Quaternionic Analysis

Theory of functions of one variable

Lorenzo Matarazzo

Linktree: https://linktr.ee/lorenzomatarazzo
Youtube: https://www.youtube.com/@Philomatha

Contents

0.1 Historical context

0.1.1 How were quaternions born?

Before diving into the theory of functions of a quaternionic variable and the algebra of the skew-field of the quaternions $\mathbb{H}$, it would be a good idea to make a quick historical digression. In fact, quaternions are, above all, a very interesting mathematical discovery from an historical perspective; their discovery predates and motivates a big amount of very important definitions and ideas from vector analysis that we commonly use nowadays.

Figure 1: Commemorative plaque on Broom bridge, the place where Hamilton first carved in stone his famous multiplication rule. [image taken from Wikipedia [Link1], uploaded by user Cone83. License: CC BY-SA 4.0].

Quaternions are attributed, according to classical mathematical historiography, to the Irish mathematician Sir William Rowan Hamilton. However, in the same period of time, an analogous theory was obtained by the French mathematician Benjamin Olinde Rodrigues, even though the latter was mostly concerned with problems of geometric nature, since it was developed precisely to tackle problems involving three-dimensional rotations [Note 1]. The discovery (or invention, depending on the philosophical views of the reader) of quaternions by Hamilton is almost legendary. It is said that the Irish mathematician had a flash of genius during a walk near the Royal Canal in Dublin with his wife Helen [Note 2]. The discovery of this number system delighted the mathematician to the point of bringing him to carve his famous ”Hamilton rule” in the stones of the Broom bridge.

0.1.2 Why quaternions?

At this point the following question arises: what motivated the development of quaternions?

To answer this question in a proficuos way it would be advisable to try to put ourselves in the shoes of a mathematician of the time.

We are in the first years of the 1840s; complex numbers have already been used as a mathematical tool for over 300 years (let's recall that they were first introduced by the Italian mathematician Gerolamo Cardano in 1545, in his famous treatise ”Ars Magna”, developed as a concept while he was working on some special cases of the Tartaglia's resolutive formula for cubic equations [Note 3]).
Not only were mathematicians of the time used to complex numbers and aware of their geometrical properties, but the first years of the 19th century were precisely the years in which the field of complex analysis had its most significant advances and most of the fundamental theorem of said branch were proven (the famous ”Cauchy-Goursat theorem”, a theorem of great importance for complex analysis, was heuristically sketched in some documents dated back to 1814, but was only fully proven in 1825 [Note 4]).
Finding ourselves in a period preceding the development of various physical theories who employed complex numbers in their mathematical formalism, we could ask ourselves what generated this strong interest for a number system that, at a first glance, might seem disconnected from reality; the motivation came mostly from their interesting geometric properties, studied in an extensive manner by mathematicians such as Wessel and Argand [Note 5]. Why was it important to talk about this? Because these were the exact same reasons for which quaternions were developed. Hamilton, initially, wanted to develop an algebra of numbers that he called ”triplets”, geometrically represented by points in space, that ideally should have let him represent 3d rotations via the operation of multiplication, analogously to the behaviour of complex numbers. However, all of his attempts were vain, and Hamilton worked for over 10 years on this problem without results. It was only the 16 of October 1843 that, during the walk we discussed earlier, he was able to develop an algebra with desireable properties that allowed to multiply, not triplets, but quadruples of numbers. The name quaternions, in fact, derives precisely from the masculine medieval latin noun ”quaternio”, that can generally be translated as ”four” or ”quadruple”. The choice of this name, at least according to various documents of his student P.G Tait [Note 6], was probably motivated by Hamilton's great interest in the pythagorical mystical symbol of the ”teractys” (that was seen by the pythagorean school as a representation of the ”kosmos”, mainly for its relationship with music and the movement of planets).
Moreover, quaternions are of extreme interest for the study of the history of mathematics, since most of the modern notation for a great amount of mathematical notions is due to the development of quaternions. In fact, words such as vector and scalar were employed for the first time in the

works of Hamilton regarding quaternions. The word vector, in fact, derives from the latin expression "vehere", that means "to transport/to guide", while the word "scalar" from the latin noun "scala", that could be translated as "scale". Also the pretty common notation used for the orthnormal basis of $\mathbb{R}^3$, $\{i, j, k\}$ to represent $\{[1,0,0]^T, [0,1,0]^T, [0,0,1]^T\}$ is due to quaternions. With this said, we will return on this historiographical matter in due time, once we will have introduced the basic notions of quaternion algebra, since by doing so it will be easier to comprehend the origin of these pieces of notation.

0.1.3 What was the reaction of the scientific community of the time?

Another important topic to explore is the reaction that the mathematicians of the time had to the discovery of quaternions. In fact, we can say that among the mathematicians of the time there was a schism between those who thought quaternions were a great discovery, and potentially the tool that could describe all of physics in an unitary language, and those who considered them to be unnecessarily complicated and not very useful. To the first group belonged mathematicians such as Hamilton, Alexander MacFarlane, Peter G.Tait, Charles Jasper Joly, Robert Ball and van Elfrinkhof; in 1899, some of the mathematicians in this group even formed the "Quaternion Society" [Note 7], a scientific society whose goal was to promote the study and the employment of quaternions and various other hypercomplex systems. In the opposite group, on the other hand, we had many other important mathematicians, such as Oliver Heaviside, the father of modern vector analysis, "freed" from its initial quaternionic component that the subject had at the beginning.

0.2 Why quaternions today?

According to what was said in the previous section, we could probably think that quaternions are quite useless, and it's thus superfluous to investigate their properties. However, even though quaternions are undoubtedly a tedious system to work with (especially with respect to algebraic computations, since the system is not commutative and computations are often long), but also for the difficulties encountered while trying to define an analysis for functions of a quaternionic variable, it is not a good idea to "throw the baby out with the bathwater". In fact, quaternions have a great amount of fields in which they're employed extensively, such as computer graphics. I'm sure some of the readers that use softwares for 3d graphics and modelling might have stumbled upon the settings "Euler angles" or "Quaternions". Quaternions are, nowadays, one of the preferred systems to represent rotations in space, since they don't present the problem of "Gimbal Lock", that happens when two rings of rotations of a gimbal become allineated, losing thus a degree of freedom; this phenomenon is typical of systems which implement rotations with euler angles. In addition, quaternions are nowadays applied in some fields of physics as well; they're useful in rigid body dynamics and they're also isomorphic to the space of hermitian matrices that act on a 2-dimensional complex Hilbert space, such as the space of a qubit or that of particles of spin $\frac{1}{2}$ (fermions). It is of great importance for physics their relationship with the Lie groups $SO(3)$ and $SU(2)$ as well; this relationship will be covered in the text in due time.

0.3 Goals of the text

In this text we will analyze the real 4-dimensional associative unitary algebra of quaternions, starting with the proofs of some fundamental algebraic facts about said structure, and ending with a focus on their metric topology induced by their canonical norm, that coincides with the canonical norm of 4-dimensional euclidean space. Once we will have done this, we will use the topological definitions of continuity and limit of a function between metric spaces in the particular case of quaternions, making some examples, and we will develop an elementary theory of functions of one quaternionic variable; this is not a very difficult endeavour up to this point. During our journey, we will encounter our first obstacles when we will try to extend differential calculus to functions of a quaternionic variable. This difficulty stems from the lack of multiplicative commutativity of quaternions. Quaternions, in fact, in contrast with complex numbers and real numbers, are just a skew-field, in other words they are a field minus the commutative property of multiplication. Let's see what happens more in depth: let's suppose we wanted to define a notion of derivative in the same way it is defined in real and complex analysis courses; we will encounter two main problems:

- The definition of "derivative of a function" would not be unique. By the non commutativity of multiplication, we will necessairily have to define a "right derivative" (where we will multiply the factor $f(x+h)-$

$f(x)$ by h^{-1} on the right) and a "left derivative" (where we will multiply by h^{-1} on the left).

- The definition of derivative we gave above would be extremely exclusive. In fact, we will prove that being a quaternionic left/right-differentiable function implies that said function is left/right linear. The class of differentiable functions under this definition is "way too small" to be of interest.

These two problems, as we will see by proceeding with our treatise, are not the end of the world. In fact, the solution will just come by defining the notion of derivative and differentiability in a different manner. The only big problem that will be persistent will be that, in contrast with real and complex analysis, here we will have various equally fruitful developments of a differential calculus of quaternionic functions. One of the most important ones was developed at the beginning of the 20th century by Fueter, Moisil, Haefeli [16] [14] [13] [15] [21] and many other mathematicians, translated and reported in the english language for the first time in the important paper by C.A Deavours "The Quaternion Calculus" [6], and then expanded in the equally important paper by the English mathematician and physicist A.Sudbery "The Quaternionic Analysis" [43].
This theory was capable of developing many quaternionic versions of impotant theorems of complex analysis, starting from the famous Cauchy's theorem (we will show the complex form on the left and the quaternionic form on the right):

$$\oint_\gamma f(z)dz = 0 \qquad \oint\!\!\!\oint\!\!\!\oint_C Dq\phi(q) = 0$$

where here $f : \mathbb{C} \to \mathbb{C}$ is a holomorphic complex function on and inside γ, that is a simple piecewise smooth closed curve, and where $\phi : \mathbb{H} \to \mathbb{H}$ is a left-regular quaternionic function in a domain $U \subset \mathbb{H}$ (we will give the meaning of regularity of quaternionic functions in due time) and C is a smooth 3-chain homologous to 0 in the singular homology of U (the text will start with a quick recap of algebraic topology, so even if the reader is not comfortable with these notions he needs not to worry). As we can see the mathematical form of these two theorems is very similar. Another result that highly resembles a result in complex analysis is the quaternionic form of Cauchy's integral formula (where we will again place the complex version on the left and the quaternionic version on the right):

$$f(z) = \frac{1}{2\pi i}\oint_\gamma \frac{f(w)}{w-z}dw \qquad n\phi(q_0) = \frac{1}{2\pi^2}\oint\!\!\!\oint\!\!\!\oint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2}Dq\phi(q)$$

Here we will omit the hypotheses for the two theorems since they're rather convoluted; we will just specify that here n denotes the winding number of the smooth chain C around q_0.
Obviously, other than just the theory previously cited, we will also cover other approaches to the analysis of functions of a quaternionic variable, such as that presented by C.G Cullen [5] [19].

0.4 Who is the target audience for this text?

As the reader probably already noticed by reading the introduction, to study quaternionic analysis a knowledge of differential geometry, general topology and algebraic topology is needed. Nevertheless I tried to make this text accessible also to people who didn't study these subjects before, by including an initial chapter in which they are exposed in such a way to just give the notions necessary to understand the text and nothing else superfluous. In this way, to read the text, you will just need good knowledge of real analysis, complex analysis and algebra (groups, vector spaces exc.). If the reader has already studied algebraic topology and differential geometry, then I recommend to skip the first chapter (in that case also some parts of the fourth chapter could feel superfluous; if that is the case I recommend to read it faster).

0.5 Note to the reader

Let us now conclude this introduction with some remarks of a non-mathematical nature. First of all I wanted to specify that, at the time of writing the text I was 17 years old; because of my young age at the time I was not part of any academic community, and for this reason I apologize in advance for any possible errors; in case the reader should come across any errors, I would be immensely grateful if they could communicate them to me. That said, let's also discuss the conventions I will adopt in this text to make its fruition clearer. I will enclose all bibliographic references in square brackets (as is customary), created using the biblatex library in LaTeX. The complete bibliography is found in the last pages of the text, and citations will be linked to specific texts there via numbers. Additionally, when necessary, I will also include the relevant page from the cited text alongside the citation. Square brackets within the text will also be used to refer the reader to the footnotes section, which is also located at the end of the text, where certain propositions from the text will be clarified, or bibliographic

references with accompanying explanations will be provided.
The majority of illustrations and graphs in this text are realized with the following libraries (or softwares): matplotlib, PGF/TikZ, GeoGebra and in some instances Blender and Gimp.
The cover of the book is a quaternionic fractal made by me with the freeware ChaosPro (download link: http://www.chaospro.de).

Chapter 1

Preliminary notions

1.1 Elements of topology

1.1.1 Why topology?

Let's begin the first section of the chapter by discussing topological spaces: what are they? It's always a good idea to answer this question by starting with an etymological analysis of the term.
The word "topology" is of Greek origin, specifically derived from the Greek words "topos" (place) and "logos" (reasoning/study/thought).
So, it can be translated into English, somewhat abruptly, as "study of place."
Before we continue, however, I suppose the reader may have the following doubt: how does topology differ from geometry? Isn't it also a study of space? The answer lies in the genesis of the discipline: the solution to the Seven Bridges of Königsberg problem, formulated and solved by Leonhard Euler in 1736, is considered a significant turning point [Note 1.1]; it is here that Euler realized that the geometric and mathematical considerations made to solve the problem were fundamentally different from those encountered in classical Euclidean geometry, and therefore, the geometry used was not "quantitative geometry," but rather qualitative geometry.
Towards the end of the 19th century, the term "topology" was increasingly used in academic language with the meaning, precisely, of qualitative geometry.
But what do we mean by "place"? Let's think of a concrete example; looking around, we realize that we live in a three-dimensional space (mathematically abstractly modelable as $\mathbb{R}^3$).
This space is a set of points, that is, triples of real values that indicate the

position of an entity in space. But now we must ask ourselves a question: is the space we live in just a set? Or is there perhaps some structure that makes it more ”interesting”? It seems very plausible to answer affirmatively to the second question; let's consider any set, for example, $A = \{\alpha, \beta, \gamma\}$. Can we say that α is close to γ? Can we calculate the distance between two elements? If we are given only this set, without any extra information, the answer is clearly no, such notions are not well-defined.
In our surrounding three-dimensional space, on the other hand, the notions just mentioned are well-defined:

- We can measure distances. For example, Marco (modeled as a point $X \in \mathbb{R}^3$, $X = (x_1, x_2, x_3)$) can say that he is at a certain distance d from Mario (modeled as another point $Y \in \mathbb{R}^3$, $Y = (y_1, y_2, y_3)$). The distance is calculated using the formula:

$$d = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} = d(X, Y)$$

  Note that this is a function that maps a pair of points in space, $(X, Y) \in \mathbb{R}^3 \times \mathbb{R}^3$, to a real number. We will call this function the Euclidean metric of $\mathbb{R}^3$, or 3-dimensional Euclidean metric; we will choose the name ”metric” because this function precisely allows us to measure distances in our space.

- Being able to measure distances precisely, the concept of ”closeness” is obviously well-defined in our space; in common language, we will say, for example, that we are ”close” or ”in the neighborhood” of something if our distance from that object is not significant (the word ”significant” in this case will be arbitrary). Therefore, to formalize what has just been said in mathematical language, we will instead say that we can say we are close to an object if there is a distance d that separates us.
  Marco from the previous example can thus say that he is close to Mario.

As we have just seen, our space $\mathbb{R}^3$ is endowed with a superstructure that defines these concepts on it; these concepts are also well-defined on the real line $\mathbb{R}$ or in the plane $\mathbb{R}^2$.
Note that it is thanks to this superstructure that we can perform ”analysis” on these spaces:
The concept of limit is closely related to the notion of distance in $\mathbb{R}$, and, in turn, so is the derivative, as it is itself a limit.

But even more significantly, the important notion of the continuity of a function depends on the concept of distance, or more generally, ”closeness.” We intuitively say that a function $f : \mathbb{R} \to \mathbb{R}$ is a continuous function if it ”has no gaps in its graph,” that is, a function is continuous at a point $x_0 \in \mathbb{R}$ if points ”sufficiently close to x_0” on the x-axis are mapped to points ”sufficiently close to $f(x_0)$” on the y-axis.
Having finished this digression, we can finally explain the purposes of general topology.

1.1.2 Topological Spaces

In the previous part of the text, we talked about some of the ”extra” properties of the set $\mathbb{R}^3$ that make it ”more than just a set.” The textbook definition (which technically provides the best idea of what is studied in a general topology course, provided you know what you are talking about) is as follows: topology is the branch of mathematics that studies topological spaces and their topological properties, defined as those properties preserved by applications called homeomorphisms.
However, this definition may seem meaningless to someone who is just getting started with the subject. Therefore, we will introduce the reader to topology in a slightly different way.
Think about what was said earlier about the space $\mathbb{R}^3$ and the arbitrary set from the previous example, $A = \{\alpha, \beta, \gamma\}$: what if we wanted to introduce a notion of ”closeness” on this set? (or even better, of ”distance,” but we will see that later when we talk about metric spaces). This is where the concept of a topological space comes into play; we want to endow an arbitrary set with a concept of closeness among its objects. Here is the definition of a topological space:

Definition 1.1 (Topological Space)**.** *A topological space consists of a set X, together with a class of subsets of X associated with each point $p \in X$. We will call the subsets associated with a point $p \in X$ neighborhoods of p. We also require that the neighborhoods defined above satisfy the following properties:*

1. *If $U \subset X$ is a neighborhood of $p \in X$, then $p \in U$.*
2. *Every subset of X that contains a neighborhood of a point $p \in X$ is also a neighborhood of p. In symbols: if $V \subset X$, then if there is a neighborhood U of p such that $U \subset V$, then V is a neighborhood of p.*

3. *If U is a neighborhood of p, then there exists a neighborhood V of p such that U is a neighborhood of every point in V.*

4. *If U and V are two neighborhoods of a point p, then their intersection $U \cap V$ is also a neighborhood of p.*

*We call a set X, together with the class $\mathscr{N} := \{\mathscr{N}_p\}_{p \in X}$ of neighborhoods for each point $p \in X$, a **topological space**, and write it as $(X, \mathscr{N})$.*

The classes of subsets of X associated with each point p are precisely how we rigorously realize our initial goal: remember that initially, the goal was to endow a set with a concept of closeness between points.
The idea is as follows: we give each point p a class of subsets that contain the points "close" to p," specifying which points are close to it, and consequently, we force this class of sets to obey certain rules for our abstract structure to be consistent with our original intuitive idea. Now let's provide some examples to clarify the abstract definition just presented:

Example 1.1. *To see if the definition "matches" with the example of $\mathbb{R}^3$ we discussed earlier, let's check if the sets of "close points" to a point in space $P \in \mathbb{R}^3$ form a set of neighborhoods.
Recalling how we defined the concept of closeness in $\mathbb{R}^3$ earlier, let's now make it rigorous by saying that a neighborhood of a point $P \in \mathbb{R}^3$ is a set containing a sphere of radius r centered at P, $B(P, r) = \{Q \mid d(P, Q) < r\}$. Finally, let's demonstrate that this class of sets forms a set of neighborhoods:*

1. *Suppose that $U \in \mathbb{R}^3$ is a neighborhood of a point $P \in \mathbb{R}^3$. By the definition given above, it follows that U contains a sphere (more commonly called a ball) of radius $r > 0$ with center P, $B(P, r)$. But this ball trivially contains P since $d(P, P) = 0 < r$, and therefore $P \in B(P, r) \subset U \implies P \in U$.*

2. *Let $S \subset \mathbb{R}^3$, and suppose that it contains a neighborhood of a point $P \in \mathbb{R}^3$, $U \subset S$. According to the definition given above, we have that the neighborhood U contains a sphere of radius r and center P, $B(P, r) \subset U$. But since $U \subset S$, we have $B(P, r) \subset S$. Hence, S is a neighborhood of P.*

3. *Let U be a neighborhood of a point $P \in \mathbb{R}^3$. Then there exists a ball $B(P, r)$ such that U contains it. Notice that all balls $B(P, \tilde{r})$ where $\tilde{r} < r$ are trivially contained in U. Therefore, the ball $B(P, \frac{r}{4})$ is contained in U. Now, let's show that U is a neighborhood of every point in $B(P, \frac{r}{4})$.*

To do this, we want to find a ball $B(Q,\epsilon)$, centered at Q with radius ϵ, contained in U.
By the definition of a ball, we have that $\tilde{Q} \in B(Q,\epsilon)$ if and only if $d(\tilde{Q},Q) < \epsilon$. But since Q was an element of $B(P,\frac{r}{4})$, we have $d(Q,P) < \frac{r}{4}$.
Applying the triangle inequality (in the case of the 3-dimensional Euclidean metric):

$$d(\tilde{Q},P) \leq d(\tilde{Q},Q) + d(Q,P) = \epsilon + \frac{r}{4}$$

If the distance $d(\tilde{Q},P) < r$, then $\tilde{Q}$ will be contained in the ball $B(P,r)$ contained in U by assumption. Therefore, we require ϵ to be a positive real number such that $\epsilon + \frac{r}{4} < r \Longrightarrow \epsilon < \frac{3r}{4}$.
We have found the desired ball, and we have shown that U is a neighborhood of every point Q in the ball $B(P,\frac{r}{4})$, as it contains an open ball around each Q.

4. *Let U and V be two neighborhoods of a point P in $\mathbb{R}^3$. By hypothesis, there exist two balls $B(P,\rho) \subset U$ and $B(P,\tilde{\rho}) \subset V$. But since the two balls have the same center, and real numbers form a totally ordered set (in simpler terms, given two real numbers $r_1, r_2 \in \mathbb{R}$, we have either $r_1 < r_2$, or $r_1 > r_2$, or $r_1 = r_2$), one of them must be contained in the other, that is:*

$$B(P,\rho) \subset B(P,\tilde{\rho}) \ \text{ or } \ B(P,\tilde{\rho}) \subset B(P,\rho)$$

In the first case, we have $B(P,\rho) \subset U$ and $B(P,\rho) \subset B(P,\tilde{\rho}) \subset V$, so $B(P,\rho) \subset U \cap V$.
In the second case, $B(P,\tilde{\rho}) \subset V$ and $B(P,\tilde{\rho}) \subset B(P,\rho) \subset U$, so $B(P,\tilde{\rho}) \subset U \cap V$; therefore, in both cases, we can find a ball contained in the intersection of the two neighborhoods: the intersection of two neighborhoods is therefore also a neighborhood.

Therefore, the set of this class of sets associated with every point $P \in \mathbb{R}^3$, called neighborhoods, $\mathscr{N}$, forms a system of neighborhoods on the set $\mathbb{R}^3$, and thus $(\mathbb{R}^3, \mathscr{N})$ forms a topology.

This definition of a neighborhood of a point, defined in an entirely analogous manner, that is, as a set containing a ball centered at the point in question and with a certain radius $r > 0$, can be extended to the n-dimensional Euclidean space $\mathbb{R}^n$. We leave it to the reader to prove that such sets form a system of neighborhoods for $\mathbb{R}^n$, as the proof is entirely analogous to the one provided earlier for $\mathbb{R}^3$.

Exercise 1.1. *Let $\mathbb{R}^n$, with $n \geq 1$, be the n-dimensional Euclidean space. Recall that on $\mathbb{R}^n$, the Euclidean metric between 2 points $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ is given by:*

$$d(P,Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = \sqrt{\sum_{m=1}^{n} (p_m - q_m)^2}$$

and that a ball with center P and radius $\rho > 0$ becomes, consequently:

$$B(P,\rho) = \left\{ Q \in \mathbb{R}^n \mid \sqrt{\sum_{m=1}^{n} (p_m - q_m)^2} = d(P,Q) < \rho \right\}$$

We call a subset of $\mathbb{R}^n$, U, a neighborhood of $P \in \mathbb{R}^n$ if it contains a ball centered at P.
Prove that the set of neighborhoods classes at all points of $\mathbb{R}^n$ forms a system of neighborhoods $\mathcal{N}$, and thus $(\mathbb{R}^n, \mathcal{N})$ forms a topology.

Exercise 1.2. *Give the set $A = \{\alpha, \beta, \gamma\}$ a topology, i.e., associate with each point of it a set of neighborhoods that satisfies the properties of definition 1.1.*

If we have two topological spaces, X and Y, we can naturally define a topology on their Cartesian product $X \times Y$ induced by the respective topologies of the original spaces.

Definition 1.2 (Product Topology)**.** *Let X and Y be two topological spaces. We call a set $S \subset X \times Y$ a neighborhood of $P = (x,y) \in X \times Y$ if there exists a neighborhood U of x in X and a neighborhood V of y in Y such that $U \times V \subset S$. This assignment of sets of neighborhoods for every point in $X \times Y$ forms a system of neighborhoods on $X \times Y$. We call the topology obtained in this way the product topology of X and Y.*

Proposition 1.1. *The product topology just defined is a topology on $X \times Y$.*

Proof. We will prove that the neighborhoods defined above satisfy the necessary axioms.

1. Let's assume that S is a neighborhood of (x,y). Then, by hypothesis, we know that there exists a neighborhood U of x in X and a neighborhood V of y in Y such that $U \times V \subset S$. But since X and Y are topological spaces, their neighborhoods satisfy the necessary axioms, and thus we know that $x \in U$ and $y \in V$. It follows that $(x,y) \in U \times V \subset S$, and therefore $(x,y) \in S$.

2. Suppose that T is a subset of $X \times Y$ containing an open neighborhood S of (x, y). By hypothesis, we have that there exist a neighborhood U of x and a neighborhood V of y such that $U \times V \subset S \subset T$, and thus, T is a neighborhood of (x, y).

3. Suppose that S is a neighborhood of (x, y). Then there exists a neighborhood U of x and a neighborhood V of y such that $U \times V \subset S$. But since X and Y are topological spaces, their neighborhoods satisfy the axioms of definition 1.1. Therefore, there will exist a neighborhood $\tilde{U}$ of x such that U is a neighborhood of every point in $\tilde{U}$, and a neighborhood $\tilde{V}$ of y such that V is a neighborhood of every point in $\tilde{V}$. It follows that S will be a neighborhood of every point in the neighborhood $\tilde{U} \times \tilde{V}$.

4. Suppose we have two neighborhoods of (x, y), S and $\tilde{S}$. By hypothesis, we know the existence of two neighborhoods U and $\tilde{U}$ of x, and two neighborhoods V and $\tilde{V}$ of y, such that $U \times V \subset S$ and $\tilde{U} \times \tilde{V} \subset \tilde{S}$. But since X and Y are topological spaces, the mentioned neighborhoods satisfy the axioms of definition 1.1. Therefore, we will have that $U \cap \tilde{U}$ is a neighborhood of x, and $V \cap \tilde{V}$ is a neighborhood of y. Now, let's consider the set $S \cap \tilde{S}$. Notice that, since $U \times V \subset S$ and $\tilde{U} \times \tilde{V} \subset \tilde{S}$, we have $(U \times V) \cap (\tilde{U} \times \tilde{V}) \subset S \cap \tilde{S}$.
   But $(U \times V) \cap (\tilde{U} \times \tilde{V}) = (U \cap \tilde{U}) \times (V \cap \tilde{V}) \subset S \cap \tilde{S}$, and therefore $S \cap \tilde{S}$ is a neighborhood of (x, y).

