SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Number systems beyond the reals

av

Eryk Longin Hinc

2022 - No K26

Number systems beyond the reals

Eryk Longin Hinc

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Sven Raum

2022

Number systems beyond the reals

Eryk L Hinc

May 18, 2022

Abstract

From childhood, people are subconsciously exposed to number systems; from counting the number of objects of something to understanding the concept of having parts of a whole or owing someone something. However, while this natural exposure to number systems helps people comprehend number systems up to the reals rather intuitively, number systems beyond the reals are more troublesome for our intuition. This paper explores complex numbers and quaternions in order to further develop understanding of how one can operate with them as well as to create some sort of an intuition for these numbers. In this paper, it is shown that complex numbers can be expressed and represented in different forms, most of which can be compared to geometrical objects; for instance, complex numbers can be represented as pairs of real numbers as well as rotation matrices. Similarly, albeit different in certain aspects, quaternions can be compared to complex numbers to find that, like the complex numbers, they can be represented as tuples of real numbers and through rotation matrices. However, they differ in that complex numbers can be represented as 2-tuples of real numbers that can describe rotations in $\mathbb{R}^2$, while quaternions can be represented as 4-tuples of real numbers that can describe rotations in $\mathbb{R}^3$. Subsequently, in this text, the history of complex numbers and quaternions is presented as it is discussed and explained in other sources, this in order to facilitate the understanding of how these number systems originated and thus strengthening the intuition. Finally, the text touches on an approach to teaching number systems; the approach in question mostly relies on teaching complex numbers and quaternions based on the pupils' prior knowledge, this can be achieved by for instance comparing complex numbers to by the pupils' previously known strategies for operating on expressions with unknown variables. One can also use the similarities that the complex numbers and the quaternions share in order to facilitate the teaching of quaternions as a mathematical object.

Acknowledgements

I would like to sincerely thank my supervisor Sven Raum for helping me whenever more clarity was needed as well as for suggesting topics and subjects to include in the paper while also initiating discussions. Without him, this paper would not have been possible.

Contents

1 Introduction

During the mandatory school years, pupils are often taught the number systems by starting with the natural numbers then expanding on those by learning about the integers followed by the rational numbers which in turn are expanded by the real numbers. Subsequently, while not mandatory in Sweden, pupils studying mathematically intensive programmes in upper secondary school have the opportunity to learn about the imaginary unit i and the complex numbers. From a pedagogical point of view, this order of learning the different number systems seems to coincide with mathematical didactics literature.

However, while the intuition relating to the number systems until the reals that one acquires by virtue of experiencing different phenomena in real life might come naturally, the same cannot necessarily be said about number systems beyond the reals. As such, this text aims to explore number systems beyond the reals. Specifically, in Section 2 of this text, the number systems up to the reals will be briefly discussed with focus on intuition, the text will then explore different representations of complex numbers in order to motivate an intuition, an exploration of quaternions then follows where complex numbers will chiefly be compared to quaternions in order to find differences and similarities between the two numbers systems. Section 3 deals with some history and background to the invention of complex numbers and quaternions. In Section 4, applications of complex numbers and quaternions are discussed. Finally, in Section 5, since this paper is written as part of a thesis for the teacher education programme, an approach to teaching higher number systems is presented with sample lesson plans.

2 Number systems and intuitiveness

2.1 The number systems until the reals

While this section of the text might appear to most as very trivial and elementary, it is crucial to understand how number systems differ and what is included in them in order to further understand how higher number systems take the lower ones into account.

Natural numbers. Starting with the natural numbers, denoted by $\mathbb{N}$, being the set of all numbers following the sequence of $0, 1, 2, 3, 4, \ldots$. The natural numbers, as the name suggest, intuitively appear very natural in the sense that strictly counting the quantity of objects is something that occurs naturally; for instance, when counting the number of apples one has where one can start with zero apples and add apples by picking them one by one.

Integers. Expanding on the natural numbers are the integers, denoted by $\mathbb{Z}$. The integers are a number system combining the natural numbers with the negative numbers, being the sequence $\ldots, -5, -4, -3, -2, -1$, together forming the sequence $\ldots, -2, -1, 0, 1, 2, \ldots$ called the integers. In terms of intuitiveness, these numbers can cause confusion, as is briefly explained by [EHHKMNPR91] when discussing numbers from a historical point of view of ancient Greece. However, returning to the example of counting apples, one can imagine with some logical leeway that if you owe someone an apple, you currently have a negative number of apples, even though it is not necessarily possible to have a negative number of any physical object per se.

Rational numbers. The following number system, called the rational numbers, is denoted by $\mathbb{Q}$ and is the set of numbers that can be written in the form $\frac{a}{b}$ where $a, b \in \mathbb{Z}$ and $b \neq 0$. This number system can, again, be motivated in a rather intuitive way if one disregards philosophical ideas of what constitutes as a whole of something. Once more using an example involving apples, one can cut an apple in half and count one of the halves as $\frac{1}{2}$ of an apple, or in three equal parts and call one part $\frac{1}{3}$ of an apple.

Real numbers. Lastly in this section, there are the real numbers which are denoted by $\mathbb{R}$. This set of numbers contains the rational numbers and the irrational numbers; irrational numbers being all numbers with continuous decimals that cannot be expressed as $\frac{a}{b}$. Some well known examples of these numbers are π, e and $\sqrt{2}$. Considering that the number of decimals that these numbers have is infinite, they appear to be as intuitive as the concept of infinity. However, unlike infinity, these numbers can always find themselves between two rational numbers. For instance, it has been proved that the number π is irrational, that is, it has an infinite number of decimals and is equal to $3.1415\ldots$. Thus, it is possible to create intervals which contain π. Crude examples are the intervals $3 < x < 4$ and $3.14 < x < 3.15$. It becomes evident here that it is possible to approximate these numbers by adding more decimals to the intervals, leading to these numbers being approximated by rational numbers. This approximation becomes especially noticeable the more decimals an approximated number has as the smallest decimals might become negligible in certain applications.

Hopefully, what has hitherto been shown is that, while some leeway may be needed to make these numbers appear intuitive, there are arguments that could be made for these number systems to appear intuitive to the human mind. Additionally, while it might be evident, it is important to note that starting with the natural numbers, each set of number systems discussed is a subset of a different set of number systems such that

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

is true.

Finally, it should also be noted that each of these number systems can be illustrated using a number system with each succeeding number system being able to illustrate more numbers on the number line with $\mathbb{R}$ spanning the entire number line.

2.2 Complex numbers

Given what has previously been presented, one might assume that there are no number systems beyond the reals as the reals cover the entire number line.

However, following the reals, are the complex numbers, denoted by $\mathbb{C}$. Before explaining how the complex numbers belong to the number systems, it is beneficial to first define these numbers. This section of the text will mostly draw on definitions and discussions from [MS20] and [EHHKMNPR91].

Definition 1. *A complex number is an expression on the form $a + ib$ where $a, b \in \mathbb{R}$ and i is a formal symbol. The set of all complex numbers is denoted by $\mathbb{C}$ and is equipped with addition, multiplication and multiplication by a real scalar c as defined by*

$$(a_1 + ib_1) + (a_2 + ib_2) = (a_1 + a_2) + i(b_1 + b_2) \tag{1}$$
$$(a_1 + ib_1)(a_2 + ib_2) = (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1) \tag{2}$$
$$c \cdot (a + ib) = (ca_1 + icb_1). \tag{3}$$

Operations 1 and 3 appear to follow the natural addition and multiplication found in real numbers. Operation 2, however, seems to significantly differ from multiplication found in the real numbers. Using the natural operations, one can observe that if the equation in 2 is to hold, then

$$\begin{aligned}(a_1 + ib_1)(a_2 + ib_2) &= a_1a_2 + i(a_1b_2 + a_2b_1) + i^2b_1b_2 \\ &= a_1a_2 + i(a_1b_2 + a_2b_1) - b_1b_2 \\ &= (a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)\end{aligned}$$

which suggests that $i^2 = -1$.

As might have been observed by the reader, a complex number $z = a + ib$ consists of two parts, the real part denoted by $Re(z)$ which equals a, and the imaginary part $Im(z)$ which equals b. In other words, if $z = a+ib$ then $Re(z) = a$ and $Im(z) = b$. The complex numbers thus consist of the real numbers and imaginary numbers, just as the previous sets of the number systems contain other sets, but one difference here is that the arbitrary complex number $z = a+ib$ with $a, b \neq 0$ is neither a real nor an imaginary number, it is a complex number. In other words, it follows that

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Additionally, for an arbitrary complex number $z = a+ib$, there exists a complex conjugate, denoted by $\overline{z}$ or sometimes z^*, of said number with $\overline{z} = a - ib$. In other words, the complex conjugate of a complex number is the complex number but with its imaginary part having the opposite sign.

