University of the Philippines Cebu

# ON QUATERNIONIC LINEAR MAPS

by

John Aaron Q. Alcoseba

An undergraduate thesis submitted to the
Sciences Cluster
University of the Philippines Cebu
Lahug, City

as partial fulfillment of the
requirements for the degree of
Bachelor of Science in Mathematics

June 2017

**On Quaternionic Linear Maps**
**by John Aaron Q. Alcoseba**
Undergraduate Thesis, Bachelor of Science in Mathematics
University of the Philippines Cebu
June 2017

Classification*: P
* I - invention or creation, P - publication, C - confidential information

| | |
|---|---|
| Available to the general public | Yes |
| Available only after consultation with author/adviser for thesis | No |
| Available only to those bound by non-disclosure or confidentiality agreement | No |

JOHN AARON Q. ALCOSEBA
Student

DR. LORNA S. ALMOCERA
Thesis Adviser

**University of the Philippines Cebu**

This is to certify that this Undergraduate Thesis entitled **"On Quaternionic Linear Maps"**, prepared and submitted by **John Aaron Q. Alcoseba** to fulfill part of the requirements for the degree of **Bachelor of Science in Mathematics**, was successfully defended and approved on May 29, 2017.

DR. LORNA S. ALMOCERA
Thesis Adviser

DR. NATIVIDAD ESTILLORE
Thesis Panel

The Sciences Cluster endorses the acceptance of this Undergraduate Thesis as partial fulfillment of the requirements for the degree of Bachelor of Science in Mathematics.

JONNIFER R. SINOGAYA, PH.D.
Chairperson
Sciences Cluster

# Table of Contents

# Acknowledgments

# Abstract

## On Quaternionic Linear Maps

John Aaron Q. Alcoseba
University of the Philippines Cebu, 2017

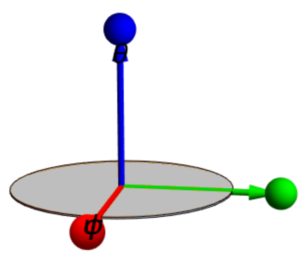
Adviser:
Dr. Lorna S. Almocera

This paper is an expository on quaternionic matrices and their implications on concepts in Linear Algebra such as the notion of vector spaces, linear maps, and determinants. We see that different definitions have been given for the latter. We choose to look into one such definition - the Study Determinant - which uses homomorphisms that represent $n \times n$ *Complex Matrices* as $2n \times 2n$ real matrices, and representing $n \times n$ *Quaternionic Matrices* as $2n \times 2n$ complex matrices.

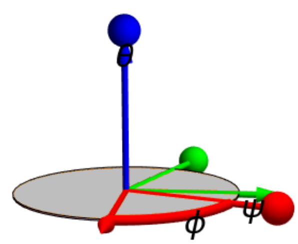
# Chapter 1

# Introduction

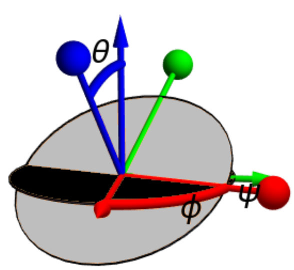
A three-dimensional rotation can be expressed as a sequence of angles - each associated with a rotation about a principal axis. We call this set of angles *Euler Angles* [11] [6] [17]. The *principal axes* are initially the $x, y$ and $z$ axes. The axes change as the rotation takes place but, nonetheless, stay perpendicular to each other.



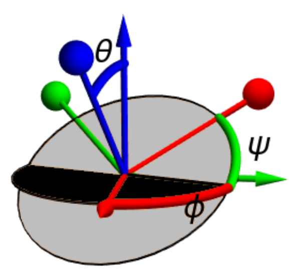

(a) Initial Principal Axes




(b) Rotation by $\phi$ about the $z$-axis (blue)




(c) Rotation by $\theta$ about the new $x$-axis (red)




(d) Rotation by $\psi$ about the new $z$-axis (blue)

**Figure 1.1**: [17] [6] Rotation by Euler Angles

As a demonstration, observe Figure 1.1 in which we have the Euler angles $\phi, \theta$ and $\psi$. Initially in 1.1a, we have the principal axes $x$ (red), $y$ (green) and $z$ (blue), each shown as arrows. We then rotate by an angle $\phi$ about the $z$-axis (blue) in 1.1b. Notice that as the rotation by the angle $\phi$ takes place, the $x$ (red) and $y$ (green) axes are displaced but remain perpendicular to $z$ and to each other. The displaced axes (each shown as a ball) now form a new set of principal axes from which we can then perform the next rotation in 1.1c.

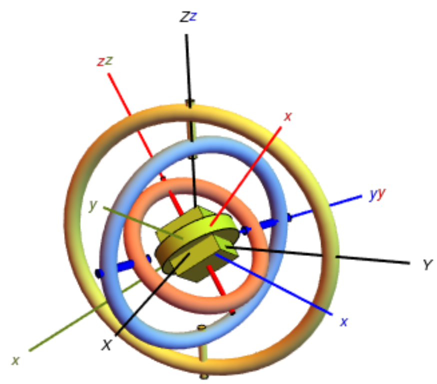



**Figure 1.2**: Gyroscope

Alternatively, we can visualize a rotation using a *gyroscope* as seen in Figure 1.2 where rotations about the $x$, $y$, and $z$ axes are represented by the red, yellow, and blue rings respectively.

Computing for the image of a point in $\mathbb{R}^3$ under a rotation using Euler angles requires $4 \times 4$ matrices like the ones shown in Figure 1.3.

$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

($a$)Rotation by $\alpha$ in the x-axis

$$R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

($b$)Rotation by $\beta$ in the y-axis

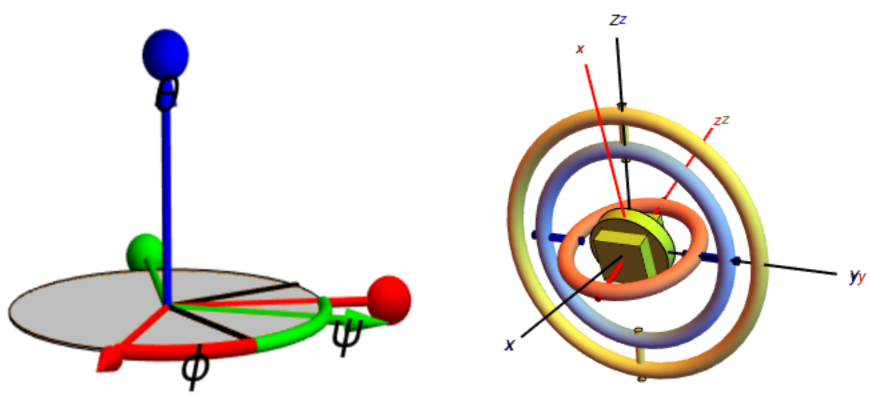
$$R_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 & 0 \\ \sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

($c$)Rotation by $\gamma$ in the z-axis

**Figure 1.3**: $4 \times 4$ Rotation Matrices about the Principal Axes

We can rotate any vector/point $\vec{v} \in \mathbb{R}^3$ by matrix multiplication $\vec{v}' = R_x(\alpha)R_y(\beta)R_z(\gamma)\vec{v}$. However, these computations are a bit tedious and require more elementary arithmetic operations [11]. It's also more difficult to determine the axis and angle of rotation [11]. Furthermore, this method is susceptible to a problem in mechanics known as the *Gimbal Lock* [8].

The gimbal lock is a phenomenon that occurs when two successive rotations are about the same axis as seen in 1.4a [6]. This means that a rotation described by $\phi$ can be described by $\psi$ too. We, therefore, have a variable that depends on another, i.e., a variable that can no longer vary independently. This indicates a loss of one degree of freedom.



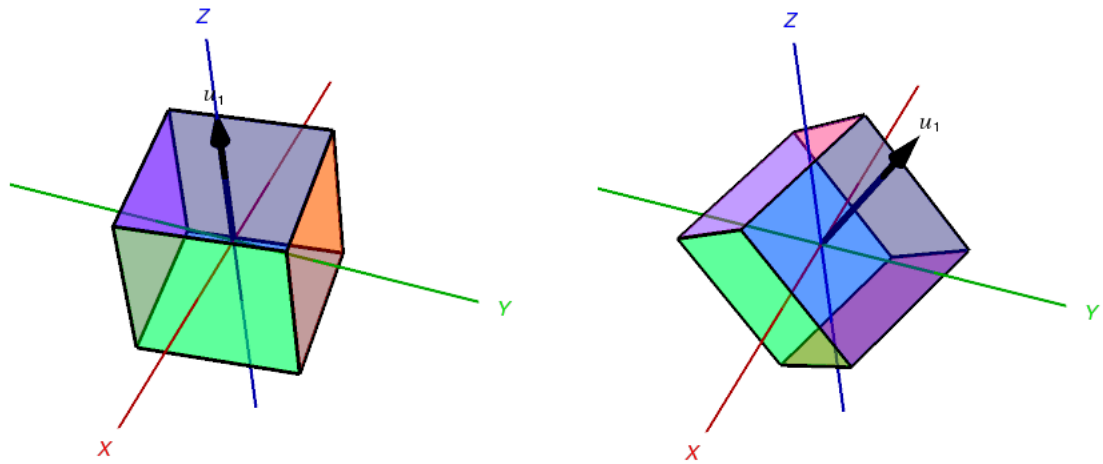

(a) A rotation by $\phi$ then $\psi$ about the same axis.

(b) The blue and yellow rings coincide resulting in a loss of one degree of freedom.

**Figure 1.4**: The Gimbal Lock

The same is true for the gyroscope example in 1.4b in which two of the rings coincide (yellow and blue). A rotation described by the red ring can also be described by the outer yellow ring. In other words, the outer rings have "locked" the rotations described by the red ring.

There is a simpler and more elegant way to describe three-dimensional rotation which involves the algebra of *Quaternions* (discovered by Rowan Hamilton in 1843). Using quaternions, one only needs to specify the angle of rotation and a three-dimensional vector about which a point in space rotates (axis of rotation).




(a) Rotation about the $z$-axis (b) Rotation about some axis

**Figure 1.5**: [2] Rotation using the quaternion $(0, \vec{u_1})$

One can write a quaternion as an ordered pair $(a, \vec{v})$ where $a$ is a scalar, and $\vec{v}$ is a three-dimensional vector [13]. In order to rotate a point/vector $\vec{v} \in \mathbb{R}^3$ about a unit vector $\vec{u_1}$ (axis of rotation), we first need to write $\vec{v}$ as a quaternion $p = (0, \vec{v})$. We then perform a conjugation on $p$ by a quaternion $q$,i.e., $p' = qpq^{-1}$ where $q = (\cos(\frac{\theta}{2}), \vec{u_1}\sin(\frac{\theta}{2}))$ [2]. The vector part of the resulting quaternion $p'$ is the vector obtained after the rotation about $\vec{u_1}$ by an angle $\theta$.

Because quaternions do not require rotations about principal axes to describe three-dimensional rotation, quaternions do not suffer from the gimbal lock. They are also found to be more compact - requiring less elementary arithmetic operations to perform rotation composition than rotation matrices [11].

Quaternions are used today in robotics, three-dimensional computer graphics, computer vision, crystallographic texture analysis, navigation, and molecular dynamics.

Mathematicians have made advancements in developing the theory of Quaternions. Notably, as one of the central points of this topic, we look into the concept of a *Quaternionic Matrix* and the implications it has on certain definitions that were already established in Linear Algebra.

One such implication is the concept of a determinant in the context of quaternionic matrices. In linear algebra, we saw that we can extend the definition of the determinant to matrices with complex entries [12]. This is possible because the complex numbers are commutative under complex multiplication [1].

Certain problems arise if we attempt to extend the classical definition to the quaternions because quaternions are not commutative under quaternion multiplication [1]. In [1], the following properties we associate with determinants are revisited and three conditions called *axioms* are given that should be satisfied in order for a definition of a determinant to be valid and useful:

- **Axiom 1** $det(A) = 0$ if and only if $A$ is singular.
- **Axiom 2** $det(AB) = det(A)det(B)$ for all quaternionic matrices $A$ and $B$.
- **Axiom 3** If $A'$ is obtained by adding a left-multiple of a row to another row or a right-multiple of a column to another column, then $det(A') = det(A)$.

Over the years, several mathematicians have come up with different ways to define a determinant for quaternionic matrices - the Cayley determinant (by Arthur Cayley in 1845), the Study determinant (by Eduard Study in 1920), the Dieudonne determinant, and Moore's determinant. Aslaksen showed whether or not the Cayley, Study, Dieudonne, and Moore's determinant satisfy the axioms [1].

In this paper, we will discuss the theory behind quaternionic linear maps and quaternionic determinants - particularly, the Study determinant.

## 1.1 List of Symbols

- $\mathrm{M}_n(\mathbb{R})$ - set of all $n \times n$ matrices with real entries.
- $\mathrm{M}_n(\mathbb{C})$ - set of all $n \times n$ matrices with complex entries.
- $\mathrm{M}_n(\mathbb{H})$ - set of all $n \times n$ matrices with quaternion entries.
- $\mathcal{D}_n(\mathbb{C})$ - set of all $n \times n$ skew-coninvolutory matrices with complex entries.
- $\mathcal{D}_n(\mathbb{H})$ - set of all $n \times n$ skew-coninvolutory matrices with quaternion entries.
- $det_{\mathbb{C}}()$ - determinant of a matrix in $\mathrm{M}_n(\mathbb{C})$.
- $det_{\mathbb{R}}()$ - determinant of a matrix in $\mathrm{M}_n(\mathbb{R})$.