□

Exercise 1.3. *Let A be the set* $A = \{\alpha, \beta, \gamma\}$*. Choose two systems of neighborhoods* $\mathcal{N}_1$ *and* $\mathcal{N}_2$ *arbitrarily in order to form the two topological spaces* $X := (A, \mathcal{N}_1)$ *and* $Y := (A, \mathcal{N}_2)$*. Then, write down the systems of neighborhoods for the product topology* $X \times Y$ *for every point in* $A \times A$*.*

1.1.3 Open Sets and Interior Points

Once we have defined a concept of ”nearness” in our space, it's time to mathematically define what it means for a point to be on the ”boundary” of a set and what it means to belong to the interior of a set.
As before, let's first reason intuitively: roughly speaking, we could say that an ”interior” point of a figure is a point immersed in it, surrounded in every direction by other points in the set. Otherwise, we would say it's a boundary point.

Now that we have rigorously introduced the concept of nearness in a topological space, we can define what it means to be ”inside” a figure as follows:

Definition 1.3. *Let S be a set in a topological space X. A point $p \in S$ is called an interior point of S if there exists a neighborhood of p, U, such that $U \subset S$. The set of all interior points of a set S is called the interior of S and is denoted by:*

$$\mathrm{int}(S) = \{p \in X \mid \exists U \in \mathscr{N}(p), \text{ such that } U \subset S\} \tag{1.1}$$

Definition 1.4 (Open Set)**.** *Let S be a subset of a topological space X. We will say that S is an open set if every point in S is an interior point, or in other words, if S is its own interior:*

$$S = \mathrm{int}(S)$$

Example 1.2. *Let $(a, b) \in \mathbb{R}$ be an open interval on the real line. Let's verify that such an open interval is indeed an open set under the definition just provided for openness.*
Let $x \in (a, b)$. We call $V_\delta(x) = (x - \delta, x + \delta)$ a δ-neighborhood centered at x, where $\delta > 0$. Visualizing it graphically:

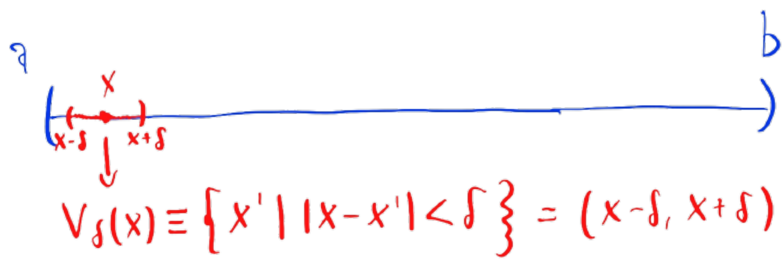


Figure 1.1: δ-neighborhood centered at x.

We need to prove that we can choose a $\delta > 0$ for which this neighborhood is contained within the set (a, b). Notice that, since $x \in (a, b)$, we have $a < x < b$. Let's find the distances from x to a and to b, and let Δ be the smaller of the two, $\Delta := \min\{|x - a|, |x - b|\}$. Graphically:

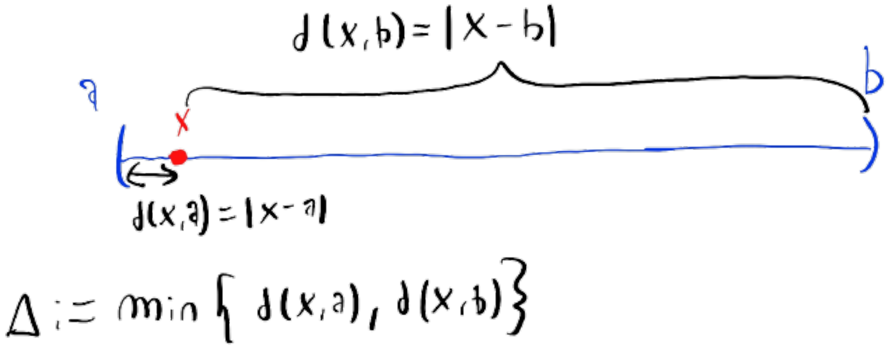


Figure 1.2: Distances of x from the endpoints a and b depicted graphically.

We will then choose δ as $\frac{\Delta}{2}$. Let's demonstrate that $V_\delta(x)$ is a subset of (a, b): if $y \in V_\delta(x)$, then $|x - y| < \delta = \frac{\Delta}{2}$. To belong to (a, b), y must be strictly less than b and strictly greater than a. Notice that, expanding $|x-y| < \delta = \frac{\Delta}{2}$, we obtain:

$$a < x - \frac{\Delta}{2} < y < \frac{\Delta}{2} + x < b$$

hence, $y \in (a, b)$, which means that (a, b) is an open set according to the topology defined earlier on $\mathbb{R}$.

Now, let's introduce an important result concerning open sets in a topological space.

Theorem 1.1. *Let X be a topological space. Then:*

1. *The union of an arbitrary collection of open sets is an open set in X, in symbols: if $\{U_\alpha\}_\alpha$ is a collection of open sets in X, then $\bigcup_\alpha U_\alpha$ is an open set.*
2. *The intersection of a finite number of open sets is an open set in X, in symbols: if $\{U_i\}_{i=1}^n$ is a finite collection of open sets in X, $\displaystyle\bigcap_{i=1}^n U_i$ is an open set.*
3. *The entire set X and the empty set $\emptyset$ are open.*

Proof. 1. Let $x \in \bigcup_\alpha U_\alpha$, then $x \in U_{\alpha^*}$ for at least one α^*. Therefore, x is an interior point of U_{α^*}, meaning there exists a neighborhood of x, V_x, contained in U_{α^*}, $V_x \subset U_{\alpha^*}$. But since it's contained in U_{α^*}, V is also contained in $\bigcup_\alpha U_\alpha$, and thus $\bigcup_\alpha U_\alpha$ is an open set.

2. Let $x \in \bigcap_{j=1}^{n} U_j$, then $x \in U_j \ \forall j \in [1, n] \cap \mathbb{N}$, $\Longrightarrow \forall j$, $\exists V_{xj} \subset U_j$, where V_{xj} is a neighborhood of x.
   Notice that, as an immediate consequence of the neighborhood axioms (just apply mathematical induction), we will have that finite intersections of neighborhoods of a point form a neighborhood, and thus $\bigcap_{j=1}^{n} V_{xj}$ is a neighborhood of x. But noticing that $\bigcap_{j=1}^{n} V_{xj} \subset \bigcap_{j=1}^{n} U_j$, we have that $\bigcap_{j=1}^{n} U_j$ is undoubtedly an open set.

3. The empty set, having no points, is trivially an open set. As for the entire set, just as trivially, we will have that every point has at least one neighborhood, and this neighborhood is contained in X.

□

Now, let's demonstrate some important facts about open sets and interior parts:

Proposition 1.2. *Let A be a set in a topological space X, then its interior part $\mathrm{int}(A)$ is open, and furthermore, every open set contained in A is contained in $\mathrm{int}(A)$.*

Proof. Let's first prove that $\mathrm{int}(A)$ is open: $p \in \mathrm{int}(A)$ means that p is an interior point of A, i.e., there exists a neighborhood U of p contained in A, $U \subset A$. Since U is a neighborhood of p, it satisfies the neighborhood axioms presented in definition 1.1, and then, by 1.1(3), we will have that there exists a neighborhood V of p such that U is a neighborhood of every point in V $\Longrightarrow V \subset \mathrm{int}(A)$. We have found a neighborhood contained in $\mathrm{int}(A)$ for every point in $\mathrm{int}(A)$. But this is precisely the definition of open.

Now, let's demonstrate the second assertion of the proposition: suppose we have an open set $\tilde{U} \subset A$. But by the definition of open set, we can find a neighborhood W of p contained in $\tilde{U}$ $\forall p \in \tilde{U}$. Consequently, $W \subset \tilde{U} \subset A$, from which it follows, by the definition of interior point, that p is an interior point of A, and thus $p \in \mathrm{int}(A)$. □

Proposition 1.3 (Alternative Characterization of a Neighborhood)**.** *U is a neighborhood of p $\Longleftrightarrow$ U contains an open set containing p.*

Proof. Let's start by proving the forward implication ($\Longrightarrow$). If U is a neighborhood of p, then for sure $p \in \mathrm{int}(U)$, as there exists a neighborhood of p contained in U, which is U itself. By the previous proposition, $\mathrm{int}(U)$ is open, and since $\mathrm{int}(U) \subset U$, we have demonstrated the forward implication.

The converse ($\Longleftarrow$) is equally straightforward: suppose that U contains an open set W containing p; then, by the definition of open set, there will be a neighborhood V of p with $V \subset W$. But then, $V \subset U$, and by the second axiom of definition 1.1, we will have that U is also a neighborhood. □

Proposition 1.4 (Alternative Characterization of an Open Set)**.** *A set U in a topological space is open if and only if it is a neighborhood of each of its points.*

Proof. Let's start with the forward implication ($\Longrightarrow$). If U is an open set, then $\forall p \in U$, there exists a neighborhood W of p such that $W \subset U$. But then, by the second axiom of definition 1.1, U is a neighborhood of each of its points.

For the reverse ($\Longleftarrow$), assuming that U is a neighborhood of each of its points, we will conclude that each of its points p is an interior point (as there exists a neighborhood $p \in U \subset U$). But then, $U = \mathrm{int}(U)$, and thus it is open. □

The results introduced above will be helpful for constructing a new definition of topology, equivalent to the one given earlier. Why would we want an alternative definition? Because the one just provided, while intuitive, is often difficult to work with. The convention is indeed to define a topology starting from the class of its open sets, and this is the convention we will use in this text from now on.

1.1.4 Alternative definition of topology

Theorem 1.2. *Let X be a set, and $\mathscr{T}$ be a class of subsets of X with the following properties:*

1. *For all $\{U_\alpha\}_\alpha \subset \mathscr{T}$, $\displaystyle\bigcup_\alpha U_\alpha \in \mathscr{T}$.*

2. *For all $\{U_j\}_{j=1}^n \subset \mathscr{T}$, $\displaystyle\bigcap_{j=1}^n U_j \in \mathscr{T}$.*

3. $X, \emptyset \in \mathscr{T}$.

Then there exists a unique way to equip X with a system of neighborhoods to make the sets in $\mathscr{T}$ the open sets of X, and it is as follows:
We will say that a set $U \subset X$ is a neighborhood of $p \in X$ if there exists a set $V \in \mathscr{T}$ such that $p \in V \subset U$.

Proof. To prove this theorem, let's first demonstrate that the neighborhoods defined in this way satisfy the axioms of definition 1.1, meaning they indeed form a system of neighborhoods for X. Then, we'll show that the open sets under the topology induced by this system of neighborhoods are precisely the elements of the set $\mathscr{T}$.

1. Let U be a neighborhood of p. Then, by the definition just given, there will exist a set $V \in \mathscr{T}$ such that $p \in V \subset U$. But then, $p \in U$.

2. Let A be a set in X, and let U be a neighborhood of p contained in A, $U \subset A$. According to the definition of neighborhood just provided, there will exist a set $V \in \mathscr{T}$ such that $p \in V \subset U \subset A$, so A is a neighborhood of p.

3. Let U be a neighborhood of p; then, there exists an element V in $\mathscr{T}$ contained in U. Notice that V is a neighborhood, as it is a set in $\mathscr{T}$ and is trivially contained in itself. Moreover, for any $q \in V$, $q \in V \subset U$, and thus U is a neighborhood of every point in V.

4. Finally, let U and $\tilde{U}$ be two neighborhoods of p. Then, there exist two elements of $\mathscr{T}$, $V, \tilde{V} \in \mathscr{T}$, such that $p \in V \subset U$ and $p \in \tilde{V} \subset \tilde{U}$. Notice that $p \in V \cap \tilde{V} \subset U \cap \tilde{U}$, and thus $U \cap \tilde{U}$ is a neighborhood of p since $V \cap \tilde{V}$ is open.

Now, let's prove that the open sets under this topology are the elements of $\mathscr{T}$.
Suppose U is an element of $\mathscr{T}$; then, for any $q \in U$, U will be a neighborhood of q since it contains a set from $\mathscr{T}$ (itself) containing q and contained in U. But then, U is open as it is a neighborhood of each of its points.
Conversely, suppose U is open; then, there exists a neighborhood Z of p for all $p \in U$. But by the definition of neighborhood given above, there will also be a set $V_p \in \mathscr{T}$ for each p, such that $p \in V_p \subset Z \subset U$. This implies that we can write our open set U as $U = \cup_{p \in U} V_p$. Since $V_p \in \mathscr{T}$ for all p, we have that $\cup_{p \in U} V_p$ is also an element of $\mathscr{T}$. Finally, Proposition 1.3 assures us that this is the unique topology where such sets are the class of open sets. ◻

Very often in general topology courses, the concept of a topological space is introduced directly with the definition provided above. Although this is the most practical set of axioms to work with, defining this concept directly in this way might cause some of the intuition behind the idea of a topological space to be lost on a newcomer, making it appear as an abstract and senseless definition. The previous discussion was mostly aimed at helping the reader understand what was hidden behind these abstract definitions in intuitive terms.

Remark 1.1. *The class of open sets $\mathscr{T}$ is often called a topology, and we will also use this convention. The open sets in the topology induced by $\mathscr{T}$ are called $\mathscr{T}$-open sets, and a set X equipped with a topology, $(X, \mathscr{T})$, is called a topological space.*

Now that we have demonstrated that these two definitions are equivalent, we will proceed to work with the latter. Before continuing with the theory, however, let's provide some illustrative examples:

Example 1.3. *Let $\mathbb{N}$ be the set of natural numbers, and $\mathscr{T}$ be the class of subsets of $\mathbb{N}$ composed of $\emptyset$ and all subsets of the form $S_n = \{n, n+1, n+2, n+3, \dots\}$:*

$$\mathscr{T} = \{\emptyset\} \cup \{S_n\}_{n\in\mathbb{N}}$$

This class of subsets forms a topology on $\mathbb{N}$; let's verify it:

1. *Trivially, as we defined $\mathscr{T}$, we have $\mathbb{N}, \emptyset \in \mathscr{T}$.*
2. *Let $\{S_n\}_{n\in I}$, $I \subset \mathbb{N}$, be a collection of subsets of the specified type above. Since $\mathbb{N}$ is a totally ordered set with a minimum element, I, being a subset, will have a minimum element that we call $m \in I$. Observing that:*
$$\bigcup_{n\in I} S_n = S_m \in \mathscr{T} \tag{1.2}$$
we conclude that the union of open sets is open.
3. *Let $S_m, S_l \in \mathscr{T}$. Then, $S_m \cap S_l = S_{\max\{m,l\}} \in \mathscr{T}$.*

Thus, the natural numbers equipped with this class of subsets form a topology, $(\mathbb{N}, \mathscr{T})$.

Example 1.4 (The Discrete Topology on X)**.** *Let X be any set, then we can define the following topology on X, called the discrete topology, by selecting the set of open sets $\mathscr{T}$ as the set of all subsets of X (called the power set and denoted by $\mathscr{P}(X)$):*

$$\mathscr{T} = \mathscr{P}(X)$$

Example 1.5 (The Indiscrete Topology on X)**.** *Let X be any set, then we can define a topology on X by selecting the set of open sets $\mathscr{T}$ as:*

$$\mathscr{T} = \{X, \emptyset\}$$

*We call this the **indiscrete topology on X**.*

Proposition 1.5. *Let $\{\mathscr{T}_\alpha\}_\alpha$ be a collection of topologies on a set X, then the intersection $\bigcap\limits_{\alpha} \mathscr{T}_\alpha$ is a topology on X.*

Proof.

1. Since $\mathscr{T}_\alpha$ are topologies, we have $X, \emptyset \in \mathscr{T}_\alpha$ for all α, and therefore $X, \emptyset \in \bigcap\limits_{\alpha} \mathscr{T}_\alpha$.

2. Let $\{U_\beta\} \subset \bigcap\limits_{\alpha} \mathscr{T}_\alpha$. Then, $\{U_\beta\}_\beta \subset \mathscr{T}_\alpha$ for all α, and thus $\bigcup_\beta U_\beta \in \mathscr{T}_\alpha$ for all α, implying that $\bigcup_\beta U_\beta \in \bigcap\limits_{\alpha} \mathscr{T}_\alpha$.

3. Let $\{U_j\}_{j=1}^n \subset \bigcap\limits_{\alpha} \mathscr{T}_\alpha$. Then, similarly to the previous case, $\{U_j\}_{j=1}^n \subset \mathscr{T}_\alpha$ for all α, and therefore $\bigcap\limits_{j=1}^{n} U_j \in \mathscr{T}_\alpha$ for all α, implying that $\bigcap\limits_{j=1}^{n} U_j \in \bigcap\limits_{\alpha} \mathscr{T}_\alpha$.

□

Exercise 1.4. *Convince yourself, by providing a counterexample, that generally, a union of topologies on a set X does **not** form a topology on X.*

Exercise 1.5. *Find all the topologies for the set $\{\alpha, \beta\}$.*

1.1.5 Accumulation points, closed sets and the closure of a set

Once the concept of nearness is introduced in our set, a similarly rigorous concept of "boundary" and closure is also convenient. We still use intuition: think of an open figure V (without a boundary) that we can think of as a subset of a larger set that we will call $\mathscr{U}$. With the concept of interior and "openness," the question arises of how to define in

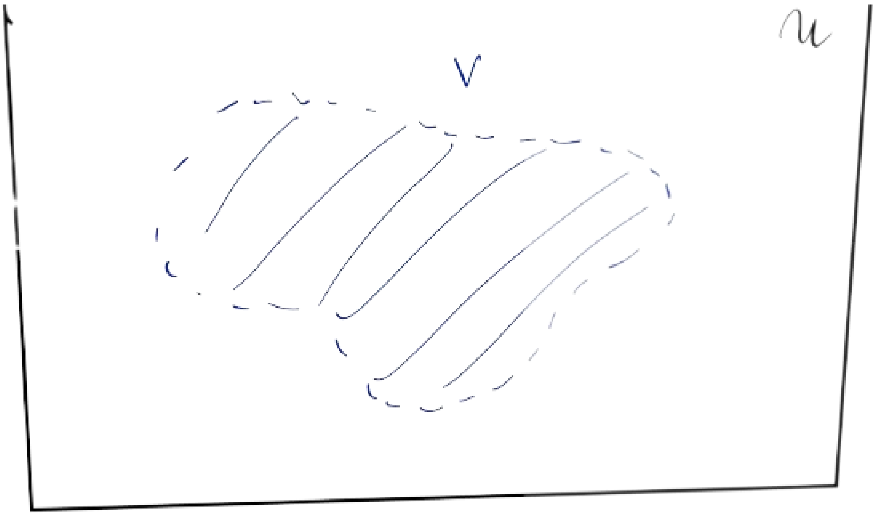
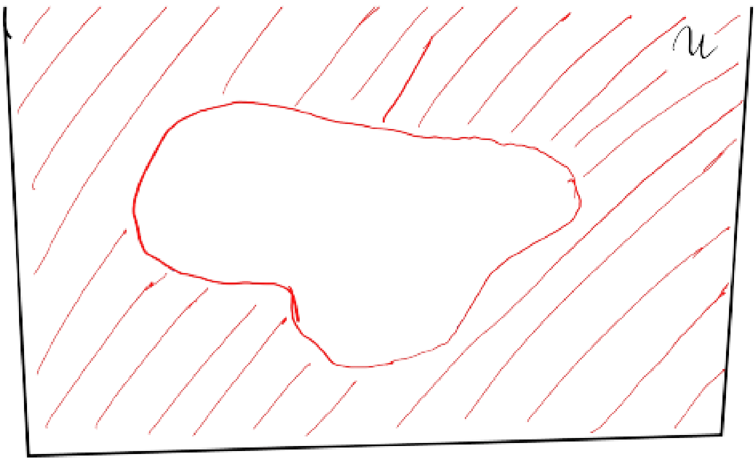


Figure 1.3: Our figure V.

We have said that an interior point is a point "surrounded" by other points of a set, and an open set is a set in which all points are interior. We can see that the complement of V in $\mathscr{U}$ will be a figure with a "boundary," and therefore, in this geometric/intuitive sense, a "closed" figure.



This preliminary reasoning motivates us to define closed sets rigorously in the following way:

Figure 1.4: Complement of V in $\mathscr{U}$.

Definition 1.5 (Closed Set). *A set $A \subset X$ is said to be closed if its complement in X, $X \setminus A = A^c$, is an open set.*

Proposition 1.6. *A set U in a topological space X is an open set if and only if its complement in X, U^c, is closed.*

Proof. This immediately follows from the definition of closed. □

Remark 1.2. *It is important not to be deceived by the fact that in common parlance the words "closed" and "open" are opposites, i.e., a door is closed $\Longleftrightarrow$ a door is not open, and vice versa. In topology, a set can be either closed or open, as well as both closed and open at the same time (in this case, it is called a "clopen" set), or neither.*

Exercise 1.6. *Convince yourself of this fact by demonstrating that, in any set X, under any topology $\mathscr{T}$, the sets $X, \emptyset \in \mathscr{T}$ are both closed and open.*

Exercise 1.7. *Let $(\mathbb{N}, \mathscr{T})$ be the topology on the natural numbers introduced earlier; determine the set of closed sets on this topology.*

Exercise 1.8. *Let $A = \{\alpha, \beta\}$. For each of the topologies on this set found earlier, find the set of closed sets.*

Now let's give another important topological definition, that of a "limit point" or "accumulation point."

Definition 1.6 (Limit Point). *Let $S \subset X$ be a set in a topological space, and $p \in X$ be a point. We will say that p is a limit point of S if every open set G_p containing p contains at least one point of S different from p. In symbols:*

$$(G_p \setminus \{p\}) \cap S \neq \emptyset$$

The set of all limit points of a set S is called the derived set of S, often denoted as $\mathrm{der}(S)$ or $\mathrm{acc}(S)$.

Given this new definition, we now prove an important characterization of closed sets:

Proposition 1.7. *A set $S \subset X$ in a topological space is closed if and only if it contains all of its limit points.*

Proof. Let's start with the direct implication ($\Longrightarrow$). Since S is closed, we know that S^c is open. Note that these two sets are disjoint, meaning $S \cap S^c = \emptyset$. Suppose, for the sake of contradiction, that a limit point $p \in \mathrm{der}(S)$ of S is not in S, $p \notin S \Longleftrightarrow p \in S^c$. But S^c is open, and $(S^c \setminus \{p\}) \cap S = \emptyset$, meaning $p \notin \mathrm{der}(S)$, a contradiction. Therefore, $\mathrm{der}(S) \subset S$.

For the converse ($\Longleftarrow$), it is enough to note that if $\mathrm{der}(S) \subset S$, then for every point p in S^c, as it is not a limit point of S, there will exist an open set G_p containing p such that $(G_p \setminus \{p\}) \cap S = \emptyset$. Since p does not belong to S, we can simply write $G_p \cap S = \emptyset$. We have found an open set disjoint from S for every point in S^c. Therefore, we can write S^c as:

$$S^c = \bigcup_{p \in S^c} G_p \tag{1.3}$$

which, by the axioms of open sets in a topology, is open. Hence, S is closed. □

Definition 1.7 (Closure of a Set)**.** *Let S be a set in a topological space X. Then the closure of S is the intersection of all closed sets that contain S. In symbols:*

$$\overline{S} = \bigcap_{C \text{ closed }, S \subset C} C$$

Proposition 1.8. *The closure of a set S, $\overline{S}$, is closed. Furthermore, if C is a closed set such that $S \subset C$, then $S \subset \overline{S} \subset C$.*

Proof. Both of these statements follow trivially from the definition of closure provided above: we can see that the closure of a set is closed because it is an arbitrary intersection of closed sets, and furthermore, being the intersection of all closed sets containing S, we have that if a closed set contains S, the intersection of the latter with potentially other sets is contained in it. □

Proposition 1.9. *Let S be a set in a topological space X; then S is closed if and only if it is equal to its closure, i.e., $S = \bar{S}$.*

Proof. The inverse implication is trivially true, as $\bar{S}$ is closed, as shown in the previous proposition. For the direct implication, it is enough to note that if S is closed, it will trivially be the "minimal" closed superset of S, and therefore $\bar{S} = \bigcap_{C \text{ closed }, S \subset C} C = S$. □

Exercise 1.9. *Prove that if A is a subset of B, then every point of accumulation of A is a point of accumulation of B, i.e.,*

$$A \subset B \Longrightarrow \mathrm{der}(A) \subset \mathrm{der}(B)$$

Theorem 1.3. *Let S be a set in a topological space, then:*

$$\overline{S} = S \cup \mathrm{der}(S) \tag{1.4}$$

Proof. To prove this result, we first prove that $S \cup \mathrm{der}(S)$ is a closed set, i.e., $(S \cup \mathrm{der}(S))^c$ is an open set. It suffices to observe that if $s \in (S \cup \mathrm{der}(S))^c$, then there exists an open set G_s containing s such that

$$G_s \cap S = \emptyset$$

since $s \notin \mathrm{der}(S)$ and $s \notin S$. But since $G_s \cap S = \emptyset$, every element g of G_s will not be a limit point of S, $g \notin \mathrm{der}(S)$, and thus $G_s \cap \mathrm{der}(S) = \emptyset$.
From this, it follows that we can express the set $(S \cup \mathrm{der}(S))^c$ as a union of open sets in the following way:

$$(S \cup \mathrm{der}(S))^c = \bigcup_{s \in (S \cup \mathrm{der}(S))^c} G_s$$

and, according to the axioms of open sets, it must be open, so $S \cup \mathrm{der}(S)$ is closed.
Now, we prove that $\overline{S} = S \cup \mathrm{der}(S)$. Since $S \cup \mathrm{der}(S)$ is a closed superset of S, we have $\overline{S} \subset S \cup \mathrm{der}(S)$.
For the reverse inclusion, it is trivial that $S \subset \overline{S}$ for S. Regarding the derived set, it suffices to observe that:

$$S \subset \overline{S} \Longrightarrow \mathrm{der}(S) \subset \mathrm{der}(\overline{S})$$

due to the result of Exercise 1.9. But $\overline{S}$ is closed and therefore contains all of its limit points; in other words:

$$\mathrm{der}(\overline{S}) \subset \overline{S}$$

and thus $\mathrm{der}(S) \subset \overline{S}$, which implies that $S \cup \mathrm{der}(S) \subset \overline{S}$, and therefore, the assertion $\overline{S} = S \cup \mathrm{der}(S)$ holds. □

Exercise 1.10. *Let R and S be two sets in a topological space. Prove that:*

$$\mathrm{der}(R \cup S) = \mathrm{der}(R) \cup \mathrm{der}(S)$$

Exercise 1.11. *Prove that, for any two sets R and S in a topological space X:*

$$\overline{R \cup S} = \overline{R} \cup \overline{S}$$

1.1.6 Internal points, boundary and the exterior

Let's redefine the concept of an interior point introduced earlier in terms of open sets:

Definition 1.8. *Let S be a set in a topological space X. A point $p \in S$ is called an interior point of S if there exists an open set U containing p that is entirely contained in S. In symbols:*

$$\exists U \in \mathcal{T} \quad p \in U \subset S$$

As before, we denote the set of interior points of S as $\mathrm{int}(S)$.