Complex numbers as pairs of real numbers. Now that complex numbers have been introduced, one might ask where on the number line the imaginary unit i belongs, and the answer is: nowhere. Real numbers can be viewed as one dimensional numbers geometrically residing only on the number line. The complex numbers, on the other hand, can be viewed as two dimensional numbers, meaning that they geometrically exist in the complex plane with one axis denoting the real numbers, i.e $Re(z)$ of all complex numbers z, and one axis denoting the imaginary numbers, i.e $Im(z)$ of all complex numbers z. Given this, complex numbers can be viewed as pairs of real numbers representing points on the complex plane. As such, one can see that if complex numbers can be represented geometrically as points in the complex plane with one coordinate corresponding to the real axis and the other to the imaginary axis, complex numbers $z = a + ib$ can be represented as $z = (a, b)$ with $Re(z) = Re((a, b)) = a$ and $Im(z) = Im((a, b)) = b$. Additionally, one can see that when represented this way, the unit element e such that $e \cdot z = z$ is $1 \equiv (1, 0)$ and if the property $i^2 = -1 = -e$ is to hold for $i \equiv (0, 1)$ then $(0, 1)$ needs to have the property that $(0, 1)^2 = (-1, 0) = -e$.

Theorem 1. *Representing complex numbers $z = a + ib$ as the ordered pairs of real numbers $z = (a, b)$ with the unit element $e = (1, 0)$ such that $e \cdot z = z$ and $(0, 1)$ with the property that $(0, 1)^2 = (-1, 0) = -e$, the map f:$\mathbb{C} \to \{(a, b) \in \mathbb{R}^2\}$ is bijective and satisfies*

$$
\begin{aligned}
f(z_1 + z_2) &= f(z_1) + f(z_2) \\
f(cz_1) &= c \cdot f(z_1) \\
f(z_1 z_2) &= f(z_1)f(z_2)
\end{aligned}
$$

for all $z_1, z_2 \in \mathbb{C}$ and $c \in \mathbb{R}$.

Proof. It can be seen in addition and multiplication by a real scalar that ordinary laws hold and thus

$$
\begin{aligned}
f((a_1 + ib_1) + (a_2 + ib_2)) &= (a_1 + a_2, b_1 + b_2) \\
&= (a_1, b_1) + (a_2, b_2) \\
&= f(a_1 + ib_1) + f(a_2 + ib_2)
\end{aligned}
$$

and

$$
\begin{aligned}
f(c(a_1 + ib_1)) &= (ca_1, cb_1) \\
&= c(a_1, b_1) \\
&= c \cdot f(a_1 + ib_1)
\end{aligned}
$$

act like operations on vectors. Using the properties $e = (1, 0)$ and $(0, 1) = (-1, 0) = -e$ as well as the addition and multiplication by a scalar used above, one can show that

$$
\begin{aligned}
f(a_1 + ib_2)f(a_2 + ib_2) &= (a_1, b_1)(a_2, b_2) \\
&= (a_1(1, 0) + b_1(0, 1))(a_2(1, 0) + b_2(0, 1)) \\
&= a_1a_2(1, 0) + (a_1b_2 + a_2b_1)(0, 1) + b_1b_2(0, 1)^2 \\
&= (a_1a_2 - b_1b_2)(1, 0) + (a_1b_2 + a_2b_1)(0, 1) \\
&= (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1) \\
&= f((a_1a_2 - b_1b_2) + i(a_1b_2 + a_2b_1)) \\
&= f((a_1 + ib_1)(a_2 + ib_2)).
\end{aligned}
$$

□

While the map in Theorem 1 is bijective and satisfies the presented operations given the special property of $(0, 1)$, it appears artificial when compared to ordinary vector operations. Additionally, using this representation of complex numbers is not advised when calculating as it is unnecessarily cumbersome; the aim of this representation is simply to show the reader that complex numbers can be represented geometrically as ordered pairs of real numbers in the complex plane.

Dimensions and rotation. Knowing that the complex numbers can be viewed and represented in different ways, it becomes easier to describe them in a geometrical and somewhat more intuitive way. For some arbitrary number $x \in \mathbb{R}$, if $x \cdot i = xi$ then that number is rotated by $\frac{\pi}{2}$ counterclockwise from the real axis to the imaginary axis, if then $xi \cdot i$, it is again rotated by $\frac{\pi}{2}$ counterclockwise back to the real axis but now being the negative number $-x$. So while the property $i^2 = -1$ is hard to understand in the one dimensional set $\mathbb{R}$ because $i \notin \mathbb{R}$, considering that $\mathbb{C}$ is a two dimensional plane, it becomes easier to understand the property $i^2 = -1$ simply as a reflection or rotation by π.

Polar form. Seeing how complex numbers can be viewed in the plane and to some extent be represented through rotation, it is also beneficial to explore the polar form representation of complex numbers.

Theorem 2. *A complex number $z = a + ib$ has the polar form $z = |z| (\cos(\theta) + i \sin(\theta))$ with $Re(z) = \cos(\theta)$, $Im(z) = \sin(\theta)$, $|z| = \sqrt{a^2 + b^2}$ and $arg(z) = \theta$ being the angle between the positive real axis and the line going through the origin and the point representing the complex number.*

Proof. The arbitrary complex number $z = a + ib$ can geometrically be represented by the point $z = (a, b)$ in the complex plane. Given this, one can draw a line with length r starting at the origin and ending at the point z. One can then form a right triangle with the hypotenuse r, height b, base a and angle θ between the line r and the positive real axis. The Pythagorean theorem then states that $r^2 = a^2 + b^2$, meaning that $r = \sqrt{a^2 + b^2}$. Using the trigonometric ratios one has

$$\cos(\theta) = \frac{a}{r} \Rightarrow r \cdot \cos(\theta) = a$$
$$\sin(\theta) = \frac{b}{r} \Rightarrow r \cdot \sin(\theta) = b.$$

Now, one can substitute a and b in z with the expressions above to get

$$\begin{aligned} z &= a + ib \\ &= r \cdot \cos(\theta) + i(r \cdot \sin\theta) \\ &= r(\cos(\theta) + i \sin(\theta)). \end{aligned}$$

Finally, because $r = \sqrt{a^2 + b^2} = |z|$ it is clear that

$$\begin{aligned} z &= a + ib \\ &= r(\cos(\theta) + i \sin(\theta)) \\ &= |z| (\cos(\theta) + i \sin(\theta)). \end{aligned}$$

$\square$

It is important to note that when viewing the complex numbers in polar form, multiplication follows multiplication theorems and identities of trigonometric functions. Multiplication of the two complex numbers $z_1 = |z_1| (\cos(\theta_1) + i \sin(\theta_1))$ and $z_2 = |z_2| (\cos(\theta_2) + i \sin(\theta_2))$ can thus be described by the following formula by applying the compound angle identity

$$\begin{aligned} z_1 z_2 = &|z_1||z_2| \big(\cos(\theta_1)\cos(\theta_2) - \sin(\theta_1)\sin(\theta_2) \\ &\qquad + i(\cos(\theta_1)\sin(\theta_2) + \cos(\theta_2)\sin(\theta_1))\big) \\ = &|z_1||z_2| (\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)) \end{aligned}$$

which in problems containing many multiplications of complex numbers seems easier to calculate.

Matrix form. Complex numbers can, thus, be represented in polar form by using trigonometric functions and multiplication with complex numbers is tied to rotation in the complex plane. A mathematical object that seems similar to this is the following 2×2 rotation matrix multiplied with the positive scalar r

$$r \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

What this matrix does is rotate a point by θ counterclockwise. It is worth noting that using matrix operations, this matrix can be written as

$$r\Big(\cos(\theta) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \sin(\theta) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \Big).$$

As can be seen here, the above expression is very similar to the complex number z written in polar form

$$z = |z| (\cos(\theta) + i \sin(\theta)).$$

By setting $r = |z| = 1$ and acknowledging that $\cos(\theta) \cdot 1 = \cos(\theta)$ with 1 being the unit element in complex numbers such that $z \cdot 1 = z$ for all complex numbers z and the identity matrix

$$I_{2\times 2} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

having the property that $AI_{2\times 2} = A$ for any 2×2 matrix A, one can see that $1 \equiv I_{2\times 2}$ in this expression. Additionally, if $\theta = \frac{\pi}{2}$, the complex number and rotation matrix equal

$$z = i \text{ and } \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \text{ respectively.}$$

If one now assumes that

$$i \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

which might be suggested by how both of these objects transform the points in the two dimensional plane, one could represent imaginary numbers as matrices. Additionally, the matrix above has the property that when squared it equals $-I_{2\times 2}$ which is essentially a translation of the $i^2 = -1$ property. Moreover, with the above comparison one could then represent the complex number $z = a + ib$ as

$$z = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

However, to prove that this is the case, it remains to verify whether complex numbers and their operations map onto matrices of this type.

Theorem 3. *The map* $f\colon \mathbb{C} \to \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}) \middle| a, b \in \mathbb{R} \right\}$ *is bijective and satisfies*

$$\begin{aligned} f(z_1 + z_2) &= f(z_1) + f(z_2) \\ f(z_1 z_2) &= f(z_1) f(z_2) \\ f(cz) &= c \cdot f(z) \\ f(z^*) &= f(z)^T \end{aligned}$$

for all $z, z_1, z_2 \in \mathbb{C}$ *and* $c \in \mathbb{R}$.