# Chapter 2

# Preliminaries

In this chapter, we present and discuss terms and known results in abstract, linear, complex and quaternion algebra which are key to understanding complex matrix representations of quaternionic linear maps.

## 2.1 Rings, Fields, and Homomorphisms

**Definition 2.1** (Ring)**.** *[5] A* ring *is a set $R$ together with two binary operations $+$ (addition) and $\cdot$ (multiplication) defined on $R$ such that the following axioms are satisfied:*

1. *$R$ under $+$ is an abelian group.*
2. *Multiplication $\cdot$ is associative.*
3. *$\forall a, b, c \in R$ the left distributive law $a \cdot (b + c) = a \cdot b + a \cdot c$ and the right distributive law $(a + b) \cdot c = a \cdot c + b \cdot c$ hold.*

The first condition essentially means that a ring must form a group under addition $+$ (i.e., it's closed and associative under $+$, it contains an identity element and has an inverse for all its elements) and that addition is commutative [5]. Rings in algebra are not to be confused with rings/tori (plural for torus) in geometry. The name *ring* originally came from the word "Zahlring" (number ring) - a term introduced by D. Hilbert in 1897 - that may have something to do with cyclical behavior of powers of algebraic integers [7]. Rings are one of the fundamental algebraic structures (groups and fields being the other two). Examples of rings are the integers $\mathbb{Z}$, the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$ and the quaternions $\mathbb{H}$.

Notice that a ring does not necessarily have a multiplicative identity. An example would be the ring of even integers $2\mathbb{Z}$ where the multiplicative identity 1 is not a member.

If a ring does have a multiplicative identity, the multiplicative identity is called a *unity* and the said ring is called a *ring with unity* [5]. In most of the rings we'll discuss, the multiplicative identity is 1.

If the non-zero elements of a ring $R$ each have multiplicative inverses (i.e., for $a \in R$ and $a \neq 0$, $\exists a^{-1} \in R$ such that $aa^{-1} = a^{-1}a = 1$) and are closed under ring multiplication, division becomes possible. Such a ring is called a *division ring* or a *skew field*.

**Definition 2.2** (Division Ring/Skew Field). *[5] A* division ring *or a* skew field *is a ring whose non-zero elements form a group under multiplication.*

Note that a division ring does not necessarily have to be commutative under multiplication. A division ring that is commutative under multiplication is called a *field*. Examples of division rings are the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, the complex numbers $\mathbb{C}$, and the quaternions $\mathbb{H}$. The integers no longer form a division ring since the only non-zero element in $\mathbb{Z}$ that has a multiplicative inverse is the unity itself.

**Definition 2.3** (Field). *[5] A* field *is a commutative division ring $R$, i.e., the non-zero elements of $R$ form an abelian group.*

Examples of fields are the rational $\mathbb{Q}$, real $\mathbb{R}$, and complex numbers $\mathbb{C}$. The quaternions $\mathbb{H}$ cannot form a field because they are, in fact, not commutative under multiplication. Rings, division rings, and fields can be summarized by the diagram shown in Figure 2.1.

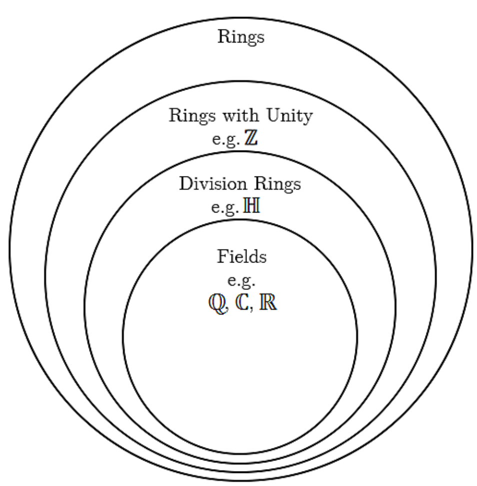



**Figure 2.1**: Diagram of Rings

**Definition 2.4** (Homomorphism)**.** *[5] A map $\phi$ of a group $G$ into a group $G'$ is a* homomorphism *if $\phi(ab) = \phi(a)\phi(b)$ $\forall a, b \in G$.*

A homomorphism is basically a structure-relating map that associates a binary operation in a group $G$ to a binary operation in another group $G'$. In other words, we say that the mapping *preserves* the binary operation that gives $G$ its group structure.

**Example 2.1.1.** Let $F$ be the additive group of continuous functions with domain $[0, 1]$. The map $\sigma : F \to \mathbb{R}$ defined by $\sigma(f) = \int_0^1 f(x)dx$ for $f \in F$ is a homomorphism since

$$\begin{aligned}\sigma(f+g) &= \int_0^1 [f(x) + g(x)]dx \\ &= \int_0^1 f(x)dx + \int_0^1 g(x)dx \\ &= \sigma(f) + \sigma(g)\end{aligned}$$

**Example 2.1.2.** Let $\mathbb{C}$ be the field of complex numbers and $\mathrm{M}_2(\mathbb{R})$ the set of $2 \times 2$ matrices. By Definition 2.3, this means that $\mathbb{C}$ is a group under multiplication. The map $\phi : \mathbb{C} \to \mathrm{M}_2(\mathbb{R})$ defined by

$$\phi(a + b\mathbf{i}) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

is a homomorphism (in fact, it is an *injective homomorphism*, i.e., it is also one-to-one) since

$$\begin{aligned}\phi[(a+b\mathbf{i})(c+d\mathbf{i})] &= \phi[(ac-bd) + (bc+ad)\mathbf{i}] \\ &= \begin{pmatrix} ac-bd & -(bc+ad) \\ bc+ad & ac-bd \end{pmatrix} \quad \text{and,} \\ \phi(a+b\mathbf{i})\phi(c+d\mathbf{i}) &= \begin{pmatrix} a & -b \\ b & a \end{pmatrix}\begin{pmatrix} c & -d \\ d & c \end{pmatrix} \\ &= \begin{pmatrix} ac-bd & -(bc+ad) \\ bc+ad & ac-bd \end{pmatrix}\end{aligned}$$

implying $\phi[(a+b\mathbf{i})(c+d\mathbf{i})] = \phi(a+b\mathbf{i})\phi(c+d\mathbf{i})$.

## 2.2 Vector Spaces and Linear Maps

**Definition 2.5** (Vector Space)**.** *[10]Let $V$ be a set on which two operations (vector addition and scalar multiplication) are defined and let $F$ be a field. If the listed axioms are satisfied $\forall \vec{u}, \vec{v}, \vec{w} \in V$ and $\forall c, d \in F$ (scalar), then $V$ is called a* vector space *over the field $F$.*

***Scalar Multiplication:***

1. $c\vec{u} \in V$
2. $c(\vec{u}+\vec{v}) = c\vec{u} + c\vec{v}$
3. $(c+d)\vec{u} = c\vec{u} + d\vec{u}$
4. $c(d\vec{u}) = (cd)\vec{u}$
5. $1\vec{u} = \vec{u}$

***Vector Addition:***

1. $\vec{u}+\vec{v} \in V$
2. $\vec{u}+\vec{v} = \vec{v}+\vec{u}$
3. $\vec{u}+(\vec{v}+\vec{w}) = (\vec{u}+\vec{v})+\vec{w}$
4. *$V$ has a zero vector $\vec{0}$ such that $\forall \vec{u} \in V$, $\vec{u}+\vec{0} = \vec{u}$*
5. *$\forall \vec{u} \in V, \exists -\vec{u} \in V$ such that $\vec{u}+(-\vec{u}) = \vec{0}$.*

A vector space is basically any algebraic structure that follows a proper notion of vector addition and scalar multiplication, i.e., the latter operations satisfy the axioms in Definition 2.5. A familiar example of a vector space would be the $n$-dimensional Euclidean space over the field $\mathbb{R}$ ($\mathbb{R}^n$) where we represent vectors as $n$-tuples with components in $\mathbb{R}$. In $\mathbb{R}^2$ and $\mathbb{R}^3$, we commonly visualize vectors as arrows in space with magnitude and direction. However, there are other vector spaces that cannot be visualized just as easily, such as the vector space of polynomials of degree $n$ ($P(n)$) [16] [10]. The polynomials form a vector space because they have a notion of vector addition (addition of like terms) and scalar multiplication (multiplication by a constant) [16] [10]. Another vector space which will be of great importance throughout the discussion is the *complex vector space* which is simply the Euclidean vector space over the field $\mathbb{C}$ ($\mathbb{C}^n$).

**Definition 2.6** (Linear Map/ Linear Transformation)**.** *[10]Let $V, W$ be vector spaces over the field $F$. The function $L : V \to W$ is called a* linear map/ linear transformation *of $V$ into $W$ when the following properties are true $\forall u, v \in V$ and $\forall c \in F$:*

1. *$L(u+v) = L(u)+L(v)$*

*2. $L(cu) = cL(u)$*

In Property 1 of Definition 2.6, we can view $L$ as having the linear map commute with vector addition, i.e., the image of the vector sum is the sum of the images. Property 2 of Definition 2.6 is interpreted as having the linear map commute with scalar multiplication. Notice that Definition 2.5 only considers scalar multiplication in the context of fields. In the case of non-commutative skew fields like the quaternions, we will have to specify the type of scalar multiplication in the vector space, i.e., either right or left scalar multiplication [18]. If we only consider right scalar multiplication in Definition 2.5, we have what is called a *right vector space* (provided that all the axioms are satisfied given the choice of scalar multiplication) [18]. An example of a right vector space is the *right quaternionic vector space*. We will later see that we can still define linear maps in quaternions provided that we only consider right vector spaces [18] [1].

We know from linear algebra that every linear transformation/linear map $L : V \to W$ (where $V$ and $W$ are vector spaces of dimension $n$ and $m$ respectively) has an $m \times n$ matrix representation [10]. Thus, we can essentially interchange the terms *matrix* and *linear map* throughout the discussion.

## 2.3 Complex Matrices

Our discussions regarding quaternionic linear maps/matrices will mostly revolve around extending properties and definitions that already hold for complex matrices. We begin by extending properties and definitions that already hold for real matrices to complex matrices.

*Complex Matrices* are matrices that have entries in $\mathbb{C}$. They are linear maps $L_{\mathbb{C}} : V \to W$ where $V$ and $W$ are *complex vector spaces*. Since its entries are complex numbers, we can take the conjugate of each of its entries. We call the resulting matrix a *conjugate matrix*.

**Definition 2.7** (Conjugate Matrix)**.** *A* conjugate matrix *is a matrix $\bar{E}$ obtained from a complex matrix $E \in M_n(\mathbb{C})$, where $M_n(\mathbb{C})$ is the set of all $n \times n$ complex matrices, by taking the complex conjugate of every entry of $E$.*

**Example 2.3.1.** Take $E = \begin{pmatrix} -\frac{13}{17} + \frac{16}{17}\mathbf{i} & -\frac{8}{17} + \frac{2}{17}\mathbf{i} \\ \frac{16}{17} - \frac{4}{17}\mathbf{i} & \frac{19}{17} + \frac{8}{17}\mathbf{i} \end{pmatrix}$. Then $\overline{E} = \begin{pmatrix} -\frac{13}{17} - \frac{16}{17}\mathbf{i} & -\frac{8}{17} - \frac{2}{17}\mathbf{i} \\ \frac{16}{17} + \frac{4}{17}\mathbf{i} & \frac{19}{17} - \frac{8}{17}\mathbf{i} \end{pmatrix}$

Operations and properties that we know hold for real matrices like matrix addition and multiplication, elementary row operations, matrix inverse, and the determinant still hold for complex matrices [1].

**Definition 2.8** (Complex Determinant)**.** *[10]The* Complex Determinant*, denoted by $det_{\mathbb{C}}$, of an $n \times n$ complex matrix $E$ is defined by,*

$$det_{\mathbb{C}}(E) = |E| = \sum_{j=1}^{n} a_{ij}c_{ij}$$

Notice that, computationally, the complex determinant is no different from the real determinant. The only difference is that computing for the determinant of a complex matrix will give us a complex number.

**Theorem 2.1.** *For a matrix $E \in M_n(\mathbb{C})$, $det_{\mathbb{C}}(\bar{E}) = \overline{det_{\mathbb{C}}(E)}$.*

*Proof.* We prove by mathematical induction.
Let $\mathrm{M}_n(\mathbb{C})$ denote the set of all $n \times n$ complex matrices. Let $n = 2$ and take $E = \begin{pmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{pmatrix}$.