Proposition 1.10. *The interior of a set S, denoted as $\mathrm{int}(S)$, is the union of all open subsets of S; in symbols:*

$$\mathrm{int}(S) = \bigcup_{U \text{ open, } U \subset S} U$$

Proof. Let's start with the "direct" inclusion ($\subset$). Suppose that $s \in \mathrm{int}(S)$. Then, by the definition given above, there exists an open set U such that $s \in U \subset S$. Let $\Sigma = \bigcup_{U \text{ open, } U \subset S} U$ be the union of all open subsets of S. Since U is an open subset of S, we have $U \subset \Sigma$, and therefore, $s \in \Sigma$.
For the reverse inclusion, assume that $s \in \Sigma$, where Σ is the same union defined earlier. However, note that Σ is itself an open subset of S, so s is an interior point, i.e., $s \in \mathrm{int}(S)$. □

Corollary 1.1. *Let S be a set in a topological space X. Then:*

1. $\mathrm{int}(S)$ *is open.*
2. *If U is an open subset of S, then $U \subset \mathrm{int}(S)$.*
3. $\mathrm{int}(S) \subset S$.
4. *S is open if and only if $S = \mathrm{int}(S)$.*

Definition 1.9 (Exterior part)**.** *Let S be a set in a topological space X. We define the exterior part of S as the interior part of its complement:*

$$\mathrm{ext}(S) = \mathrm{int}(X \setminus S) = \mathrm{int}(S^c)$$

Definition 1.10 (Boundary of a set)**.** *Let S be a set in a topological space X. We call the boundary of S, denoted as ∂S, the set of points that belong to neither the interior nor the exterior of S:*

$$\partial S = X \setminus (\mathrm{int}(S) \cup \mathrm{ext}(S)) = (\mathrm{int}(S) \cup \mathrm{ext}(S))^c$$

We conclude this subsection with one last important result:

Theorem 1.4. *Let S be a set in a topological space X. Then:*

$$\overline{S} = \mathrm{int}(S) \cup \partial S \tag{1.5}$$

Proof. Observing that $X = \mathrm{int}(S) \cup \partial S \cup \mathrm{ext}(S)$, we have $(\mathrm{int}(S) \cup \partial S)^C = \mathrm{ext}(S)$. To prove the theorem, we only need to show that $(\overline{S})^c = \mathrm{ext}(S)$. For the direct inclusion, observe that if $p \in (\overline{S})^c$, then $p \notin \mathrm{der}(S)$, so there exists an open set U containing p such that:

$$(U \setminus \{p\}) \cap S = \emptyset$$

Since $p \notin S$, we also have that U and S are disjoint. Therefore, by the definition of the exterior of a set, $p \in \mathrm{ext}(S)$, i.e., $(\overline{S})^c \subset \mathrm{ext}(S)$.
For the reverse inclusion, suppose that p belongs to $\mathrm{ext}(S)$ and, using the definition of the exterior as given earlier, deduce that there exists an open set U such that:

$$p \in U \subset S^c$$

It immediately follows that $U \cap S \neq \emptyset$, so p is not a limit point of S, i.e., $p \notin \mathrm{der}(S)$. However, since U is contained in S^c, we have $p \notin S$, and therefore, $p \notin \overline{S}$, i.e., $\mathrm{ext}(S) \subset (\overline{S})^c$.
Hence, $(\overline{S})^c = \mathrm{ext}(S) = (\mathrm{int}(S) \cup \partial S)^c$. □

1.1.7 Bases

In this subsection, we introduce the concept of a **base for a topological space**. By base, we mean a collection of open sets that can be used to generate our space.

Definition 1.11 (Base for a topological space)**.** *Let $(X, \mathscr{T})$ be a topological space. A class $\mathscr{B}$ of open sets of X is called a* ***base*** *for the topology $\mathscr{T}$ if every open set $U \in \mathscr{T}$ can be expressed as the union of members of $\mathscr{B}$.*

Exercise 1.12. *Prove that, for a topological space $(X, \mathscr{T})$, $\mathscr{B}$ is a base for $\mathscr{T}$ if and only if for every point p in an open set U, $p \in U$, there exists an element of the base $B \in \mathscr{B}$ such that:*

$$p \in B \subset U$$

Now, given a set X (without a topological structure) and a class of subsets of it, we ask under what conditions this class forms a base for a topology on X. The conditions are provided by the following theorem:

Theorem 1.5. *Let $\mathscr{B}$ be a class of subsets of a non-empty set X. Then, $\mathscr{B}$ will be a base for a topology on X if and only if it has the following properties:*

1. *X is given as the union of members of $\mathscr{B}$, i.e.:*

$$X = \bigcup_{B \in \mathscr{B}} B$$

2. *For every $B, \overline{B} \in \mathscr{B}$, $B \cap \overline{B}$ is the union of members of $\mathscr{B}$, or alternatively, if $p \in B \cap \overline{B}$, then there exists a set $B_p \in \mathscr{B}$ such that $p \in B_p \subset B \cap \overline{B}$.*

Proof. Let's start by proving the direction ($\Longrightarrow$). Suppose that $\mathscr{B}$ is a base for a topology $\mathscr{T}$ on X; by the definition of a base, every $\mathscr{T}$-open set $U \subset X$ can be expressed as a union of elements of $\mathscr{B}$. Therefore, since X is an open set, we have $X = \bigcup_{B \in \mathscr{B}} B$. For the second point, observe that the elements of the base $\mathscr{B} \subset \mathscr{T}$ are open, so given two elements of the base $B, \overline{B} \in \mathscr{B}$, $B \cap \overline{B}$ will be open by the properties of open sets. Being open, it can be expressed as a union of elements of $\mathscr{B}$. This completes the proof of the forward direction.

Now, suppose that a class of subsets $\mathscr{B}$ of X satisfies properties 1 and 2. We want to prove that the class of sets $\mathscr{T}$ formed by all subsets of X that can be expressed as unions of elements of $\mathscr{B}$ is indeed a topology on X. If this is the case, then we have obtained a topology of which $\mathscr{B}$ is trivially a base.

Let's start by proving property 1.2-1 (i.e., property no. 1 of theorem 1.2). We can see that this property is true for our class $\mathscr{T}$: $\emptyset$, being the empty union, can be expressed as the union of elements of $\mathscr{B}$, and therefore, $\emptyset \in \mathscr{T}$. As for X, it is expressible as a union of elements of $\mathscr{B}$ by assumption (property 1), so $X \in \mathscr{T}$.

Let $\{U_\alpha\}_\alpha$ be a collection of elements of $\mathscr{T}$. By how we defined $\mathscr{T}$, for every α, the set U_α can be expressed as a union of elements of $\mathscr{B}$, which means that the union $\bigcup\limits_\alpha U_\alpha$ can also be represented as a union of elements of $\mathscr{B}$, i.e., $\bigcup\limits_\alpha U_\alpha \in \mathscr{T}$. This demonstrates property 1.2-2.

Finally, consider two sets U and V in $\mathscr{T}$. According to the definition we provided for $\mathscr{T}$, both U and V can be expressed as unions of elements of $\mathscr{B}$:

$$U = \bigcup_\alpha B_\alpha \quad V = \bigcup_\beta B_\beta \;\; ; \;\; B_\alpha, B_\beta \in \mathscr{B} \quad \forall \alpha, \beta$$

Using the properties of unions and intersections, we have:

$$U \cap V = \bigcup_{\alpha,\beta} B_\alpha \cap B_\beta$$

But according to property 2 of our class $\mathscr{B}$, we know that $B_\alpha \cap B_\beta$ is a union of elements of $\mathscr{B}$, which implies that $U \cap V$ is also a union of elements of $\mathscr{B}$, thus demonstrating property 1.2-3. □

1.1.8 Continuous functions and homeomorphisms

We conclude this section on general topology by introducing the concept of continuity. This concept, as the reader may have deduced from the name, is nothing more than a generalization to more general spaces of the concept of continuity of real functions. The underlying idea is the same: ideally, we would like a function $f : X \to Y$ between topological spaces that maps points in the domain "close" to a point $x \in X$ to points in the codomain "close" to $f(x) \in Y$." Now that we have a rigorous definition of "closeness" in our spaces, we can introduce this notion rigorously as follows:

Definition 1.12 (Continuous function at x)**.** *Let $f : X \to Y$ be a function between topological spaces $(X, \mathscr{T})$ and $(Y, \mathscr{T}')$. We will say that f is* ***continuous at*** *x, where here x is a point of X, if the inverse image (also called pre-image) of a neighborhood of $f(x)$ is a neighborhood of x, i.e.:*

$$V \in \mathscr{N}_{f(x)} \Longrightarrow f^{-1}(V) \in \mathscr{N}_x$$

The definition given in terms of neighborhoods is local. It is possible to provide a logically equivalent global definition of continuity, this time in terms of open sets, as follows:

Definition 1.13 (Continuous function)**.** *Let $f : X \to Y$ be a function between topological spaces $(X, \mathscr{T})$ and $(Y, \mathscr{T}')$. We will say that f is* ***continuous*** *(with respect to the topologies $\mathscr{T}$ and $\mathscr{T}'$) if the pre-image $f^{-1}(V)$ of every $\mathscr{T}'$-open set V of Y is an $\mathscr{T}$-open set of X, i.e.:*

$$V \in \mathscr{T}' \Longrightarrow f^{-1}(V) \in \mathscr{T}$$

Exercise 1.13. *Prove that a function $f : X \to Y$ is continuous if and only if the preimage of every element of a base $\mathscr{B}$ of the topology on Y is an open set in X.*

Of great importance is the notion of homeomorphism: the reader may have encountered concepts such as isomorphisms of vector spaces, groups, or rings along the way. In the same way that the aforementioned concepts preserved the algebraic structure of these structures, thus representing an algebraic equivalence, homeomorphisms of topological spaces preserve the topological properties (we will define some shortly) of a space and represent a topological equivalence. Therefore, let us give the definition of homeomorphism:

Definition 1.14 (Homeomorphism of topological spaces)**.** *Let $(X, \mathscr{T})$ and $(Y, \mathscr{T}')$ be two topological spaces. We will say that these two spaces are **homeomorphic**, and write $X \cong Y$, if there exists a bijective map $\varphi : X \to Y$ such that both φ and φ^{-1} are continuous functions. We call the map φ a **homeomorphism**.*

Exercise 1.14. *Prove that the relation "being homeomorphic" forms an equivalence relation among topological spaces, i.e., prove that this relation satisfies the following properties:*

- *Reflexive: $X \cong X$.*
- *Symmetric: $X \cong Y \implies Y \cong X$.*
- *Transitive: $X \cong Y$ and $Y \cong Z \implies X \cong Z$.*

Homeomorphisms of topological spaces preserve some properties of the spaces; these properties are called topological properties for this reason. Let us list some of them (we will not prove that they are topological properties, as that would take much time, and this introduction aims to provide only the essentials).

Definition 1.15 (Connectedness)**.** *We will say that a space X is **disconnected** if it is the union of two disjoint non-empty open sets, in symbols:*

$$X = U \sqcup V \quad U, V \text{ open} \quad U, V \neq \emptyset$$

If a space is not disconnected, then we will say that it is connected.

Definition 1.16 (Hausdorff Spaces)**.** *We will say that a topological space $(X, \mathscr{T})$ is a Hausdorff space (or T-2 space) if for every two distinct points $p, q \in X$, there exist disjoint open sets containing them. In symbols:*

$$\forall p, q \in X \quad \exists U_p, U_q \in \mathscr{T} \text{ such that } \quad U_p \cap U_q = \emptyset \quad \text{with } p \in U_p \text{ , } q \in U_q$$

1.2 Elements of algebraic topology

Now, in this section, we introduce the necessary definitions and concepts of algebraic topology to address this text, making it accessible even to readers who have not previously studied this subject. Algebraic topology, as the name suggests, uses algebraic tools (groups, rings, etc.) to study topological spaces and their properties. In particular, the goal is to associate an algebraic structure to a space that is topologically invariant. This means that for two topological spaces that are homeomorphic, denoted as X and Y, with associated algebraic structures A_X and A_Y, respectively, we want these two structures A_X and A_Y to be isomorphic. We will often see that such structures are not only invariant with respect to homeomorphisms but also with respect to other, more ”broad” equivalence relations on topological spaces, such as homotopy equivalence.

We begin this section by introducing the concept of homotopy and the fundamental group.

1.2.1 Homotopy

Let's imagine we have a disk and an annular region (see the figure below).

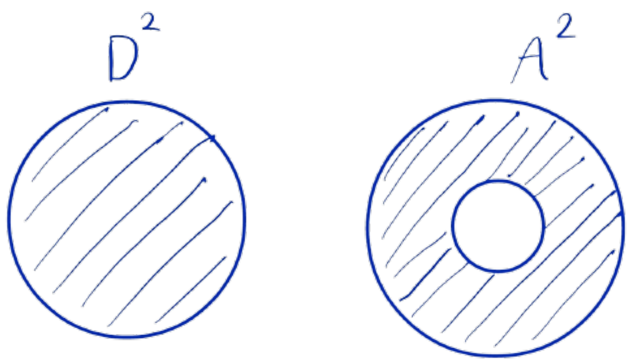


Figure 1.5: A disk D^2 and an annular region A^2.

We can explicitly verify that it would be very difficult to prove whether or not these two spaces are homeomorphic by examining their known topological properties (for example, both spaces are Hausdorff and connected).

Therefore, we need to find a distinction of a different nature. The immediate intuitive response would be to say that one space has a hole while the other does not. However, how can we mathematically formalize this statement? It is for this need that we introduce the concept of "homotopy" and "homotopy paths."

First, let's define the concept of a "path" in a general topological space X:

Definition 1.17 (Path in a topological space). *We will call a continuous function $\gamma : [0,1] \to X$, meaning a continuous function from the unit interval $I := [0,1] \subset \mathbb{R}$ to X, a path (or curve) in a topological space X. We will denote $\gamma(0) = x \in X$ and $\gamma(1) = y \in X$ as the endpoints of γ. We will say that the path is **closed** if the two endpoints coincide, i.e., if $\gamma(0) = \gamma(1)$.*

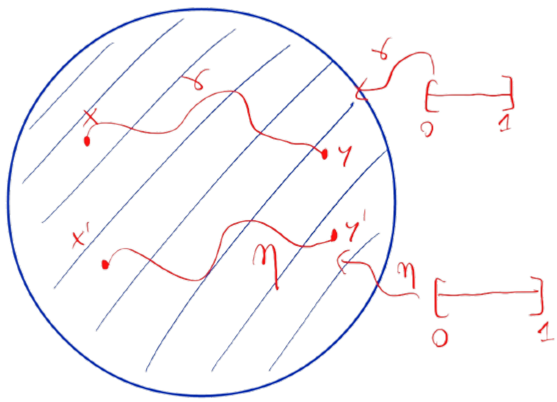


Figure 1.6: Visualization of paths in a space.

We are now ready to develop a strategy to "identify holes in a space" using the concept of "deformability" of one curve into another, which we will introduce below.

First, let's define the concept of "homotopy" for paths in a space:

Definition 1.18 (Homotopy of paths). *Let γ and η be two paths in X, meaning two continuous functions from the interval $[0,1]$ to X. We will say that these two paths are homotopic in X if there exists a continuous function $F : I^2 \to X$, where $I^2 := [0,1] \times [0,1] = \{(s,t) \; ; \; 0 \leq s \leq 1 \; , \; 0 \leq t \leq 1\}$, such that $F|_{s\times\{0\}} = F(s,0) = \gamma(s)$ and $F|_{s\times\{1\}} = F(s,1) = \eta(s)$. In this case, we write $\gamma \simeq \eta$.*

The geometric intuition behind the idea of homotopy is to continuously deform one curve into another using a continuous function $F(s,t)$, where we can intuitively view the second parameter t as a time parameter. This function, at time $t = 0$, equals the first curve $\gamma(s)$, and at time $t = 1$, it equals the second curve $\eta(s)$. When it takes values strictly between 0 and 1, $0 < t' < 1$, we can visualize $F(s,t')$ as an "intermediate" curve between γ and η, as illustrated in the following image:

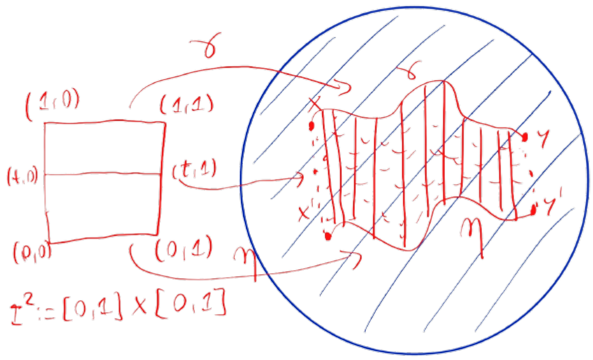


Figure 1.7: Homotopy of paths.

The idea of homotopy can also be extended to two general continuous functions as follows:

Definition 1.19 (Homotopy of two continuous functions)**.** *Let X and Y be two topological spaces, and let f and $g : X \to Y$ be two continuous functions from X to Y. We will call f and g homotopic functions on X (and write $f \simeq g$) if there exists a continuous function $F : X \times [0,1] \to Y$, $(x,t) \to F(x,t)$, such that $F(x,0) = f(x)$ and $F(x,1) = g(x)$.*

The idea is the same as before (continuously deform one function into another using a time parameter t), and we can see that homotopy of paths is "contained" in this definition. Let's now work on a practical example to familiarize the reader with this newly introduced concept.

Example 1.6. *Prove that for all continuous functions $f, \phi : [0,1] \to \mathbb{R}^n$, $f \simeq \phi$, meaning that for any two paths f and ϕ in Euclidean n-dimensional space, they are homotopic.*
Proof*: Let's explicitly express the coordinates of the images of these functions, i.e.,*

$$f(s) = \langle f_1(s), f_2(s), \ldots, f_n(s) \rangle$$

$$\phi(s) = \langle \phi_1(s), \phi_2(s), \dots, \phi_n(s) \rangle$$

Now, let's construct the following function F, which we define as:

$$F(s,t) := \langle (1-t)f_1(s) + t\phi_1(s), (1-t)f_2(s) + t\phi_2(s), \dots, (1-t)f_n(s) + t\phi_n(s) \rangle$$

Here, s and t are in the interval $[0,1]$. It is straightforward to verify that $F(s,0) = f(s)$ and $F(s,1) = \phi(s)$. Furthermore, it is equally straightforward to observe that this function is continuous (taking into account the notion of continuity for functions from $\mathbb{R}^2$ to $\mathbb{R}^n$ and utilizing the continuity assumed for f and ϕ).

Proposition 1.11. *Homotopy of paths is an equivalence relation between paths on a topological space X.*

Proof. Let's enumerate the properties to prove, starting with reflexivity:

- For reflexivity, given a $\gamma : [0,1] \to X$, we can simply set $F(s,t) = \gamma(s)$, where here $0 \leq t \leq 1$. We see that the defined F will be continuous because γ is continuous by assumption, and we trivially have $F(s,0) = \gamma(s) = F(s,1)$, from which it follows that $\gamma \simeq \gamma$.

- For symmetry, given two paths γ and η such that $\gamma \simeq \eta$, we want to prove that $\eta \simeq \gamma$:
By hypothesis, we know the existence of a continuous function $F : I^2 \to X$, such that $F(s,0) = \gamma(s)$ and $F(s,1) = \eta(s)$. We obtain the desired homotopy $\eta \simeq \gamma$ with the continuous function $\tilde{F} : I^2 \to X$ defined as $\tilde{F}(s,t) := F(s,1-t)$.

- Finally, for transitivity, let's assume we have three paths γ, η, and $\nu : [0,1] \to X$, homotopic in pairs, i.e., $\gamma \simeq \eta$ and $\eta \simeq \nu$. By assumption, we have a continuous function F providing the homotopy between γ and η, and another continuous function G providing the homotopy between η and ν. To construct a homotopy between γ and ν, we'll use the following function:

$$H(s,t) := \begin{cases} F(s,2t) & 0 \leq t \leq \frac{1}{2} \\ G(s,2t-1) & \frac{1}{2} \leq t \leq 1 \end{cases}$$

We observe that $H(s,0) = F(s,0) = \gamma(s)$ and $H(s,1) = G(s,1) = \nu(s)$. The only thing left to verify is the continuity of H, which we can do at $t = \frac{1}{2}$.

Let U be a neighborhood of $H(s, \frac{1}{2})$, where here s is a generic point in $[0, 1]$. By the continuity of F, there exists a $\delta_1 > 0$ such that:

$$|s' - s| < \delta_1 \quad \text{and} \quad |2t' - 1| < 2\delta_1 \Longrightarrow F(s', 2t') \in V_1$$

where here V_1 is a neighborhood of $F(s, 1)$. But $F(s, 1) = H(s, \frac{1}{2})$, so $F(s', 2t') \in U$. Similarly, for the continuity of G, there exists a $\delta_2 > 0$ such that:

$$|s' - s| < \delta_2 \quad \text{and} \quad |(2t' - 1) - 0| < 2\delta_2 \Longrightarrow G(s', 2t' - 1) \in V_2$$

where here V_2 is a neighborhood of $G(s, 0)$. But $G(s, 0) = H(s, \frac{1}{2})$, so $G(s', 2t' - 1) \in U$.
Now, by setting $\delta := \min\{\delta_1, \delta_2\}$, we have:

$$|s' - s| < \delta \quad \text{and} \quad |t' - \frac{1}{2}| < \delta \Longrightarrow H(s', t') \in U$$

Thus, the function H is continuous.

□

Returning to the previous example, however, we realize that in order to use this notion for our geometric purposes, we need to somehow "fix" the starting and ending points of the two curves and make them coincide. This motivates us to introduce the following concept:

Definition 1.20 (Homotopy with fixed endpoints)**.** *Let γ and $\eta : [0, 1] \to X$ be two paths such that $\gamma(0) = \eta(0) = x$ and $\gamma(1) = \eta(1) = y$. We say that γ and η are homotopic paths with fixed endpoints x and y if there exists a continuous function $F : I^2 \to X$ such that $F(s, 0) = \gamma(s)$, $F(s, 1) = \eta(s)$ for all $s \in [0, 1]$, and $F(0, t) = x$, $F(1, t) = y$ for all $t \in [0, 1]$.*

Geometrically, our function F will map the horizontal line $s = 0$ to γ, the horizontal line $s = 1$ to η, and the two vertical lines $t = 0$ and $t = 1$ to x and y, respectively, as illustrated in the figure below.

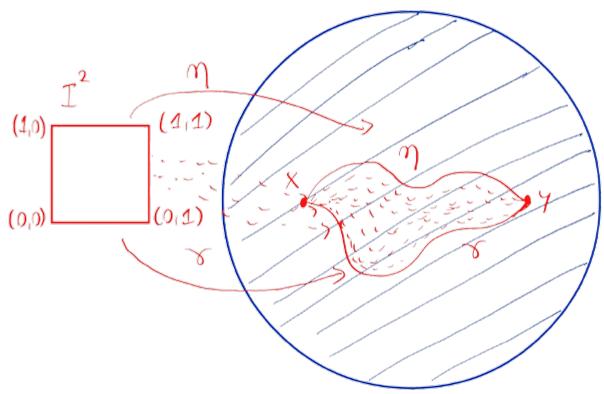


Figure 1.8: Homotopy of paths with fixed endpoints x and y.

Similarly, we can define homotopies of closed curves with respect to a single base point, which we will call a "base point":

Definition 1.21 (Homotopy of two closed curves with respect to the base point x_0)**.** *Let γ and η be two closed curves on a space X both starting from a point $x_0 \in X$, i.e., two curves such that $\gamma(0) = \gamma(1) = \eta(0) = \eta(1) = x_0$. We call these closed curves (or closed paths) based on the point x_0. We also call γ and η curves homotopic with respect to the base point x_0 if there exists a continuous function $F : I^2 \to X$ such that $F(s, 0) = \gamma(s)$ and $F(s, 1) = \eta(s)$ for all $s \in [0, 1]$, and $F(0, t) = F(1, t) = x_0$ for all $t \in [0, 1]$.*

Using a procedure similar to that of Proposition 1.11, it can be proven that both homotopy with fixed endpoints and homotopy of closed paths with respect to a single base point are equivalence relations. Therefore, we can consider equivalence classes of these relations. In particular, we are interested in equivalence classes of closed paths based on a point x_0 with respect to homotopy with the base point x_0.

The reason why this set of equivalence classes (which we will call **homotopy classes**) is so interesting is that it allows us to construct (using an operation we will define) a group associated with a point in a topological space, which we will call the **fundamental group**.

Definition 1.22 (Product of closed paths based on a point x_0)**.** *Given two closed paths based on a point x_0, which we will call γ and η, we can define a notion of product as follows:*

$$\gamma\eta(s) := \begin{cases} \gamma(2s) & 0 \leq s \leq \frac{1}{2} \\ \eta(2s-1) & \frac{1}{2} \leq s \leq 1 \end{cases} \tag{1.6}$$

It is straightforward to verify that this product is also a closed path based on x_0 (i.e., this binary operation is closed). The identity with respect to this product is given by the constant function $i(s) = x_0$ for all s. For every closed path γ, there exists its inverse $\overline{\gamma}(s) := \gamma(1-s)$, a path such that $\gamma\overline{\gamma} = \overline{\gamma}\gamma = i(s) = x_0$. Furthermore, this product is associative. The proofs of these facts are not particularly enlightening and will be omitted here to keep this introduction from becoming excessively long.