Proof. Starting with addition, it follows that

$$\begin{aligned} f((a_1 + ib_1) + (a_2 + ib_2)) &= \begin{pmatrix} a_1 + a_2 & -(b_1 + b_2) \\ b_1 + b_2 & a_1 + a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} + \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} \\ &= f(a_1 + ib_1) + f(a_2 + ib_2). \end{aligned}$$

Additionally, it is clear that the same applies to multiplication

$$\begin{aligned} f(a_1 + ib_1) f(a_2 + ib_2) &= \begin{pmatrix} a_1 & -b_1 \\ b_1 & a_1 \end{pmatrix} \begin{pmatrix} a_2 & -b_2 \\ b_2 & a_2 \end{pmatrix} \\ &= \begin{pmatrix} a_1 a_2 - b_1 b_2 & -(a_1 b_2 + b_1 a_2) \\ a_1 b_2 + b_1 a_2 & a_1 a_2 - b_1 b_2 \end{pmatrix} \\ &= f((a_1 a_2 - b_1 b_2) + i(a_1 b_2 + b_1 a_2)) \\ &= f((a_1 + ib_1)(a_2 + ib_2)) \end{aligned}$$

as well as scalar multiplication with $c \in \mathbb{R}$

$$\begin{aligned} c \cdot f(a+ib) &= c \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \\ &= \begin{pmatrix} ca & -cb \\ cb & ca \end{pmatrix} \\ &= f(ca + icb) \\ &= f(c(a+ib)). \end{aligned}$$

Finally, as can be derived from the above as well as matrix operations, it becomes apparent that

$$\begin{aligned} f((a+ib)^*) &= f(a-ib) \\ &= \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \\ &= \begin{pmatrix} a & -b \\ b & a \end{pmatrix}^T \\ &= f(a+ib)^T. \end{aligned}$$

□

Thus, as has been illustrated, while the complex numbers may seem unnatural and counter-intuitive, they can be expressed and viewed in many different ways. It is true that imaginary and complex numbers are confusing when viewed in relation to the number line, but that is because doing so, while it might appear natural, is incorrect because a complex number $z = a + ib$ with $b \neq 0$ is in the complex plane, not on the number line. However, as has been shown, the complex numbers can be viewed and expressed as pairs of real numbers, points in a plane and real matrices. Complex and imaginary numbers can also be used to describe the rotation in the plane, similar to rotation matrices.

2.3 Quaternions

Considering how the complex numbers can be motivated and represented as objects in the two dimensional space. This suggests that there might be numbers that can, in a similar way, be motivated and represented as objects in spaces of higher dimensions. This, in fact, happens to indeed be the case. Following the complex numbers are the quaternions and this section of the text will mostly draw on [K99]

Definition 2. *A quaternion is a number that can be written as the expression* $a+ib+jc+kd$ *where* $a, b, c, d \in \mathbb{R}$ *and the rule* $i^2 = j^2 = k^2 = ijk = -1$ *applies.*

Within number systems, the set containing all quaternions is denoted by $\mathbb{H}$. Addition of quaternions and multiplication with a real scalar as operations do not differ much from how they are treated in previously discussed number systems. That is, for the two quaternions $q = q_0 + iq_1 + jq_2 + kq_3$ and $p = p_0 + ip_1 + jp_2 + kp_3$ and real scalar c, addition and multiplication by a real scalar are defined as

$$\begin{aligned} cq &= cq_0 + icq_1 + jcq_2 + kcq_3 \\ q + p &= q_0 + p_0 + i(q_1 + p_1) + j(q_2 + p_2) + k(q_3 + p_3) \end{aligned}$$

and addition is also associative and commutative.

Quaternion multiplication. However, quaternion multiplication differs quite substantially from multiplication discussed in previous sections. Quaternion multiplication is associative and distributive but it is generally not commutative meaning that in general for the quaternions q and p, one has $qp \neq pq$. This can be derived from the defining equations for the quaternion, namely $ijk = -1$. Multiplying this identity by $-k$ from the right, we obtain $ij = k$. On the other hand, using this identitiy, we find that

$$ji = 1 \cdot ji = (-k)(ij)ji = -k(-1)^2 = -k.$$

From here it is possible to continue multiplying in similar ways to describe all possible products of the elements i, j, k in the quaternions, but it is also possible to use substitution instead to find the same identities. The following lemma shows a table containing the possible combinations.

Lemma 1. *Through the rule $i^2 = j^2 = k^2 = ijk = -1$ the factors i, j and k in $ijk = -1$ can be expressed as a product of the remaining two factors. All possible combinations are shown in the table below with the rows being the first factor and columns being the second factor.*

·	1	i	j	k
1	1	i	j	k
i	i	-1	k	-j
j	j	-k	-1	i
k	k	j	-i	-1

. (4)

Proof. First, 1 is the multiplicative unit element e in quternions such that, for all quaternions q one has $e \cdot q = q$ meaning that the second row and column are automatically filled out by the first row and column respectively. Second, the main diagonal corresponds to the $i^2 = j^2 = k^2 = -1$ part of the rule and is thus also automatically filled out. Last, using the rule $i^2 = j^2 = k^2 = ijk = -1$, it is possible to multiply the equation $ijk = -1$ by k from both sides to get $ijk^2 = -k$ which is equivalent to $ij = k$. From here one can use substitution to obtain the remaining products

$$
\begin{aligned}
ij &= k \\
ik &= i(ij) = -j \\
jk &= (-ik)k = i \\
ji &= j(jk) = -k \\
ki &= (-ji)i = j \\
kj &= (ij)j = -i.
\end{aligned}
$$

$\square$

Theorem 4. *For two arbitrary quaternions $q = q_0 + iq_1 + jq_2 + kq_3$ and $p = p_0 + ip_1 + jp_2 + kp_3$ the general formula for quaternion multiplication is*

$$
\begin{aligned}
qp =& (q_0 + iq_1 + jq_2 + kq_3)(p_0 + ip_1 + jp_2 + kp_3) \\
=& q_0p_0 + iq_0p_1 + jq_0p_2 + kq_0p_3 \\
&+ iq_1p_0 - q_1p_1 + kq_1p_2 - jq_1p_3 \\
&+ jq_2p_0 - kq_2p_1 - q_2p_2 + iq_2p_3 \\
&+ kq_3p_0 + jq_3p_1 - iq_3p_2 - q_3p_3.
\end{aligned} \tag{5}
$$

The above formula can be derived from the distributive property of quaternions. Additionally, the non-commutativity of quaternion multiplication can be recovered by comparing Equation 5 to

$$
\begin{aligned}
pq =& (p_0 + ip_1 + jp_2 + kp_3)(q_0 + iq_1 + jq_2 + kq_3) \\
=& q_0p_0 + iq_0p_1 + jq_0p_2 + kq_0p_3 \\
&+ iq_1p_0 - q_1p_1 - kq_1p_2 + jq_1p_3 \\
&+ jq_2p_0 + kq_2p_1 - q_2p_2 - iq_2p_3 \\
&+ kq_3p_0 - jq_3p_1 + iq_3p_2 - q_3p_3
\end{aligned}
$$

which further indicates that, in general, $qp \neq pq$.

Conjugate and inverse. Similarly to the complex numbers, a quaternion $q = q_0 + iq_1 + jq_2 + kq_3$ can be represented as consisting of two parts, the real part q_0 and the vector part $iq_1 + jq_2 + kq_3$, a quaternion whose real part is equal to zero is called a pure quaternion. The conjugate of a quaternion q is denoted by q^* and $q^* = q_0 - iq_1 - jq_2 - kq_3$, that is, the conjugate of a quaternion is itself but with a vector part containing opposite signs. Given this, while relatively tedious, it is easy to show that

$$qq^* = q^*q = q_0^2 + q_1^2 + q_2^2 + q_3^2 \tag{6}$$

$$q + q^* = 2q_0 = 2Re(q) \tag{7}$$

$$\sqrt{q^*q} = |q| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2} \tag{8}$$

using quaternion multiplication. Using quaternion multiplication, it is also fairly easy to show that

$$(qp)^* = p^*q^* \tag{9}$$

Additionally, non-zero quaternions have an inverse q^{-1} such that

$$qq^{-1} = q^{-1}q = 1.$$

Specifically, using the property in Equation 6, the inverse of an arbitrary quaternion q can be described by

$$\begin{aligned} q^{-1}qq^* &= q^*qq^{-1} = q^* \\ \Rightarrow q^{-1} &= \frac{q^*}{|q|^2} \end{aligned}$$

with the special case of $q^{-1} = q^*$ when $|q|^2 = 1$, that is, when dealing with a quaternion of length 1, or, in other words, a unit quaternion.

4-tuple representation. As was the case with complex numbers, quaternions can be represented as vectors or points in a space. However, it is vital to note that, while there may be some similarities in representing complex numbers as points in space and representing quaternions as points in space, complex numbers are 2-tuples and quaternions are 4-tuples which makes the two number systems vastly differ.