Then, $det_{\mathbb{C}}(\bar{E}) = \begin{vmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{vmatrix} = \overline{ad} - \overline{bc} = \overline{ad - bc} = \overline{det_{\mathbb{C}}(E)}$

Suppose $\overline{det_{\mathbb{C}}(E)} = det_{\mathbb{C}}(\bar{E})$ holds for $E \in \mathrm{M}_k(\mathbb{C})$ for some $k \geq 2$ $(*)$.
Let $X \in \mathrm{M}_{k+1}(\mathbb{C})$. Then,

$$\overline{det_{\mathbb{C}}(X)} = \overline{\sum_{j=1}^{k+1} a_{ij}c_{ij}} = \sum_{j=1}^{k+1} \overline{a_{ij}}\ \overline{c_{ij}}$$

Note that $\overline{c_{ij}} = (-1)^{i+j}\overline{M_{ij}}$ where $\overline{M_{ij}}$ is the determinant of the $k \times k$ matrix obtained by deleting the $i^{th}$ row and the $j^{th}$ column of the original matrix.
By $(*)$, $\overline{M_{ij}}$ is the determinant of a $k \times k$ conjugate matrix. Thus, we see that we are computing for the determinant of a $(k+1) \times (k+1)$ conjugate matrix. $\square$

Theorem 2.1 states that computing for the determinant commutes with conjugation, i.e., the determinant of the conjugate matrix is the conjugate of the determinant.

**Definition 2.9** (Coninvolutory Matrix)**.** *A matrix is said to be* coninvolutory *if* $E\bar{E} = I_n$ *for* $E \in M_n(\mathbb{C})$ *where* $M_n(\mathbb{C})$ *denotes the set of all* $n \times n$ *complex matrices.*

By manipulation, we obtain $E^{-1} = \bar{E}$. Hence, a matrix whose inverse is its own conjugate matrix is a coninvolutory matrix.

**Example 2.3.2.** Consider the complex matrix, $E = \begin{pmatrix} -\frac{13}{17} + \frac{16}{17}\mathbf{i} & -\frac{8}{17} + \frac{2}{17}\mathbf{i} \\ \frac{16}{17} - \frac{4}{17}\mathbf{i} & \frac{19}{17} + \frac{8}{17}\mathbf{i} \end{pmatrix}$

We see that

$$\begin{aligned} E\bar{E} &= \begin{pmatrix} -\frac{13}{17} + \frac{16}{17}\mathbf{i} & -\frac{8}{17} + \frac{2}{17}\mathbf{i} \\ \frac{16}{17} - \frac{4}{17}\mathbf{i} & \frac{19}{17} + \frac{8}{17}\mathbf{i} \end{pmatrix} \begin{pmatrix} -\frac{13}{17} - \frac{16}{17}\mathbf{i} & -\frac{8}{17} - \frac{2}{17}\mathbf{i} \\ \frac{16}{17} + \frac{4}{17}\mathbf{i} & \frac{19}{17} - \frac{8}{17}\mathbf{i} \end{pmatrix} \\ &= \frac{1}{17}\frac{1}{17} \begin{pmatrix} -13 + 16\mathbf{i} & -8 + 2\mathbf{i} \\ 16 - 4\mathbf{i} & 19 + 8\mathbf{i} \end{pmatrix} \begin{pmatrix} -13 - 16\mathbf{i} & -8 - 2\mathbf{i} \\ 16 + 4\mathbf{i} & 19 - 8\mathbf{i} \end{pmatrix} \\ &= \frac{1}{289} \begin{pmatrix} 289 & 0 \\ 0 & 289 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= I_2 \end{aligned}$$

We see that $E$ is a coninvolutory matrix. *Remark:* We've obtained $E$ using the factorization provided in Theorem 2.3 found in [12].

*Remark.* If we take the concept of coninvolutory matrices in the context of real matrices, we get $EE = E^2 = I_n$ since for any $n \times n$, $E = \overline{E}$. This is what we call an *Involutory Matrix*, i.e., a matrix whose inverse is itself.

**Definition 2.10** (Involutory Matrix)**.** *An* $n \times n$ *real matrix* $X$ *is called an* involutory matrix *if* $X^2 = I_n$ *where* $I_n$ *is the* $n \times n$ *identity matrix.*

Because linear maps that represent a 180° rotation or a reflection are of order 2 (i.e. applying their associated actions twice on a point/vector leaves the point/vector unchanged), they are involutory matrices.

**Example 2.3.3.** Take the rotation matrix $R_x(\alpha)$ in Chapter 1. If $\alpha = \pi$, we have a rotation matrix that represents a 180° rotation given by,

$$R_x(\pi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\pi & -\sin\pi & 0 \\ 0 & \sin\pi & \cos\pi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

. We now see that,

$$(R_x(\pi))^2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I_4$$

**Definition 2.11** (Skew-Coninvolutory Matrix)**.** *A matrix is said to be* skew-coninvolutory *if* $E\bar{E} = -I_n$ *for* $E \in M_n(\mathbb{C})$.

We see that a matrix whose inverse is the negative of its own conjugate matrix is a skew-coninvolutory matrix.

*Remark.* If we take this in the context of real matrices, we get $EE = E^2 = -I_n$. Notice how this closely resembles a property of the complex number $\mathbf{i}$, i.e., $\mathbf{i}^2 = -1$. In fact, we have a special name for these linear maps. We call them *complex structures* [19].

**Definition 2.12** (Complex Structure)**.** *A* complex structure *of a vector space* $V$ *is defined by the linear map (linear transformation)* $J : V \to V$ *such that* $J^2 = -I$*, where* $I$ *is the identity map. [19]*

*Remark.* If a vector space $V$ has a complex structure (i.e. if we can define a complex structure in the vector space), it means that $V$ can essentially "mimic" a complex vector space [14]. To see how this works, observe Example 2.3.4.

**Example 2.3.4.** The matrix $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ gives a complex structure in $\mathbb{R}^2$ since

$$J^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2.$$

If we can think of $a + b\mathbf{i} \in \mathbb{C}$ as a 2-dimensional vector $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ and multiply $\vec{v}$ by $J$, we get,

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$$

which corresponds to the product of $\mathbf{i}$ and $a + b\mathbf{i}$, $-b + a\mathbf{i}$.

There are vector spaces in which we cannot define a complex structure. For instance, the vector space $\mathbb{R}^1$ or simply $\mathbb{R}$. Since this is a 1-dimensional vector space, we have $1 \times 1$ matrices to represent linear maps $L_1 : \mathbb{R} \to \mathbb{R}$. Linear maps like $L_1$ are identical to scalar multiplication. Since there is no element $x \in \mathbb{R}$ such that $x^2 = -1$, we cannot define a complex structure in $\mathbb{R}$.

We can observe a pattern in the dimensions of the vector spaces where a complex structure can be defined. We see that only vector spaces (specifically real vector spaces) with even dimensions can have a complex structure defined. This is proven in the context of skew-coninvolutory complex matrices as shown in Theorem 2.2.

**Theorem 2.2.** *[12] Let $\mathcal{D}_n(\mathbb{C})$ denote the set of all $n \times n$ skew-coninvolutory complex matrices. Then $\mathcal{D}_n(\mathbb{C})$ is empty when $n$ is odd.*

*Proof.* If $E \in \mathcal{D}_n(\mathbb{C})$ then $E\bar{E} = -I_n$.
Taking the determinant of both sides,

$$\begin{aligned} det_{\mathbb{C}}(E\bar{E}) &= det_{\mathbb{C}}(-I_n) \\ det_{\mathbb{C}}(E)det_{\mathbb{C}}(\bar{E}) &= (-1)^n \\ det_{\mathbb{C}}(E)\overline{det_{\mathbb{C}}(E)} &= (-1)^n, \text{ by Theorem 2.1} \\ |det_{\mathbb{C}}(E)|^2 &= (-1)^n, \text{ since } det_{\mathbb{C}} \text{ yields a complex number.} \end{aligned}$$

Since $|det_{\mathbb{C}}(E)|^2 > 0$, $(-1)^n > 0$. Hence, $n$ must be even. $\square$

*Remark.* Theorem 2.2 puts a restriction on the dimension of complex matrices that are skew-coninvolutory. In the context of real matrices, this means that $E^2 = -I_n$ only holds if $E$ is a $2n \times 2n$ real matrix, i.e., a complex structure only exists for real matrices with even dimensions.

**Example 2.3.5.** We see that in the $1 \times 1$ case, $\bar{\mathbf{i}} = -\mathbf{i}$ but $\mathbf{i}\bar{\mathbf{i}} = -\mathbf{i}^2 = 1 \neq -1$. Whereas in the $2\times 2$ case, consider $E = \begin{pmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{pmatrix}$. Then $E\bar{E} = \begin{pmatrix} 0 & \mathbf{i} \\ -\mathbf{i} & 0 \end{pmatrix} \begin{pmatrix} 0 & -\mathbf{i} \\ \mathbf{i} & 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

We will see more manifestations of Theorem 2.2 in later discussions especially when we represent complex matrices as real matrices.

## 2.4 Basics of Quaternion Algebra

In this section, we introduce properties and operations associated with quaternions such as addition, multiplication, conjugation, norm, and inverse.

**Definition 2.13** (Quaternion)**.** *[12] The four-dimensional algebra of* Quaternions*, denoted by* $\mathbb{H}$*, is generated by the basis elements* $\{1, \boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}\}$ *such that* $\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = \boldsymbol{ijk} = -1$*.* $\mathbb{H} := \{a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} | a, b, c, d \in \mathbb{R}\}$*.*

Examples of quaternions are $1 + 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$, $-2 + 5\mathbf{i} - 6\mathbf{k}$, and $3\mathbf{i} - \frac{1}{2}\mathbf{j} + \sqrt{2}\mathbf{k}$. As seen in Chapter 1, we can also write a quaternion $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ as an ordered pair $(a, \vec{v})$ where $a$ is a scalar and $\vec{v}$ is a 3-dimensional vector corresponding to the coefficients of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ respectively, i.e., the vector $(b, c, d)$. For instance, the quaternion $-2 + 5\mathbf{i} - 6\mathbf{k}$ can be written as $(-2, (5, 0, -6))$ and the quaternion $3\mathbf{i} - \frac{1}{2}\mathbf{j} + \sqrt{2}\mathbf{k}$ can be written as $(0, (3, -\frac{1}{2}, \sqrt{2}))$. We call a quaternion whose scalar part $a = 0$, a *pure quaternion* [11] [13]. Notice that $\mathbb{C} \subseteq \mathbb{H}$ since a quaternion becomes a complex number if $c, d = 0$.

### 2.4.1 Addition and Multiplication

**Addition.** Adding quaternions is straightforward and follows component-wise addition.

**Definition 2.14** (Quaternion Addition)**.** *[11] For quaternions* $q_1 = a_1 + b_1\boldsymbol{i} + c_1\boldsymbol{j} + d_1\boldsymbol{k}$ *and* $q_2 = a_2 + b_2\boldsymbol{i} + c_2\boldsymbol{j} + d_2\boldsymbol{k}$*,*

$$q_1 + q_2 = (a_1 + a_2) + (b_1 + b_2)\boldsymbol{i} + (c_1 + c_2)\boldsymbol{j} + (d_1 + d_2)\boldsymbol{k}.$$

Alternatively, if $q_1 = (a_1, \vec{v_1})$ and $q_2 = (a_2, \vec{v_2})$ where $\vec{v_1}$ and $\vec{v_2}$ are $(b_1, c_1, d_1)$ and $(b_2, c_2, d_2)$ respectively, then

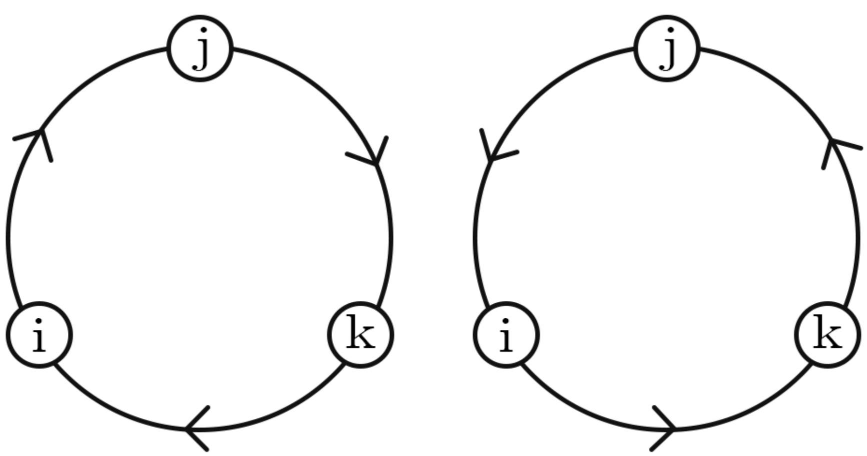
$$(a_1, \vec{v_1}) + (a_2, \vec{v_2}) = (a_1 + a_2, \vec{v_1} + \vec{v_2})$$

The quaternions form an *abelian group* under addition (i.e., they form a commutative group under addition) because: (1) quaternion addition is associative, (2) the additive identity 0 is a quaternion, (3) the additive inverse of a quaternion $q$ is $-q$ which is also a quaternion, and (4) quaternion addition is commutative since the scalars $a, b, c$ and $d$ are in $\mathbb{R}$.

**Multiplication.** Quaternion multiplication is governed by Definition 2.13, where $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -1$. We can then derive the following:

$$\begin{aligned} \mathbf{jk} &= \mathbf{i} \qquad & \mathbf{kj} &= -\mathbf{i} \\ \mathbf{ki} &= \mathbf{j} \qquad & \mathbf{ik} &= -\mathbf{j} \\ \mathbf{ij} &= \mathbf{k} \qquad & \mathbf{ji} &= -\mathbf{k} \end{aligned}$$

It can be useful to remember the multiplication of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ using Figure 2.2. We see that going about the circle clockwise starting at $\mathbf{i}$ and going to $\mathbf{j}$ gives us a positive $\mathbf{k}$. On the other hand, going about the circle counter-clockwise starting at $\mathbf{k}$ and going to $\mathbf{j}$ gives us a negative $\mathbf{i}$.