By using this notion of path product, we will define the product of homotopy classes of paths as follows:

Definition 1.23. *Let γ and η be two closed paths based on a point x_0, and let $[\gamma]$ and $[\eta]$ be their respective homotopy classes. We define the product of these two classes as follows:*

$$[\gamma][\eta] = [\gamma\eta]$$

In other words, the product of the equivalence classes of γ and η is the equivalence class of their product path $\gamma\eta$ (It should be verified that this product is well-defined, but the interested reader can read more in the cited references in the bibliography [23] [49]).

We observe that, by the properties previously observed for the product of closed paths based on a point, the product of homotopy classes is also closed (since the homotopy class of $\gamma\eta$ is still the homotopy class of a closed path), associative, has an identity, and has an inverse for each element. Therefore, the set of homotopy classes of closed paths based on a point x_0 forms a group structure with respect to the defined multiplication. We call this group the **fundamental group of X at** x_0, denoted as $\pi_1(X, x_0)$.

It can be shown that for path-connected spaces (i.e., a space X in which there exists a path connecting any two points $x, y \in X$), the fundamental group is independent of the choice of a base point x_0, and we can simply write $\pi_1(X)$.

The fundamental group of path-connected spaces is an invariant under the homotopy equivalence of spaces, which we define below.

Definition 1.24 (Homotopy Equivalence of Spaces)**.** *We will say that two topological spaces X and Y are homotopy equivalent if there exist two continuous maps $f : X \to Y$ and $g : Y \to X$ such that the composition $f \circ g$ is homotopic to the identity id_Y on Y ($f \circ g \simeq id_Y$), and the composition $g \circ f$ is homotopic to the identity id_X on X ($g \circ f \simeq id_X$).*

Exercise 1.15. *Prove that the relation "being homotopy equivalent" is an equivalence relation between topological spaces.*

Exercise 1.16. *Prove that two homeomorphic topological spaces X and Y are homotopy equivalent. [Hint: Recall that you have a bijective and continuous function (i.e., it is its own inverse, and both are continuous functions) $\varphi : X \to Y$ between the two spaces, and remember the reflexive property of homotopy of functions.]*

We are now ready to answer the question we posed at the beginning of this subsection. In an algebraic topology course, it is customary to show that the disk D^2 is a contractible space (technically, a contractible space is defined as a ”space that is a deformation retract of a point”; intuitively, it can be interpreted as a space that is ”continuously” deformable into a point). Contractible spaces have a trivial fundamental group ($\pi_1(X) \cong 1$). It is also known that the fundamental group of our annular region A is an infinitely cyclic group (i.e., isomorphic to $\mathbb{Z}$). Therefore, the two fundamental groups are not isomorphic, which implies that the two spaces are not homotopy equivalent, which in turn implies that the two spaces are not homeomorphic.

1.2.2 Simplicial homology

In this subsection, we will introduce the concept of homology. Homology is one of the three central concepts of algebraic topology, along with homotopy (discussed earlier) and cohomology (which can be viewed as a dual concept to homology). Unlike homotopy, understanding and visualizing homology is less immediate, but we will attempt to explain its goals and motivations.

The objective, as before, is to associate topologically invariant groups with a space that, in a certain sense, can classify and identify its ”holes.” Before delving into the details of how to achieve this, let’s take a brief digression to motivate the introduction of the concept of a **simplicial complex** (which we will see shortly).

Imagine having a surface that is easily visualized in three-dimensional space, such as the 2-torus $\mathbb{T}^2$, shown below:

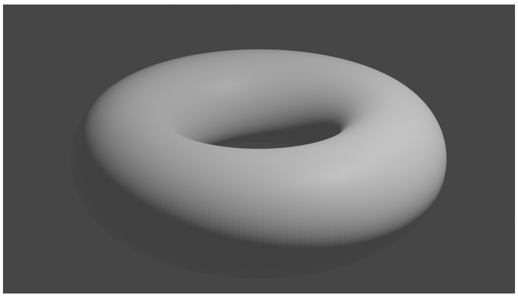


Figure 1.9: 2-torus $\mathbb{T}^2$ [rendered with Blender].

The torus is a somewhat ”complex” space to study, and ideally, we would like to replace it with something more ”simple” that is homeomorphic to it. One solution is to ”triangulate” it, i.e., to reconstruct its surface by ”gluing” together various triangles, as shown below (we would need to prove which spaces are triangulable and which are not, but since this is merely a preparatory introduction to the text, we will not do so as it would require a considerable amount of time).

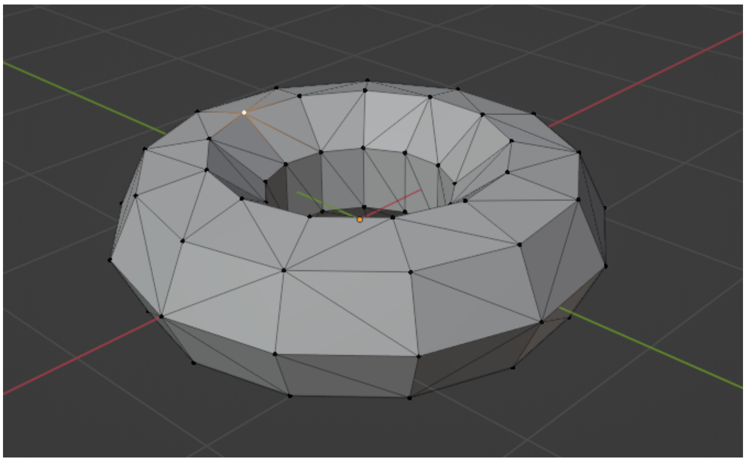


Figure 1.10: Triangulated 2-torus $\mathbb{T}^2$ [rendered with Blender].

To put it formally, triangulating our torus means finding a space formed by glued triangles that is homeomorphic to $\mathbb{T}^2$. We will rigorously define the concept of a "space formed by glued triangles" by introducing the notion of a **simplicial complex**.

From this point onwards, we can study and construct our simpler space using these glued triangles. In this context, we will talk about simplicial homology. With that said, let's give the necessary definitions for developing the concept of a simplicial complex:

Definition 1.25 (Simplex)**.** *Let $x_0, x_1, \ldots, x_m$ be $(m+1)$ linearly independent points in $\mathbb{R}^n$. We will call the Euclidean m-simplex (denoted as $(x_0, x_1, \ldots, x_m)$) the set of all points in space $(y_j)_{j=1}^n \in \mathbb{R}^n$ such that:*

$$y_j = \sum_{i=0}^{m} \lambda_i x_{ij} \quad \text{with} \quad \sum_{i=0}^{m} \lambda_i = 1 \quad \text{and} \quad \lambda_i \geq 0 \quad \forall i \tag{1.7}$$

(where x_{ij} is the j-th coordinate of x_i). We will also call the coefficients λ_i the "barycentric coordinates" of y_j in the simplex, and we will call the points $x_0, x_1, \ldots, x_m$ the vertices of the simplex.

The concept of a simplex is a generalization of the n-dimensional triangle. A 0-simplex is a point, a 1-simplex is a line segment, a 2-simplex is a triangle, a 3-simplex is a tetrahedron, and so on.

To define a simplex, we used a non-negative linear combination whose sum is equal to 1. This sum has a name that we will use frequently from now on, as it will be essential to the topic:

Definition 1.26 (Convex Combination)**.** *Let $x_0, x_1, \ldots, x_m \in \mathbb{R}^n$ be $(m+1)$ points in Euclidean space $\mathbb{R}^n$. We call the convex combination of $x_0, x_1, \ldots, x_m$ the set of points $y \in \mathbb{R}^n$ of the form:*

$$y = \sum_{i=0}^{m} \lambda_i x_i$$

where here

$$\sum_{i=0}^{m} \lambda_i = 1 \quad \text{and} \quad \lambda_i \geq 0 \quad \forall i$$

.

It can be shown (but since this is not a specifically algebraic topology book, we will not prove it) that every m-simplex in $\mathbb{R}^n$ is homeomorphic. Consequently, we can work with an arbitrary simplex that is simpler to use for our purposes. This simplex will be what we call the Euclidean standard m-simplex, denoted as Δ^m. Let's define it explicitly:

Definition 1.27 (Euclidean Standard Simplex)**.** *The Euclidean m-simplex in $\mathbb{R}^n$, denoted as Δ^m, is the m-simplex whose vertices are the points $\overline{x}_i = (\delta_{ij})_{j=1}^n$, where here δ_{ij} is the Kronecker delta defined as:*

$$\delta_{ij} := \begin{cases} 1 & \text{if} \quad i = j \\ 0 & \text{if} \quad i \neq j \end{cases}$$

i.e., the points $\overline{x}_i$ have coordinates equal to 0 everywhere except in their i-th position, $\overline{x}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$.

Now that we have introduced the concept of simplices, we need a way to rigorously define the idea of "gluing simplices" together. To do this, we first define the concept of a "face" of a simplex:

Definition 1.28 (Face of a Simplex)**.** *We will call a face of a simplex the convex combination of a non-empty subset of its vertices.*

With this, we are now ready to define what a simplicial complex is:

Definition 1.29 (Simplicial Complex)**.** *A simplicial complex K is a finite set of simplices such that:*

1. *If $A \in K$ is a simplex, and α is a face of A, then $\alpha \in K$.*

2. *For all $A, B \in K$, $A \cap B$ is either a common face of both simplices or the empty set, i.e.:*

$$A \cap B = \emptyset \quad \text{or} \quad A \cap B = \alpha$$

   where α is a face of both A and B.

Orientation is another important concept when dealing with simplices. You are probably familiar with the notion of orientation in other contexts, such as curves. For simplices, the concept is analogous; choosing an orientation means selecting the direction to traverse them. We can induce an orientation on a simplex by specifying an order for its vertices, as shown in the figure below:

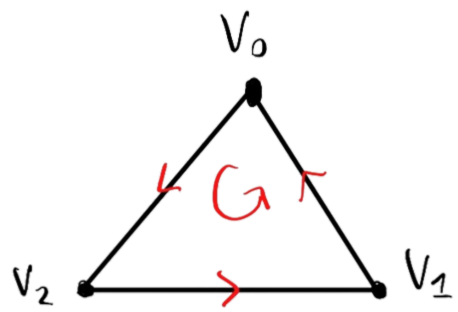


Figure 1.11: Orientation of a simplex

It's essential to note that an orientation on a simplex also induces an orientation on its faces, as illustrated in the figure above.

Now, let's imagine concatenating simplices in space and represent their geometric concatenation using a formal sum. Here we introduce the abelian group of "concatenations of simplices," which we will call the chain group C_n:

Definition 1.30 (Chain Group)**.** *Let $A_1^m, A_2^m, \ldots, A_k^m$ be k m-simplices with orientations in a simplicial complex K. We will call an m-chain the following finite formal sum with integer coefficients:*

$$\alpha = \lambda_1 A_1^m + \lambda_2 A_2^m + \ldots + \lambda_k A_k^m = \sum_{i=1}^{k} \lambda_i A_i^m$$

where $\lambda_i \in \mathbb{Z}$ for all i. The set of m-chains on a simplex, denoted as C_n, has an abelian group structure under the formal addition of these chains.

Of particular importance for homology theory is the concept of the boundary of an m-simplex A^m, obtained by summing all the $(m-1)$-dimensional faces of A^m (with the orientation induced by A^m). This notion aligns with our geometric intuition of the "boundary" of a geometric figure.

Definition 1.31 (Boundary of a Simplex)**.** *Let A^m be an m-simplex in a simplicial complex K. We will call the boundary of A^m the following $(m-1)$-chain in K given by:*

$$\partial A^m := A_0^{m-1} + A_1^{m-1} + \ldots + A_m^{m-1} \tag{1.8}$$

where here A_i^{m-1} are the $(m+1)$ $(m-1)$-faces of A^m. This notion extends to an m-chain $\alpha := \sum_i \lambda_i A_i^m$ as follows:

$$\partial \alpha = \sum_i \lambda_i \partial A_i^m \tag{1.9}$$

It can be proven that the boundary operator ∂ applied twice to any m-chain α results in zero, i.e., $\partial^2 \alpha = 0$ for all $\alpha \in C_m$. We can write this more concisely as $\partial^2 = 0$.

Proposition 1.12. *The boundary operator ∂ applied 2 times to any m-chain α will yield 0, i.e $\partial^2 \alpha = 0$ $\forall \alpha \in C_m$. We will write more concisely:*

$$\partial^2 = 0$$

Now, let's introduce the concept of cycles on a simplicial complex:

Definition 1.32 (Cycles on a Simplicial Complex)**.** *Let $\alpha \in C_m$ be an m-chain in a simplicial complex K. We will call α an m-cycle on K if its boundary is zero, i.e., $\partial \alpha = 0$. We will denote the set of m-cycles on a space as Z_m or as $Ker(\partial)$ (as it is indeed the kernel of the ∂ operator). This set, with the operation inherited from the group C_m, forms a subgroup of C_m.*

Now, we move on to the definition of boundaries on a simplicial complex:

Definition 1.33 (Boundaries on a Simplicial Complex). *Let $\alpha \in C_m$ be an m-chain in a simplicial complex K. Then, we will call α an m-boundary on K if it can be expressed as the boundary of an $(m+1)$-chain β, i.e.:*

$$\alpha = \partial\beta$$

We denote the set of boundaries on a simplicial complex as B_m or alternatively as $\mathrm{Im}(\partial)$ (since it is the image of the ∂ operator). This set is a subgroup of the abelian group Z_m. However, since every subgroup of an abelian group is a normal subgroup, we can define the quotient group Z_m/B_m. This quotient group will be called the m-th homology group and is denoted as H_m. Let's define it more precisely:

Definition 1.34 (m-th Homology Group on a Simplicial Complex). *Let K be a simplicial complex, and let Z_m and B_m (according to the notation introduced earlier) be the groups of m-cycles and m-boundaries on K, respectively. Then, we call the m-th homology group on K, denoted as H_m, the quotient group of Z_m by B_m:*

$$H_m := Z_m/B_m = \mathrm{Ker}(\partial)/\,\mathrm{Im}(\partial) \tag{1.10}$$

The sequence of homology groups (indexed by the natural number m) just constructed is called **simplicial homology**, as this sequence is associated with a simplicial complex K. The geometric purpose of the m-th homology group of a simplicial complex is to identify the m-dimensional holes within it. This sequence of groups can also be used to classify holes in more general spaces (recall the example of the torus mentioned at the beginning of this subsection) by constructing what is called the triangulation of a space (which we can now define rigorously: it is an isomorphism between a topological space and a simplicial complex).

Another important concept is that of homologous chains:

Definition 1.35 (Homologous Chains). *Let α and β be two m-chains on a simplicial complex K. We will then say that these two chains are homologous, and we will write $\alpha \simeq \beta$, if their difference is the boundary of an $(m+1)$-chain γ. In symbols, we can write:*

$$\alpha - \beta = \partial\gamma$$

Exercise 1.17. *Verify that the relation "α is homologous to β" is an equivalence relation on the set of m-chains on K.*

Definition 1.36 (Homologous m-Cycles to 0)**.** *Let α be an m-cycle on a simplicial complex K. We will then say that α is homologous to 0 if it is the boundary of an $(m+1)$-chain γ in K. In symbols:*

$$\alpha = \partial\gamma$$

1.2.3 Singular homology

The homology groups developed earlier were only associated with simplicial complexes. We will now develop a new type of homology that can be associated with more general topological spaces (which we will see is equivalent to simplicial homology in the case of triangulable spaces).

First, let's define the concept of an m-simplex on a space, which, to distinguish it from a classical simplex, we will call a **singular m-simplex**. Intuitively, the concept will be defined as that of a curve on a space, meaning that a simplex on a space is a continuous function from an m-simplex in $\mathbb{R}^n$ to a topological space X.

Definition 1.37 (Singular m-Simplex)**.** *Let X be a topological space, then we will call a continuous map $\sigma : \Delta^m \to X$ a **singular m-simplex** in X, where here Δ^m is the standard Euclidean m-simplex.*

Using the concept of a singular simplex on a space, we can also define the concept of chains on a space:

Definition 1.38 (Singular m-Chains on a Topological Space X)**.** *We will call **singular m-chains** on X, denoted as $C_m(X)$, the abelian group generated by the set of singular m-simplices on X.*

Similarly, we will define an operator of boundary on this newly constructed structure, which will allow us to define a sequence of homology groups:

Definition 1.39 (Boundary of an m-Simplex)**.** *Let σ be a singular m-simplex on X. We define the boundary of σ as follows:*

$$\partial\sigma := \sum_{k=0}^{m} (-1)^k \sigma|_{e_0,e_1,\ldots,e_{k-1},e_{k+1},\ldots e_m}$$

where here $e_0, e_1, \ldots, e_m$ are the vertices of the standard m-simplex Δ^m.

Given an m-singular chain $\alpha = \sum_i \lambda_i \sigma_i$, concatenating the simplices σ_i with weights $\lambda_i \in \mathbb{Z}$, the boundary of α, $\partial\alpha$, is defined similarly as the sum of the boundaries of the simplices σ_i, $\partial\sigma_i$, weighted by their respective λ_i:

$$\partial\alpha := \sum_{i=1}^{n} \lambda_i \partial\sigma_i$$

At this point, we are ready to define the m-th singular homology group on a topological space X:

Definition 1.40 (m-th Singular Homology Group on X)**.** *We will call the m-th singular homology group on X, denoted as $H_m(X)$, the quotient of the kernel of the m-th boundary operator by the image of the (m+1)-th boundary operator:*

$$H_m(X) = \mathrm{Ker}(\partial_m) / \mathrm{Im}(\partial_{m+1})$$

The last important facts that a reader who is determined to read this monograph, in case they have not yet encountered algebraic topology on their journey, should know about singular homology are that it is a topological invariant. In other words, given two homeomorphic topological spaces X and Y:

$$X \cong Y$$

their respective m-th singular homology groups will also be isomorphic:

$$H_m(X) \cong H_m(Y)$$

Another important fact is the equivalence between simplicial homology and singular homology for triangulable spaces. That is, for a topological space X homeomorphic to a simplicial complex K, its m-th simplicial homology group, denoted here as $H_m^{\Delta}(X)$, is isomorphic to its m-th singular homology group, denoted as $H_m^{\Sigma}(X)$:

$$H_m^{\Delta}(X) \cong H_m^{\Sigma}(X)$$

This motivates our initial intuition to understand the concepts of simplicial homology and simplicial complexes.

1.3 Elements of differential geometry

Of great importance for the comprehension of this monograph will be some concepts of differential geometry. As already mentioned, since this introduction is nothing more than a quick recap aimed at making this text accessible to a broader audience, we won't be able to delve deeply into the concepts we introduce, but we will provide just the minimum necessary to make the contents of this monograph understandable even to a reader without prior preparation in the mathematical disciplines of topology, differential geometry, and algebraic topology. Therefore, as we did in the section on algebraic topology, we will skip some proofs as they would require a much larger amount of time than we have available for proper introduction, and often a large number of equally lengthy preparatory lemmas.

1.3.1 Differentiable Manifolds

Let's begin by introducing the central concept of differential geometry, that of **differentiable manifold**. The fundamental idea behind the notion of a manifold is as follows: we, as humans, live on what can approximately be represented in a theoretical model by a sphere, S^2. However, in our everyday life, we perceive the surface we live on as flat, i.e., as a region of a two-dimensional plane. With this, we have just exemplified, in simple terms, the concept of a locally Euclidean space, i.e., a space that **locally** "resembles" Euclidean space $\mathbb{R}^n$ (in the case of the sphere, we have a Euclidean space of dimension 2, i.e., a space locally similar to $\mathbb{R}^2$). But how can we express this notion more rigorously? To do so, we will introduce the concepts of chart and atlas on a topological space X:

Definition 1.41. *Let X be a topological space, $U \subset X$ an open set in X, and $\tilde{U} \subset \mathbb{R}^n$ an open set in $\mathbb{R}^n$. We will call a homeomorphism $\varphi : U \to \tilde{U}$ an n-chart on X.*

Definition 1.42 (Smooth Atlas on a Topological Space)**.** *We will call an n-atlas C^∞ on X (or smooth atlas) a collection of n-charts $\mathscr{A} = \{\varphi_\alpha : U_\alpha \to \tilde{U}_\alpha \,;\, U_\alpha \text{ open in } X \,,\, \tilde{U}_\alpha \text{ open in } \mathbb{R}^n\}$ such that:*

1. *$\{U_\alpha\}$ is an open cover of X, i.e., $\bigcup_\alpha U_\alpha = X$.*

2. *If $U_\alpha \cap U_\beta \neq \emptyset$, then the function (which we will call a transition function)*

$$\varphi_\beta \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap U_\beta) \to \varphi_\beta(U_\alpha \cap U_\beta)$$

is a C^∞ function in the classical sense of real analysis (since it is a function from $\mathbb{R}^n$ to $\mathbb{R}^n$, we can apply the usual definitions of real analysis to it for one and more variables).

Remark 1.3. *It is possible to "relax" the second postulated property of an atlas, allowing, for example, the transition function to be simply of class C^1 (i.e., differentiable once): in this case, we will simply refer to them as differentiable atlases (although sometimes this name is also given to C^∞ atlases) or C^1 atlases.*

Definition 1.43 (Compatibility of Atlases)**.** *Two n-atlases $\mathscr{A}$ and $\mathscr{A}'$ are said to be compatible if their union is, in turn, a smooth atlas.*

Exercise 1.18. *Prove that the relation "being compatible" on a set of n-atlases on a space is an equivalence relation.*

At this point, we are ready to give the definition of a smooth manifold:

Definition 1.44 (Smooth Manifold)**.** *A smooth n-manifold M is a Hausdorff topological space with a countable base, associated with an equivalence class of smooth n-atlases with respect to the compatibility relation defined earlier.*

An equivalence relation of atlases with respect to the "being compatible" relation is often called a **differential structure**. Moreover, by a topological space with a countable base, we simply mean a topological space whose topology has a countable base (i.e., a finite or countably infinite base).

However, the reader may now wonder: what does all of this mean intuitively? Let us set aside the assumption that M is Hausdorff and has a countable base since the reasons for defining a manifold this way depend on various more technical reasons that we will not cover in this appendix (such as the problem of the existence of partitions of unity). Therefore, let us focus only on the second part of the definition. Let's start by looking at how the concept of an n-chart on a manifold can be geometrically interpreted. Let's consider the torus $\mathbb{T}^2$ again:

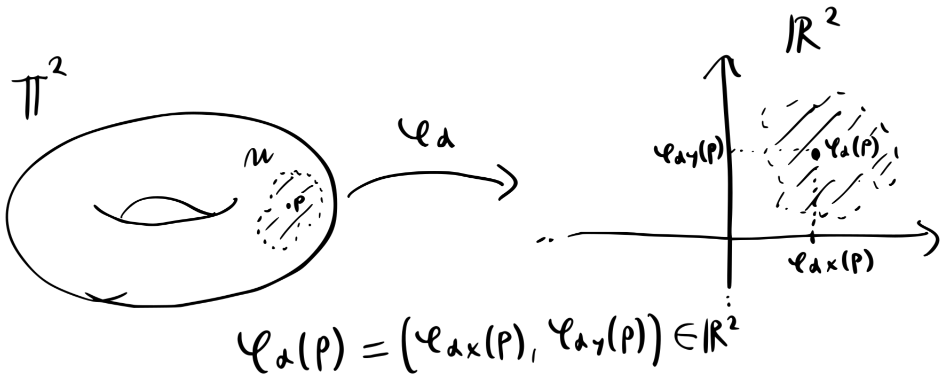


Figure 1.12: Chart on a 2-torus.

As we can see, a chart φ_α maps an open neighborhood U of a point p in our manifold (in this case, $\mathbb{T}^2$) to an open neighborhood $\varphi_\alpha(U)$ of $\varphi_\alpha(p)$ in $\mathbb{R}^2$. In other words, a chart (φ_α, U) provides coordinates for every point in the open set U of $\mathbb{T}^2$. Now, suppose we have another chart on our torus, $\psi : V \to \psi(V) \subset \mathbb{R}^2$:

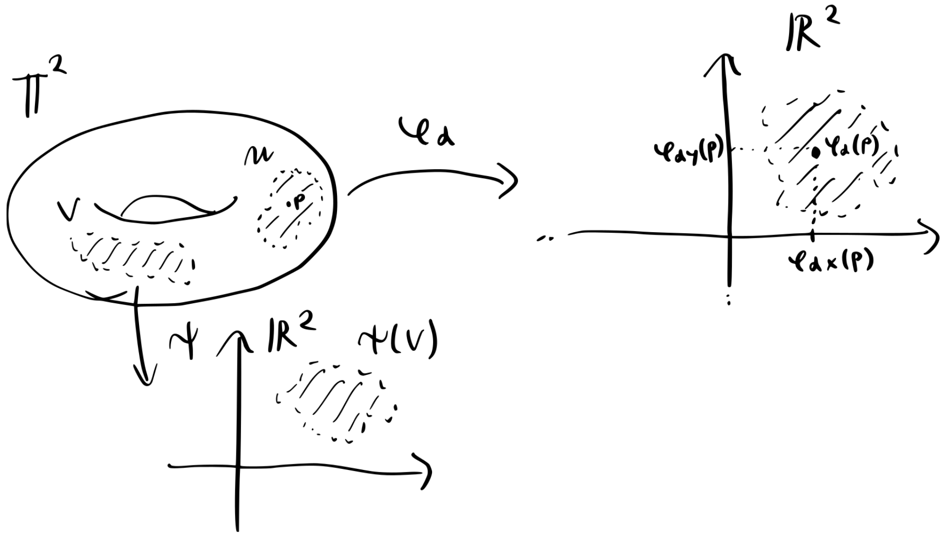


Figure 1.13: Charts on a 2-torus.

