.

*Proof.* In the proof of Lemma 6.18, we have shown that for any linearly independent vectors $x$ and $y$ in $V_0$, there exists an element $A \in \mathbf{Sp}(2,1)$ such that $Ae_1 = x$ and $Ae_3 = y$, which implies the transitivity of the $\mathbf{Sp}(2,1)$-action on $V_0 \times V_0 \setminus \Delta$. □

# Bibliography


[1] N. Jacobson, *Basic Algebra I,* W. H. Freeman. (1974).

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

[3] R. E. Johnson, *On the equation $\chi\alpha = \gamma\chi + \beta$ over an algebraic division ring,* Bull. Amer. Math. Soc. 50, 202–207 (1944).

[4] P. G. Tait *An Elementary Treatise on Quaternions*, Cambridge University Press (1890).

[5] K. Shoemake, *Animating Rotation with Quaternion Curves,* Computer Graphics. 19(3), 245–254 (1985).

[6] J. J. Shu & L. S. Ouw, *Pairwise alignment of the DNA sequence using hypercomplex number representation,* Bulletin of Mathematical Biology. 66(5), 1423–1438 (2004).

[7] J. J. Shu & Y. Li, *Hypercomplex cross-correlation of DNA sequences,* Journal of Biological Systems. 18(4), 711–725 (2010).

[8] J. L. Brenner, *Matrices of quaternions,* Pacific J. Math. 1, 329–335 (1951).

[9] Y. H. Au-Yeung, *On the convexity of numerical range in quaternionic Hilbert spaces*, Linear and Multilinear Algebra. 16, 93–100 (1984).

## Acknowledgments in Korean

2년의 대학원 공부가 굉장히 어려웠습니다. 교수님들 덕분에 제가 공부를 못 한 것에 비해 수월하게 2년을 보냈습니다. 감사합니다. 특히, 너무 많이 모자란 저를 논문 제출이라는 단계까지 오게 해 주신 김성운 교수님께 감사합니다.

# Curriculum Vitae in Korean

이　　　름: 양 수 연

생 년 월 일: 1993년 10월 02일

## 학　　력

2009. 3. – 2012. 2.　제주중앙여자고등학교

2012. 3. – 2018. 2.　제주대학교 수학과 (학사)