(a) Positive (b) Negative

**Figure 2.2**: Visualizing the multiplication of $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$

From elements $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ alone, we can already see that the quaternions are not commutative under multiplication. However, quaternion multiplication is associative [11]. Furthermore, quaternions commute with a scalar in $\mathbb{R}$ [11]

**Definition 2.15** (Quaternion Multiplication)**.** *[11] For quaternions* $q_1 = a_1 + b_1\boldsymbol{i} + c_1\boldsymbol{j} + d_1\boldsymbol{k}$ *and* $q_2 = a_2 + b_2\boldsymbol{i} + c_2\boldsymbol{j} + d_2\boldsymbol{k}$*,*

$$\begin{aligned} q_1q_2 &= (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)\boldsymbol{i} \\ &\quad + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)\boldsymbol{j} + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)\boldsymbol{k} \end{aligned}$$

**Example 2.4.1.** Take the quaternions $q_1 = 1 + 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$ and $q_2 = -2 + 5\mathbf{i} - 6\mathbf{k}$. Using the formula given by Definition 2.15, we have,

$$\begin{aligned} q_1q_2 &= [(1)(-2) - (2)(5) - (3)(0) - (4)(-6)] + [(1)(5) + (2)(-2) + (3)(-6) - (4)(0)]\mathbf{i} \\ &\quad + [(1)(0) - (2)(-6) + (3)(-2) + (4)(5)]\mathbf{j} + [(1)(-6) + (2)(0) - (3)(5) + (4)(-2)]\mathbf{k} \\ &= 12 - 17\mathbf{i} + 26\mathbf{j} - 29\mathbf{k}. \end{aligned}$$

*Remark.* Alternatively, by viewing quaternions as ordered pairs $q_1 = (a_1, \vec{v_1})$ and $q_2 = (a_2, \vec{v_2})$ we can restate Definition 2.15 as,

$$q_1q_2 = (a_1a_2 - \vec{v_1} \cdot \vec{v_2},\ a_1\vec{v_2} + a_2\vec{v_1} + \vec{v_2} \times \vec{v_1}) \tag{2.1}$$

where $\vec{v_1} \cdot \vec{v_2}$ and $\vec{v_2} \times \vec{v_1}$ are the vector *dot product* and *cross product* respectively [11] [13].

**Example 2.4.2.** We can compute for the product of the quaternions given in Example 2.4.1 using Equation 2.1. For $q_1 = (1, (2, 3, 4))$ and $q_2 = (-2, (5, 0, -6))$, we have

$$\begin{aligned} q_1q_2 &= ((1)(-2) - (10 - 24),\ (5, 0, -6) - 2(2, 3, 4) + (-18, 32, -15)) \\ &= (12, (-17, 26, -29)) \end{aligned}$$

**Theorem 2.3.** *For quaternions* $q, p$ *and* $r$*,* $q(p + r) = qp + qr$ *and* $(p + r)q = pq + rq$*, i.e., the left and right distributive laws hold.*

*Proof.* Let $q = (q_0, \vec{q})$, $p = (p_0, \vec{p})$, and $r = (r_0, \vec{r}) \in \mathbb{H}$. Then,

$$\begin{aligned}
q(p+r) &= (q_0, \vec{q})[(p_0, \vec{p}) + (r_0, \vec{r})] \\
&= (q_0, \vec{q})((p_0 + r_0), \vec{p} + \vec{r}) \text{ by Definition 2.14} \\
&= (q_0(p_0 + r_0) - \vec{q} \cdot (\vec{p} + \vec{r}), q_0(\vec{p} + \vec{r}) + (p_0 + r_0)\vec{q} + (\vec{p} + \vec{r}) \times \vec{q}\ ) \\
&\text{by Definition 2.15} \\
&= (q_0p_0 + q_0r_0 - \vec{q} \cdot \vec{p} - \vec{q} \cdot \vec{r}, q_0\vec{p} + q_0\vec{r} + p_0\vec{q} + r_0\vec{q} + \vec{p} \times \vec{q} + \vec{r} \times \vec{q}\ ) \\
&\text{since the distributive laws hold for dot products and cross products [10].} \\
&= ((q_0p_0 - \vec{q} \cdot \vec{p}) + (q_0r_0 - \vec{q} \cdot \vec{r}), (q_0\vec{p} + p_0\vec{q} + \vec{p} \times \vec{q}) + (q_0\vec{r} + r_0\vec{q} + \vec{r} \times \vec{q})\ ) \\
&= (q_0p_0 - \vec{q} \cdot \vec{p}, q_0\vec{p} + p_0\vec{q} + \vec{p} \times \vec{q}\ ) + (q_0r_0 - \vec{q} \cdot \vec{r}, q_0\vec{r} + r_0\vec{q} + \vec{r} \times \vec{q}\ ) \\
&= qp + qr.
\end{aligned}$$

The right distributive law can also be proven in a similar manner. □

*Remark.* Since (1) the quaternions form an abelian group under quaternion addition, (2) quaternion multiplication is associative and (3) the left and right distributive laws hold, by Definition 2.1 the set of quaternions $\mathbb{H}$ forms a ring.

### 2.4.2 Other Operations and Properties

**Definition 2.16** ($\mathbb{H}$-Conjugate)**.** *The $\mathbb{H}$-Conjugate of a quaternion $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ is $\bar{q} = a - b\boldsymbol{i} - c\boldsymbol{j} - d\boldsymbol{k}$. Alternatively, the $\mathbb{H}$-conjugate of a quaternion $(a, \vec{v})$ where $\vec{v} = (b, c, d)$ is $(a, -\vec{v})$.*

**Example 2.4.3.** Take the quaternion $q = 1 + 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$. Then $\bar{q} = 1 - 2\mathbf{i} - 3\mathbf{j} - 4\mathbf{k}$.

*Remark.* Notice that $q\bar{q} = (a, \vec{v})(a, -\vec{v}) = (a^2 - \vec{v} \cdot (-\vec{v}),\ - a\vec{v} + a\vec{v} + \vec{v} \times \vec{v})$. But we know from linear algebra that, $\vec{v} \cdot \vec{-v} = -\vec{v} \cdot \vec{v} = -|v|^2$ and $\vec{v} \times \vec{v} = 0$, hence $q\bar{q} = (a^2 + |v|^2, (0, 0, 0))$. We also know that $|v|^2 = b^2 + c^2 + d^2$ hence $q\bar{q} = a^2 + b^2 + c^2 + d^2$.

**Definition 2.17** ($\mathbb{H}$-Norm)**.** *The $\mathbb{H}$-Norm of a quaternion $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ is $|q| = \sqrt{q\bar{q}} = \sqrt{a^2 + b^2 + c^2 + d^2}$*

**Example 2.4.4.** The $\mathbb{H}$-norm of the quaternion $q = 1+2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$ is $|q| = \sqrt{1^2+2^2+3^2+4^2}$ $= \sqrt{30}$.

*Remark.* Since $\mathbb{C} \subseteq \mathbb{H}$, we see that definition 2.16 reduces to the definition of a complex conjugate when $c, d = 0$. In a similar manner, the definition of an $\mathbb{H}$-norm 2.17 reduces to the definition of a modulus in $\mathbb{C}$.

**Definition 2.18** (Inverse)**.** *The inverse of a quaternion $q$ is a quaternion $q^{-1}$ such that $q^{-1}q = qq^{-1} = 1$.*

**Theorem 2.4.** *If $q \in \mathbb{H}$ and $q \neq 0$, then $q^{-1} = \bar{q}/|q|^2$*

**Example 2.4.5.** The inverse of the quaternion $q = 1 + 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$ is

$$q^{-1} = \frac{1 - 2\mathbf{i} - 3\mathbf{j} - 4\mathbf{k}}{\sqrt{30}} = \frac{\sqrt{30}}{30} - \frac{\sqrt{30}}{15}\mathbf{i} - \frac{\sqrt{30}}{10}\mathbf{j} - \frac{2\sqrt{30}}{15}\mathbf{k}$$

*Remark.* We can see that the set of non-zero quaternions form a group under multiplication since: (1) quaternion multiplication is associative, (2) the multiplicative identity 1 is also a quaternion and (3) every non-zero quaternion has an inverse given by Theorem 2.4. Hence, by Definition 2.2, the set of quaternions $\mathbb{H}$ forms a division ring or a skew field. Since quaternion multiplication is not commutative, $\mathbb{H}$ cannot be a field.

**Theorem 2.5.** *[11] Let $q, p, r \in \mathbb{H}$. Then we have the following properties:*

1. $\overline{qp} = \bar{p}\bar{q}$.

2. $(qp)^{-1} = p^{-1}q^{-1}$ *provided that the inverses of $p$ and $q$ exist.*

Notice how the properties in Theorem 2.5 follow directly from the fact that the quaternions form a skew field. In abstract algebra, we see that the properties in Theorem 2.5 were already shown to hold for any group [5].

**Theorem 2.6.** *[1] For $z \in \mathbb{C}$, $z\boldsymbol{j} = \boldsymbol{j}\bar{z}$. Alternatively, $\boldsymbol{j}z = \bar{z}\boldsymbol{j}$. [1]*

We can now write $(c\mathbf{j}+d\mathbf{k})$ as $\mathbf{j}(c-d\mathbf{i}) = (c+d\mathbf{i})\mathbf{j}$. Hence, a quaternion $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k}$ can be written as $(a+b\mathbf{i})+\mathbf{j}(c-d\mathbf{i}) = (a+b\mathbf{i})+(c+d\mathbf{i})\mathbf{j}$.

Recall that a complex number is written as $a + b\mathbf{i}$ where $a, b \in \mathbb{R}$, i.e., we can view the complex numbers as a *2-dimensional algebra* over $\mathbb{R}$. Similarly, a quaternion is written as $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ where $a, b, c, d \in \mathbb{R}$, i.e., we can view the quaternions as a 4-dimensional algebra over $\mathbb{R}$. Now, since we can write a quaternion as $(a+b\mathbf{i})+\mathbf{j}(c-d\mathbf{i}) = (a+b\mathbf{i})+(c+d\mathbf{i})\mathbf{j}$ for $a+b\mathbf{i}, c+d\mathbf{i}, c-d\mathbf{i} \in \mathbb{C}$, we can, therefore, view $\mathbb{H}$ as a 2-dimensional algebra over $\mathbb{C}$ [12].

## 2.5 Quaternionic Linear Maps

The non-commutativity of quaternion multiplication presents a problem in defining *Quaternionic Linear Maps*, specifically on Property 2 of Definition 2.6, i.e., $L(cv) = cL(v)$ where we take $c$ as a scalar in $\mathbb{H}$. Let $L_{\mathbb{H}} : V \to W$ be a quaternionic linear map where $m$ and $n$ are the dimensions of $V$ and $W$ respectively. Then, $L_{\mathbb{H}}$ has an $m \times n$ matrix representation $A$ [10] (Note that $A$ has quaternions as entries).

Let's first consider the case where $V$ and $W$ are *left vector spaces*, i.e., $V$ and $W$ are vector spaces defined by left scalar multiplication [18]. Then, for $L_{\mathbb{H}}(v) = Av$, we see that $cAv = cL_{\mathbb{H}}(v)$ for $c \in \mathbb{H}$. However, $L_{\mathbb{H}}(cv) = Acv$ and we know that $Acv \neq cAv$ because $\mathbb{H}$ is not commutative under quaternion multiplication. This means that $L_{\mathbb{H}}(cv) \neq cL_{\mathbb{H}}(v)$, thus, Property 2 of Definition 2.6 fails [18]. Therefore, we cannot define a quaternionic linear map for left vector spaces.

However, if we consider the case where $V$ and $W$ are *right vector spaces*, i.e., $V$ and $W$ are vector spaces defined by right scalar multiplication [18] [1], we have, $L_{\mathbb{H}}(vc) = Avc = L_{\mathbb{H}}(v)c$. Thus, Property 2 of Definition 2.6 holds if we consider the case where $V$ and $W$ are right vector spaces. This means that we can only define quaternionic linear maps for right vector spaces.

Before we define a *quaternionic structure* on a complex vector space, we first introduce a different kind of map over complex vector spaces.