Theorem 5. *Quaternions $q = a + ib + jc + kd$ can be written in the form $q = (a, b, c, d)$ in the set $\mathbb{R}^4$ of all ordered 4-tuples of real numbers. When represented in this form, the two quaternions $q = q_0 + q_1 + q_2 + q_3$ and $p = p_0 + p_1 + p_2 + p_3$ have addition, multiplication and multiplication by a real scalar c defined as*

$$\begin{aligned} q + p &= (q_0 + p_0, q_1 + p_1, q_2 + p_2, q_3 + p_3) \\ cq &= (cq_0, cq_1, cq_2, cq_3) \\ qp &= (q_0p_0 - q_1p_1 - q_2p_2 - q_3p_3, \\ &\quad q_0p_1 + q_1p_0 + q_2p_3 - q_3p_2, \\ &\quad q_0p_2 - q_1p_3 + q_2p_0 + q_3p_1, \\ &\quad q_0p_3 + q_1p_2 - q_2p_1 + q_3p_0). \end{aligned} \tag{10}$$

Matrix representation. Similarly to the complex numbers, quaternions can be represented using matrices with all real values. One such matrix, namely a 4×4 matrix can be obtained by commuting the terms in Equation 10:

$$\begin{aligned} qp = (&q_0p_0 - q_1p_1 - q_2p_2 - q_3p_3, \\ &q_1p_0 + q_0p_1 - q_3p_2 + q_2p_3, \\ &q_2p_0 + q_3p_1 + q_0p_2 - q_1p_3, \\ &q_3p_0 - q_2p_1 + q_1p_2 + q_0p_3) \end{aligned}$$

and from here it is clear that the expression visually resembles that of the matrix multiplication

$$\begin{pmatrix} q_0p_0 & -q_1p_1 & -q_2p_2 & -q_3p_3 \\ q_1p_0 & q_0p_1 & -q_3p_2 & q_2p_3 \\ q_2p_0 & q_3p_1 & q_0p_2 & -q_1p_3 \\ q_3p_0 & -q_2p_1 & q_1p_2 & q_0p_3 \end{pmatrix} = \begin{pmatrix} q_0 & -q_1 & -q_2 & -q_3 \\ q_1 & q_0 & -q_3 & q_2 \\ q_2 & q_3 & q_0 & -q_1 \\ q_3 & -q_2 & q_1 & q_0 \end{pmatrix} \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{pmatrix}.$$

Theorem 6. *The map* $f{:}\mathbb{H} \to \left\{ \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} \in \mathrm{M}_4(\mathbb{R}) \middle| a, b, c, d \in \mathbb{R} \right\}$ *is bijective and satisfies*

$$\begin{aligned} f(q+p) &= f(q) + f(p) \\ f(qp) &= f(q)f(p) \\ f(sq) &= s \cdot f(q) \\ f(q^*) &= f(q)^T \end{aligned}$$

for all $q, p \in \mathbb{H}$ *and* $s \in \mathbb{R}$.

While tedious, the reader can, using matrix operations, examine that the matrix representation satisfies these operations. This is essentially having the same result as the bijection of complex numbers onto a set of 2×2 real matrices.

Rotation. Similarly to how the complex numbers could be used to describe rotations, quaternions can also be used to describe rotation. However, while complex numbers are 2-tuples describing rotation in the plane, quaternions are 4-tuples that can be used to describe rotations in three dimensional space. As has been shown, quaternions can be represented as vectors with four coordinate values, however, if the real part of a quaternion is set to 0, that is, when dealing with pure quaternions, it can be treated as a vector with three coordinates. Geometrically such a vector can be represented in a complex room in three dimensional space with the three axes corresponding to i, j and k. Thus, a pure quaternion $v = iv_1 + jv_2 + kv_3$ can geometrically be represented as a vector in the three dimensional space being v_1 units along the i-axis, v_2 units along the j-axis, and v_3 units along the k-axis. Since quaternions can be used to represent rotation in the three dimensional space, there should exist quaternions that, when multiplied with, rotate the vector around some axis in the three dimensional space. However, by multiplying an arbitrary quaternion $q = q_0 + iq_1 + jq_2 + kq_3$ with v one gets

$$\begin{aligned} qv &= (q_0, q_1, q_2, q_3)(0, v_1, v_2, v_3) \\ &= (- q_1v_1 - q_2v_2 - q_3v_3, \\ &\quad q_1v_1 + q_2v_3 - q_3v_2, \\ &\quad q_0v_2 - q_1v_3 + q_3v_1, \\ &\quad q_0v_3 + q_1v_2 - q_2v_1) \end{aligned}$$

which is generally no longer a pure quaternion. However, by then multiplying with q^* from the right side, the real part of the new quaternion cancels out and the resulting quaternion qvq^* is a pure quaternion. Thus, if v is treated as a vector in three dimensional space, it has now been transformed into another vector qvq^* in three dimensional space by the double quaternion multiplication.

Theorem 7. *For the arbitrary quaternion $q = q_0 + q_1 + q_2 + q_3$ and pure quaternion $v = v_1 + v_1 + v_3$ the product of the multiplication qvq^* is a pure quaternion. That is, $Re(qvq^*) = 0$ if $Re(v) = 0$.*

Proof. By using the distributive property of quaternions and the Equations 6, 7 and 9, one can show that

$$\begin{aligned} 2Re(qvq^*) &= qvq^* + (qvq^*)^* = qvq^* + q^{**}v^*q^* \\ &= qvq^* + qv^*q^* = q(v + v^*)q^* \\ &= q(2Re(v))q^* = 0. \end{aligned}$$

□

Because it is very tedious to perform the qvq^* multiplication, it is useful to have a general formula for the quaternion resulting from the multiplication.

Lemma 2. *The quaternion $w = qvq^*$ with $q = q_0 + iq_1 + jq_2 + kq_3$, $v = iv_1 + jv_2 + kv_3$ and $|q| = 1$ can be obtained from the formula*

$$\begin{aligned} w = qvq^* =& i((2q_0^2 - 1 + 2q_1^2)v_1 + (2q_1q_2 - 2q_0q_3)v_2 + (2q_1q_3 + 2q_0q_2)v_3) \\ &+ j((2q_1q_2 + 2q_0q_3)v_1 + (2q_0^2 - 1 + 2q_2^2)v_2 + (2q_2q_3 - 2q_0q_1)v_3) \\ &+ k((2q_1q_3 - 2q_0q_2)v_1 + (2q_2q_3 + 2q_0q_1)v_2 + (2q_0^2 - 1 + 2q_3^2)v_3). \end{aligned}$$

Because the calculation needed to arrive at the product above is relatively easy but extremely tedious, no proof will be presented. However, a sketch of a proof will now follow.

Sketch of a proof. To facilitate the calculations needed for this proof, one can perform the needed operations in steps. First, to arrive at the product of qvq^* one can calculate qv by following the general formula for quaternion multiplication as shown in Theorem 4 and then using the distributive and associative laws, one can group all real terms and all terms containing i, j and k to get

$$\begin{aligned} qv =& (q_0 + q_1 + q_2 + q_3)(0 + v_1 + v_2 + v_3) \\ &= x_0 + ix_1 + jx_2 + kx_3 \\ &= -q_1v_1 - q_2v_2 - q_3v_3 && (11) \\ &+ i(q_1v_1 + q_2v_3 - q_3v_2) && (12) \\ &+ j(q_0v_2 - q_1v_3 + q_3v_1) && (13) \\ &+ k(q_0v_3 + q_1v_2 - q_2v_1) && (14) \end{aligned}$$

with x_0, x_1, x_2 and x_3 being equal to the expressions in 11, 12, 13 and 14 respectively. One can then, similarly, perform the multiplication

$$\begin{aligned} qv \cdot q^* &= (x_0 + ix_1 + jx_2 + kx_3)(q_0 - iq_1 - jq_2 - kq_3) \\ &= x_0q_0 + x_1q_1 + x_2q_2 + x_3q_3 \\ &+ i(x_1q_0 - x_0q_1 + x_3q_2 - x_2q_3) \\ &+ j(x_2q_0 - x_3q_1 - x_0q_2 + x_1q_3) \\ &+ k(x_3q_0 + x_2q_1 - x_1q_2 - x_0q_3). \end{aligned}$$

At this point one can simply calculate each term independently and then combine all terms. □

Angle and directional vector. Additionally, to facilitate the intuition of rotating in three dimensional space by completing the double quaternion multiplication discussed above, if q is set to be a unit quaternion, that is $|q| = 1$, then it becomes clear that there is a relationship between

$$q = q_0^2 + (q_1^2 + q_2^2 + q_3^2) = 1$$

and

$$\cos^2\theta + \sin^2\theta = 1$$

where there exists an angle θ such that

$$\cos^2\theta = q_0^2$$

and

$$\sin^2\theta = (q_1^2 + q_2^2 + q_3^2).$$

However, because θ exists in an infinite domain, it can be defined uniquely by restricting its domain to $-\pi < \theta \leq \pi$. One can then also see that

$$\cos\theta = q_0$$

and

$$\sin\theta = \sqrt{(q_1^2 + q_2^2 + q_3^2)}$$

for some angle θ. Furthermore, in order to effectively be able to use the relationship above, it would be beneficial if one could write a quaternion q in terms of trigonometric functions. This is possible by noting that for a unit quaternion q

$$\begin{aligned} q &= q_0 + iq_1 + jq_2 + kq_3 \\ &= \cos\theta + \frac{(iq_1 + jq_2 + kq_)}{\sin\theta}\sin\theta \\ &= \cos\theta + \frac{(iq_1 + jq_2 + kq_)}{\sqrt{(q_1^2 + q_2^2 + q_3^2)}}\sin\theta. \end{aligned}$$