Since U and V are disjoint on $\mathbb{T}^2$, we don't have to worry much because the two charts map separate points on our manifold to specific points in Euclidean space. However, let's ask ourselves the following question: what if the two open sets U and V had intersected? Now the need to "agree" the charts of a manifold arises, i.e., to have a way to change from one coordinate system to another so that the concept of a coordinate chart of a neighborhood of a point makes sense. To illustrate this, let's look at the following figure:

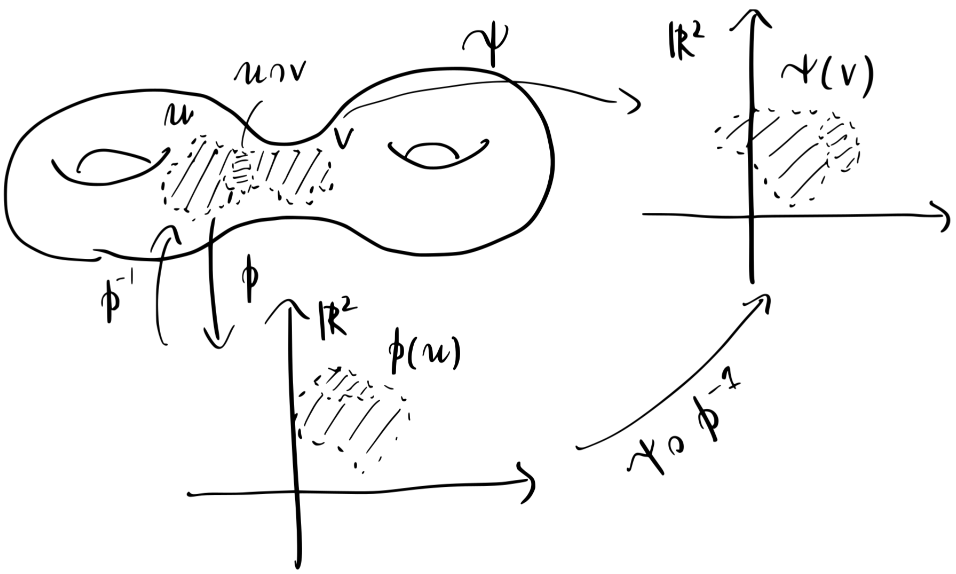


Figure 1.14: Charts on a genus 2 surface.

Here, the map $\psi \circ \phi^{-1}$ is precisely the one presented earlier and is the map that allows us to change coordinates in the region common to both U and V. Our assumption for a smooth manifold is that this transition function should be smooth and continuous so that any change of coordinates "behaves well" in mathematical terms.

Now, let's see some examples of smooth manifolds:

Example 1.7. *Consider the complex n-dimensional space $\mathbb{C}^n$: this topological space forms a 2n-smooth manifold with respect to the atlas $\mathscr{A} = \{(\mathbb{C}^n, \phi)\}$, where here we define $\phi : \mathbb{C}^n \to \mathbb{R}^{2n}$ to be:*

$$\phi(z_1, z_2, \dots, z_n) = (\mathfrak{R}(z_1), \mathfrak{I}(z_1), \mathfrak{R}(z_2), \mathfrak{I}(z_2), \dots, \mathfrak{R}(z_n), \mathfrak{I}(z_n))$$

where $\mathfrak{R}$ and $\mathfrak{I}$ here indicate the real and imaginary parts of a complex number.

We will call manifolds whose atlas is formed by a single chart **global manifolds**.

Exercise 1.19. *Prove that Euclidean space $\mathbb{R}^n$ is a global manifold.*

Exercise 1.20. *Prove that the 2-sphere S^2, with respect to the atlas $\mathscr{A} = \{(U_1^+,\phi_1^+),(U_2^+,\phi_2^+),(U_3^+,\phi_3^+),(U_1^-,\phi_1^-)(U_2^-,\phi_2^-),(U_3^-,\phi_3^-)\}$, forms the structure of a 2-manifold, where here the sets $U_i^\pm$ and the charts $\phi_i^\pm$ ($i = 1,2,3$) are defined as:*

$$U_i^+ = \{x = (x_1,x_2,x_3) \in S^2 \mid x_i > 0\} \quad U_i^- = \{x = (x_1,x_2,x_3) \in S^2 \mid x_i < 0\}$$

$$\begin{cases} \phi_1^\pm = (x_2,x_3) \\ \phi_2^\pm = (x_1,x_3) \\ \phi_3^\pm = (x_1,x_2) \end{cases}$$

1.3.2 Smooth maps and diffeomorphisms

Another important concept in differential geometry is that of smooth maps and diffeomorphisms (in category theory, we would call them morphisms and isomorphisms of the category of manifolds).

Definition 1.45 (Smooth Maps)**.** *Let M and N be two smooth manifolds with atlases $\mathscr{A}_M = \{(U_\alpha,\varphi_\alpha)\}$ and $\mathscr{A}_N = \{(V_\beta,\psi_\beta)\}$. We call a continuous function $f : M \to N$ a **smooth map** if for all α,β such that:*

$$f^{-1}(V_\beta) \cap U_\alpha \neq \emptyset$$

the composition of functions:

$$\psi_\beta \circ f \circ \varphi_\alpha^{-1} : \varphi_\alpha(U_\alpha \cap f^{-1}(V_\beta)) \to \psi_\beta(V_\beta)$$

is C^∞. We denote the set of all smooth maps from M to N by $C^\infty(M,N)$. If the set $N = \mathbb{R}$, then we simply write $C^\infty(M)$.

Exercise 1.21. *Prove that constant functions are smooth maps between smooth manifolds.*

Exercise 1.22. *Prove that the definition of a smooth map given above does not depend on the choice of atlases for M and N within their equivalence classes.*

Definition 1.46 (Diffeomorphism of Smooth Manifolds)**.** *Let M and N be, as before, two smooth manifolds. We call a homeomorphism $f : M \to N$ smooth with a smooth inverse f^{-1} a **diffeomorphism**. If there exists a diffeomorphism between two manifolds, we say that they are **diffeomorphic**.*

1.3.3 Tangent Spaces

Imagine having a manifold and a point p on it. We want to consider the vectors in the "tangent plane" at p," which we can interpret geometrically as vectors "coming out" of the point p, as shown in the following figure:

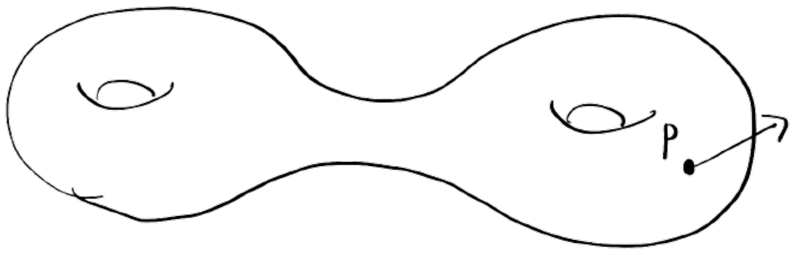


Figure 1.15: Tangent vector at a point p on a genus 2 surface.

The idea is to think of a vector as a directional derivative, as a directional derivative provides exactly the two pieces of information we need: the point from which our vector should "emerge" and its direction. However, the concept of a directional derivative is specific to functions from $\mathbb{R}^n$ to $\mathbb{R}$. To extend it to more general spaces, we need to create a more general mathematical entity. This motivates the definition of a **tangent vector**.

Definition 1.47 (Tangent Vector)**.** *Let M be a smooth manifold, and let $p \in M$ be one of its points. We will call an $\mathbb{R}$-linear map $v : C^\infty(M) \to \mathbb{R}$ that satisfies the so-called Leibniz property a* ***tangent vector to M at p****. In other words, a tangent vector v is a map from the set of smooth functions from M to $\mathbb{R}$ to $\mathbb{R}$ with the following properties: it is $\mathbb{R}$-linear, meaning:*

$$v(\lambda f + \mu g) = \lambda v(f) + \mu v(g) \quad \forall \lambda, \mu \in \mathbb{R} \ , \ \forall f, g \in C^\infty(M)$$

and it is a derivation, meaning:

$$v(fg) = f(p)v(g) + g(p)v(f) \quad \forall f, g \in C^\infty(M)$$

The set of tangent vectors at a point in a manifold forms a vector space, denoted as the **tangent space at p**, represented as T_pM.

For $\mathbb{R}^n$, a basis of $T_p\mathbb{R}^n$ consists of vectors $\{\frac{\partial}{\partial x_i}|_p \quad ; \quad i \in \mathbb{N} \cap [1,n]\}$, which aligns with our initial intuition about directional derivatives and the concept of tangent vectors. Therefore, $\dim(T_p\mathbb{R}^n) = n = \dim(\mathbb{R}^n)$. For a more general manifold M, given a chart $\varphi : U \to \mathbb{R}^n$, denoted as $\varphi = (r_1 \circ \varphi, r_2 \circ \varphi, \ldots, r_n \circ \varphi)$, where r_i is the i-th coordinate function, the basis of T_pM will be determined by the following tangent vectors induced by the local coordinates at a point:

$$\frac{\partial}{\partial x_i}|_p(f) := \frac{\partial}{\partial r_i}|_{\varphi(p)}(f \circ \varphi^{-1})$$

.

1.3.4 Differentials and Cotangent Spaces

Definition 1.48 (Differential of a Smooth Map)**.** *Let $f : M \to N$ be a smooth map between two manifolds, and let $p \in M$ be a point on M. We will call the differential of f at p the following map:*

$$df_p : T_pM \to T_{f(p)}N$$

defined as:

$$df_p(v)(h) := v(h \circ f) \quad \forall v \in T_pM \;\; \forall h \in C^\infty(N)$$

For the case of a map from M to $\mathbb{R}$, its differential

$$df_p : T_pM \to \mathbb{R}$$

will be, according to the definition of the dual space of a vector space, an element of T_p^*M, the dual space of T_pM, defined as:

$$T_p^*M := \mathrm{Hom}(T_pM; \mathbb{R})$$

We will call this space the **cotangent space of M at p**. For a function of the type $f : M \to \mathbb{R}$ seen before, for $v \in T_pM$, its differential will be equal to:

$$df_p(v) = v(f)$$

1.3.5 Tangent Bundles, Vector Fields, and Cotangent Bundles

We introduced the concept of tangent space at a point p on a manifold a couple of subsections ago. The tangent bundle of a manifold M is the disjoint union of tangent spaces at every point of M:

Definition 1.49 (Tangent Bundle)**.** *We define the tangent bundle TM of a manifold M as:*

$$TM := \bigsqcup_{p\in M} T_pM = \bigcup_{p\in M} \{p\} \times T_pM \tag{1.11}$$

Therefore, we can view an element of the tangent bundle as a pair (p, v), where $p \in M$ is a point of the manifold M, and $v \in T_pM$ is a tangent vector of M at p.

This set can be given a topology and a smooth manifold structure induced by the smooth manifold structure of M. The construction proceeds as follows: given a chart $\varphi = (x_1, x_2, \ldots, x_n) : U \to \mathbb{R}^n$ on M, we define a chart on TM, denoted as $\overline{\varphi}$, as follows:

$$\overline{\varphi} := (x_1 \circ \pi, x_2 \circ \pi, \ldots, x_n \circ \pi, dx_1, \ldots, dx_n) : TU \to \mathbb{R}^n \times \mathbb{R}^n \tag{1.12}$$

Here, $x_i \circ \pi$ is the composition of the i-th coordinate function of $\mathbb{R}^n$ and the projection of the tangent bundle, and $dx_i : TU \to \mathbb{R}$ is a map that maps $(p, v) \in TU$ to $(dx_i)_p(v)$.

We define a topology on TM in which we declare sets $V \subset TM$ open if, for every chart on M, $\varphi : U \to \mathbb{R}^n$, $\overline{\varphi}(V \cap TU) \subset \mathbb{R}^n \times \mathbb{R}^n$ is open.

With respect to this topology, the charts $(TU, \overline{\varphi})$ on TM induced by the charts (U, φ) on M form an atlas, and with respect to this smooth structure, the tangent bundle becomes a smooth manifold of double the dimension of M, $\dim(TM) = 2n$, where n is the dimension of the manifold M.

Similarly, we will define the concept of a **cotangent bundle**:

Definition 1.50 (Cotangent Bundle of a Smooth Manifold M)**.** *Let M be a smooth manifold; we will define the cotangent bundle T^*M as the disjoint union of cotangent spaces at every point p of M, i.e.:*

$$T^*M = \bigsqcup_{p\in M} T^*_pM = \bigcup_{p\in M} \{p\} \times T^*_pM \tag{1.13}$$

Similarly, we can induce a smooth manifold structure on the cotangent bundle T^*M, of double dimension with respect to M, induced by the smooth manifold structure on M.

The final important concept we will introduce in this subsection is that of a vector field. The intuition behind the idea is precisely what is often learned in an elementary course in multivariable calculus; a vector field is a function that associates a vector to every point in space, with the vectors ”emerging” from that point.

We can formalize this idea with the following definition:

Definition 1.51 (Smooth Vector Field)**.** *Let M be a smooth manifold, and let X be a smooth function from M to TM, $X : M \to TM$, such that the composition:*

$$\pi \circ X = id$$

is the identity. In other words, for every $p \in M$, $X(p) \in T_pM$.

1.3.6 Tensor Product of Vector Spaces

In this subsection, we will give the definition of the tensor product and demonstrate some of its properties.

Definition 1.52 (Tensor Product)**.** *Given two vector spaces V and W, we will call the* ***tensor product of V and W*** *a vector space $V \otimes W$ together with a bilinear map $\otimes : V \times W \to V \otimes W$ (of which we denote the image of an element of $V \times W$ as $\otimes((v, w)) = v \otimes w$) such that for every bilinear map $b : V \times W$, there exists a unique linear map $\overline{b} : V \otimes W \to U$ such that $\overline{b}(v \otimes w) = b(v, w)$ $(v \in V, w \in W)$. This property is called the* ***universal property of the tensor product****.*
Graphically, we are saying that the following diagram commutes:

$$\begin{tikzcd} V \times W \arrow[r, "b"] \arrow[d, "\otimes"] & U \\ V \otimes W \arrow[ur, dotted, "\overline{b}"] & \end{tikzcd}$$

Proposition 1.13 (Uniqueness of the Tensor Product)**.** *If there exist two tensor products of V and W, then they are isomorphic.*

Proof. Firstly, suppose that such tensor products exist, which we denote as $V \otimes_1 W$ and $V \otimes_2 W$. The two tensor products, by assumption, will have two bilinear maps $\otimes_1 : V \times W \to V \otimes_1 W$ and $\otimes_2 : V \times W \to V \otimes_2 W$, and they satisfy the universal property. Representing all this with diagrams:

$$\begin{tikzcd} V \times W \arrow[r, "b_1"] \arrow[d, "\otimes_1"] & U_1 \\ V \otimes_1 W \arrow[ur, dotted, "\overline{b_1}"] & \end{tikzcd} \qquad \begin{tikzcd} V \times W \arrow[r, "b_2"] \arrow[d, "\otimes_2"] & U_2 \\ V \otimes_2 W \arrow[ur, dotted, "\overline{b_2}"] & \end{tikzcd}$$

Now, let $U_2 = V \otimes_1 W$, $b_2 = \otimes_1$, and $U_1 = V \otimes_2 W$, $b_1 = \otimes_2$. We obtain the following commutative diagrams:

$$\begin{tikzcd} V \times W \arrow[r, "\otimes_2"] \arrow[d, "\otimes_1"] & V \otimes_2 W \\ V \otimes_1 W \arrow[ur, dotted, "\overline{\otimes_2}"] & \end{tikzcd} \qquad \begin{tikzcd} V \times W \arrow[r, "\otimes_1"] \arrow[d, "\otimes_2"] & V \otimes_1 W \\ V \otimes_2 W \arrow[ur, dotted, "\overline{\otimes_1}"] & \end{tikzcd}$$

Now, let $T_1 = \overline{\otimes_1} \circ \overline{\otimes_2} : V \otimes_1 W \to V \otimes_1 W$, and $T_2 = \overline{\otimes_2} \circ \overline{\otimes_1} : V \otimes_2 W \to V \otimes_2 W$. By the fundamental property of the tensor product, we obtain the following commutative diagrams:

$$\begin{tikzcd} V \times W \arrow[r, "\otimes_1"] \arrow[d, "\otimes_1"] & V \otimes_1 W \\ V \otimes_1 W \arrow[ur, dotted, "T_1"] & \end{tikzcd} \qquad \begin{tikzcd} V \times W \arrow[r, "\otimes_2"] \arrow[d, "\otimes_2"] & V \otimes_2 W \\ V \otimes_2 W \arrow[ur, dotted, "T_2"] & \end{tikzcd}$$

We observe that the identity maps $id_1 : V \otimes_1 W \to V \otimes_1 W$ and $id_2 : V \otimes_2 W \to V \otimes_2 W$ make the same diagrams commute:

$$\begin{tikzcd} V \times W \arrow[r, "\otimes_1"] \arrow[d, "\otimes_1"] & V \otimes_1 W \\ V \otimes_1 W \arrow[ur, dotted, "id_1"] & \end{tikzcd} \qquad \begin{tikzcd} V \times W \arrow[r, "\otimes_2"] \arrow[d, "\otimes_2"] & V \otimes_2 W \\ V \otimes_2 W \arrow[ur, dotted, "id_2"] & \end{tikzcd}$$

Hence, by the assumed uniqueness in the definition of the tensor product, we have $T_1 = id_1$ and $T_2 = id_2$. In other words, $\overline{\otimes_1}$ and $\overline{\otimes_2}$ are inverses of each other. Thus, we have constructed an isomorphism from $V \otimes_1 W$ to $V \otimes_2 W$, proving the claim. □

Now, we provide a constructive proof of the existence of the tensor product (now we can properly say ”the” instead of ”a” since we have shown its uniqueness up to isomorphism).

Theorem 1.6. *The tensor product between two vector spaces V and W, denoted as $V \otimes W$, exists.*

Proof. Let's begin by considering what we will call the **formal product space of V and W**, which we define as:

$$V * W := \text{span}_{\mathbb{R}}\{v * w \mid v \in V, w \in W\}$$

Furthermore, consider the subspace I of $V * W$ defined as follows:

$$I := \text{span}_{\mathbb{R}} \begin{Bmatrix} \begin{cases} (v_1 + v_2) * w - v_1 * w - v_2 * w \\ v * (w_1 + w_2) - v * w_1 - v * w_2 \\ \lambda(v * w) - (\lambda v) * w \\ \lambda(v * w) - v * (\lambda w) \end{cases} & ; v, v_1, v_2 \in V \ , \ w, w_1, w_2 \in W \ \lambda \in \mathbb{R} \end{Bmatrix} \tag{1.14}$$

We claim that the quotient space of $V * W$ by the subspace I is our tensor product, i.e.:

$$V \otimes W = V * W/I$$

Define the map $\otimes : V \times W \to V \otimes W$ as the composition of the inclusion $\imath : V \times W \to V * W$ and the quotient map $q : V * W \to V * W/I$. Since we have "quotiented out" the subspace I from the larger space $V * W$, everything in I collapses to the identity in $V * W/I$. Therefore, the defined map $\otimes$ maps all expressions of the form (1.14) to the identity, and it is thus bilinear.
Now, we verify that the universal property is satisfied: let $b : V \times W \to U$ be a bilinear map. This map induces another linear map (and the unique one) $\overline{b} : V * W \to U$ given by $\overline{b}(v * w) = b(v, w)$. Since b is bilinear, $\overline{b}$ reduces to the identity over our subspace I defined in equation (1.14). Thus, we obtain a linear map $\tilde{b} : V \otimes W \to U$ with $\tilde{b}(v \otimes w) = \overline{b}(v * w) = b(v, w)$. Since vectors of the form $v \otimes w$ with $v \in V$ and $w \in W$ span $V \otimes W$, we have $\tilde{b}$ is unique. Therefore, the universal property is satisfied, and the claim is proven. □

The tensor product has many other important properties, which we will list below without providing proofs due to time constraints (although the reader can try to prove them, as they only require knowledge of certain facts of linear algebra).

Proposition 1.14 (Properties of the Tensor Product)**.** *The tensor product satisfies the following properties:*

- $\mathrm{Hom}(V \otimes W, \mathbb{R}) \cong \mathrm{Mult}(V \times W, \mathbb{R})$*, where here* $\mathrm{Mult}(V \times W, \mathbb{R})$ *is the vector space of bilinear maps from* $V \times W$ *to* $\mathbb{R}$*.*
- $\dim(V \otimes W) = \dim V \dim W$ *(an immediate consequence of the first fact).*
- $V \otimes W \cong W \otimes V$ *(commutative).*
- $V \otimes (U \otimes W) \cong (V \otimes U) \otimes W$ *(associative).*

Now, let's introduce the important definition of a tensor algebra, but first, let's define what it means to take the "n-th tensor power" of a vector space.

Definition 1.53 (Tensor Power)**.** *Let* V *be a vector space. Then, we will define the n-th tensor power of* V*, denoted as* $V^{\otimes_n}$*, as:*

$$V^{\otimes_n} := \bigotimes_{i=0}^{n} V \tag{1.15}$$

Based on the definition of tensor power, we define the **tensor algebra of a vector space V** as follows:

Definition 1.54 (Tensor Algebra of a Vector Space V)**.** *Let V be a vector space. We will call the **tensor algebra**, denoted as $\mathfrak{T}(V)$, of V the following vector space:*

$$\mathfrak{T}(V) := \bigoplus_{n=0}^{\infty} V^{\otimes_i} \tag{1.16}$$

Definition 1.55 (Graded Algebra)**.** *We say that an algebra $\mathbb{A}$ is graded if and only if:*

$$\mathbb{A} = \bigoplus_{n=0}^{\infty} A_n \tag{1.17}$$

and for all $a \in A_i$, $b \in A_j$, we have:

$$a \cdot b \in A_{i+j}$$

*Here, $\cdot$ is the multiplication with respect to which $\mathbb{A}$ forms an algebra. We call the elements $a \in A_i$ **i-th-degree elements of** $\mathbb{A}$ or **elements of degree i**.*

Exercise 1.23. *Prove that the tensor algebra $\mathfrak{T}(V)$ of a vector space V is a graded algebra with respect to the tensor product $\otimes$.*

1.3.7 Differential forms and the Grassmann exterior algebra

We are now ready to define the Grassmann exterior algebra of a vector space, which will be used to rigorously define mathematical entities called "differential forms," which will be of fundamental importance for understanding this monograph. The definition of the exterior algebra can be given in terms of a categorical universal property, similar to what was done with the tensor product. However, in this text, we will opt for a direct definition in terms of a quotient of algebras, which is much more common.

Definition 1.56 (Grassmann Exterior Algebra of a Vector Space)**.** *We will call the **Grassmann exterior algebra of V**, denoted by $\bigwedge^*(V)$, the following algebra:*

$$\bigwedge^*(V) := \mathfrak{T}(V)/I \tag{1.18}$$

where here I is the ideal of the tensor algebra $\mathfrak{T}(V)$ generated by the set $\{x \otimes x \ : \ x \in V\}$.

This algebra inherits from the tensor algebra $\mathfrak{T}(V)$ a product, which we will call the **exterior product** and denote it as $\wedge$. For every $v \in V$, we have $v \wedge v = 0$ in the exterior algebra $\bigwedge^*(V)$ (this immediately follows from the definition: by taking a quotient of algebras, we are "intuitively" reducing I to the identity). From this property, it immediately follows that, for given $v_1, v_2 \in V$:

$$(v_1 + v_2) \wedge (v_1 + v_2) = v_1 \wedge v_1 + v_1 \wedge v_2 + v_2 \wedge v_1 + v_2 \wedge v_2 = 0$$

which implies $v_1 = v_2 = -v_2 \wedge v_1$, i.e., the exterior product is anti-commutative.

The exterior algebra itself is a graded algebra since it is the quotient of a graded algebra ($\mathfrak{T}(V)$) by one of its graded ideals (I). Regarding $\mathfrak{T}(V)$, it is trivially true that it is a graded algebra, as seen in exercise 1.23. For the ideal I, we notice that we can write it as the direct sum:

$$I = \bigoplus_{n=0}^{\infty}(I \cap \mathfrak{T}^n(V)) = \bigoplus_{n=0}^{\infty}(I \cap V^{\otimes_n}) = \bigoplus_{n=2}^{\infty}(I \cap V^{\otimes_n})$$

(the last equality in the series is true because $V \cap I = V^{\otimes_1} \cap I = 0$ and $\mathbb{R} \cap I = V^{\otimes_0} \cap I = 0$).

By defining $\bigwedge^n(V) = \mathfrak{T}^n(V)/(\mathfrak{T}^n(V) \cap I)$ (where here $\mathfrak{T}^n(V)$ represents the elements of degree n in the tensor algebra $\mathfrak{T}(V)$, i.e., $V^{\otimes_n}$), we can write our Grassmann exterior algebra as:

$$\bigwedge^*(V) = \bigoplus_{n=0}^{\infty} \bigwedge^n(V) \tag{1.19}$$

We will call the space $\bigwedge^n(V)$ the space of the nth Grassmann exterior product of V. The dimension of the components in the "direct sum decomposition" of the graded algebra $\bigwedge^*(V)$ is given by:

$$\dim \bigwedge^n(V) = \binom{\dim V}{n} = \frac{(\dim V)!}{n!(\dim V - n)!} \tag{1.20}$$

Given a vector space V with m dimensions and a basis $\{v_1, \dots, v_m\}$, we have that the set $\{v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_n}\ ;\ 1 \leq i_1 \leq \cdots \leq i_n \leq m\}$ is a basis for $\bigwedge^n(V)$. The cardinality of this set is precisely $\binom{\dim V}{n}$, in accordance with what was said before.

At this point, we are ready to define the concept of a differential form:

Definition 1.57 (Differential Form on a Manifold M)**.** *We will call a **differential form** on a manifold M (or sometimes a smooth form) a smooth* *mapping $\omega : M \to \bigwedge^n(T^*M)$ such that $\pi \circ \omega : M \to M$ is the identity (where here π is the projection of the cotangent bundle introduced two subsections ago). This condition can also be expressed as:*

$$\omega_p \in \bigwedge^n(T_p^*M) \quad \forall p \in M$$

where here ω_p is the image of p under the mapping ω.

We will denote the space of n-forms on M as $\Omega^n(M)$ and denote the space of all differential forms on M as:

$$\Omega(M) = \bigoplus_{n=0}^{\infty} \Omega^n(M)$$

It is possible to multiply differential forms on M using the exterior product on the space $\bigwedge^n(T^*M)$. Let $\omega \in \Omega^n(M)$ and $\eta \in \Omega^l(M)$ be two differential forms on M. We define their $(n+l)$-product form as:

$$(\omega \wedge \eta)_p := \omega_p \wedge \eta_p \tag{1.21}$$

which means that for every point $p \in M$, it associates the exterior product of the images of ω and η: $\omega_p \wedge \eta_p$.

1.3.8 Pullback and Integration

We conclude this introduction with a discussion on the integration of differential forms. Before we can talk about integrals of differential forms, we need to digress on the concept of **pullback**.