**Definition 2.19** (Antilinear Map)**.** *[15] An* antilinear map *or antilinear operator is a map $\Psi : V \to V$ where $V$ is a complex vector space, such that for $c_1, c_2 \in \mathbb{C}$ and $v_1, v_2 \in V$, $\Psi(c_1v_1 + c_2v_2) = \bar{c_1}\Psi(v_1) + \bar{c_2}\Psi(v_2)$.*

**Example 2.5.1.** We define the map $\Psi : \mathbb{C}^2 \to \mathbb{C}^2$ by $\Psi(v) = \overline{Jv}$ where $v \in \mathbb{C}^2$ and

$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. We then see that for $c_1, c_2 \in \mathbb{C}$ and $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$,

$$\Psi(v_1) = \overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}} = \begin{pmatrix} -\bar{y_1} \\ \bar{x_1} \end{pmatrix}, \Psi(v_2) = \overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}} = \begin{pmatrix} -\bar{y_2} \\ \bar{x_2} \end{pmatrix} \text{ and,}$$

$$\begin{aligned}
\Psi(c_1 v_1 + c_2 v_2) &= \overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} c_1 x_1 + c_2 x_2 \\ c_1 y_1 + c_2 y_2 \end{pmatrix}} = \overline{\begin{pmatrix} -(c_1 y_1 + c_2 y_2) \\ c_1 x_1 + c_2 x_2 \end{pmatrix}} = \begin{pmatrix} -\overline{(c_1 y_1 + c_2 y_2)} \\ \overline{c_1 x_1 + c_2 x_2} \end{pmatrix} \\
&= \begin{pmatrix} -(\bar{c_1}\bar{y_1} + \bar{c_2}\bar{y_2}) \\ \bar{c_1}\bar{x_1} + \bar{c_2}\bar{x_2} \end{pmatrix} = \begin{pmatrix} -\bar{c_1}\bar{y_1} \\ \bar{c_1}\bar{x_1} \end{pmatrix} + \begin{pmatrix} -\bar{c_2}\bar{y_2} \\ \bar{c_2}\bar{x_2} \end{pmatrix} = \bar{c_1} \begin{pmatrix} -\bar{y_1} \\ \bar{x_1} \end{pmatrix} + \bar{c_2} \begin{pmatrix} -\bar{y_2} \\ \bar{x_2} \end{pmatrix} \\
&= \bar{c_1}\Psi(v_1) + \bar{c_2}\Psi(v_2)
\end{aligned}$$

**Definition 2.20** (Quaternionic Structure)**.** *[14] A* quaternionic structure *on a complex vector space $V$ is an antilinear map $S : V \to V$ such that $S^2(v) = -v$.*

*Remark.* A complex vector space that has a quaternionic structure can essentially "mimic" a quaternionic vector space (specifically, a right quaternionic vector space) [14] - similar to how a real vector space with a complex structure can "mimic" a complex vector space.

**Example 2.5.2.** The antilinear map $\Psi$ in Example 2.5.1 actually gives a quaternionic structure in the complex vector space $\mathbb{C}^2$ since for $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2$,

$$\Psi^2(v) = \Psi(\Psi(v)) = \overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -\bar{y} \\ \bar{x} \end{pmatrix}} = \overline{\begin{pmatrix} -\bar{x} \\ -\bar{y} \end{pmatrix}} = -\begin{pmatrix} x \\ y \end{pmatrix} = -v$$

Since by Theorem 2.6 we can express a quaternion $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ as $q = (a + b\mathbf{i}) + \mathbf{j}(c - d\mathbf{i})$. We can then think of the quaternion $q$ as a 2-dimensional complex vector $\vec{v} = \begin{pmatrix} a + b\mathbf{i} \\ c - d\mathbf{i} \end{pmatrix}$ and take the image of $\vec{v}$ under $\Psi$. We then get,

$$\overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a + b\mathbf{i} \\ c - d\mathbf{i} \end{pmatrix}} = \overline{\begin{pmatrix} -(c - d\mathbf{i}) \\ a + b\mathbf{i} \end{pmatrix}} = \begin{pmatrix} -(c + d\mathbf{i}) \\ a - b\mathbf{i} \end{pmatrix}$$

which corresponds to multiplying $q$ by $\mathbf{j}$ on the right, i.e., $q\mathbf{j} = a\mathbf{j} + b\mathbf{k} - c - d\mathbf{i} = -(c + d\mathbf{i}) + \mathbf{j}(a - b\mathbf{i})$.

*Remark.* From Example 2.5.2, we see that the quaternionic structure $\Psi$ corresponds to right scalar multiplication by the quaternion $\mathbf{j}$ [1]. In Chapter 3, when we discuss how to represent quaternions as complex matrices, we will see that right multiplication by $\mathbf{j}$ has no complex matrix representation [1]. This is problematic because all quaternionic linear maps (matrices) can only commute with right scalar multiplication [18] [1]. We will, therefore, have to use $\Psi$. Also note that for $\Psi$ to correspond to right multiplication by $\mathbf{j}$, we have to use $(a+b\mathbf{i})+\mathbf{j}(c-d\mathbf{i})$ to express a quaternion as a 2-dimensional algebra over $\mathbb{C}$ instead of $(a+b\mathbf{i})+(c+d\mathbf{i})\mathbf{j}$.

Before we proceed, let's revisit Property 2 of Definition 2.6, i.e. $L(cv) = cL(v)$ where $c$ is a scalar. Recall that this property requires a linear map $L$ to commute with scalar multiplication. We can represent scalar multiplication as a linear map, with the matrix representation $cI_n = I_n c$. We can now restate $L(cv) = cL(v)$ as having the matrix representation $A$ of the linear map $L$ commute with the matrix $cI_n$, i.e., $A(cI_n) = (cI_n)A$. Therefore, in order for a mapping to satisfy Property 2, its matrix representation must commute with $cI_n$.

A complex linear map, for instance, commutes with scalar multiplication, especially with scalar multiplication by $\mathbf{i}$ since for any complex matrix $E$, $E(\mathbf{i}I_n) = (\mathbf{i}I_n)E$ [1]. A quaternionic linear map commutes with right scalar multiplication especially by $\mathbf{j}$ (however, again as we'll see in Chapter 3, there is no matrix representation for right scalar multiplication), i.e., for quaternionic linear map $L_{\mathbb{H}}$ and $v$ a quaternionic vector $L_{\mathbb{H}}(v\mathbf{j}) = L_{\mathbb{H}}(v)\mathbf{j}$. Therefore, if we are to represent a complex linear map as a real linear map, the said map has to *commute with the complex structure* defined in the corresponding real space. Likewise, if we are to represent a quaternionic linear map as a complex linear map, such a linear map also has to *commute with the quaternionic structure* defined in the corresponding complex space [1].

*Remark.* What it means for a complex linear map $N$ to commute with the quaternionic structure is having $N\Psi(v) = N\overline{Jv} = \overline{JNv}$. This means that taking the image of $v$ under $\Psi$ first before multiplying by $N$ is the same as multiplying $v$ by $N$ first before taking the image of $Nv$ under $\Psi$ [1].

### 2.5.1 Quaternionic Matrices

*Quaternionic Matrices* are matrices that have entries in $\mathbb{H}$. They are linear maps $L_{\mathbb{H}} : V \to W$ where $V$ and $W$ are *right quaternionic vector spaces.* Like complex matrices we can also take the $\mathbb{H}$-conjugate of each of the entries. The resulting matrix is called a *conjugate quaternionic matrix.*

Theorems 2.7 and 2.8 are extremely useful for multiplying square matrices especially those that, in turn, have square matrices as entries.

**Theorem 2.7.** *For matrices $A, B, C \in M_n(\mathbb{H})$, $A(B+C) = AB + AC$.*

*Proof.* Let $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}] \in \mathrm{M}_n(\mathbb{C})$. Then $B + C = [b_{ij} + c_{ij}]$ and

$$\begin{aligned} A(B+C) &= [\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})] = [\sum_{k=1}^{n}(a_{ik}b_{kj} + a_{ik}c_{kj})] \\ &= [\sum_{k=1}^{n} a_{ik}b_{kj} + \sum_{k=1}^{n} a_{ik}c_{kj}] = [\sum_{k=1}^{n} a_{ik}b_{kj}] + [\sum_{k=1}^{n} a_{ik}c_{kj}] = AB + AC \end{aligned} \tag{2.2}$$

□

The same method of proof can be used for the right distributive law. Furthermore, since $\mathbb{R} \subseteq \mathbb{C} \subseteq \mathbb{H}$, Theorem 2.7 holds for matrices in $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$.

**Theorem 2.8.** *For matrices $A_{ij}, B_{ij} \in M_n(\mathbb{H})$ where $i, j = 1, 2, ..., m$,*

$$\begin{pmatrix} A_{11} & \dots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \dots & A_{mm} \end{pmatrix} \begin{pmatrix} B_{11} & \dots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{m1} & \dots & B_{mm} \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{m} A_{1k}B_{k1} & \cdots & \sum_{k=1}^{m} A_{1k}B_{km} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} A_{mk}B_{k1} & \cdots & \sum_{k=1}^{m} A_{mk}B_{km} \end{pmatrix}$$

*Proof.*

$$\begin{pmatrix} A_{11} & \dots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \dots & A_{mm} \end{pmatrix} \begin{pmatrix} B_{11} & \dots & B_{1m} \\ \vdots & \ddots & \vdots \\ B_{m1} & \dots & B_{mm} \end{pmatrix}$$

$$= \begin{pmatrix} a_{1111} & \dots & a_{111n} & & a_{1m11} & \dots & a_{1m1n} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ a_{11n1} & \dots & a_{11nn} & & a_{1mn1} & \dots & a_{1mnn} \\ & \vdots & & \ddots & & \vdots & \\ a_{m111} & \dots & a_{m11n} & & a_{mm11} & \dots & a_{mm1n} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ a_{m1n1} & \dots & a_{m1nn} & & a_{mmn1} & \dots & a_{mmnn} \end{pmatrix} \begin{pmatrix} b_{1111} & \dots & b_{111n} & & b_{1m11} & \dots & b_{1m1n} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ b_{11n1} & \dots & b_{11nn} & & b_{1mn1} & \dots & b_{1mnn} \\ & \vdots & & \ddots & & \vdots & \\ b_{m111} & \dots & b_{m11n} & & b_{mm11} & \dots & b_{mm1n} \\ \vdots & \ddots & \vdots & \dots & \vdots & \ddots & \vdots \\ b_{m1n1} & \dots & b_{m1nn} & & b_{mmn1} & \dots & b_{mmnn} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{m}[\sum_{l=1}^{n} a_{1kil}b_{k1lj}] & \dots & \sum_{k=1}^{m}[\sum_{l=1}^{n} a_{1kil}b_{kmlj}] \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{m}[\sum_{l=1}^{n} a_{mkil}b_{k1lj}] & \dots & \sum_{k=1}^{m}[\sum_{l=1}^{n} a_{mkil}b_{kmlj}] \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^{m} A_{1k}B_{k1} & \cdots & \sum_{k=1}^{m} A_{1k}B_{km} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{m} A_{mk}B_{k1} & \cdots & \sum_{k=1}^{m} A_{mk}B_{km} \end{pmatrix}$$

□

Theorem 2.8 also holds for matrices in $\mathrm{M}_n(\mathbb{R})$ and $\mathrm{M}_n(\mathbb{C})$ because $\mathbb{R} \subseteq \mathbb{C} \subseteq \mathbb{H}$.

# Chapter 3

# Results and Discussions

In this chapter, we discuss the theory behind quaternionic determinants - specifically, the Study Determinant - and the homomorphisms involved.

## 3.1 The Cayley Determinant and Aslaksen's Axioms

In 1845, 2 years after William Rowan Hamilton discovered quaternions, Arthur Cayley attempted to define the determinant of a quaternionic matrix using the usual formula (we denote the Cayley determinant by $Cdet$). Note that the following definition can be extended for $n \times n$ quaternionic matrices.

**Definition 3.1** (2 × 2 Cayley Determinant)**.** *[1] For a* $2 \times 2$ *quaternionic matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $Cdet(A) = ad - cb$ *for* $a, b, c, d \in \mathbb{H}$.

Notice that the order in which the entries are multiplied matters. We can see that for a $3 \times 3$ quaternionic matrix

$$B = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}, Cdet(B) = (aei + bfg + cdh) - (gec + hfa + idb).$$

**Example 3.1.1.** [1] Let $M = \begin{pmatrix} \mathbf{k} & \mathbf{j} \\ \mathbf{i} & 1 \end{pmatrix}$. Then, $Cdet(M) = \mathbf{k} - \mathbf{ij} = \mathbf{k} - \mathbf{k} = 0$. Hence, we can say that by the Cayley determinant, $M$ is singular.

**Example 3.1.2.** [1] Consider the transpose of the matrix $M$, $M^T = \begin{pmatrix} \mathbf{k} & \mathbf{i} \\ \mathbf{j} & 1 \end{pmatrix}$. Then $Cdet(M^T) = \mathbf{k} - (\mathbf{ji}) = \mathbf{k} - (-\mathbf{k}) = 2\mathbf{k}$. Hence, we can say that by the Cayley determinant, $M^T$ is invertible.

Taking into account the fact that the quaternions are non-commutative, one might ask whether or not this determinant behaves the way we expect - Will it really determine whether or not a quaternionic matrix is singular or not? Will the properties of the determinant still hold? Will the determinant still be a map from $M_n(G) \rightarrow G$ (in this case, $G = \mathbb{H}$) where $M_n(G)$ is the set of all $n \times n$ matrices over the elements of $G$? The last question comes from the fact that the determinant of a complex matrix yields a complex number.

### 3.1.1 Brenner's Determinant Function and Aslaksen's Axioms

We take a step back and revisit what it means for a mapping to be a determinant. J.L. Brenner [9] and Helmer Aslaksen [1] offer different approaches to this, however, we will see how we can both arrive at the same conclusions. The following definition is taken from [9].