Theorem 8. *A quaternion $q = q_0+iq_1+jq_2+kq_3$ with $|q| = 1$ can be represented in its polar form as*

$$q = \cos\theta + \sin\theta \cdot u$$

where

$$u = \frac{(iq_1 + jq_2 + kq_)}{\sqrt{(q_1^2 + q_2^2 + q_3^2)}}.$$

The unit quaternion above can be represented by the angle θ and the unit vector u. Subsequently, when using this unit quaternion to rotate an arbitrary pure quaternion, the rotation can be understood as being the rotation around an axis with the same direction as the vector u by an angle of 2θ. It has previously been shown that when multiplying a pure quaternion by only one quaternion, the product was generally no longer a pure quaterion. However, the resulting quaternion would still have rotated by θ around u. Thus, when performing the second multiplication by the conjugate of q, the real part cancels out while the vector is rotated once more. In other words, since the quaternion multiplication happens twice, in order to not transform the pure quaternion into a quaternion whose real part is not equal to zero, the vector representing the pure quaternion is rotated by θ twice.

To further illustrate that the operation qvq^* can in fact be seen as a rotation in the three dimensional space, one can, by looking at the quaternion w in Lemma 2, see that the product qvq^* for an arbitrary unit quaternion q and pure quaternion v resembles and can be represented by the matrix transformation

$$Qv = \begin{pmatrix} 2(q_0^2+q_1^2)-1 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & 2(q_0^2+q_2^2)-1 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & 2(q_0^2+q_3^2)-1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$

In other words, it follows that $w = qvq^* = Qv$. However, it remains to test whether this matrix transformation is a rotation by acknowledging that a rotation matrix Q_{nxn} is a rotation matrix if and only if $Q^TQ = I$ and $\det(Q) = 1$.

Lemma 3. *The matrix*

$$Q = \begin{pmatrix} 2(q_0^2+q_1^2)-1 & 2(q_1q_2-q_0q_3) & 2(q_1q_3+q_0q_2) \\ 2(q_1q_2+q_0q_3) & 2(q_0^2+q_2^2)-1 & 2(q_2q_3-q_0q_1) \\ 2(q_1q_3-q_0q_2) & 2(q_2q_3+q_0q_1) & 2(q_0^2+q_3^2)-1 \end{pmatrix}$$

is a rotation matrix in three dimensional space representing the rotation performed on the pure quaternion v in qvq^*.

The proof of this lemma consists of a long calculation, making use of the fact that $|q| = 1$ in order to simplify expressions at key moments. The following sketch aims to demonstrate the principle of how a proof to Lemma 3 can be obtained by presenting the calculations for $(Q^TQ)_{11}$ and $(Q^TQ)_{12}$ as well as $\det(Q)$.

Sketch of a proof. If Q is a rotation matrix then $Q^TQ = I$ and $\det(Q) = 1$. Starting with $(Q^TQ)_{11}$ one has

$$\begin{aligned}
(Q^TQ)_{11} =& (2(q_0^2 + q_1^2) - 1)^2 + (2(q_1q_2 + q_0q_3))^2 + (2(q_1q_3 - q_0q_2))^2 \\
=& 4(q_0^2q_0^2 + 2q_0^2q_1^2 + q_1^2q_1^2) - 4(q_0^2 + q_1^2) + 1 \\
& + 4(q_1^2q_2^2 + 2q_0q_1q_2q_3 + q_0^2q_3^2) \\
& + 4(q_1^2q_3^2 - 2q_0q_1q_2q_3 + q_0^2q_2^2) \\
=& 4(q_0^2q_0^2 + q_1^2q_1^2 + 2q_0^2q_1^2 + q_0^2q_2^2 + q_0^2q_3^2 + q_1^2q_2^2 + q_1^2q_3^2 - q_0^2 - q_1^2) + 1 \\
=& 4(q_0^2q_0^2 + q_1^2q_1^2q_0^2(1 - q_0^2) + q_1^2(1 - q_1^2) - q_0^2 - q_1^2) + 1 \\
=& 4(q_0^4 + q_1^4 + q_0^2 - q_0^4 + q_1^2 - q_1^4 - q_0^2 - q_1^2) + 1 \\
=& 1.
\end{aligned}$$

Subsequently, one can show that

$$\begin{aligned}
(Q^TQ)_{12} =& (2(q_0^2 + q_1^2) - 1) \cdot 2(q_1q_2 - q_0q_3) \\
& + 2(q_1q_2 + q_0q_3)(2(q_0^2 + q_2^2) - 1) \\
& + 2(q_1q_3 - q_0q_2) \cdot 2(q_2q_3 + q_0q_1) \\
=& 4(q_0^2q_1q_2 - q_0^2q_0q_3 + q_1^2q_1q_2 - q_1^2q_0q_3) - 2(q_1q_2 - q_0q_3) \\
& + 4(q_0^2q_1q_2 + q_0^2q_0q_3 + q_2^2q_0q_3 + q_2^2q_1q_2) - 2(q_1q_2 + q_0q_3) \\
& + 4(q_3^2q_1q_2 + q_1^2q_0q_3 - q_2^2q_0q_3 - q_0^2q_1q_2) \\
=& 4(q_1^2q_1q_2 + q_2^2q_1q_2 + q_0^2q_1q_2 + q_3^2q_1q_2) - 4q_1q_2 \\
=& 4(q_1q_2(q_0^2 + q_1^2 + q_2^2 + q_3^2) - q_1q_2) \\
=& 0.
\end{aligned}$$

Similar calculations can be done for the remaining positions of the resulting matrix Q^TQ leading to $Q^TQ = I$ with I being the identity matrix. Next, it is also necessary to show that $\det(Q) = 1$. Because the calculations are very long, this will be done in segments by noting that

$$\begin{aligned}
\det(Q) =& a + b + c \\
=& (2(q_0^2 + q_1^2) - 1) \begin{vmatrix} 2(q_0^2 + q_2^2) - 1 & 2(q_2q_3 - q_0q_1) \\ 2(q_2q_3 + q_0q_1) & 2(q_0^2 + q_3^2) - 1 \end{vmatrix} & (15) \\
& - 2(q_1q_2 - q_0q_3) \begin{vmatrix} 2(q_1q_2 + q_0q_3) & 2(q_2q_3 - q_0q_1) \\ 2(q_1q_3 - q_0q_2) & 2(q_0^2q_3^2) - 1 \end{vmatrix} & (16) \\
& + 2(q_1q_3 + q_0q_2) \begin{vmatrix} 2(q_1q_2 + q_0q_3) & 2(q_0^2q_2^2) - 1 \\ 2(q_1q_3 - q_0q_2) & 2(q_2q_3 + q_0q_1) \end{vmatrix} & (17)
\end{aligned}$$

with a, b and c being equal to the expressions in 15, 16 and 17 respectively.

Starting with the expression in 15, one has

$$\begin{aligned}
a =&(2(q_0^2+q_1^2)-1)\\
&\quad\cdot((2(q_0^2+q_2^2)-1)(2(q_0^2+q_3^2)-1)-2(q_2q_3-q_0q_1)\cdot 2(q_2q_3+q_0q_1))\\
=&(2(q_0^2+q_1^2)-1)\\
&\quad\cdot(4(q_0^2q_0^2+q_0^2q_3^2+q_0^2q_2^2+q_2^2q_3^2)-2(q_0^2+q_2^2)-2(q_0^2+q_3^2)+1\\
&\quad-4(q_2^2q_3^2+q_0q_1q_2q_3-q_0q_1q_2q_3-q_0^2q_1^2))\\
=&(2(q_0^2+q_1^2)-1)\\
&\quad\cdot(-2(q_2^2+q_3^2)+1)\\
=&-4(q_0^2q_2^2+q_0^2q_3^2+q_1^2q_2^2+q_1^2q_3^2)+2(q_0^2+q_1^2+q_2^2+q_3^2)-1\\
=&-4(q_0^2q_2^2+q_0^2q_3^2+q_1^2q_2^2+q_1^2q_3^2)+1\\
=&4(-q_0^2q_2^2-q_0^2q_3^2-q_1^2q_2^2-q_1^2q_3^2)+1.
\end{aligned}$$