Definition 1.58 (Pullback of Smooth Functions)**.** *Let $F : M \to N$ be a smooth map between two manifolds M and N, and let $f \in C^\infty(N)$ be a smooth function on N. The map F induces what we call the **pullback application**, $F^* : C^\infty(N) \to C^\infty(M)$, defined as:*

$$F^*f := f \circ F \tag{1.22}$$

Given a smooth function $F : M \to N$, we can extend the definition of pullback to differential forms: let ω be an n-form, $p \in M$, and $v_1, \ldots, v_n \in T_pM$. Then:

$$(F^*\omega)_p(v_1, \ldots, v_n) := \omega_{F(p)}(dF_pv_1, \ldots, dF_pv_n) \tag{1.23}$$

Next, we list some important properties of pullbacks that will be very useful later:

Proposition 1.15 (Properties of Pullbacks)**.** *Let $F : M \to N$ and $G : N \to Z$ be two smooth functions, and let $\omega \in \Omega^n(M)$, $\eta \in \Omega^l(M)$ be two forms on M. Then the following expressions are valid:*

1. $(G \circ F)^* \omega = F^*(G^* \omega)$.

2. $(F^* \omega) \wedge (F^* \eta) = F^*(\omega \wedge \eta)$ *(Pullbacks "preserve" the exterior product).*

3. $F^* df = d(f \circ F)\ \forall f \in C^\infty(N)$.

Let's now finally talk about the integration of differential forms on a manifold. Due to time constraints (an introduction is too small to contain a proper and rigorous treatment of all these facts), we won't demonstrate the necessary details (such as the "well-definedness" of the notion of such an integral, its properties, or the intuition behind the idea). Let's first introduce the concept of the support of a form:

Definition 1.59 (Support of an n-form)**.** *Let $\omega \in \Omega^n(M)$ be an n-form on M: we will call the **support of** ω the closure of the set of points where ω is non-zero, i.e:*

$$\operatorname{supp} \omega := \overline{\{p \in M\, ;\ \omega_p \neq 0\}} \tag{1.24}$$

*We will say that a form has **compact support** if its support $\operatorname{supp}\omega$ is compact.*

Another preliminary definition that we will need is that of an **oriented manifold**:

Definition 1.60. *Let M be an m-manifold and let $\{\varphi_\alpha : U_\alpha \to \mathbb{R}^m\}$ be its atlas. We will say that M is orientable if $\forall \alpha, \beta$ we have:*

$$\det(d(\varphi_\alpha \circ \varphi_\beta^{-1})) > 0 \tag{1.25}$$

$\forall p \in \varphi_\beta(U_\alpha \cap U_\beta)$. The choice of such an atlas is called the orientation of M.

Proposition 1.16. *An m-manifold M is orientable if and only if it has an m-form everywhere non-zero, which we will call **volume form**, often indicated with ν.*

Now, we can finally define the concept of **integral of an n-form on an m-manifold**.

Definition 1.61 (Integral of an n-form on a Manifold)**.** *Let M be an orientable m-dimensional manifold, and let $\omega \in \Omega^m(M)$ be an m-form with compact support. Then, we call the integral of ω on M the following expression:*

$$\int_M \omega = \int_{\varphi(U)} (\varphi^{-1})^* \omega \tag{1.26}$$

It is also possible to define a theory of integration on smooth chains (by smooth chains, we simply mean chains on a space whose simplices are all smooth functions).

Let γ be an m-chain in a manifold M, and let ω be an (m-1)-form on M. Then we define the integral of ω over γ as follows:

$$\int_\gamma \omega = \sum_i a_i \int_{\Delta^m} \sigma_i^* \omega \tag{1.27}$$

Here, $\sigma_i^* \omega$ is the pullback of ω with respect to the smooth function $\sigma_i : \Delta^m \to M$, and Δ^m denotes the m-dimensional standard Euclidean simplex.

1.3.9 Exterior Differentiation

Theorem 1.7. *There exists a unique $\mathbb{R}$-linear map from $\Omega^n(M)$ to $\Omega^{n+1}(M)$, denoted by $d : \Omega^n(M) \to \Omega^{n+1}(M)$, with the following properties:*

1. *df is the differential of f for a 0-form f.*
2. *$d^2 = 0$, i.e., $d(df) = 0$ for all 0-forms f.*
3. *$d(\omega \wedge \eta) = d\omega \wedge \eta + (-1)^m (\omega \wedge \eta)$, where $\omega \in \Omega^m(M)$.*

We call this map the ***exterior differential****.*

Given a local coordinate system in an m-manifold, $(x_1, \ldots, x_m)$, and an n-form $\omega \in \Omega^n(M)$

$$\omega = w dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_n} = w dx_I$$

We can explicitly compute the exterior derivative of ω as follows:

$$d\omega = \sum_i \frac{\partial w}{\partial x_i} dx_i \wedge dx_I \tag{1.28}$$

Here, $dx_I = dx_{i_1} \wedge dx_{i_2} \wedge \cdots \wedge dx_{i_n}$.

1.3.10 Stokes' Theorem and Integral Properties

We conclude this preparatory section with a list of some important theorems concerning integrals of forms on manifolds (or chains) that have just been defined. We begin with the most important one, Stokes' theorem on manifolds with boundaries:

Theorem 1.8 (Stokes' Theorem on Manifolds with Boundaries)**.** *Let M be an orientable m-manifold with a boundary, where the $(m-1)$-manifold ∂M is given the orientation induced by M, and let ω be an $(m-1)$-form with compact support. Then:*

$$\int_M d\omega = \int_{\partial M} \omega \tag{1.29}$$

For manifolds with boundaries, we refer to a topological space in which each point has a neighborhood homeomorphic to an open subset of the half-space $\mathbb{R}^{n+} = \{(x_i)_{i=1}^n \ ; \ x_1 \geq 0\}$.

There exists a version of the theorem for the integration of forms on chains, known as the "Stokes' theorem for chains":

Theorem 1.9 (Stokes' Theorem for Chains)**.** *Let γ be an m-chain in a manifold M, and let ω be an $(m-1)$-form on M. Then:*

$$\int_{\partial\gamma} \omega = \int_\gamma d\omega \tag{1.30}$$

As an immediate corollary of Stokes' theorem for chains, we have the following fact:

Proposition 1.17. *Let ω be a closed m-form (i.e., a form whose exterior derivative is zero, $d\omega = 0$), and let γ and η be two homologous m-chains ($\gamma \simeq \eta$). Then:*

$$\int_\gamma \omega = \int_\eta \omega$$

We can concisely state the content of this proposition with the following statement: "The integrals of closed forms over homologous chains are equal."

Proof. By the definition of homologous chains, we will have $\gamma - \eta = \partial\nu$ for an $(m+1)$-chain ν. We will thus get:

$$\int_\gamma \omega - \int_\eta \omega = \int_{\partial\nu} \omega = \int_\nu d\omega = 0$$

□

Chapter 2

Algebra of quaternions

In this chapter, we will study quaternions from an algebraic perspective. First, we will construct them, and then we will observe and prove their most important properties.

2.1 Construction of the Quaternion Algebra

2.1.1 Cayley-Dickson Construction of $\mathbb{C}$

First, let's construct our quaternion algebra. To do this, we will use the Cayley-Dickson construction, which allows us to iteratively generate algebras over the real numbers that are "larger" (with double the dimension) than the previous one. We define a product on the Cartesian product of the previous algebra with itself. Let's recall the Cayley-Dickson construction of the 2-dimensional associative unitary $\mathbb{R}$-Algebra of Complex Numbers $\mathbb{C}$:

We define a complex number $z \in \mathbb{C}$ as a pair of real numbers (a, b) with $a, b \in \mathbb{R}$. We also define what it means for two complex numbers to be "equal":

Definition 2.1. *Two complex numbers $z_1, z_2 \in \mathbb{C}$; $z_1 := (\alpha_1, \beta_1), z_2 := (\alpha_2, \beta_2)$ with $\alpha_i, \beta_i \in \mathbb{R}$, i=1,2, are said to be equal if and only if $\alpha_1 =_{\mathbb{R}} \alpha_2$ and $\beta_1 =_{\mathbb{R}} \beta_2$, where here $=_{\mathbb{R}}$ denotes the canonical equivalence relation on $\mathbb{R}$. If two complex numbers are equal, we write: $z_1 =_{\mathbb{C}} z_2$.*

This relation is an equivalence relation on $\mathbb{C}$. The properties that make it an equivalence relation (transitive, symmetric, and reflexive) follow trivially from the fact that the equivalence relation on the real numbers, on which this newly introduced relation is based, also satisfies these properties. We will omit the proof to avoid making the text lengthy.

To simplify notation, from now on, we will represent the newly defined canonical equivalence relation on complex numbers simply with the symbol "=".

Now, let's define algebraic operations on this structure:

- We define component-wise addition of two complex numbers $z_1 = (a_1, b_1)$ and $z_2 = (a_2, b_2)$ as:
$$z_1 + z_2 = (a_1 + a_2, b_1 + b_2) \tag{2.1}$$
- We define the multiplication of a complex number z by a scalar $\lambda \in \mathbb{R}$ as:
$$\lambda z = (\lambda a, \lambda b) \tag{2.2}$$
- We define the multiplication of two complex numbers z_1 and $z_2 \in \mathbb{C}$ as:
$$(a, b) \cdot (c, d) = (ac - db, ad + cb) \tag{2.3}$$
- Finally, we define the conjugation of a complex number $z \in \mathbb{C}$ as:
$$(a, b)^* = (a, -b) \tag{2.4}$$

It is immediately verifiable that $\mathbb{C}$ forms a vector space isomorphic to $\mathbb{R}^2$ with component-wise addition and scalar multiplication.

We now have a 2-dimensional vector space over $\mathbb{R}$ with a multiplicative structure defined by the formula above and a unary operation called "conjugation." We need to verify that this multiplicative structure satisfies the necessary axioms to be called an associative $\mathbb{R}$-Algebra.

Proposition 2.1. *The previously defined product on the set of pairs (a,b) with a,b $\in \mathbb{R}$ satisfies the following properties for $z_1, z_2, z_3 \in \mathbb{C}$ and $\lambda \in \mathbb{R}$:*

1. $z_1(z_2z_3) = (z_1z_2)z_3$ *(Associative).*
2. $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$ *(Right Distributive).*
3. $(z_1 + z_2)z_3 = z_1z_3 + z_2z_3$ *(Left Distributive).*
4. $\lambda(z_1z_2) = (\lambda z_1)(z_2) = (z_1)(\lambda z_2)$ *(Associative Property of Scalar Multiplication).*

Furthermore, the structure has a multiplicative identity, the pair (1,0) = $1_{\mathbb{C}}$. Let $\mathbb{C}$ = {(a,b) | a,b $\in \mathbb{R}$} be the set of pairs of real numbers. In other words, the vector space that $\mathbb{C}$ forms with component-wise addition and scalar multiplication forms an associative unitary $\mathbb{R}$-Algebra when considered with the previously defined multiplicative structure.

Proof. Let $z_1 = (\alpha_1, \beta_1)$, $z_2 = (\alpha_2, \beta_2)$, $z_3 = (\alpha_3, \beta_3)$ with $\alpha_i, \beta_i \in \mathbb{R}$, $i = 1, 2, 3$.

1. (Associative Property)

$$\begin{aligned} z_2 z_3 &= (\alpha_2\alpha_3 - \beta_2\beta_3, \alpha_2\beta_3 + \alpha_3\beta_2) \\ z_1(z_2 z_3) &= (\alpha_1(\alpha_2\alpha_3 - \beta_2\beta_3) - \beta_1(\alpha_2\beta_3 + \alpha_3\beta_2), \\ &\quad \alpha_1(\alpha_2\beta_3 + \alpha_3\beta_2) + \beta_1(\alpha_2\alpha_3 - \beta_2\beta_3)) \\ &= (\alpha_1\alpha_2\alpha_3 - \alpha_1\beta_2\beta_3 - \beta_1\alpha_2\beta_3 - \beta_1\alpha_3\beta_2, \\ &\quad \alpha_1\alpha_2\beta_3 + \alpha_1\alpha_3\beta_2 + \beta_1\alpha_2\alpha_3 - \beta_1\beta_2\beta_3) \\ &= (z_1 z_2) z_3 \end{aligned}$$

2. (Right Distributive Property)

$$\begin{aligned} z_1(z_2 + z_3) &= (\alpha_1, \beta_1)(\alpha_2 + \alpha_3, \beta_2 + \beta_3) \\ &= (\alpha_1(\alpha_2 + \alpha_3) - \beta_1(\beta_2 + \beta_3), \beta_1(\alpha_2 + \alpha_3) + \alpha_1(\beta_2 + \beta_3)) \\ &= (\alpha_1\alpha_2 + \alpha_1\alpha_3 - \beta_1\beta_2 - \beta_1\beta_3, \beta_1\alpha_2 + \beta_1\alpha_3 + \alpha_1\beta_2 + \alpha_1\beta_3) \\ &= z_1 z_2 + z_1 z_3 \end{aligned}$$

3. (Left Distributive Property, similar to Right Distributive)

4. (Associative Property of Scalar Multiplication)

$$\begin{aligned} \lambda(z_1 z_2) &= \lambda(\alpha_1\alpha_2 - \beta_1\beta_2, \alpha_1\beta_2 + \beta_1\alpha_2) \\ &= (\lambda\alpha_1\alpha_2 - \lambda\beta_1\beta_2, \lambda\alpha_1\beta_2 + \lambda\beta_1\alpha_2) \\ &= \lambda(z_1 z_2) \end{aligned}$$

□

However, the algebra we have just considered has some extra properties that make it a field. In fact, it is commutative, and every nonzero complex number $z = (\alpha, \beta)$ has a multiplicative inverse $z^{-1} = (\frac{\alpha}{\alpha^2+\beta^2}, \frac{-\beta}{\alpha^2+\beta^2})$.

Finally, let's define the concept of the norm of a complex number:

$$|z| = \sqrt{zz^*} = \sqrt{\alpha^2 + \beta^2} \tag{2.5}$$

This coincides with the 2-dimensional Euclidean norm. This norm allows us to induce a metric:

$$\delta(z_1, z_2) = |z_1 - z_2| = \sqrt{(\alpha_1 - \alpha_2)^2 - (\beta_1 - \beta_2)^2} \tag{2.6}$$

Thus, it provides complex numbers with a topological structure of a metric space, which we will use to study classical complex analysis.

Now, we need to repeat a somewhat similar procedure to construct the skew-field of quaternions and endow it with a metric that will allow us to give it a topological structure for the study of analysis on quaternions.

2.1.2 Cayley-Dickson Construction of Quaternions

We now define quaternions $\mathbb{H} = \{(u, v)\ ;\ u, v \in \mathbb{C}\}$ as a pair of complex numbers. We define what we will call the canonical equivalence relation on quaternions.

Definition 2.2. *Two quaternions $q_1, q_2 \in \mathbb{H}$; $q_1 = (\alpha_1, \beta_1), q_2 = (\alpha_2, \beta_2)$ with $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{C}$ are said to be equal if and only if $\alpha_1 =_{\mathbb{C}} \alpha_2$ and $\beta_1 =_{\mathbb{C}} \beta_2$, where here $=_{\mathbb{C}}$ denotes the canonical equivalence relation defined earlier on complex numbers. If two quaternions are equal, we write: $q_1 =_{\mathbb{H}} q_2$.*

The verification that this relation is an equivalence relation on the set of quaternions is straightforward, and we leave it to the reader.

For simplicity, we will write this newly introduced equivalence relation as "=" from now on. Let $\lambda \in \mathbb{R}$ and $q_1, q_2 \in \mathbb{H}$, with $q_1 = (u_1, v_1)$ and $q_2 = (u_2, v_2)$, where $u_1, u_2, v_1, v_2 \in \mathbb{C}$. We will define component-wise addition and scalar multiplication as follows:

$$q_1 + q_2 = (u_1 + u_2, v_1 + v_2) \tag{2.7}$$

$$\lambda q_1 = (\lambda u_1, \lambda v_1) \tag{2.8}$$

Under these operations, $\mathbb{H}$ forms a 4-dimensional real vector space. To make it an associative $\mathbb{R}$-Algebra, we need to introduce multiplication and an operation called quaternionic conjugation:

$$(u_1, v_1) * (u_2, v_2) = (u_1 u_2 - v_2 v_1^*, u_1^* v_2 + u_2 v_1) \tag{2.9}$$

$$(u, v)^* = (u^*, -v) \tag{2.10}$$

Proposition 2.2. *$\mathbb{H}$, equipped with the operations of component-wise addition, scalar multiplication, and multiplication $*$ defined above, forms an associative $\mathbb{R}$-Algebra. In other words, multiplication $*$ satisfies the following axioms:*

1. $q_1(q_2 q_3) = (q_1 q_2) q_3$ *(Associative).*
2. $q_1(q_2 + q_3) = q_1 q_2 + q_1 q_3$ *(Right Distributive).*
3. $(q_1 + q_2) q_3 = q_1 q_3 + q_2 q_3$ *(Left Distributive).*
4. $\lambda(q_1 q_2) = (\lambda q_1)(q_2) = (q_1)(\lambda q_2)$ *(Associative Property of Scalar Multiplication).*

Proof. Siano $q_1 = (\alpha_1, \beta_1)$, $q_2 = (\alpha_2, \beta_2)$, $q_3 = (\alpha_3, \beta_3)$, con $\alpha_i, \beta_i \in \mathbb{C}$; $i = 1, 2, 3$.

1. $q_1q_2 = (\alpha_1\alpha_2 - \beta_2\beta_1^*, \alpha_1^*\beta_2 + \alpha_2\beta_1) \implies$
   $(q_1q_2)q_3 = ((\alpha_1\alpha_2-\beta_2\beta_1^*)\alpha_3-(\alpha_1^*\beta_2+\alpha_2\beta_1)^*\beta_3, (\alpha_1\alpha_2-\beta_2\beta_1^*)^*\beta_3+\alpha_3(\alpha_1^*\beta_2+\alpha_2\beta_1)) = (\alpha_1\alpha_2\alpha_3 - \beta_2\beta_1^*\alpha_3 - \alpha_1\beta_2^*\beta_3 - \alpha_2^*\beta_1^*\beta_3, \alpha_1^*\alpha_2^*\beta_3 - \beta_2^*\beta_1\beta_3 + \alpha_3\alpha_1^*\beta_2 + \alpha_3\alpha_2\beta_1)$
   $q_2q_3 = (\alpha_2\alpha_3 - \beta_3\beta_2^*, \alpha_2^*\beta_3 + \beta_2\alpha_3) \implies$
   $q_1(q_2q_3) = (\alpha_1(\alpha_2\alpha_3-\beta_3\beta_2^*)-\beta_1^*(\alpha_2^*\beta_3+\beta_2\alpha_3), \alpha_1^*(\alpha_2^*\beta_3+\beta_2\alpha_3)+\beta_1(\alpha_2\alpha_3 - \beta_3\beta_2^*)) = (\alpha_1\alpha_2\alpha_3 - \alpha_1\beta_3\beta_2^* - \beta_1^*\alpha_2^*\beta_3 - \beta_1^*\beta_2\alpha_3, \alpha_1^*\alpha_2^*\beta_3 + \alpha_1^*\beta_2\alpha_3 + \beta_1\alpha_2\alpha_3 - \beta_1\beta_3\beta_2^*) \implies$
   $(q_1q_2)q_3 = q_1(q_2q_3)$

2. $q_2 + q_3 = (\alpha_2 + \alpha_3, \beta_2 + \beta_3) \implies$
   $q_1(q_2 + q_3) = (\alpha_1(\alpha_2 + \alpha_3) - \beta_1^*(\beta_2 + \beta_3), \alpha_1^*(\beta_2 + \beta_3) + \beta_1(\alpha_2 + \alpha_3)) = (\alpha_1\alpha_2 + \alpha_1\alpha_3 - \beta_1^*\beta_2 - \beta_1^*\beta_3, \alpha_1^*\beta_2 + \alpha_1^*\beta_3 + \beta_1\alpha_2 + \beta_1\alpha_3)$
   $q_1q_2 = (\alpha_1\alpha_2 - \beta_1^*\beta_2, \alpha_1^*\beta_2 + \beta_1\alpha_2) \wedge q_1q_3 = (\alpha_1\alpha_3 - \beta_1^*\beta_3, \alpha_1^*\beta_3 + \beta_1\alpha_3) \implies$
   $q_1q_2 + q_1q_3 = (\alpha_1\alpha_2 - \beta_1^*\beta_2 + \alpha_1\alpha_3 - \beta_1^*\beta_3, \alpha_1^*\beta_2 + \beta_1\alpha_2 + \alpha_1^*\beta_3 + \beta_1\alpha_3) = q_1(q_2 + q_3)$

3. The proof is completely analogous to the one before.

4. $\lambda(q_1q_2) = \lambda(\alpha_1\alpha_2 - \beta_2\beta_1^*, \alpha_1^*\beta_2 + \alpha_2\beta_1) = (\lambda\alpha_1\alpha_2 - \lambda\beta_2\beta_1^*, \lambda\alpha_1^*\beta_2 + \lambda\alpha_2\beta_1)$
   $(\lambda q_1)(q_2) = (\lambda\alpha_1, \lambda\beta_1)(\alpha_2, \beta_2) = ((\lambda\alpha_1)\alpha_2-(\lambda\beta_1)^*\beta_2, (\lambda\alpha_1)^*\beta_2+\alpha_2(\lambda\beta_1)) = (\lambda\alpha_1\alpha_2 - \lambda\beta_2\beta_1^*, \lambda\alpha_1^*\beta_2 + \lambda\alpha_2\beta_1) = \lambda(q_1q_2)$
   $q_1(\lambda q_2) = (\alpha_1, \beta_1)(\lambda\alpha_2, \lambda\beta_2) = (\alpha_1(\lambda\alpha_2) - \beta_1^*(\lambda\beta_2), \alpha_1^*(\lambda\beta_2) + (\lambda\alpha_2)\beta_1) = (\lambda\alpha_1\alpha_2 - \lambda\beta_2\beta_1^*, \lambda\alpha_1^*\beta_2 + \lambda\alpha_2\beta_1) \implies$
   $\lambda(q_1q_2) = (\lambda q_1)(q_2) = q_1(\lambda q_2)$

□

Quaternions also have a multiplicative identity, denoted as $1_{\mathbb{H}} = (1, 0)$. For simplicity, we will refer to it simply as 1. Quaternions form a 4-dimensional associative unitary $\mathbb{R}$-Algebra. Since quaternions are defined as a pair of complex numbers, and complex numbers are, in turn, a pair of real numbers, we can represent a quaternion as a quadruple of real numbers: $q = (a, b, c, d)$, where $a, b, c, d \in \mathbb{R}$. From now on, we will use the following notation: we represent a quaternion $q = (a, b, c, d) \in \mathbb{H}$ as $q = a+bi+cj+dk$, where $a, b, c, d \in \mathbb{R}$, and where $i = (i, 0)$, $j = (0, 1)$, and $k = (0, i)$.

Theorem 2.1 (Hamilton's Rule)**.** *Let i, j, and k be defined as above. Then the following equation holds:*

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

Proof. $i^2 = (i, 0)(i, 0) = (-1 - 0, (-i)(0) + (0)(i)) = (-1, 0) = -1$
$j^2 = (0, 1)(0, 1) = ((0)(0) - (1)^*(1), 0 + 0) = (-1, 0) = -1$

$k^2 = (0,i)(0,i) = (0-(-i)(i),0) = (-1,0) = -1$
$ijk = (i,0)(0,1)(0,i) = (0,-i)(0,i) = (-1,0) = -1$ □

Corollary 2.1. *Let i, j, and k be defined as above. Then the following equations hold:*

- $ij = k = -ji$
- $jk = i = -kj$
- $ki = j = -ik$

The corollary follows immediately from the theorem with a simple algebraic manipulation of the equations.

We can summarize the results obtained with the following diagram:

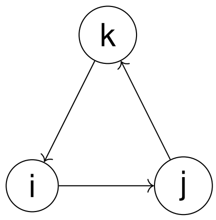


Figure 2.1: Multiplicative graph of imaginary units of quaternions.

The key to interpreting the graph is as follows: we move between two units (represented as vertices) by following the arrows that connect them. If we follow the direction of the arrow, we obtain the next element; if we move in the opposite direction to the arrow, we obtain the negation of the next element (i.e. with the opposite sign).

2.1.3 Basic Algebraic Operations of Quaternions in Terms of Their 4 Components

Once quaternions are expressed in this "operational" form, which is certainly more convenient than the algebraic construction seen earlier, we derive formulas for various algebraic operations of quaternions in terms of their 4 components a, b, c, d. Let's start with binary operations.

Let $q_1 = a_1 + b_1 i + c_1 j + d_1 k$ and $q_2 = a_2 + b_2 i + c_2 j + d_2 k$ be two quaternions, where $a_i, b_i, c_i, d_i \in \mathbb{R}$ for $i = 1, 2$.

Addition is straightforward:

- $q_1 + q_2 = (a_1 + a_2) + (b_1 + b_2)i + (c_1 + c_2)j + (d_1 + d_2)k$

Subtraction between q_1 and q_2 is simply the addition of the additive inverse of q_2 to q_1, which exists due to the underlying structure of vector space:

- $q_1 - q_2 = (a_1 - a_2) + (b_1 - b_2)i + (c_1 - c_2)j + (d_1 - d_2)k$

Multiplication, on the other hand:

- $q_1q_2 = (a_1 + b_1i + c_1j + d_1k)(a_2 + b_2i + c_2j + d_2k) = a_1a_2 + a_1b_2i + a_1c_2j + a_1d_2k + b_1a_2i + b_1b_2i^2 + b_1c_2ij + b_1d_2ik + c_1a_2j + c_1b_2ji + c_1c_2j^2 + c_1d_2jk + d_1a_2k + d_1b_2ki + d_1c_2kj + d_1d_2k^2$

Simplifying, it becomes:

- $q_1q_2 = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_1b_2 + b_1a_2 - d_1c_2 + c_1d_2)i + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k$

For the unary operation of quaternionic conjugation:

- $q_1^* = a_1 - b_1i - c_1j - d_1k$

Exercise 2.1. *Let q_1, q_2 be two quaternions. Calculate their commutator $[q_1, q_2] = q_1q_2 - q_2q_1$.*

Exercise 2.2. *Let $q_1 = 2j - 3k$, $q_2 = 3 - i - j + k$, and $q_3 = i$ be quaternions; calculate:*

a $q_1 - 5q_2$

b q_1q_2

c q_2q_1

d $q_1q_3 - q_2$

e $(q_3q_2)^*$

f $q_2^*q_3^*$

The last two questions in the previous exercise should have noticed something: they are equal. Is this true in general? The answer is affirmative. Motivated by this observation, let's now examine some properties of quaternionic conjugation.