**Definition 3.2** (Determinant Function). *For a field $F$, a determinant over the matrices of $M_n(F)$ is a function det from $M_n(F)$ into $F$ such that*

$$det(AB) = det(A)det(B) = det(B)det(A) \tag{3.1}$$

*holds either*

1. $\forall A, B \in M_n(F)$ *or*

2. $\forall$ *invertible* $A, B \in M_n(F)$

.

Aslaksen, on the other hand, uses a more axiomatic approach in defining a determinant, presenting three determinant *axioms* which a determinant definition must satisfy in order for it to behave the way we expect, i.e., it has the properties we associate with determinants.

- **Axiom 1.** $det(A) = 0$ if and only if $A$ is singular.
- **Axiom 2.** $det(AB) = det(A)det(B)$ for all quaternionic matrices $A$ and $B$.

- **Axiom 3.** If $A'$ is obtained by adding a left-multiple of a row to another row or a right-multiple of a column to another column, then $det(A') = det(A)$ (as we have already encountered in linear algebra, this operation can be described by an elementary matrix [1]).

Notice that a determinant is essentially a function that:

1. Maps to 0 if a matrix is singular
2. Preserves multiplication and,
3. Remains unchanged after applying the elementary operation of adding a left/right-multiple of one row/column to another row/column respectively.

Also notice that Alaksen's second axiom is the first condition in Brenner's determinant function.

We can define the determinant simply as a function that constantly maps to 0 for all singular matrices and 1 for all non-singular matrices [9]. This will satisfy the above axioms [9] [1], however, we will mostly deal with determinants that are non-trivial (for instance the $det_{\mathbb{R}}$, $det_{\mathbb{C}}$, and $Cdet$).

**Theorem 3.1.** *[9] If det is not constantly equal to 1 or 0 (i.e., det is not a mapping $det : M_n(F) \to F$ where $F$ is a field with two elements), then $det(B) = 0$ for all singular matrices.*

Theorem 3.1 not only shows how the determinant function in 3.2 holds for the non-trivial case, it also shows that conditions (1) and (2) of Definition 3.2 are essentially equivalent [9]. This means that a determinant should map all singular matrices to 0.

**Theorem 3.2.** *[1] Let $M_n(\mathbb{H})$ be the set of all $n \times n$ quaternionic matrices. If det satisfies all of Aslaksen's axioms, then $det(M_n(\mathbb{H}))$ is a commutative subset of $\mathbb{H}$.*

Notice how Theorem 3.2 is already implied in Brenner's determinant function in which it is already deemed necessary for the images to commute. By Theorem 3.2, we see that *det* satisfying Aslaksen's axioms must only map to a commutative subset of $\mathbb{H}$, which is the complex numbers. Therefore, *det* cannot be a mapping onto $\mathbb{H}$. Since $Cdet$ is onto $\mathbb{H}$, by contrapositive of theorem 3.2, $Cdet$ does not satisfy at least one of the axioms [1].

**Example 3.1.3.** [1] As an illustration, we show that $Cdet$ doesn't satisfy Axiom 1. Recall examples 3.1.1 and 3.1.2. Notice that,

$$
\begin{aligned}
M\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\implies \begin{pmatrix} \mathbf{k} & \mathbf{j} \\ \mathbf{i} & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\implies & \mathbf{k}x + \mathbf{j}y = 0 \text{ and } \mathbf{i}x + y = 0 \\
\implies & x = 0 \text{ and } y = 0. \\
\implies & M \text{ is invertible. This contradicts with the fact that } Cdet(M) = 0.
\end{aligned}
$$

Also notice that,

$$
M^T\begin{pmatrix} -1 \\ \mathbf{j} \end{pmatrix} = \begin{pmatrix} \mathbf{k} & \mathbf{i} \\ \mathbf{j} & 1 \end{pmatrix}\begin{pmatrix} -1 \\ \mathbf{j} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
$$

$$
\implies M^T \text{ is singular. This contradicts with the fact that } Cdet(M^T) = 2\mathbf{k}.
$$

In [1], it is also shown that $Cdet$ doesn't satisfy Axioms 2 and 3.

## 3.2 Matrix Homomorphisms

Because the Cayley determinant fails to satisfy Aslaksen's axioms, the classical definition of the determinant cannot be extended to quaternionic matrices. It was not until 1920, that a new approach in defining a quaternionic determinant was presented in [1]. The idea is to transform quaternionic matrices into complex matrices from which one could then just simply take the determinant [1]. The method involves homomorphisms between quaternionic, complex, and real matrices.

In this section, we take a closer look into these homomorphisms - first discussing the motivation behind them and then the theory.

### 3.2.1 Representing Complex Numbers as Real Matrices

Recall in Chapter 2 that $\mathbb{C}$ forms a complex vector space. For us to represent a complex number $a + b\mathbf{i}$ as a real matrix, we first have to express it as a linear map $f : \mathbb{C} \to \mathbb{C}$, i.e., a $1 \times 1$ complex matrix $[a + b\mathbf{i}]$. Thus, we have $f(z) = [a + b\mathbf{i}]z$ which is identical to complex multiplication by $a + b\mathbf{i}$. We see that the images of 1 and $\mathbf{i}$ under $f$ are $a + b\mathbf{i}$ and $-b + a\mathbf{i}$ respectively. We know from abstract algebra that we can define a bijection from the field of complex numbers to the 2D-plane ($\mathbb{R}^2$) - a mapping $\Theta : \mathbb{C} \to \mathbb{R}^2$ where a complex number $a + b\mathbf{i}$ is mapped to a vector/point $(a, b)$ in the 2D-plane. Under the function $\Theta$ (in which case 1 is mapped to $(1, 0)$ while $\mathbf{i}$ is mapped to $(0, 1)$), we seek a $2 \times 2$ real matrix that maps $(1, 0)$ to $\Theta(a+b\mathbf{i}) = (a, b)$ and $(0, 1)$ to $\Theta(-b+a\mathbf{i}) = (-b, a)$. Let this matrix be $F = \begin{pmatrix} \alpha & \beta \\ \chi & \delta \end{pmatrix}$ where $\alpha, \beta, \chi,$ and $\delta \in \mathbb{R}$. Then,

$$\begin{pmatrix} \alpha & \beta \\ \chi & \delta \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \implies \begin{pmatrix} \alpha \\ \chi \end{pmatrix} = \begin{pmatrix} a \\ b \end{pmatrix} \implies \alpha = a; \chi = b \text{ and}$$

$$\begin{pmatrix} \alpha & \beta \\ \chi & \delta \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix} \implies \begin{pmatrix} \beta \\ \delta \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix} \implies \beta = -b; \delta = a$$

Therefore, $F = \begin{pmatrix} \alpha & \beta \\ \chi & \delta \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

*Remark.* The matrix $F$ commutes with the complex structure in $\mathbb{R}^2$, $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, since,

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} -b & -a \\ a & -b \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

Therefore, the matrix $F$ is a real linear map that represents the complex linear map $f$.

**Example 3.2.1.** Take the complex number $z = -3 + 2\mathbf{i}$. Then its real matrix representation is $\begin{pmatrix} -3 & -2 \\ 2 & -3 \end{pmatrix}$. We can use this real matrix representation to multiply $z$ with another complex number, say $1 + 2\mathbf{i}$. $\begin{pmatrix} -3 & -2 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -7 \\ -4 \end{pmatrix}$ which corresponds to the complex number obtained by multiplying $z(1 + 2\mathbf{i}) = (1 + 2\mathbf{i})z = -7 - 4\mathbf{i}$.

It is important to note that $F$ in general, doesn't represent the complex number $a+b\mathbf{i}$ itself but the linear map associated with multiplying a complex number by $a+b\mathbf{i}$. In this case, $F$ can represent both left and right multiplication because complex numbers are commutative under multiplication.

In representing the complex number $a+b\mathbf{i}$ as a real matrix, we associate it with the real matrix $F$, i.e., we define a map $\phi : \mathbb{C} \to \mathrm{M}_2(\mathbb{R})$ where $\mathrm{M}_2(\mathbb{R})$ is the set of all $2 \times 2$ real matrices, such that $a+b\mathbf{i} \mapsto \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$

In order for a real matrix representation to represent one and only one complex number $a+b\mathbf{i}$, $\phi$ must be *injective.* $\phi$ must also preserve the structure that complex multiplication gives to $\mathbb{C}$, i.e., $\phi$ must also be a homomorphism. We've already shown in Chapter 2 Example 2.1.2 that $\phi$ is indeed a homomorphism. In the next subsection, we show that $\phi$ is, in fact, an *injective homomorphism* in the general case.

### 3.2.2 Homomorphisms from $\mathbf{M}_n(\mathbb{C})$ to $\mathbf{M}_{2n}(\mathbb{R})$

Let $\mathrm{M}_n(\mathbb{C})$ denote the set of all $n \times n$ complex matrices and $\mathrm{M}_{2n}(\mathbb{R})$ denote the set of all $2n \times 2n$ real matrices. When we represent complex matrices as real matrices, we are essentially representing complex linear maps as real linear maps. Recall in Chapter 2, that in order for a real linear map to represent a complex linear map, the said real map must commute with the complex structure defined in the corresponding real space $\mathbb{R}^{2n}$. We can define a matrix

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$$

We see that,

$$J^2 = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} = \begin{pmatrix} -I_n & 0 \\ 0 & -I_n \end{pmatrix} = -I_{2n}$$

Therefore, by Definition 2.12 $J$ gives a *complex structure* in $\mathbb{R}^{2n}$. Let $\phi : \mathrm{M}_n(\mathbb{C}) \to \mathrm{M}_{2n}(\mathbb{R})$ such that the image of a complex matrix under $\phi$ is its real matrix representation. Then $\phi(\mathrm{M}_n(\mathbb{C})) = \{P \in \mathrm{M}_{2n}(\mathbb{R}) | JP = PJ\}$, i.e., the real matrix representations of complex matrices are those $2n \times 2n$ real matrices that commute with $J$ [1].

The question now is which $2n \times 2n$ real matrices matrices commute with the complex structure in $\mathbb{R}^2$. Notice that every complex matrix can be represented as the sum of a real matrix and a purely imaginary matrix, i.e., for an $n \times n$ complex matrix $Z$, $Z = A + B\mathbf{i}$ where $A, B \in \mathrm{M}_n(\mathbb{R})$ [1]. We define a mapping

$$[1]\phi(A + B\mathbf{i}) = \begin{pmatrix} A & -B \\ B & A \end{pmatrix} \tag{3.2}$$

Notice that $F$ in 3.2.1 is a special case of 3.2 where $n = 1$. If we can show that $\phi$ is an injective homomorphism (i.e., it preserves multiplication and is one-to-one), we can essentially represent any complex matrix as a $2n \times 2n$ real matrix. Also notice how the dimension of the real matrix is necessarily even. This is a direct consequence of Theorem 2.2.

**Theorem 3.3.** *Let $\phi : M_n(\mathbb{C}) \to M_{2n}(\mathbb{R})$ such that $C+D\boldsymbol{i} \mapsto \begin{pmatrix} C & -D \\ D & C \end{pmatrix}$ where $C+D\boldsymbol{i} \in M_n(\mathbb{C})$. Then $\phi$ is an injective homomorphism.*

*Proof.* First, we show that $\phi$ is injective. Let $A + B\mathbf{i}, C + D\mathbf{i} \in \mathrm{M}_n(\mathbb{C})$. Then,

$$\begin{aligned}
&\phi(A + B\mathbf{i}) = \phi(C + D\mathbf{i}) \\
\implies &\begin{pmatrix} A & -B \\ B & A \end{pmatrix} = \begin{pmatrix} C & -D \\ D & C \end{pmatrix} \\
\implies &A = C \text{ and } B = D \text{ by Matrix Equality} \\
\implies &A + B\mathbf{i} = C + D\mathbf{i} \\
\implies &\phi \text{ is injective.}
\end{aligned}$$

We now show that $\phi$ is a homomorphism. Let $A + B\mathbf{i}$, $C + D\mathbf{i} \in \mathrm{M}_n(\mathbb{C})$. Then

$$\begin{aligned}
\phi[(A + B\mathbf{i})(C + D\mathbf{i})] &= \phi[(A + B\mathbf{i})C + (A + B\mathbf{i})D\mathbf{i}] \text{ by Theorem 2.7} \\
&= \phi[AC + BC\mathbf{i} + AD\mathbf{i} - BD] \\
&= \phi[(AC - BD) + (BC + AD)\mathbf{i}] \\
&= \begin{pmatrix} AC - BD & -(BC + AD) \\ BC + AD & AC - BD \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}\phi[(A+B\mathbf{i})]\phi[(C+D\mathbf{i})] &= \begin{pmatrix} A & -B \\ B & A \end{pmatrix}\begin{pmatrix} C & -D \\ D & C \end{pmatrix} \\ &= \begin{pmatrix} AC-BD & -(BC+AD) \\ BC+AD & AC-BD \end{pmatrix} \text{ by Theorem 2.8.}\end{aligned}$$

Therefore, $\phi[(A+B\mathbf{i})(C+D\mathbf{i})] = \phi[(A+B\mathbf{i})]\phi[(C+D\mathbf{i})]$ implying that $\phi$ is a homomorphism.