For the expression in 16, one has

$$\begin{aligned}
b =&-2(q_1q_2-q_0q_3)\\
&\quad\cdot(2(q_1q_2+q_0q_3)(2(q_0^2+q_3^2)-1)-2(q_2q_3-q_0q_1)\cdot 2(q_1q_3-q_0q_2))\\
=&-2(q_1q_2-q_0q_3)\\
&\quad\cdot(4(q_0^2q_1q_2+q_0^2q_0q_3+q_3^2q_1q_2+q_3^2q_0q_3)-2(q_1q_2+q_0q_3)\\
&\quad-4(q_3^2q_1q_2-q_2^2q_0q_3-q_1^2q_0q_3+q_0^2q_1q_2))\\
=&-2(q_1q_2-q_0q_3)\\
&\quad\cdot(4(q_0^2q_0q_3+q_1^2q_0q_3+q_2^2q_0q_3+q_3^2q_0q_3)-2(q_1q_2+q_0q_3))\\
=&-2(q_1q_2-q_0q_3)(4q_0q_3-2(q_1q_2+q_0q_3))\\
=&-8(q_0q_1q_2q_3-q_0^2q_3^2)+4(q_1^2q_2^2+q_0q_1q_2q_3-q_0q_1q_2q_3-q_0^2q_3^2)\\
=&8(q_0^2q_3^2-q_0q_1q_2q_3)+4(q_1^2q_2^2-q_0^2q_3^2).
\end{aligned}$$

For the expression in 17, one has

$$\begin{aligned}
c =&2(q_1q_3+q_0q_2)\\
&\quad\cdot(2(q_1q_2+q_0q_3)\cdot 2(q_2q_3+q_0q_1)-(2(q_0^2+q_2^2)-1)\cdot 2(q_1q_3-q_0q_2))\\
=&2(q_1q_3+q_0q_2)\\
&\quad\cdot(4(q_2^2q_1q_3+q_1^2q_0q_2+q_3^2q_0q_2+q_0^2q_1q_3)\\
&\quad-4(q_0^2q_1q_3-q_0^2q_0q_2+q_2^2q_1q_3-q_2^2q_0q_2)+2(q_1q_3-q_0q_2))\\
=&2(q_1q_3+q_0q_2)\\
&\quad\cdot(4q_0q_2(q_0^2+q_1^2+q_2^2+q_3^2)+2(q_1q_3-q_0q_2))\\
=&2(q_1q_3+q_0q_2)(4q_0q_2+2(q_1q_3-q_0q_2))\\
=&8(q_0q_1q_2q_3+q_0^2q_2^2)+4(q_1^2q_3^2-q_0q_1q_2q_3+q_0q_1q_2q_3-q_0^2q_2^2)\\
=&8(q_0q_1q_2q_3+q_0^2q_2^2)+4(q_1^2q_3^2-q_0^2q_2^2).
\end{aligned}$$

Now all that remains is to calculate $\det(Q)$.

$$
\begin{aligned}
\det(Q) =& a + b + c \\
=& 4(-q_0^2q_2^2 - q_0^2q_3^2 - q_2^2q_2^2 - q_1^2q_3^2) + 1 \\
& + 8(q_0^2q_3^2 - q_0q_1q_2q_3) + 4(q_1^2q_2^2 - q_0^2q_3^2) \\
& + 8(q_0q_1q_2q_3 + q_0^2q_2^2) + 4(q_1^2q_3^2 - q_0^2q_2^2) \\
=& 1.
\end{aligned}
$$

$\square$

While the loss of commutativity in quaternions and the idea that quaternions are a 4-tuple that can be used to describe rotations in three dimensions might appear artificial and counterintuitive at first, quaternions can be represented in terms of mathematical objects with real values. Furthermore, the non-commutativity property of a mathematical object that can be used to describe three dimensional rotation should appear natural and intuitive not when compared to the preceding number systems but rather in conjunction with how three dimensional rotation can be described in the real world. For instance, if a book is rotated by $90°$ around some axis and then $90°$ around another axis orthogonal to the first while a different book is first rotated by $90°$ along the second axis and then $90°$ along the first axis. These two books, starting with the same orientation will not be parallel after the rotations have been performed. This property of general non-commutativity can be observed in other mathematical objects describing three dimensional rotation as well, for instance rotation matrices.

2.4 Number systems of higher dimensions

If viewing the aforementioned number systems as sets in domains of different dimensions, it becomes interesting to consider whether there are number systems beyond quaternions and if there is a limit to how many number systems there are. By using the Cayley-Dickson construction, it is possible to produce an infinite number of number systems[Wiki1]. However, by carrying out the Cayley-Dickson construction, the newly produced number systems lose certain properties; the complex numbers lose order, quaternions, as has been discussed, lose commutativity, the succeeding number system octonions lose associativity, following the octonions are the sedenions which together with the all subsequent number systems produced by the Cayley-Dickson construction lose alternativity.

3 History of complex numbers and quaternions

This section of the text aims to briefly deal with the historical background to the invention of complex numbers and quaternions. As such, motivations for the invention of these mathematical objects as well as some words on intuition will be presented as accounted for in the literature used.

3.1 Complex numbers

This subsection will draw on discussions in [EHHKMNPR91] and [M06] when presenting the historical background to complex numbers.

It is said that a problem that mathematicians of the 16th century tried to find a solution to is the general cubic equation

$$x^3 + ax^2 + bx + c = 0.$$

According to [M06], the Italian mathematician and professor Scipione del Ferro first discovered the solution to solving the cubic equation. Subsequently, on his deathbed, del Ferro passed down this solution to his pupil who in turn challenged Italian mathematician Niccolò Fontana Tartaglia to a mathematical contest in which Tartaglia was challeneged to solve cubic equations with imaginary roots. Tartaglia discovered the formula for solving cubic equations and won the contest. He then told the formula to mathematician Gerolamo Cardano who later published the solution.

As is stated in [EHHKMNPR91], while Cardano was already exploring complex numbers in his book, he still seemed to not entirely grasp how these numbers behave and as such, he incorrectly wrote in his book that the equation

$$x(10 - x) = 40$$

has the two solutions

$$x = 5 \pm \sqrt{-15}$$

suggesting that while he was working with complex numbers, he possibly still struggled and thus made such incorrect statements. He then later in his book provides his formula

$$x = \sqrt[3]{\frac{q}{2} + \sqrt{d}} + \sqrt[3]{\frac{q}{2} - \sqrt{d}} \qquad \text{with} \qquad d = \frac{q}{2}^2 - \frac{p}{3}^3$$

to the equation $x^3 = px + q$.

Thus, it can be seen that the discovery of the complex numbers, while not necessarily yet as developed as the number system is in contemporary mathemathics, occured as a result of trying to solve the general cubic equation. They are also, as is well known, used to solve quadratic equations such as $x^2 = -1$. However, [EHHKMNPR91] write that complex numbers were still, for a long time, used incorrectly and they were also largely considered to not be useful and thus not treated equally to other mathematical objects. This reputation started changing when mathematicians started representing complex numbers geometrically as the numbers on the axis of a plane. Subsequently, the French mathematician Augustin-Louis Cauchy expressed complex numbers as "any symbolic expression of the form $a + b\sqrt{-1}$, where a, b denote two real quantities" (as translated from French by [EHHKMNPR91]) giving the complex numbers a purely algebraic interpretation. Furthermore, to completely seal the geometrical representation of complex numbers, Irish mathematician Sir William Rowan

Hamilton formally defined the complex numbers as represented as an ordered pair of real numbers equipped with addition and multiplication such that the associative, commutative and distribute laws hold.

3.2 Quaternions

It is explained in [EHHKMNPR91] that Hamilton who defined complex numbers as pairs of real numbers in the plane $\mathbb{R}^2$ became interested in discovering whether a number system in the three dimensional room $\mathbb{R}^3$ equipped with operations analogous to those of the complex numbers could exist. He first tried with $a+ib+jc$ equipped with the rule $i^2 = j^2 = -1$. To test whether such a system could work, he used the principle that the length of the product of two vectors should be equal to the length of the product of their individual lengths as is the case in complex numbers $z = a + ib$:

$$\begin{aligned} z^2 =& (a+ib)^2 = a^2 - b^2 + 2iab \\ \Rightarrow \left|z^2\right| =& \sqrt{Re(z)^2 + Im(z)^2} \\ =& \sqrt{a^4 + b^4 + 2a^2b^2} \\ =& a^2 + b^2 \\ =& |z|^2 . \end{aligned}$$

However, when, in a similar way, performing this test on his number system corresponding to vectors in $\mathbb{R}^3$, Hamilton realized that the vector $a + ib + jc$ squared was equal to

$$(a + ib + jc)^2 = a^2 - b^2 - c^2 + 2iab + 2jac + 2ijbc \tag{18}$$

while

$$|a + ib + jc|^4 = (a^2 - b^2 - c^2)^2 + (2ab)^2 + (2ac)^2.$$

Meaning that the principle discussed above would hold for $ij = 0$. However, he did not like this outcome and realized that instead of writing $2ij$ in Eqaution 18 he should write $ij + ji$ which in turn caused him to sacrifice the commutative law. He later discusses, in a letter, that he preferred setting $ij = -ji$ to satisfy $ij + ji = 0$. He then set $ij = k$ and $ji = -k$ and realized that he could try to create a system within $\mathbb{R}^4$ by clearly defining k as

$$k^2 = (ij)(ij) = i(ji)j = i(-ij)j = -i^2j^2 = -1.$$

Having come to this realization on a walk, he carved the quaternion rule

$$i^2 = j^2 = k^2 = ijk = -1$$

on a bridge in Dublin.