Proposition 2.3. *Let q_1 and q_2 be two quaternions, and λ be a real number. Then:*

- $(q_1 \pm q_2)^* = q_1^* \pm q_2^*$ *(additive linearity).*
- $(q_1^*)^* = q_1$ *(* is an involution).*
- $(q_1q_2)^* = q_2^*q_1^*$.
- $(\lambda q_1)^* = \lambda q_1^*$ *(linearity with respect to scalar multiplication).*
- $q_1^*q_1 = q_1q_1^* = a_1^2 + b_1^2 + c_1^2 + d_1^2 \in \mathbb{R}^+$, *different from zero for every* $q_1 \neq 0$.
- $q_1 + q_1^* = 2a_1 \in \mathbb{R}$.

Proof.
- $q_1 \pm q_2 = (a_1 \pm a_2) + (b_1 \pm b_2)i + (c_1 \pm c_2)j + (d_1 \pm d_2)k \Longrightarrow (q_1 \pm q_2)^* = (a_1 \pm a_2) - (b_1 \pm b_2)i - (c_1 \pm c_2)j - (d_1 \pm d_2)k = (a_1 \pm a_2) + (-b_1 \mp b_2)i + (-c_1 \mp c_2)j + (-d_1 \mp d_2)k = q_1^* \pm q_2^*$.
- $q_1^* = a_1 - b_1 i - c_1 j - d_1 k \Longrightarrow (q^*)^* = a_1 + b_1 i + c_1 j + d_1 k$.
- $(q_1q_2)^* = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) - (a_1b_2 + b_1a_2 - d_1c_2 + c_1d_2)i - (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j - (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k$
  $q_2^*q_1^* = (a_2 - b_2 i - c_2 j - d_2 k)(a_1 - b_1 i - c_1 j - d_1 k) = a_2a_1 - a_2b_1 i - a_2c_1 j - a_2d_1 k - b_2a_1 i - b_1b_2 + b_2c_1 ij + b_2d_1 ik - c_2a_1 j + c_2b_1 ji - c_2c_1 + c_2d_1 jk - d_2a_1 k + d_2b_1 ki + d_2c_1 kj - d_2d_1 = a_2a_1 - a_2b_1 i - a_2c_1 j - a_2d_1 k - b_2a_1 i - b_1b_2 + b_2c_1 k - b_2d_1 j - c_2a_1 j - c_2b_1 k - c_2c_1 + c_2d_1 i - d_2a_1 k + d_2b_1 j - d_2c_1 i - d_2d_1 = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) - (a_1b_2 + b_1a_2 - d_1c_2 + c_1d_2)i - (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j - (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k = (q_1q_2)^*$.
- $(\lambda q_1)^* = (\lambda a_1 + \lambda b_1 i + \lambda c_1 j + \lambda d_1 k)^* = \lambda a_1 - \lambda b_1 i - \lambda c_1 j - \lambda d_1 k = \lambda q_1^*$.
- To simplify the notation, we will denote q_1 simply as $q = a + bi + cj + dk$. $qq^* = (a + bi + cj + dk)(a - bi - cj - dk) = a^2 - abi - acj - adk + abi + b^2 - bcij - bdik + caj - cbji + c^2 - cdjk + dak - dbki - dckj + d^2 = a^2 + b^2 + c^2 + d^2 - bck + bdj + cbk - cdi - dbj + dci = a^2 + b^2 + c^2 + d^2$, which is a non-negative real number.
  $q^*q = (a - bi - cj - dk)(a + bi + cj + dk) = a^2 + abi + acj + adk - abi + b^2 - bcij - bdik - caj - cbji + c^2 - cdjk - dak - dbki - dckj + d^2 = a^2 + b^2 + c^2 + d^2 - bck + bdj + cbk - cdi - dbj + dci = a^2 + b^2 + c^2 + d^2 = qq^*$.
- $q + q^* = a + bi + cj + dk + a - bi - cj - dk = 2a$.

□

We notice that quaternionic conjugation is an anti-automorphism of rings, i.e., it is a bijective function that preserves addition but reverses the order of multiplication.

Definition 2.3. *We say that an algebra $\mathbb{A}$ over the real numbers equipped with an involutive linear mapping $^* : \mathbb{A} \longrightarrow \mathbb{A}$ such that for all $a, b \in \mathbb{A}$, $(ab)^* = b^*a^*$ is well-normed if, for all $a \in \mathbb{A}$ other than the additive identity:*

1. $aa^* = a^*a$ *is a real number* > 0.

2. $a + a^* \in \mathbb{R}$.

It immediately follows as a corollary of Proposition 2.3 that the algebra of quaternions is a well-normed algebra.

Exercise 2.3. *Prove the following facts:*

- *The associative unitary $\mathbb{R}$-algebra of complex numbers is well-normed.*
- *Complex conjugation is an automorphism of the field $\mathbb{C}$.*

Quaternionic conjugation allows us to define a concept of length and distance on quaternions (a norm and a metric induced by it), which coincides exactly with 4-dimensional Euclidean length and distance.

Definition 2.4 (Norm of a quaternion)**.** *Let $q = a+bi+cj+dk$ be a quaternion. We call $||q|| = \sqrt{qq^*} = \sqrt{q^*q} = \sqrt{a^2 + b^2 + c^2 + d^2}$ the norm of quaternion q.*

It is immediately verifiable that this forms a norm on the vector space under consideration.

Historical Note 2.1. *In his original works, Hamilton defined the norm of a quaternion $||q||$ with the alternative name of "tensor" of q, using the notation Tq [Note 2.1]. It is worth noting that here the word "tensor" has nothing to do with what we now understand as tensors today but simply indicates a non-negative real number, which Hamilton defined as a "number without a sign."*

Now, let's look at some properties of quaternionic norm.

Proposition 2.4. *For $q_1, q_2 \in \mathbb{H}$, the following hold:*

- $||q_1q_2|| = ||q_1||||q_2||$
- $||q_1^*|| = || - q_1|| = ||q_1||$

Proof. We will prove only the first property since the second one is trivially true because negating or taking the conjugate of a quaternion does not change the value of their square.

Let $\alpha = a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2$, $\beta = a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2$, $\gamma = a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2$, and $\delta = a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2$.

Then, $||q_1q_2|| = \sqrt{\alpha^2 + \beta^2 + \gamma^2 + \delta^2}$.

Now, observe that:

$$
\begin{aligned}
\alpha^2 &= a_1^2a_2^2 + b_1^2b_2^2 + c_1^2c_2^2 + d_1^2d_2^2 - 2a_1a_2b_1b_2 - 2a_1a_2c_1c_2 \\
&\quad - 2a_1a_2d_1d_2 + 2b_1b_2c_1c_2 + 2b_1b_2d_1d_2 + 2c_1c_2d_1d_2 \\
\beta^2 &= -2a_2b_1c_2d_1 + 2a_2b_1c_1d_2 - 2a_1b_2c_2d_1 + 2a_1b_2c_1d_2 + a_2^2b_1^2 \\
&\quad + 2a_1a_2b_2b_1 + a_1^2b_2^2 + c_2^2d_1^2 + c_1^2d_2^2 - 2c_1c_2d_1d_2 \\
\gamma^2 &= 2a_2b_2c_1d_1 - 2a_2b_1c_1d_2 + 2a_1b_2c_2d_1 - 2a_1b_1c_2d_2 + a_2^2c_1^2 \\
&\quad + 2a_1a_2c_2c_1 + a_1^2c_2^2 + b_2^2d_1^2 + b_1^2d_2^2 - 2b_1b_2d_1d_2 \\
\delta^2 &= -2a_2b_2c_1d_1 - 2a_1b_2c_1d_2 + 2a_2b_1c_2d_1 + 2a_1b_1c_2d_2 + a_2^2d_1^2 \\
&\quad + a_1^2d_2^2 + 2a_1a_2d_1d_2 + b_2^2c_1^2 - 2b_1b_2c_2c_1 + b_1^2c_2^2.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\alpha^2 + \beta^2 + \gamma^2 + \delta^2 \\
&\quad = a_1^2a_2^2 + b_1^2b_2^2 + c_1^2c_2^2 + d_1^2d_2^2 + a_1^2b_2^2 + b_1^2a_2^2 + c_1^2d_2^2 + d_1^2c_2^2 \\
&\quad\quad + a_1^2c_2^2 + b_1^2d_2^2 + c_1^2a_2^2 + d_1^2b_2^2 + a_1^2d_2^2 + b_1^2c_2^2 + c_1^2b_2^2 + d_1^2a_2^2 \\
&\quad = (a_1^2 + b_1^2 + c_1^2 + d_1^2)(a_2^2 + b_2^2 + c_2^2 + d_2^2).
\end{aligned}
$$

Hence, $||q_1q_2|| = \sqrt{(a_1^2 + b_1^2 + c_1^2 + d_1^2)(a_2^2 + b_2^2 + c_2^2 + d_2^2)} = ||q_1||||q_2||$. □

Historical Note 2.2. *The first part of the assertion in Proposition 2.4 has an interesting connection to geometry and number theory and was discovered nearly 100 years before the discovery of quaternions by Euler. It is known as Euler's "Four-Square Identity," which Euler found while investigating the properties of natural numbers expressible as the sum of 4 squares (of natural numbers). Euler was interested in this class of numbers because he wanted to prove that every natural number could be written as the sum of the squares of 4 natural numbers; this result was later proven by Lagrange [Note 2.2]. The result in question, now known as the "Four-Squares Theorem," was already conjectured in Diofantus's important work "Arithmetica" (unfortunately not fully preserved) [Note 2.3].*

Proposition 2.5 (Inverse)**.** *For $q \in \mathbb{H}$, different from 0, there exists a quaternion $q^{-1} = \frac{q^*}{||q||^2}$ such that $qq^{-1} = q^{-1}q = 1$.*

Exercise 2.4. *Prove Proposition 2.5.*

This last property of quaternions confirms that they form a skew-field, which is a field except for the commutative multiplication axiom.

Definition 2.5. *For $q_1, q_2 \in \mathbb{H}$, we call $q_1 q_2^{-1}$ the right quotient and $q_2^{-1} q_1$ the left quotient.*

Returning to the norm we have just defined on quaternions, it allows us to define a metric, i.e., a concept of distance between two quaternions.

$$\delta(q_1, q_2) = \|q_1 - q_2\| = \sqrt{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2 + (d_1 - d_2)^2} \quad (2.11)$$

Proposition 2.6. *The function $\delta : \mathbb{H} \times \mathbb{H} \longrightarrow \mathbb{R}$ just defined forms a metric on $\mathbb{H}$, i.e., it satisfies the following axioms for all $q_1, q_2, q_3 \in \mathbb{H}$:*

1. $\delta(q_1, q_2) \geq 0$.
2. $\delta(q_1, q_2) = 0 \Longleftrightarrow q_1 = q_2$.
3. $\delta(q_1, q_2) = \delta(q_2, q_1)$ *(Symmetry).*
4. $\delta(q_1, q_2) \leq \delta(q_1, q_3) + \delta(q_3, q_2)$ *(Triangle Inequality).*

Proof. The first two properties follow as an immediate corollary of the fact that $\mathbb{H}$ is a well-normed algebra. The third property is also immediate. The triangle inequality follows from the Cauchy-Schwarz inequality as follows: Let r, s be two quaternions, expandable into components as $r = r_1 + r_2 i + r_3 j + r_4 k$ and $s = s_1 + s_2 i + s_3 j + s_4 k$, such that $\|r + s\| \neq 0$. Observe that:

$$\begin{aligned}
\|r + s\|^2 &= |r_1 + s_1|^2 + |r_2 + s_2|^2 + |r_3 + s_3|^2 + |r_4 + s_4|^2 \\
&= \sum_{i=1}^{4} |r_i + s_i|^2,
\end{aligned}$$

but by the properties of the canonical norm on real numbers $|\ | : \mathbb{R} \to \mathbb{R}^+$, we know that $|r_i + s_i| \leq |r_i| + |s_i|$ for all $i \in \{1, 2, 3, 4\}$. Therefore,

$$\begin{aligned}
\|r + s\|^2 &\leq \sum_{i=1}^{4} |r_i + s_i|(|r_i| + |s_i|) \\
&= \sum_{i=1}^{4} |r_i + s_i||r_i| + \sum_{i=1}^{4} |r_i + s_i||s_i|,
\end{aligned}$$

and by the Cauchy-Schwarz inequality:

$$||r+s||^2 \leq ||r+s|| \cdot ||r|| \quad \text{and}$$
$$||r+s||^2 \leq ||r+s|| \cdot ||s||,$$

which implies:

$$||r+s|| \leq ||r|| + ||s||.$$

To obtain part d of the proposition, set $r = q_1 - q_3$ and $s = q_3 - q_2$. □

Historical Note 2.3. *A quaternion with a length equal to 1, i.e., a quaternion $q \in \mathbb{H}$ with $|q| = 1$, was named by Hamilton as a "versor" [Note 2.4]. I am sure that this term is not new to the reader; you may have encountered it in an introductory course on linear algebra. Well, the modern usage of this term in the context of vector algebra is precisely attributed to Hamilton and originated during the writing of his texts about the emerging quaternionic algebra.*

2.2 Vector Part and Scalar Part of a Quaternion

Considering the substructure of a vector space of quaternions analyzed so far, we notice that it is a direct sum of two vector spaces:

$$\mathbb{H} \cong \mathbb{R} \oplus \mathrm{P} \tag{2.12}$$

Where P is a real 3-dimensional Euclidean vector space. Choosing the positively oriented orthonormal basis $\{i, j, k\}$, we can write an element v $\in$ P as v $= xi + yj + zk$, where $x, y, z \in \mathbb{R}$.

What we observed previously suggests us to view a quaternion as an object composed of a real number and a vector in $\mathbb{R}^3$; in other words, it consists of what we will call a "scalar part" and a "vector part."

Historical Note 2.4. *As mentioned earlier, in Hamilton's works, we can find the first documented instance of the use of the words "scalar" and "vector" in the English language in academic mathematical literature [Note 2.5]. Nowadays, we use the word "vector" to refer to an element of a vector space. On the other hand, when Hamilton used this word, he had a strictly geometric idea in mind; a vector was simply understood as a triplet, visually represented as an "arrow" in space. In particular, he noticed that he could write a quaternion q as $q = Sq + Vq$, where S and V respectively denote its scalar and vector parts. He also noted that a quaternion could be written as a product, $q = TqUq$, where Tq denotes the tensor of q and Uq its versor (its normalized version).*

Definition 2.6. *Let $q = a + bi + cj + dk$ be a quaternion. Then, we will say that $\mathrm{Sc}(q) = \Re(q) = a = \frac{1}{2}(q + q^*)$ is its scalar (or real) part, while we will say that $\mathrm{Vec}(q) = \Im(q) = bi + cj + dk = \frac{1}{2}(q - q^*)$ is its vector (or imaginary) part.*

Motivated by physics, we can interpret the scalar part of a quaternion as its "temporal" part, and the imaginary part as its "spatial" part, writing:

$$q = t + xi + yj + zk \tag{2.13}$$

with $t, x, y, z \in \mathbb{R}$.

We have seen in the previous section how laborious the algebraic operations with quaternions can be. Sometimes, we will need to write quaternions and their algebraic operations in a more compact and concise form. This need motivates us to use the Einstein summation convention for the "spatial" dimension of quaternions, which is widely used in tensor analysis for a similar reason (tensor expressions quickly become full of summations, and this convention allows us to omit them, only mentioning indices repeated twice, which we will call "dummy indices"). We can then write a quaternion $q \in \mathbb{H}$ as:

$$q = t + xi + yj + zk = t + \sum_{i=1}^{3} x_i e_i = t + x_i e_i \tag{2.14}$$

where $x_1 = x, x_2 = y, x_3 = k$, and $e_1 = i, e_2 = j, e_3 = k$.

This convention simplifies many equations. We can write the quaternion conjugate of q as:

$$q^* = t - x_i e_i \tag{2.15}$$

The sum (or subtraction) as:

$$q_1 \pm q_2 = t_1 \pm t_2 + (x_{1i} \pm x_{2i})e_i \tag{2.16}$$

Hamilton's rule becomes:

$$e_i e_j = -\delta_{ij} + \epsilon_{ijk} e_k \tag{2.17}$$

where δ_{ij} is the Kronecker delta and ϵ_{ijk} is the Levi-Civita symbol.

Finally, the product:

$$q_1 q_2 = t_1 t_2 + t_1 x_{2j} e_j + t_2 x_{1i} e_i + x_{1i} x_{2j} \epsilon_{ijk} e_k - x_{1i} x_{2j} \delta_{ij} \tag{2.18}$$

where here, as before, $q_1 = t_1 + x_{1i} e_i = t_1 + x_{11} i + x_{12} j + x_{13} k$, and $q_2 = t_2 + x_{21} i + x_{22} j + x_{23} k$.

2.2.1 Vector Operations with Quaternions

Let us now define, in quaternionic terms, some important operations that can be performed with vectors:

Definition 2.7 (Scalar Product of Quaternions)**.** *We define the scalar product of 2 quaternions* $q_1, q_2 \in \mathbb{H}$ *as:* $\langle q_1|q_2\rangle = a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 = \mathrm{Sc}(q_1)\,\mathrm{Sc}(q_2) + \langle \mathrm{Vec}(q_1)|\,\mathrm{Vec}(q_2)\rangle$, *where in the second expression* $\langle\cdot|\cdot\rangle$ *denotes the canonical scalar product on* $\mathbb{R}^3$*.*

Exercise 2.5. *Prove that* $\langle q_1|q_2\rangle = \mathrm{Sc}(q_1q_2^*) = \mathrm{Sc}(q_2q_1^*)$.

Using the Einstein convention, we can write the scalar product as:

$$\langle q_1|q_2\rangle = t_1t_2 + x_{1i}x_{2i} \tag{2.19}$$

Definition 2.8 (Cross Product of Two Pure Quaternions)**.** *Let* q_1, q_2 *be two pure quaternions, that is two quaternions whose real part is equal to 0, of the form* $q_1 = x_1i + y_1j + z_1k, q_2 = x_2i + y_2j + z_2k$*. We define the cross product as:*

$$q_1 \times q_2 = \mathrm{Vec}(q_1q_2) = (y_1z_2 - y_2z_1)i + (x_2z_1 - x_1z_2)j + (x_1y_2 - x_2y_1)k \tag{2.20}$$

Sometimes, by abusing notation, we recall the formula of a cross product of two pure quaternions q_1 and q_2 with the determinant of the following formal matrix:

$$q_1 \times q_2 = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} \tag{2.21}$$

Using the Einstein convention, we can write the cross product more compactly using the Levi-Civita symbol:

$$q_1 \times q_2 = \epsilon_{ijk}x_{1j}x_{2k}e_i \tag{2.22}$$

Where here $x_{11} = x_1, x_{12} = y_1, x_{13} = z_1, x_{21} = x_2, x_{22} = y_2, x_{23} = z_2$.
Geometrically, the cross product of 2 pure quaternions is another pure quaternion, with magnitude equal to the area of the parallelogram determined by the 2 quaternions in question, direction perpendicular to the plane determined by the pair of quaternions, and orientation either outward or inward from the plane of the 2 quaternions.

Proposition 2.7. *Let* $q_1, q_2, q_3 \in \mathrm{P}$ *be 3 pure quaternions, then:*

$$q_1 \times (q_2 \times q_3) = \langle q_1|q_3\rangle q_2 - \langle q_1|q_2\rangle q_3 \tag{2.23}$$

Proof. $q_1 \times (q_2 \times q_3) = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ y_2z_3 - z_2y_3 & z_2x_3 - x_2z_3 & x_2y_3 - y_2x_3 \end{vmatrix} =$
$(x_2y_3y_1 - y_2x_3y_1 - z_1z_2x_3 + z_1x_2z_3)i + (z_1y_2z_3 - z_2y_3z_1 - x_1x_2y_3 + x_1y_2x_3)j + (x_1z_2x_3 - x_1x_2z_3 - y_1y_2z_3 + z_2y_3y_1)k$
$\langle q_1|q_3\rangle q_2 = (x_1x_3 + y_1y_3 + z_1z_3)(x_2i + y_2j + z_2k) = x_1x_3x_2i + x_1x_3y_2j + x_1x_3z_2k + y_1y_3x_2i + y_1y_2y_3j + y_1y_3z_2k + z_1z_3x_2i + z_1z_3y_2j + z_1z_3z_2k$
$\langle q_1|q_2\rangle q_3 = (x_1x_2 + y_1y_2 + z_1z_2)(x_3i + y_3j + z_3k) = x_1x_2x_3i + x_1x_2y_3j + x_1x_2z_3k + y_1y_2x_3i + y_1y_2y_3j + y_1y_2z_3k + z_1z_2x_3i + z_1z_2y_3j + z_1z_2z_3k \implies$
$\langle q_1|q_3\rangle q_2 - \langle q_1|q_2\rangle q_3 = (x_2y_3y_1 - y_2x_3y_1 - z_1z_2x_3 + z_1x_2z_3)i + (z_1y_2z_3 - z_2y_3z_1 - x_1x_2y_3 + x_1y_2x_3)j + (x_1z_2x_3 - x_1x_2z_3 - y_1y_2z_3 + z_2y_3y_1)k = q_1 \times (q_2 \times q_3)$ □

Being completely analogous to that on $\mathbb{R}^3$, it has the same properties probably already known to the reader, which we list below:

Proposition 2.8. *Let $\times$ be the cross product of 2 pure quaternions just defined. Then it has the following properties:*

- $q_1 \times q_1 = 0$ *(Idempotence).*
- $q_1 \times q_2 = -q_2 \times q_1$ *(Anticommutativity).*
- $q_1 \times (q_2 + q_3) = q_1 \times q_2 + q_1 \times q_3$ *(Distributive).*
- $q_1 \times (q_2 \times q_3) + q_2 \times (q_3 \times q_1) + q_3 \times (q_1 \times q_2) = 0$ *(Jacobi Identity).*

Proof.
- $q \times q = (yz - yz)i + (xz - xz)j + (xy - xy)k = 0.$
- $q_1 \times q_2 = (y_1z_2 - y_2z_1)i + (x_2z_1 - x_1z_2)j + (x_1y_2 - x_2y_1)k.$
  $q_2 \times q_1 = (y_2z_1 - y_1z_2)i + (x_1z_2 - x_2z_1)j + (x_2y_1 - x_1y_2)k = -q_1 \times q_2.$
- $q_1 \times q_3 = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_3 & y_3 & z_3 \end{vmatrix} = (y_1z_3 - y_3z_1)i + (x_3z_1 - x_1z_3)j + (x_1y_3 - x_3y_1)k$
  $q_1 \times q_2 = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{vmatrix} = (y_1z_2 - y_2z_1)i + (x_2z_1 - x_1z_2)j + (x_1y_2 - x_2y_1)k$
  $q_1 \times (q_2 + q_3) = \begin{vmatrix} i & j & k \\ x_1 & y_1 & z_1 \\ x_2 + x_3 & y_2 + y_3 & z_2 + z_3 \end{vmatrix} = (y_1z_2 + y_1z_3 - z_1y_2 - z_1y_3)i + (z_1x_2 + z_1x_3 - x_1z_2 - x_1z_3)j + (x_1y_2 + x_1y_3 - y_1x_2 - y_1x_3)k = q_1 \times q_2 + q_1 \times q_3.$
- Using the result of proposition 2.7 3 times we get: $q_1 \times (q_2 \times q_3) + q_2 \times (q_3 \times q_1) + q_3 \times (q_1 \times q_2) = \langle q_1|q_3\rangle q_2 - \langle q_1|q_2\rangle q_3 + \langle q_2|q_1\rangle q_3 - \langle q_2|q_3\rangle q_1 + \langle q_3|q_2\rangle q_1 - \langle q_3|q_1\rangle q_2 = 0.$

□

Proposition 2.9. *Let $\times$ be the cross product of 2 pure quaternions just defined. Then:*

- $i \times j = k$.
- $j \times k = i$.
- $k \times i = j$.

Proof.

- $i \times j = \begin{vmatrix} i & j & k \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = (0(0) - 0(1))i - (1(0) - (0)(0)) + (1)k = k.$
- $j \times k = \begin{vmatrix} i & j & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = (1 - 0)i - (0(1) - 0(0))j + (0(0) - 1(0)k = i.$
- $k \times i = \begin{vmatrix} i & j & k \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix} = (0(0) - 1(0))i - (0(0) - 1)j + (0(0) - 1(0)k = j.$

□

Exercise 2.6. *Let $q_1 = j - k$, $q_2 = -i + k$, $q_3 = -2k$:*

1. *Calculate:* $q_1 \times q_2$
2. *Calculate* $q_3 \times q_2$
3. *Calculate* $\langle q_1 | q_3 \rangle$
4. *Verify the Jacobi identity for* q_1, q_2 *and* q_3

The algebraic notions just introduced also allow us to re-express the elementary algebraic operations of quaternions in terms of operations on their scalar and vector parts. Let q be a quaternion $q \in \mathbb{H}$; we can see it as a pair, this time composed of, as originally called by Hamilton, a scalar $\mathrm{Sc}(q)$, and a vector $\mathrm{Vec}(q) \in P$: $q = (\mathrm{Sc}(q), \mathrm{Vec}(q))$.
In this form, we can rewrite the multiplication of two quaternions $q_1 = (a_1, v_1)$, $q_2 = (a_2, v_2)$ as:

$$q_1 q_2 = (a_1 a_2 - \langle v_1 | v_2 \rangle, a_1 v_2 + a_2 v_1 + v_1 \times v_2) \tag{2.24}$$

and their sum, trivially, as:

$$q_1 + q_2 = (a_1 + a_2, v_1 + v_2) \tag{2.25}$$

Let's now define other operations on quaternions:

Definition 2.9. *Let q_1 and q_2 be two quaternions. We then call:*

$$(q_1, q_2)_\varepsilon = \frac{q_1^* q_2 + q_2^* q_1}{2} \tag{2.26}$$

the Euclidean outer product of 2 quaternions.

Proposition 2.10. *Let q_1 and q_2 be two quaternions, then:*

$$(q_1, q_2)_\varepsilon = \mathrm{Sc}(q_1)\,\mathrm{Vec}(q_2) - \mathrm{Sc}(q_2)\,\mathrm{Vec}(q_1) - \mathrm{Vec}(q_1) \times \mathrm{Vec}(q_2) \tag{2.27}$$

Proof. Direct calculation, left as an exercise to the reader. □

The last proposition makes us notice a fact: the Euclidean outer product of 2 quaternions lives in the space of pure quaternions P.

Definition 2.10. *Let q_1 and q_2 be two quaternions. We call:*

$$(q_1, q_2)_{even} = \frac{q_1 q_2 + q_2 q_1}{2} \tag{2.28}$$

the even product of two quaternions or, alternatively, the Grassmann inner product.