Thus, $\phi$ is an injective homomorphism. □

**Example 3.2.2.** Take the complex matrix

$$Z = \begin{pmatrix} 1+2\mathbf{i} & 3\mathbf{i} \\ 1 & -2+\mathbf{i} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & -2 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}\mathbf{i}. \text{ Then, } \phi(Z) = \begin{pmatrix} 1 & 0 & -2 & -3 \\ 1 & -2 & 1 & -1 \\ 2 & 3 & 1 & 0 \\ 0 & 1 & 1 & -2 \end{pmatrix}$$

We know that $\phi(Z)$ is the only real matrix representation of $Z$ and that $\phi(Z)$ corresponds to the complex map $Z$ in the complex vector space $\mathbb{C}^2$ since $\phi$ is an injective homomorphism.

### 3.2.3 Representing Quaternions as Complex Matrices

Recall in Chapter 2 that we can view the quaternions as a 2-dimensional algebra over $\mathbb{C}$ by having $a+b\mathbf{i}+c\mathbf{j}+d\mathbf{k} = (a+b\mathbf{i})+\mathbf{j}(c-d\mathbf{i})$. In general, we can write any quaternion as $x+\mathbf{j}y$ where $x, y \in \mathbb{C}$. Because of this, we see that we can define a bijection $\Omega : \mathbb{H} \to \mathbb{C}^2$ where a quaternion $q = x+\mathbf{j}y$ is mapped to $(x, y) \in \mathbb{C}^2$. In representing a quaternion $q = x+\mathbf{j}y$ as a complex matrix, we first have to express it as a linear map $s : \mathbb{H} \to \mathbb{H}$, i.e., a $1\times 1$ quaternionic matrix $[x+\mathbf{j}y]$. Thus, we have $s(q) = [x+\mathbf{j}y]q$ which is identical to left quaternionic multiplication by $x+\mathbf{j}y$. We see that,

$$\begin{aligned} s(1) &= x+\mathbf{j}y \\ s(\mathbf{i}) &= (x+\mathbf{j}y)\mathbf{i} = x\mathbf{i}+\mathbf{j}(y\mathbf{i}) \\ s(\mathbf{j}) &= (x+\mathbf{j}y)\mathbf{j} = x\mathbf{j}+\mathbf{j}y\mathbf{j} = \mathbf{j}\bar{x}+\mathbf{j}^2\bar{y} = -\bar{y}+\mathbf{j}\bar{x} \\ s(\mathbf{k}) &= (x+\mathbf{j}y)\mathbf{k} = x\mathbf{k}+\mathbf{j}y\mathbf{k} = -x\mathbf{j}\mathbf{i}+\bar{y}\mathbf{j}\mathbf{k} = \bar{y}\mathbf{i}-\mathbf{j}(\bar{x}\mathbf{i}) \end{aligned}$$

Under the function $\Omega$, we see that the images of 1, $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$ are $(1,0), (\mathbf{i},0), (0,1)$ and $(0,-\mathbf{i})$ respectively and the images of $s(1), s(\mathbf{i}), s(\mathbf{j})$ and $s(\mathbf{k})$ are $(x,y), (x\mathbf{i}, y\mathbf{i}), (-\bar{y}, \bar{x})$ and $(\bar{y}\mathbf{i}, -\bar{x}\mathbf{i})$. We seek a matrix in $\mathrm{M}_2(\mathbb{C})$ such that $(1,0) \mapsto (x,y)$, $(\mathbf{i},0) \mapsto (x\mathbf{i}, y\mathbf{i})$, $(0,1) \mapsto (-\bar{y}, \bar{x})$, and $(0,-\mathbf{i}) \mapsto (\bar{y}\mathbf{i}, -\bar{x}\mathbf{i})$.

Let this matrix be $S = \begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix}$, where $\kappa, \lambda, \mu$, and $\nu \in \mathbb{C}$. Then,

$$\begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{pmatrix} \kappa \\ \mu \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \implies \kappa = x \text{ and } \mu = y$$

$$\begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix}\begin{pmatrix} \mathbf{i} \\ 0 \end{pmatrix} = \begin{pmatrix} x\mathbf{i} \\ y\mathbf{i} \end{pmatrix} \implies \begin{pmatrix} \kappa\mathbf{i} \\ \mu\mathbf{i} \end{pmatrix} = \begin{pmatrix} x\mathbf{i} \\ y\mathbf{i} \end{pmatrix} \implies \kappa = x \text{ and } \mu = y$$

$$\begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\bar{y} \\ \bar{x} \end{pmatrix} \implies \begin{pmatrix} \lambda \\ \nu \end{pmatrix} = \begin{pmatrix} -\bar{y} \\ \bar{x} \end{pmatrix} \implies \lambda = -\bar{y} \text{ and } \nu = \bar{x}$$

$$\begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix}\begin{pmatrix} 0 \\ -\mathbf{i} \end{pmatrix} = \begin{pmatrix} \bar{y}\mathbf{i} \\ -\bar{x}\mathbf{i} \end{pmatrix} \implies \begin{pmatrix} -\lambda\mathbf{i} \\ -\nu\mathbf{i} \end{pmatrix} = \begin{pmatrix} \bar{y}\mathbf{i} \\ -\bar{x}\mathbf{i} \end{pmatrix} \implies \lambda = -\bar{y} \text{ and } \nu = \bar{x}$$

Hence, $S = \begin{pmatrix} \kappa & \lambda \\ \mu & \nu \end{pmatrix} = \begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix}$.

*Remark.* The matrix $S$ commutes with the quaternionic structure $\Psi$ in Chapter 2 Example 2.5.2 since for a complex vector $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{C}^2$,

$$\begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix}\overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}} = \begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix}\begin{pmatrix} -\bar{y} \\ \bar{x} \end{pmatrix} = \begin{pmatrix} -x\bar{y} - \bar{y}\bar{x} \\ -y\bar{y} + \bar{x}^2 \end{pmatrix} \text{ and}$$

$$\overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}} = \overline{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x^2 - \bar{y}y \\ yx + \bar{x}y \end{pmatrix}} = \begin{pmatrix} -x\bar{y} - \bar{y}\bar{x} \\ -y\bar{y} + \bar{x}^2 \end{pmatrix}$$

Hence, the matrix $S$ is a complex linear map that represents the quaternionic linear map $s$.

**Example 3.2.3.** Take the quaternion $q = -2 + \mathbf{i} - 5\mathbf{j} + 2\mathbf{k} = -2 + \mathbf{i} + \mathbf{j}(-5 - 2\mathbf{i})$. Then its complex matrix representation is $\begin{pmatrix} -2 + \mathbf{i} & 5 - 2\mathbf{i} \\ -5 - 2\mathbf{i} & -2 - \mathbf{i} \end{pmatrix}$. We can use this matrix to multiply $q$ with another quaternion, say $1 + \mathbf{i} + 3\mathbf{j} - 4\mathbf{k} = 1 + \mathbf{i} + \mathbf{j}(3 + 4\mathbf{i})$. We have, $\begin{pmatrix} -2+\mathbf{i} & 5-2\mathbf{i} \\ -5-2\mathbf{i} & -2-\mathbf{i} \end{pmatrix} \begin{pmatrix} 1+\mathbf{i} \\ 3+4\mathbf{i} \end{pmatrix} = \begin{pmatrix} 20+13\mathbf{i} \\ -5-18\mathbf{i} \end{pmatrix}$ which corresponds to the quaternion obtained by multiplying $q(1+\mathbf{i}+3\mathbf{j}-4\mathbf{k}) = 20+13\mathbf{i}-5\mathbf{j}+18\mathbf{k} = 20+13\mathbf{i}+\mathbf{j}(-5-18\mathbf{i})$ on the left.

Notice that $S$ only represents left multiplication in the quaternions. One would probably ask whether or not we can have a complex matrix representation for right multiplication.

Consider the quaternionic function defined by multiplying a quaternion $q$ by some quaternion $x + \mathbf{j}y$ on the right, $s_R(q) = q(x + \mathbf{j}y)$ such that $y \neq 0$. We see that,

$$\begin{aligned} s_R(1) &= x + \mathbf{j}y \\ s_R(\mathbf{i}) &= \mathbf{i}x - \mathbf{j}\mathbf{i}y \\ s_R(\mathbf{j}) &= -y + \mathbf{j}x \\ s_R(\mathbf{k}) &= -\mathbf{i}y - \mathbf{j}\mathbf{i}x \end{aligned}$$

We use the same method in obtaining the complex matrix representation. We let $S_R = \begin{pmatrix} \kappa_R & \lambda_R \\ \mu_R & \nu_R \end{pmatrix}$ be the complex matrix representation of this function.

$$\begin{pmatrix} \kappa_R & \lambda_R \\ \mu_R & \nu_R \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \implies \begin{pmatrix} \kappa_R \\ \mu_R \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \implies \kappa_R = x \text{ and } \mu_R = y$$

$$\begin{pmatrix} \kappa_R & \lambda_R \\ \mu_R & \nu_R \end{pmatrix} \begin{pmatrix} \mathbf{i} \\ 0 \end{pmatrix} = \begin{pmatrix} \mathbf{i}x \\ -\mathbf{i}y \end{pmatrix} \implies \begin{pmatrix} \kappa_R\mathbf{i} \\ \mu_R\mathbf{i} \end{pmatrix} = \begin{pmatrix} \mathbf{i}x \\ -\mathbf{i}y \end{pmatrix} \implies \kappa_R = x \text{ and } \mu_R = -y$$

$$\begin{pmatrix} \kappa_R & \lambda_R \\ \mu_R & \nu_R \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \implies \begin{pmatrix} \lambda_R \\ \nu_R \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} \implies \lambda_R = -y \text{ and } \nu_R = x$$

$$\begin{pmatrix} \kappa_R & \lambda_R \\ \mu_R & \nu_R \end{pmatrix} \begin{pmatrix} 0 \\ -\mathbf{i} \end{pmatrix} = \begin{pmatrix} -\mathbf{i}y \\ -\mathbf{i}x \end{pmatrix} \implies \begin{pmatrix} -\lambda_R\mathbf{i} \\ -\nu_R\mathbf{i} \end{pmatrix} = \begin{pmatrix} -\mathbf{i}y \\ -\mathbf{i}x \end{pmatrix} \implies \lambda_R = y \text{ and } \nu_R = x$$

This implies that $\mu_R = \lambda_R = 0$. Hence, the complex matrix representation of right multiplication by a quaternion is $\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$. However, we see that, $\begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$

which doesn't correspond to $s_R(1) = x + \mathbf{j}y$. Thus, we see that there is no matrix representation for right multiplication [1]. This is because right scalar multiplication itself, if expressed as a mapping $L_R$, does not satisfy Property 2 of Definition 2.6, i.e., $L_R(cv) \neq cL_R(v)$ since quaternion multiplication is not commutative. In representing a quaternion $x + \mathbf{j}y$ as a complex matrix, we associate it with the complex matrix $S$, i.e., we define a map $\psi : \mathbb{H} \to \mathrm{M}_2(\mathbb{C})$ where $\mathrm{M}_2(\mathbb{C})$ is the set of all $2 \times 2$ complex matrices, such that $x + \mathbf{j}y \mapsto \begin{pmatrix} x & -\bar{y} \\ y & \bar{x} \end{pmatrix}$.

In order for a complex matrix representation to represent one and only one quaternion $x + \mathbf{j}y$, $\psi$ must be injective. $\psi$ must also preserve the structure that right quaternion multiplication gives to $\mathbb{H}$, i.e., $\psi$ must be a homomorphism. In the next subsection, we show that $\psi$ is indeed an injective homomorphism.

### 3.2.4 Homomorphisms from $\mathbf{M}_n(\mathbb{H})$ to $\mathbf{M}_{2n}(\mathbb{C})$

Let $\mathrm{M}_n(\mathbb{H})$ be the set of all $n \times n$ quaternionic matrices and $\mathrm{M}_{2n}(\mathbb{C})$ be the set of all $2n \times 2n$ complex matrices. When we represent quaternionic matrices as complex matrices, we are essentially representing quaternionic linear maps as complex linear maps. Recall in Chapter 2, that in order for a complex linear map to represent a quaternionic linear map, the said complex map must commute with the quaternionic structure defined in the corresponding complex space $\mathbb{C}^{2n}$.

For $v = \begin{pmatrix} \chi \\ \gamma \end{pmatrix} \in \mathbb{C}^{2n}$ where $\chi = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $\gamma = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, we can define a mapping,

$$\Psi(v) = \overline{\begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} \chi \\ \gamma \end{pmatrix}} = \begin{pmatrix} -\overline{\gamma} \\ \overline{\chi} \end{pmatrix}$$

We see that $\Psi$ satisfies Definition 2.20, i.e., $\Psi$ gives a *quaternionic structure* in $\mathbb{C}^{2n}$ since,

$$\Psi^2(v) = \Psi(\Psi(v)) = \overline{\begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} -\overline{\gamma} \\ \overline{\chi} \end{pmatrix}} = \begin{pmatrix} -\chi \\ -\gamma \end{pmatrix} = -v.$$

Let $\psi : \mathrm{M}_n(\mathbb{H}) \to \mathrm{M}_{2n}(\mathbb{C})$ such that the image of a quaternionic under $\psi$ is its complex matrix representation. Then the complex matrix representations of quaternionic matrices are those $2n \times 2n$ complex matrices $N$ that commute with $\Psi$, i.e.,

$$N\overline{Jv} = \overline{JNv}. \tag{3.3}$$

We can simplify Equation 3.3. We know that a complex matrix commutes with conjugation, i.e., $N\overline{v} = \overline{Nv}$. In other words, conjugating before multiplying by a complex matrix $N$ is the same as first multiplying by the complex matrix $N$ and then conjugating. Thus, $N\overline{Jv} = \overline{NJv}$. Hence, $N\overline{Jv} = \overline{JNv}$ if and only if $\overline{NJv} = J\overline{Nv}$, which means that $NJ = J\overline{N}$. Notice that Theorem 2.6 in Chapter 2 (i.e. for $z \in \mathbb{C}$, $z\mathbf{j} = \mathbf{j}\bar{z}$) is a special case of $NJ = J\overline{N}$ where $n = 1$. Hence, the complex matrices that commute with the quaternionic structure are those $2n \times 2n$ complex matrices $N$, such that $NJ = J\overline{N}$, i.e., $\psi(\mathrm{M}_n(\mathbb{H})) = \{N \in \mathrm{M}_{2n}(\mathbb{C}) | NJ = J\overline{N}\}$.