Thus, as opposed to the complex numbers, the quaternions were not invented with the intention to solve an already existing problem, but rather to investigate whether it was possible to expand the complex number system representation of $\mathbb{R}^2$ into a different system that would be represented as 3-tuples of real numbers in $\mathbb{R}^3$. Failing to do so, Hamilton instead invented quaternions which can be represented as 4-tuples in $\mathbb{R}^4$ while simultaneously being able to describe vector transformations in $\mathbb{R}^3$.

4 Applications of complex numbers and quaternions

Problems in real life can often be solved in a myriad of different manners. Therefore, the intention behind this section will be to briefly discuss some applications of complex numbers and quaternions as well as generally illustrate areas where it might be possible to use other measures to solve certain problems, but where using complex numbers and quaternions might provide benefits which other measures might not.

4.1 Complex numbers in physics and engineering

Complex numbers have been discussed in Section 2.2 where varying representations of complex numbers were made with one intent being to motivate some sort of intuition that one might derive from these representations and which might be more difficult to do given different representations. However, a question remains of the applications of complex numbers.

Complex numbers are used in a handful of functions within multiple fields of study. [RTWLTGAN21] briefly mention that complex numbers can often be used in electromagnetism to simplify calculations of electromagnetic waves. Additionally, complex numbers are used in the mathematical transform called the Fourier transform, which can be used by electrical engineers in signal processing in order to analyse, identify and modify components of signals [Wiki2].

Furthermore, complex numbers are famously used in wave functions in quantum physics. [RTWLTGAN21] discuss and describe that within quantum physics, it is possible to motivate both a real quantum theory and a complex quantum theory. However, it is extensively argued in the article that in recent years, some scientists have started to shift from a view of complex numbers being used as tools to facilitate computing in quantum physics to acknowledging that certain problems require complex numbers in order to be solved. [RTWLTGAN21] specifically give several examples of problems that can be described using both real quantum theory and complex quantum theory, but they also provide an example involving two entangled particles which they argue cannot be described using real quantum theory.

4.2 Quaternions in spatial rotation

As has been discussed in Section 2.3, quaternions can be used to describe rotation in three dimensional space. However, there are many other mathematical objects that can be used to describe rotation in three dimensional space as well [H06]. One such mathematical object are the Euler angles; the Euler angles are sets of three fixed axes in a coordinate system that can be used to describe three dimensional rotation by rotating about one axis at a time. This manner of three dimensional rotation does, however, come with certain issues[H06]. An issue that Euler angles have is that each of the sets of axes can be sequentially rotated in such a way that one loses one degree of freedom in the rotation. In other words, one axis of rotation becomes parallel to another leading to both of them describing the same three dimensional rotation, this phenomenon is called gimbal lock[H06]. However, quaternions avoid gimbal lock by being able to rotate along any axis, instead of only being able to rotate along three fixed axes.

As such, professions dealing with three dimensional computer graphics and engineers working with rotation in three dimensional space can use quaternions to describe said rotation instead of Euler angles in order to avoid issues of gimbal lock[H06].

5 Teaching higher number systems

It is generally agreed among researchers and teachers that the prior knowledge that students of mathematics carry should be used to facilitate the teaching of new mathematical concepts or objects [P10][PB17][PLvB19]. Similarly, it is often discussed that when students are in the process of creating a deeper understanding of mathematical concepts or objects, it is crucial that they make connections between the objects, formal definitions as well as individual understanding of the behaviours of said objects [LR12]. One way to make such connections is to connect a given concept to mathematical symbols and objects that are often central when discussing the concept. Furthermore, connections can be made between different representations of the same concept or object [P10][PB17][PLvB19]. For instance, a function as a mathematical object maps elements from a set X to a set Y so that each element in X is mapped to exactly one element in Y. Functions can, however, be represented by graphs, tables, algebraic expressions etc. and being able to make connections between the different representations and the function as an object contributes to a deeper understanding of the concept of function [PST13].

Thus, when teaching complex numbers to students, it is vital to try to draw on their prior knowledge of mathematics. This can of course be done in many different manners. It is, however, important to note that in order to teach complex numbers to students, one also needs to clearly assert that complex numbers do differ from what the students have learned so far. For example, the rule $i^2 = -1$ is one that might seem very alien to new learners and thus solving a multiplication of two given complex numbers might seem confusing to said students. But in such a scenario, the teacher can highlight the similarities that complex numbers have to real numbers in that both are distributive and commutative and that i can just be treated as some unknown variable x with the only exception that, unlike a real variable, $i^2 = -1$. This kind of explanation draws on students' prior knowledge of multiplication and algebra while introducing one new rule, thus creating connections between variables and the imaginary number i.

However, there are dangers in carelessly drawing similarities to objects such as variables because it can cause students to try to view i as a variable that they need to solve for. Therefore, it can be beneficial to represent complex numbers as objects which might appear more intuitive to the students. This can for instance be done by representing complex numbers geometrically in the complex plane as pairs of two real numbers, thus eliminating the possibly confusing i while also drawing on their prior knowledge of points and vectors in two dimensional space.

Given 45 minutes to present complex numbers to a group of upper secondary school pupils with an adequate and somewhat homogeneous understanding of concepts that are usually discussed prior to complex numbers, one could structure a lesson plan similar to the following.

Complex numbers lesson plan:

5 min repetition of the number line and introduction of the imaginary unit i and the $i^2 = -1$ property of i.

10-15 min example of adding and multiplying complex numbers.

5-10 min comparing previous results to similar real polynomials.

10-15 min comparing the number line with the complex plane and placing complex numbers in the plane.

5 min explaining $i^2 = -1$ geometrically as a 90° rotation twice.

The main teaching approach argued for in this text revolves around drawing on student's prior knowledge; it is thus important to have a brief repetition on real numbers and, since the plan is to later show the complex plane to hopefully facilitate the students in understanding the geometrical relationship between real and complex numbers, this can be done by having a brief repetition of the number line. It is also important to introduce the imaginary unit i and its square rule prior to introducing other complex numbers.

Once the rule $i^2 = -1$ has been introduced, it would be beneficial to show some simple examples of addition and multiplication with i. These examples should start as very simple multiplications and additions of small real numbers.

For instance, the teacher could show examples such as

$$
\begin{aligned}
i + i &= 2i \\
0 + i &= i \\
2 + i &= i + 2 \\
3 \cdot i &= 3i \\
2i \cdot i &= -2 \\
0 \cdot i &= 0.
\end{aligned}
$$

Here, the teacher could remind the pupils that these results are not too dissimilar from how addition and multiplication operate with variables; that is, real numbers do not mix with variables when adding $2 + x = x + 2$ and that adding two of the same variable gives $x + x = 2x$. It is however vital to constantly remind the pupils that variable multiplication vastly differs from multiplication of the imaginary number i. Subsequently, the teacher should explain, with the use of the previously shown operations that an expression written on the form $a + ib$ is a complex number. Once these examples are shown and discussed, the teacher can further show examples of addition and multiplication of complex numbers with both real and imaginary components such as

$$
\begin{gathered}
(2 + 2i) + (i + 3) = (2 + 3) + (2i + i) = 5 + 3i \\
(1 + i)i = i + i^2 = i - 1 \\
(2 + i)(2 + i) = 2(2 + i) + i(2 + i) = 4 + 4i + i^2 = 4 + 4i - 1 = 4i + 3.
\end{gathered}
$$

While this might appear as slightly more advanced to the pupils depending on the group, the teacher can again draw similarities to polynomial addition and multiplication with the only difference being that any occurrence of i^2 should be replaced with -1 and here it is also important to explicitly state and stress that unlike problems containing variables, the pupils should never try to solve for i as it is not a variable and instead a number. By doing so, the teacher can show and even explicitly express that complex numbers distribute just like real numbers do and that performing the operations is per se not any different from how they are used to doing it with real numbers and variables.

To further illustrate the point that i is not to be solved for and that it is a number, the teacher can then refer back to the number line and make it into a complex plane in order to show that like numbers, one can graphically pinpoint multiples of imaginary numbers on the imaginary axis. Furthermore, the teacher can then, in order to show the geometrical representation of complex numbers, mark certain coordinates on the complex plane and have the students try to figure out what complex number is marked. This would again draw on the pupils' knowledge of planes, axes, coordinates, points and vectors.

Finally, to truly motivate the rule $i^2 = -1$, the teacher can show what happens to complex numbers as represented in the complex plane when they are multiplied by i; that is, that there is a 90° rotation counterclockwise from the given number. This can easily be done by first showing

$$\begin{aligned} 1 \cdot i &= i = i^1 \\ i \cdot i &= -1 = i^2 \\ -1 \cdot i &= -i = i^3 \\ -i \cdot i &= 1 = i^4 = i^0 \end{aligned}$$

and then possibly showing how this transforms other complex numbers in the same manner.

As is the case with any teaching, however, it is crucial to consider the group and the current situation in a broader context before teaching complex numbers. What has hitherto been proposed is simply just an example of how such a lesson plan could potentially be structured. There are many other, equally yielding ways of teaching complex numbers, one can for instance do so by expanding on solving polynomial equations without real roots. However, the main argument made here is to use the pupils' prior knowledge and make connections to different representations and objects when teaching new mathematical concepts.