The question now is which $2n \times 2n$ complex matrices $N$ satisfy $NJ = J\overline{N}$. Notice that every quaternionic matrix can be represented as the sum $Q = X + \mathbf{j}Y$ where $X, Y \in \mathrm{M}_n(\mathbb{C})$ [1]. We define a mapping

$$[1]\psi(X + \mathbf{j}Y) = \begin{pmatrix} X & -\overline{Y} \\ Y & \overline{X} \end{pmatrix} \tag{3.4}$$

Notice that the matrix $S$ is a special case of 3.4. Again, if we can show that $\psi$ is an injective homomorphism, we can essentially represent any quaternionic matrix as a $2n \times 2n$ complex matrix.

**Theorem 3.4.** *Let $\psi : M_n(\mathbb{H}) \to M_{2n}(\mathbb{C})$ such that $X + \boldsymbol{j}Y \mapsto \begin{pmatrix} X & -\overline{Y} \\ Y & \overline{X} \end{pmatrix}$ where $X + \boldsymbol{j}Y \in M_n(\mathbb{H})$. Then $\psi$ is an injective homomorphism.*

*Proof.* We first show that $\psi$ is injective. Let $X + \mathbf{j}Y, V + \mathbf{j}W \in \mathrm{M}_n(\mathbb{H})$. Then,

$$\begin{aligned}
&\psi(X + \mathbf{j}Y) = \psi(V + \mathbf{j}W) \\
&\implies \begin{pmatrix} X & -\overline{Y} \\ Y & \overline{X} \end{pmatrix} = \begin{pmatrix} V & -\overline{W} \\ W & \overline{V} \end{pmatrix} \\
&\implies X = V \text{ and } Y = W \text{ by Matrix Equality} \\
&\implies X + \mathbf{j}Y = V + \mathbf{j}W \\
&\implies \psi \text{ is injective.}
\end{aligned}$$

We now show that $\psi$ is a homomorphism. Let $X+\mathbf{j}Y,\ V+\mathbf{j}W \in \mathrm{M}_n(\mathbb{H})$. Then

$$\begin{aligned}
\psi[(X+\mathbf{j}Y)(V+\mathbf{j}W)] &= \psi[X(V+\mathbf{j}W)+\mathbf{j}Y(V+\mathbf{j}W)] \text{ by Theorem 2.7}\\
&= \psi[XV+X\mathbf{j}W+\mathbf{j}YV+\mathbf{j}Y\mathbf{j}W]\\
&= \psi[XV+\mathbf{j}\overline{X}W+\mathbf{j}YV+\mathbf{j}^2\overline{Y}W]\\
&= \psi[(XV-\overline{Y}W)+\mathbf{j}(\overline{X}W+YV)]\\
&= \begin{pmatrix} XV-\overline{Y}W & -\overline{(\overline{X}W+YV)}\\ \overline{X}W+YV & \overline{XV-\overline{Y}W}\end{pmatrix}\\
&= \begin{pmatrix} XV-\overline{Y}W & -X\overline{W}-\overline{YV}\\ \overline{X}W+YV & \overline{XV}-Y\overline{W}\end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\psi[(X+\mathbf{j}Y)]\psi[(V+\mathbf{j}W)] &= \begin{pmatrix} X & -\overline{Y}\\ Y & \overline{X}\end{pmatrix}\begin{pmatrix} V & -\overline{W}\\ W & \overline{V}\end{pmatrix}\\
&= \begin{pmatrix} XV-\overline{Y}W & -X\overline{W}-\overline{YV}\\ \overline{X}W+YV & \overline{XV}-Y\overline{W}\end{pmatrix} \text{ by Theorem 2.8.}
\end{aligned}$$

□

**Example 3.2.4.** Take the quaternionic matrix

$$Q=\begin{pmatrix} 1+2\mathbf{i}-3\mathbf{j}+\mathbf{k} & 2\mathbf{i}+5\mathbf{k}\\ 1-\mathbf{i} & 3+\mathbf{j}+\mathbf{k}\end{pmatrix}=\begin{pmatrix} 1+2\mathbf{i} & 2\mathbf{i}\\ 1-\mathbf{i} & 3\end{pmatrix}+\mathbf{j}\begin{pmatrix} -3-\mathbf{i} & -5\mathbf{i}\\ 0 & 1-\mathbf{i}\end{pmatrix}.$$

Hence, $\psi(Q)=\begin{pmatrix} 1+2\mathbf{i} & 2\mathbf{i} & 3-\mathbf{i} & -5\mathbf{i}\\ 1-\mathbf{i} & 3 & 0 & -1-\mathbf{i}\\ -3-\mathbf{i} & -5\mathbf{i} & 1-2\mathbf{i} & -2\mathbf{i}\\ 0 & 1-\mathbf{i} & 1+\mathbf{i} & 3\end{pmatrix}$. We know that $\psi(Q)$ is the only complex matrix representation of $Q$ and that $\psi(Q)$ corresponds to the quaternionic map $Q$ in the right quaternionic vector space $\mathbb{H}^2$ since $\psi$ is an injective homomorphism.

## 3.3 The Study Determinant

The *Study Determinant* uses the homomorphisms $\phi$ and $\psi$ to transform quaternionic matrices into complex or real matrices. We can then compute for the determinant of the complex or real matrix obtained by applying $\phi$ and $\psi$.

**Definition 3.3** (Study Determinant)**.** *For $M \in M_n(\mathbb{H})$, the* Study Determinant *is defined by $Sdet(M) = det_{\mathbb{C}}(\psi(M)) = \sqrt{det_{\mathbf{R}}(\phi(\psi(M)))}$.*

Notice that we can compute for the Study determinant in two different ways:

1. Simply getting the complex matrix representation of the quaternionic matrix and proceeding to take its complex determinant or;
2. Getting the real matrix representation of the quaternionic matrix by composing $\phi$ and $\psi$, and then taking the square root of the real determinant.

In the following example, we take the Study determinant of the quaternionic matrix in Example 3.1.1 for which the Cayley determinant failed to identify as non-singular.

**Example 3.3.1.** Let $M = \begin{pmatrix} \mathbf{k} & \mathbf{j} \\ \mathbf{i} & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ \mathbf{i} & 1 \end{pmatrix} + \mathbf{j} \begin{pmatrix} -\mathbf{i} & 1 \\ 0 & 0 \end{pmatrix}$. Then,

1. $Sdet(M) = det_{\mathbb{C}}(\psi(M)) = det_{\mathbb{C}} \begin{pmatrix} 0 & 0 & -\mathbf{i} & -1 \\ \mathbf{i} & 1 & 0 & 0 \\ -\mathbf{i} & 1 & 0 & 0 \\ 0 & 0 & -\mathbf{i} & 1 \end{pmatrix} = 4.$

2. We see that $\psi(M) = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix} \mathbf{i}.$

$$\sqrt{det_{\mathbf{R}}(\phi(\psi(M)))} = \left( det_{\mathbf{R}} \begin{pmatrix} 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \right)^{1/2} = \sqrt{16} = 4$$

Hence, we can see that by the Study determinant, the matrix $M$ is invertible which it really is as shown in section 3.1.

In [1], it's been shown that the Study Determinant satisfies all of Aslaksen's axioms. For instance, Axiom 2 which states that $det(AB) = det(A)det(B)$ is satisfied since the mappings $\phi$ and $\psi$ are homomorphisms.

# Chapter 4

# Summary and Recommendations

In this chapter, we present a summary of what we've discussed in Chapters 2 and 3. We also present some recommendations for further study.

In Chapter 2, we introduced the notion of a complex vector space. We saw that the concept of a complex linear map and the concept of a complex determinant still hold. We also saw the concept of a *complex structure* in a real vector space which allows us to mimic a complex vector space in a real vector space. We saw that the complex structure $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ in $\mathbb{R}^2$ corresponds to multiplication by $\mathbf{i}$ in the complex vector space $\mathbb{C}$. We saw that in order for a real linear map to correspond to a complex linear map, the said real linear map has to commute with the complex structure defined in its corresponding real space. We also saw that we can only define complex structures for real vector spaces of even dimensions since the set of all $n \times n$ skew-coninvolutory matrices is empty when $n$ is odd.

We also introduced the notion of a right quaternionic vector space. We deal with right quaternionic vector spaces because we saw that we cannot define a quaternionic linear map for left quaternionic vector spaces. We introduced the concept of a *quaternionic structure* in a complex vector space which allows us to mimic a right quaternionic vector space in a complex vector space. We saw that a quaternionic structure $\Psi(v) = \overline{Jv}$ in $\mathbb{C}^2$ where $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponds to right multiplication by $\mathbf{j}$ in the right quaternionic vector space $\mathbb{H}$. We saw that in order for a complex linear map correspond to a quaternionic linear map, the said complex linear map has to commute with the quaternionic structure defined in its corresponding complex space.

In Chapter 3, we saw a problem in defining a determinant for quaternionic matrices. We saw that in order for a determinant to behave as expected, it must satisfy all three of Aslaksen's axioms and the consequence of which is that the determinant should map onto a commutative subset of $\mathbb{H}$. We also saw that the Cayley determinant did not satisfy any of Aslaksen's axioms especially on the axiom concerning the determinant to be 0 for singular quaternionic matrices and that another approach was to be considered. The Study Determinant was one such approach that was shown in [1] to satisfy all of the three axioms.

In Chapter 3, we also introduced the matrix homomorphisms $\phi$ which allows us to represent $n \times n$ complex matrices as $2n \times 2n$ real matrices and $\psi$ which allows us to represent $n \times n$ quaternionic matrices as $2n \times 2n$ complex matrices. We also saw how $\phi$ and $\psi$ are used to define the Study Determinant.

For further study, we recommend the following:

- Expose the other quaternionic determinants, e.g., the Dieudonne determinant and Moore's determinant. There are other quaternionic determinants recommended in [1].
- Show that the set of all $n \times n$ skew-coninvolutory matrices is empty when $n$ is odd. Since the Study Determinant implies that $Sdet(-I_n) = (-1)^{2n} = 1^n$, we cannot use it to prove the latter. Use another quaternionic determinant instead.
- See if we can define a quaternionic structure in a complex vector space with odd dimensions. If we cannot, show that a quaternionic structure can only be defined on vector spaces with even dimensions.

# List of References


[1] H. Aslaksen, *Quaternionic determinants*, The Mathematical Intelligencer, (1996).

[2] G. Balmes, *Rotating a cube using quaternions.* Wolfram Demonstrations Project, February 2016.

[3] ——, *Rotating a unit vector in 3d using quaternions.* Wolfram Demonstrations Project, February 2016.

[4] J. Bell, *Complexification, complex structures, and linear ordinary differential equations*, April 2014.

[5] J. B. Fraleigh, *A First Course in Abstract Algebra, 7th Ed.*, Pearson, 2002.

[6] H. Goldstein, C. Poole, and J. Safko, *Classical Mechanics.*

[7] B. D. (https://math.stackexchange.com/users/242/bill dubuque), *Why are rings called rings?* Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/61505 (version: 2013-12-12).

[8] Y.-B. Jia, *Rotation in the space*, September 2016.

[9] B. J.L., *Applications of the dieudonne determinant*, Linear Algebra and Its Applications, (1968).

[10] R. Larson, *Elementary Linear Algebra.*

[11] A. Lerios, *Rotations and quaternions*, December 1995.

[12] J. P. B. S. Maria, *The sum of coninvolutory matrices*, Master's thesis, University of the Philippines, Diliman, Quezon City, 2012.

[13] Mathoma, *3d rotations in general: Rodrigues rotation formula and quaternion exponentials.*

[14] J. W. Robbin, *Real, complex, and quaternionic representations.* URL:http://www.math.wisc.edu/ robbin/angelic/RCH-G.pdf.

[15] J. Sakurai, *Modern Quantum Mechanics*, 1994.

[16] G. Sanderson, *Abstract vector spaces - essence of linear algebra chapter 11.*

[17] F. W. Strauch, *Euler angles.* Wolfram Demonstrations Project, August 2011.

[18] M. van Leeuwen (https://math.stackexchange.com/users/18880/marc-van leeuwen), *On complexifying vector spaces.* Mathematics Stack Exchange. URL:https://math.stackexchange.com/q/85198 (version: 2011-11-24).

[19] E. W. Weisstein, *Complex structure, from mathworld - a wolfram web resource.*