Teaching quaternions in the upper secondary school context is rare and, in Sweden to the very least, does not fall under any of the regular mathematics courses. There are however situations where teaching of more advanced mathematical concepts happen. There can for instance exist additional optional activities for those who perhaps are interested and passionate about mathematics. Additionally, some schools offer optional courses that teach some sort of specialized fields of mathematics. I have in my experience as a teacher student witnessed, during one of my internships in an upper secondary school in Sweden, a mathematical course that was offered for last year pupils that have shown adequate results in mathematics as well as in general subjects. This course revolved around mathematics and programming surrounding artificial intelligence and the teacher of the course was part of a mathematical community collaborating with universities. Among other concepts, this course included set theory as well as proof by mathematical induction, both of which are generally treated as being of a level more advanced than upper secondary school mathematics. It is possible that higher number systems could be taught in some optional course similar to this. As such, the following proposal of teaching quaternions is based on the assumption that the group being taught consists of last year upper secondary school pupils in Swedish programmes that teach more mathematics, such as the natural sciences. Assuming that a lesson is around 45 minutes long one could structure it as follows

Quaternions lesson plan:

5 min repetition of complex numbers on Cartesian form.

5 min compare the Cartesian form to complex numbers represented as 2-tuples.

5 min introduce quaternions as 4-tuples.

5-10 min reconstruct quaternions on the form $a + ib + jc + kd$ from the 4-tuple.

5 min show the $i^2 = j^2 = k^2 = ijk = -1$ rule and compare it to complex numbers.

10 min discuss addition and multiplication with real scalar.

10 min discuss the non-commutative property in quaternion multiplication and create a multiplication table.

Teaching quaternions as a number system can be confusing and unintuitive. However, in this regard, quaternions have many similarities with complex numbers. Complex numbers can be represented as a 2-tuple of real numbers and quaternions can equally be represented as a 4-tuple of real numbers, thus, drawing on the pupils prior knowledge of complex numbers would be something that can easily be done given that they have a chance to first have a brief repetition of complex numbers. By having a repetition of complex numbers as 2-tuples and as the expression $a + ib$, it further allows the pupils to make representational connections within the concept of complex numbers. This connection that hopefully establishes in the pupils can then be translated to quaternions.

Introducing quaternions as a number system beyond the complex numbers and representing it as a 4-tuple makes possible for the pupils to make the connection between complex numbers and quaternions as being similar with one existing in higher dimensions. Now using this 4-tuple, one can try to reconstruct the quaternion on the $a + ib + jc + kd$ form. It can be beneficial to show an example with real coefficients such as first showing $3 + 2i$ which can be represented as $(3, 2)$ and then showing the 4-tuple $(4, 3, 1, 2)$ and reconstructing it into $4 + 3i + j + 2k$ in order to illustrate that complex numbers and quaternions can be expressed in similar forms.

It is then beneficial to remind the pupils that the unit number in complex numbers has to follow the $i^2 = -1$ property, from there the teacher can build upon how in a similar way both j^2 and k^2 equal -1 with the addition that ijk also equals -1 giving the rule $i^2 = j^2 = k^2 = ijk = -1$.

However, because quaternions lose the commutative property in multiplication, it becomes vital to first show where similarities between quaternions and complex numbers lie, by illustrating addition and multiplication by real scalars, and subsequently explicitly state and discuss the crucial differences of these two number systems. This can also be done through examples such as

$$(2 + 4j + k) + (1 + 3i + 2k) = 2 + 1 + 3i + 4j + (1 + 2)k = 3 + 3i + 4j + 3k$$
$$2 \cdot (1 + 2i + 3j + 4k) = 2 + 4i + +6j + 8k.$$

Lastly, the arguably most unique property of quaternions compared to previous number systems, is that quaternion multiplication is non-commutative. This needs to clearly be communicated to the pupils and it can seem difficult if they have not previously considered that mathematical objects can in fact be non-commutative. An example of this can be shown if the pupils seem to fully grasp everything that has been discussed so far and if the teacher is aware of the precise proficiency level of the pupils, but a proof of this can also be saved for a later date if need be. What needs to be communicated to the pupils is that quaternions are generally not commutative and explain how left and right side multiplication is performed and subsequently it might be useful to, together with the pupils, create a table such as the one in Lemma 1 by using the rule $i^2 = j^2 = k^2 = ijk = -1$.

Finally, again, teaching a new concept to a class or group of pupils can be done in many more ways than what has just been presented here. For instance, if the pupils for some reason already have worked with mathematical objects such as matrices, which are generally not commutative as well, it might be more helpful for the teacher to draw on this knowledge to facilitate the understanding of quaternion multiplication.

The main point of this entire section is that learning entirely new mathematical concepts, such as the two that have been covered here, can cause difficulty and confusion in pupils and this issue can, to some degree at least, be eliminated by, instead of teaching something in a completely new way, teaching a new concept by drawing on the pupils' prior knowledge and thus effectively only making a small portion of the new concept something that appears as new to the pupils. Having the pupils be familiar with most operations and properties of new concepts helps them in only focusing on the concepts that they might lack prior knowledge to combat, thus reducing their workload.

6 Conclusion

The aim of this text was to explore the complex numbers as well as the quaternions; that is, to explore how to compute using them, as well as create some sort of intuition on them as number systems and mathematical objects. This was done by representing both the complex numbers and the quaternions in terms of expressions only containing real values that can geometrically be tied to $\mathbb{R}^2$ and $\mathbb{R}^3$ as well as motivating their discovery. Hopefully, the reader can now, however elementary, grasp that there is intuition to be recovered even from number systems that do not necessarily appear as natural as others. Additionally, as a consequence of this paper being part of the teacher education programme, an approach to teaching higher number systems based on prior knowledge hopefully also illustrated that teaching and learning can be facilitated by considering intuition in the form of prior knowledge.

References

[EHHKMNPR91] Ebbinghaus, H.-D, Hermes, H., Hirzebruch, F., Koecher, M., Mainzer, K., Neukirch, J., Prestel, A. & Remmert, R. (1991). *Numbers.* Springer, New York.

[H06] Hanson, A.J. (2006). *Visualizing quaternions.* Morgan Kaufmann Publishers Inc., San Francisco.

[RTWLTGAN21] Renou, M.-O., Trillo, D., Weilenmann, M., Le, T. P., Tavakoli, A., Gisin, N., Acin, A. Navascues, M.(2021). Quantum theory based on real numbers can be experimentally falsified. *Nature, 600*, 625-629.

[M06] Merino, O. (2006). *A short history of complex numbers.* Retrieved May 8, 2022, from `https://www.math.uri.edu/~merino/spring06/mth562/ShortHistoryComplexNumbers2006.pdf`.

[MS20] Mokhithi, M. Shock, J. (2020). *Introduction to complex numbers.* Retrieved May 8, 2022, from `https://open.uct.ac.za/bitstream/handle/11427/31769/Mokhithi%20%26%20Shock%20_%20Complex%20Numbers%20v1.pdf?sequence=1`

[K99] Kuipers, J. B. (1999). *Quaternions rotation sequences: A primer with applications to orbits, aerospace and virtual reality.* Princeton University Press, Princeton.

[LR12] Larsson, M. Ryve, A. (2012). Balancing on the edge of compentency-oriented versus procedural-oriented practices: Orchestrating whole-class discussions of complex mathematical problems. *Math Ed Res J, 24*, 447-465.

[PB17] Pettersson, K. Brandell, G. (2017). *Att utveckla elevers begreppsförmåga.* Skolverket. Retrieved May 8, 2022, from `https://larportalen.skolverket.se/LarportalenAPI/api-v2/document/path/larportalen/material/inriktningar/1-matematik/Gymnasieskola/446-hogkoleforberedandematematikundervisning%20GY/3_attutvecklaeleversbegreppsformaga/Material/Flik/Del_03_MomentA/Artiklar/HFBGy_03A_01_attutvecklaeleversbegreppsformaga.docx`.

[PLvB19] Pettersson, A., Liljekvist, Y. van Bommel, J. (2019). Studying concept elements as a way to trace students' conceptual understanding. *Nordic Studies in Mathematics Education, 24* (1), 5-26.

[PST13] Pettersson, K., Stadler, E. Tambour, T. (in press). Development of students' understanding of the threshold concept of function. *Proceedings of the Eight Congress of the European Society for Research in Mathematics Education CERME8*, 6-10 February 2013, Antalaya, Turkey.

[P10] Persson, P.-E. (2010). *Räkna med bokstäver: En longitudinell studie av vägar till en förbättrad algebraundervisning på gymnasienivå.* [Doctoral dissertation, Luleå Tekniska Universitet]. Luleå Tekniska Universitet, Luleå.

[Wiki1] Wikipedia page on Cayley-Dickson construction. Retrieved May 8, 2022, from https://en.wikipedia.org/wiki/Cayley%E2%80%93Dickson_construction

[Wiki2] Wikipedia page on singal processing. Retreieved May 8, 2022, from https://en.wikipedia.org/wiki/Signal_processing.