The just defined product can alternatively be defined as half of the anticommutator $\{q_1, q_2\}$ of two quaternions, i.e:

$$(q_1, q_2)_{even} = \frac{1}{2}\{q_1, q_2\}$$

2.3 Polar Form of a Quaternion

Similarly to the complex case, we can express quaternions in a polar form, that is, in a form that explicitly states their norm, their argument, and, for quaternions, also their sign (which we will define shortly). [Note: from now on, since the ambiguity present at the beginning of this chapter between the absolute value of $\mathbb{R}$ and the quaternionic norm $\|\ \|$ no longer exists, we will interchangeably use the symbols $\|\ \|$ and $|\ |$ to indicate the norm of a quaternion, preferring the latter to make our expressions more "clean"].
Let q be a quaternion with a non-zero vector part; $q = \mathrm{Sc}(q) + \mathrm{Vec}(q) = |q|\Big(\frac{\mathrm{Sc}(q)}{|q|} + \frac{|\mathrm{Vec}(q)|}{|q|}\frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|}\Big)$.
Since $|q| = |q| \cdot \Big|\frac{\mathrm{Sc}(q)}{|q|} + \frac{\mathrm{Vec}(q)}{|q|}\Big|$, it follows that $\Big|\frac{\mathrm{Sc}(q)}{|q|} + \frac{\mathrm{Vec}(q)}{|q|}\Big| = 1$, hence:

$$\left(\frac{\mathrm{Sc}(q)}{|q|}\right)^2 + \left(\frac{|\mathrm{Vec}(q)|}{|q|}\right)^2 = 1 \tag{2.29}$$

Moreover:

$$\frac{\mathrm{Sc}(q)}{|q|} \leq 1 \tag{2.30}$$

$$\frac{|\mathrm{Vec}(q)|}{|q|} \leq 1 \tag{2.31}$$

Thus, there exists an angle θ such that:

$$\cos(\theta) = \frac{\mathrm{Sc}(q)}{|q|}$$

and

$$\sin(\theta) = \frac{|\mathrm{Vec}(q)|}{|q|}$$

From this, it follows that we can write a quaternion with a non-zero vector part as:

$$q = |q|\left(\cos(\theta) + \frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|}\sin(\theta)\right) \tag{2.32}$$

We call this the **polar form of a quaternion**.
Let's now introduce some notation; let's call the angle θ the argument of q, written $\theta = \arg(q)$, let's also call $\frac{q}{|q|}$ the sign of q, or alternatively the unit vector of q, written as $\mathrm{sgn}(q) = Uq = \frac{q}{|q|}$.
We can therefore rewrite equation (2.32) as:

$$q = |q|(\cos(\arg(q)) + \mathrm{sgn}(\mathrm{Vec}(q))\sin(\arg(q))) \tag{2.33}$$

Similarly to what we did for complex numbers, we will call the restriction of the argument between 0 and π the principal argument of q, which we denote with the symbol $\mathrm{Arg}(q)$.

Exercise 2.7. *Find the argument and write the polar representation of the following quaternions:*

1. $1 + 4j - 2k$
2. $i - j$
3. $1 + i + j + k$
4. $-1 - i - j$

Remember that for complex numbers the following relation is valid:

$$zw = |z||w|(\cos(\arg(z) + \arg(w)) + i\sin(\arg(z) + \arg(w)))$$

i.e the product of complex numbers z and w has as its norm (geometrically the radial component) the product of the norms of z and w, and as argument the sum of the arguments. This is also valid for quaternions, with one exception: the quaternions we multiply must have proportional vector parts, i.e their vector parts must have the same direction. In the polar form of a quaternion, in fact, we will have that $\sin(\arg(q))$ will scale an arbitrary unit vector living on the sphere S^2 (the unit vector $\mathrm{sgn}(\mathrm{Vec}(q))$, it lives on the sphere $S^2 \subset \mathbf{P}$ as it is a normal vector of the three-dimensional vector space of pure quaternions), and therefore we must also worry about the direction of the latter. Let's now prove this fact: let $r = |r|(\cos\theta + \vec{u}\sin\theta)$ and $s = |s|(\cos\phi + \vec{u}\sin\phi)$ be two quaternions with the same sign (that is, two quaternions with proportional vector parts), then:

$$rs = |r||s|(\cos\theta + \vec{u}\sin\theta)(\cos\phi + \vec{u}\sin\phi) =$$

$$|r||s|(\cos\theta\cos\phi - \sin\theta\sin\phi + (\sin\theta\cos\phi + \sin\phi\cos\theta)\vec{u})$$

which, using the sum formulas for sine and cosine, becomes:

$$rs = |r||s|(\cos(\theta + \phi) + \vec{u}\sin(\theta + \phi))$$

As an immediate corollary of this relationship, provable by mathematical induction, there is a quaternionic analogue of de Moivre's theorem, which we state here:

Corollary 2.1 (De Moivre's Theorem)**.** *Let q be a quaternion with a non-null vector part and with $\arg(q) = \theta$, then, for $n \in \mathbb{Z}$ we will have:*

$$(\cos\theta + \mathrm{sgn}(\mathrm{Vec}(q))\sin\theta)^n = \cos(n\theta) + \mathrm{sgn}(\mathrm{Vec}(q))\sin(n\theta) \tag{2.34}$$

Having said this, we can now move on to a discussion of polynomials and polynomial equations in $\mathbb{H}$.

2.4 Quaternionic polynomials

The goal of this section is to define a concept similar to that of real/complex polynomial for the quaternionic skew-field.
The main problem we encounter for quaternions is the following: quaternions are not a commutative ring, hence we cannot construct the ring of polynomials with coefficients in the aforementioned set in the canonical way. Nevertheless, it will be beneficial to recall the construction of the aforementioned for the field of complex numbers, and then see how we can tackle the problem in the case of quaternions.

2.4.1 Construction of the Polynomial Ring $\mathbb{C}[x]$

As proved at the beginning of this chapter, $(\mathbb{C}, +, *)$ is a field.
We call $\mathbb{C}^\omega$ the set of all sequences with elements in $\mathbb{C}$. We define a canonical equivalence relation on it

Definition 2.11 (Canonical equivalence relation for sequences in $\mathbb{C}$)**.** *Two sequences $(z_i)_{i=0}^\infty$ and $(w_i)_{i=0}^\infty$ of complex numbers are said to be equal if and only if $z_i = w_i$ for all i.*

Finally, we define two binary operations: let $(z_i)_{i=0}^\infty$ and $(w_i)_{i=0}^\infty$ be two sequences, then:

$$(z_i)_{i=0}^\infty + (w_i)_{i=0}^\infty = (z_i + w_i)_{i=0}^\infty \tag{2.35}$$

Multiplication, on the other hand:

$$(z_i)_{i=0}^\infty (w_i)_{i=0}^\infty = \left(\sum_{i+j=k} z_i w_j \right)_{k=0}^\infty \tag{2.36}$$

$\mathbb{C}^\omega$ forms a commutative unitary ring with respect to these two binary operations. A proof of this fact is generally provided, for the more general case of a set of sequences with values in a commutative unitary ring, in courses on algebra and ring theory.

Definition 2.12. *Let $(z_i)_{i=0}^\infty$ be a sequence. We will say that it is almost everywhere zero if there exists a natural number n such that $z_i = 0$ for all $i \geq n$.*

We will call the set of almost everywhere zero sequences with values in the complex numbers complex-valued polynomials. Extending the operations defined above on the latter, we observe that it is closed with respect to the aforementioned. Thus, the polynomials with values in $\mathbb{C}$ are a sub-ring of the ring of sequences with values in $\mathbb{C}$. We can express elements of this sub-ring in the form most congenial to the reader as:

$$p(z) = \zeta_0 + \zeta_1 z + \cdots + \zeta_n z^n = \sum_{i=0}^{n} \zeta_i z^i$$

where here $z = (0, 1, 0, \dots)$ is the sequence of all zeros except for the position $i = 1$.

2.4.2 Construction of the Polynomial Ring R[x] for a Unitary Ring R

So far so good, at this point we might wonder what could be the problems encountered defining a similar structure on a non-commutative ring. Indeed, we will encounter no problem in defining an entirely analogous notion, but we will realize that such a structure will have some undesired implications.
Let R be a ring (not necessarily commutative) with a multiplicative unit 1_R. We call R^ω the set of sequences with values in R, indicated as $(a_i)_{i=0}^\infty$ with $a_i \in R$ for all $i \in \mathbb{N}$. We define a canonical equivalence relation on this set:

Definition 2.13 (Canonical equivalence relation for sequences in R)**.** *Two sequences $(a_i)_{i=0}^\infty$ and $(b_i)_{i=0}^\infty$ with values in R are said to be equal if and only if $a_i = b_i$ for all i.*

Let $(a_i)_{i=0}^\infty$, $(b_i)_{i=0}^\infty$ be two sequences in R^ω. We equip this set with 2 binary operations; an addition and a multiplication:

$$(a_i)_{i=0}^\infty + (b_i)_{i=0}^\infty = (a_i + b_i)_{i=0}^\infty \tag{2.37}$$

$$(a_i)_{i=0}^\infty (b_i)_{i=0}^\infty = \left(\sum_{i+j=k} a_i b_j \right)_{k=0}^\infty \tag{2.38}$$

Proposition 2.11. *The set R^ω forms a unitary ring under the operations just defined.*

Proof.
- R^ω is closed with respect to both operations. This follows from the closure of R, being a ring itself (i.e., $a + b \in R$, $ab \in R$ for all $a, b \in R$).
- The additive identity is $0_{R^\omega} = (0, 0, 0, \dots)$.

- The associativity and commutativity of addition follow from the same properties valid in the base ring R. Indeed, let $(a_i), (b_i), (c_i)$ be 3 sequences with values in R:
  $(a_i) + (b_i) = (a_i + b_i)$, but $a_i + b_i = b_i + a_i$ for all i as the addition in R is commutative, thus; $(a_i) + (b_i) = (a_i + b_i) = (b_i + a_i) = (b_i) + (a_i)$.
  Similarly, for the associative:
  $((a_i) + (b_i)) + (c_i) = ((a_i + b_i) + c_i)$, but $(a_i + b_i) + c_i = a_i + (b_i + c_i)$ for all i as the addition in R is associative. Thus:
  $((a_i) + (b_i)) + (c_i) = ((a_i + b_i) + c_i) = (a_i + (b_i + c_i)) = a_i + ((b_i) + (c_i))$.

- Let $a = (a_i)_{i=0}^{\infty}$ in R. Its additive inverse is $-a = (-a_i)$, defined thanks to the existence of additive inverses for each element of the sequence, guaranteed by the ring structure of R.

- The multiplicative identity of the ring is $1_R = (1, 0, 0, 0, \dots)$.

- Now the associativity of multiplication: $((a_i)(b_i))(c_i) = (\sum_{i+j=k} a_i b_j)(c_i) = (\sum_{r+s=l}(\sum_{i+j=r} a_i b_j)c_s)$, which by the distributive property of multiplication on the base ring, $(\sum_{r+s=l} \sum_{i+j=r} a_i b_j c_s)$ and
  $(a_i)((b_i)(c_i)) = (a_i)(\sum_{i+j=k} b_i c_j) = (\sum_{r+s=l} a_r(\sum_{i+j=s} b_i c_j)) =$
  $(\sum_{r+s=l} \sum_{i+j=s} a_r b_i c_j)$, which by reindexing we can show to be equal to $((a_i)(b_i))(c_i)$.

- Finally the distributive, on the right:
  $((a_i)+(b_i))(c_i) = (\sum_{i+j=k}(a_i+b_i)c_j) = (\sum_{i+j=k} a_i c_j + \sum_{i+j=k} b_i c_j) = (\sum_{i+j=k} a_i c_j) + (\sum_{i+j=k} b_i c_j) = (a_i)(c_i) + (b_i)(c_i)$.
  And on the left:
  $(c_i)((a_i)+(b_i)) = (\sum_{i+j=k} c_i(a_j+b_j)) = (\sum_{i+j=k} c_i a_j + \sum_{i+j=k} c_i b_j) = (\sum_{i+j=k} c_i a_j) + (\sum_{i+j=k} c_i b_j) = (c_i)(a_i) + (c_i)(b_i)$.

$\square$

We call the set of almost everywhere zero sequences with values in R the set of polynomials with values in R. Such a set, with the operations inherited above, forms a unitary subring of R^{ω}. Moreover, to simplify the notation, we will henceforth denote almost everywhere zero sequences of the type $(a, 0, 0, \dots)$, thus with only $a_0 \neq 0$ and all others equal to 0, as simply $a = (a, 0, 0, \dots)$.
We call X the sequence $(0, 1, 0, \dots)$. It is easily shown that for all $n \in \mathbb{N}$:

$$X^n = (\delta_{in})_{i=0}^{\infty} \tag{2.39}$$

Where here, as usual, δ_{in} indicates the Kronecker delta. We call this sequence an indeterminate over R. So far, the construction has been entirely

analogous to that for polynomials with coefficients in a field or in a commutative ring (minus the commutativity of polynomial multiplication, obviously).
However, it is at this point that we encounter the first difficulty: for all $n \in \mathbb{N}$, X^n commutes with R (where here by R we mean the set of sequences $(a, 0, 0, \dots)$ varying a, equipped with the operations defined previously, which is isomorphic to R). That is, let $a \in R$, for $n = 1$:

$$aX = (a, 0, 0, \dots)(0, 1, 0, \dots) = (0, a, 0, \dots) = (0, 1, 0, \dots)(a, 0, 0, \dots) = Xa$$

and for any n:

$$aX^n = (a, 0, \dots)(\delta_{in}) = (a\delta_{in}) = X^n a \tag{2.40}$$

From this, it follows that we can write a polynomial $\phi(X) \in R[X]$ as:

$$\phi(X) = \sum_{i=0}^{n} a_i X^i = \sum_{i=0}^{n} X^i a_i \tag{2.41}$$

This causes us quite a few problems; indeed, our interpretation that we would ideally have wanted to have of an indeterminate over R collapses. Formally, in fact, we would have wanted it to behave algebraically like a generic element of a non-commutative ring, and then define the classic notions of polynomial value, evaluation function, zeros, etc.
Taking an example in the quaternion skew-field; in $\mathbb{H}[X]$ such identity is true: $iX = Xi$, imagining wanting to substitute a quaternion for X, such as j, we immediately verify a discrepancy: $ji \neq ij$.

Proposition 2.12. *Let $a \in \mathbb{H}$. Then: $qa = aq$ for all $q \in \mathbb{H} \iff a \in \mathbb{R}$, or in other words; $Z(\mathbb{H}) = \mathbb{R}$, the center (Z from the German zentrum) of $\mathbb{H}$ is $\mathbb{R}$.*

Therefore it would have been acceptable to let the indeterminate commute with the coefficients only if the latter were real numbers. But this is very restrictive.
To adequately study polynomials with quaternionic coefficients, we must divide them into 3 categories:

1. Left polynomials: which we will formally indicate as $f(q) = \sum_{i=o}^{n} a_i q^i$ with $a_i \in \mathbb{H}$ for all i.

2. Right polynomials: which we will formally indicate as $f(q) = \sum_{i=o}^{n} q^i a_i$ with $a_i \in \mathbb{H}$ for all i.

3. General polynomials: the general form of a quaternion polynomial is quite difficult to provide. In this text, we will only be interested in a subclass of this type of polynomial, called **linear general polynomials**, of the form $f(q) = a_o q a_1 q \dots q a_n + \phi(q)$, where $\phi(q)$ is a finite sum of similar monomials $a_0 q a_1 q \dots q a_k$ with $k < n$.

We will use the just-constructed ring of polynomials for the development of the theory of right and left polynomials, defining them starting from elements of $\mathbb{H}[X]$. Let us first give the definition of **value of a quaternion polynomial**.

Definition 2.14 (Left value of a polynomial)**.** *Let $\phi(q) \in \mathbb{H}[X]$ be a quaternion coefficient polynomial, $\phi(q) = \sum_{i=0}^{n} a_i q^i$, and let $\alpha \in \mathbb{H}$ be a quaternion. We call the* ***left value*** *of ϕ at α, written $\phi^{\leftarrow}(\alpha)$, the quaternion:*

$$\phi^{\leftarrow}(\alpha) = \sum_{i=0}^{n} a_i \alpha^i \tag{2.42}$$

we also call the function $\mathrm{ev}_{\phi}^{\leftarrow} : \mathbb{H} \to \mathbb{H}$ defined as:

$$\mathrm{ev}_{\phi}^{\leftarrow}(\alpha) := \sum_{i=0}^{n} a_i \alpha^i$$

the ***left evaluation function*** *of ϕ.*

The definition of the right value of a polynomial is entirely analogous, we report it only for completeness even if it may seem redundant:

Definition 2.15 (Right value of a quaternion polynomial)**.** *Let $\phi(q) \in \mathbb{H}[X]$ be a quaternionic polynomial, $\phi(q) = \sum_{i=0}^{n} a_i q^i$, and let $\alpha \in \mathbb{H}$ be a quaternion. We call the* ***right value*** *of ϕ at α, written $\phi^{\rightarrow}(\alpha)$, the quaternion:*

$$\phi^{\rightarrow}(\alpha) = \sum_{i=0}^{n} \alpha^i a_i \tag{2.43}$$

we also call the function $\mathrm{ev}_{\phi}^{\rightarrow} : \mathbb{H} \to \mathbb{H}$ defined as:

$$\mathrm{ev}_{\phi}^{\rightarrow}(\alpha) := \sum_{i=0}^{n} \alpha^i a_i$$

the ***right evaluation function*** *of ϕ.*

Given these two definitions, we are now ready to define with greater rigor the concept of right and left polynomial on quaternions.

Definition 2.16 (Left quaternionic coefficient polynomial)**.** *We will call a* ***left quaternionic coefficient polynomial*** *the association of a left evaluation function with a quaternionic coefficient polynomial $\phi \in \mathbb{H}[X]$, which we will indicate with the following pair:*

$$(\phi(q), \mathrm{ev}_{\phi}^{\leftarrow})$$

Definition 2.17 (Right Quaternionic Coefficient Polynomial)**.** *We shall call a* ***right quaternionic coefficient polynomial*** *the association of a right evaluation function to a quaternionic coefficient polynomial $\phi \in \mathbb{H}[X]$, which we indicate with the following pair:*

$$(\phi(q), \mathrm{ev}_{\phi}^{\rightarrow})$$

From now on, to simplify notation and make it more suggestive, we will formally denote left quaternionic polynomials as:

$$\phi(q) = \sum_{i=0}^{n} a_i q^i$$

and right quaternionic polynomials as:

$$\phi(q) = \sum_{i=0} q^i a_i$$

and, having thus eliminated the ambiguity in how they are computed, we will refer to their evaluation functions simply as ev_ϕ.
We now extend the important concept of the **zero of a polynomial** to the newly defined left/right quaternionic coefficient polynomials.

Definition 2.18 (Zero/root of a left/right quaternionic polynomial)**.** *Let $\phi(q) \in \mathbb{H}[x]$ be a left/right quaternionic polynomial. Let $\zeta \in \mathbb{H}$, then we say that ζ is a zero/root of $\phi(q)$ if the value of ϕ at ζ is equal to 0, or in symbolic terms:*

$$\phi(\zeta) = 0 \tag{2.44}$$

In other words, the zeros of a polynomial are the elements of the kernel of its evaluation function, $\zeta \in \ker(\mathrm{ev}_\phi)$.

We now introduce the first important result of the theory of polynomials with coefficients in $\mathbb{H}$:

Theorem 2.2 (Fundamental Theorem of Algebra for a Left/Right Quaternionic Polynomial)**.** *Every left/right quaternionic polynomial of positive degree has a quaternionic root.*

Proof. Let f be a left quaternionic coefficient polynomial, let's write it by explicitly showing its components; $f(q) = a(q) + ib(q) + jc(q) + kd(q)$. Multiplying it by its quaternionic conjugate we obtain: $ff^* = a(q)^2 + b(q)^2 + c(q)^2 + d(q)^2$, a real coefficient polynomial. Let r be a complex root of ff^* and let: $a(r) = a_1 + ia_2$, $b(r) = b_1 + ib_2$, $c(r) = c_1 + ic_2$, $d(r) = d_1 + id_2$:
$ff^*(r) = 0 = (a_1 + ia_2)^2 + (b_1 + ib_2)^2 + (c_1 + ic_2)^2 + (d_1 + id_2)^2 = a_1^2 - a_2^2 + 2a_1a_2i + b_1^2 - b_2^2 + 2b_1b_2i + c_1^2 - c_2^2 + 2c_1c_2i + d_1^2 - d_2^2 + 2d_1d_2i = (a_1^2 + b_1^2 + c_1^2 + d_1^2 - a_2^2 - b_2^2 - c_2^2 - d_2^2) + 2(a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2)i$
From the definition of equality of two complex numbers it follows that:
$a_1a_2 + b_1b_2 + c_1c_2 + d_1d_2 = 0$ and
$a_1^2 + b_1^2 + c_1^2 + d_1^2 - a_2^2 - b_2^2 - c_2^2 - d_2^2 = 0$
We now want to find a non-zero quaternion $\gamma = \gamma_1 + \gamma_2 i + \gamma_3 j + \gamma_4 k$ such that $\Phi(q) = f(q)\gamma$ has r as a root. Calculating, $\Phi(q)$ becomes:
$\Phi(q) = \gamma_1 a(q) - \gamma_2 b(q) - \gamma_3 c(q) - \gamma_4 d(q) + i(\gamma_2 a(q) + \gamma_1 b(q) + \gamma_4 c(q) - \gamma_3 d(q)) + j(\gamma_3 a(q) + \gamma_1 c(q) + \gamma_2 d(q) - \gamma_4 b(q)) + k(\gamma_4 a(q)\gamma_1 d(q) + \gamma_3 b(q) - \gamma_2 c(q))$. For r to be a root of Φ, γ must be a solution of the following linear system:

$$\begin{bmatrix} a_1 - b_2 & -a_2 - b_1 & d_2 - c_1 & -c_2 - d_1 \\ a_2 + b_1 & a_1 - b_2 & -d_1 - c_2 & c_1 - d_2 \\ c_1 + d_2 & d_1 - c_2 & a_1 + b_2 & -b_1 + a_2 \\ c_2 - d_1 & d_2 + c_1 & a_2 - b_1 & -b_2 - a_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 0_{\mathbb{R}^4} \tag{2.45}$$

A direct calculation shows that the determinant of the matrix of coefficients of the system is equal to 0, and therefore there exists a non-zero solution. The remainder theorem guarantees us that $\Phi(q) = g(q)(q - r)$.
Thus, $f(q) = g(q)\gamma^{-1}\gamma(q - r)\gamma^{-1}$. Let $G(q) = g(q)\gamma^{-1}$, then, due to the properties of quaternionic rotations (functions of the type $\rho_q(p) = qpq^{-1}$) which we will prove later, $f(q) = G(q)(q - \zeta_r)$, where $\zeta_r = \gamma r \gamma^{-1}$. This proves the assertion. □

The proof for a right quaternionic polynomial is entirely analogous. However, unlike the complex and real cases, left/right quaternionic coefficient polynomials can have more solutions than their degree, and potentially even infinite solutions.

Example 2.1 (Quaternionic Square Roots of -1)**.** *Consider the second-degree equation:*

$$q^2 + 1 = 0 \tag{2.46}$$

Let's solve this equation directly; substituting $q = a+bi+cj+dk$ *into equation* (2.46) *we obtain:*

$$a^2 - b^2 - c^2 - d^2 + 2abi + 2acj + 2adk = -1 \tag{2.47}$$

Applying the definition of equality of quaternions:

$$\begin{cases} a^2 - b^2 - c^2 - d^2 = -1 \\ 2ab = 0 \\ 2ac = 0 \\ 2ad = 0 \end{cases} . \tag{2.48}$$

The last 3 equations of the system require that either a *is equal to 0, or alternatively* $b, c, d = 0$. *The latter route is not feasible, as it would imply that* a, *a real number, has a square equal to -1* $a^2 = -1$. *Therefore* $a = 0$.
It follows that the general form of a solution of equation (2.48) *is a quaternion* $q = xi + yj + zk$ *with:*

$$x^2 + y^2 + z^2 = 1 \tag{2.49}$$

The solutions of the equation thus live on the sphere $S^2 \in \mathbb{R}^3 \cong$ ***P***.

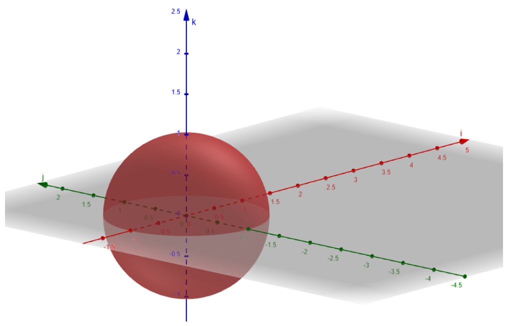


Figure 2.2: Visualization of the set of quaternionic square roots of -1.

We are not, therefore, faced with finite sets of zeros, as in the case of polynomial equations in $\mathbb{R}$ or $\mathbb{C}$, nor infinitely countable, but rather with continua of solutions. At this point, our goal is to classify such sets from a topological point of view. In the previous case, the set of zeros was homeomorphic to S^2; we want to explore the topological properties of the sets of zeros of left/right polynomials in more general cases. Before starting, however, let's give another example:

Example 2.2. *Now consider the equation $q^2 - 1 = 0$; similarly to what was done before, we substitute $q = a + bi + cj + dk$ and solve for the components of the quaternion:*

$$a^2 - b^2 - c^2 - d^2 + 2abi + 2acj + 2adk = 1 \tag{2.50}$$

thus:

$$\begin{cases} a^2 - b^2 - c^2 - d^2 = 1 \\ 2ab = 0 \\ 2ac = 0 \\ 2ad = 0 \end{cases} . \tag{2.51}$$

In this case, however, we will necessarily have $b = c = d = 0$, as if on the contrary we have $a = 0$, it will result that $-b^2 - c^2 - d^2$, a negative real quantity, is equal to 1.
The only two quaternionic roots of 1 will therefore be: $q = \pm 1$. In this case, however, the set of solutions is a set of two points, thus a discrete set.

2.4.3 Topological Properties of $\ker(\mathrm{ev}_\phi)$

Now, as mentioned before, we investigate the topological properties of the zero sets of left/right quaternionic polynomials, i.e., the kernel of the evaluation functions of the right/left quaternionic polynomials.
Before investigating the topological properties of the zero sets of a quaternionic equation, let us introduce some preliminary definitions:

Definition 2.19 (Centralizer of a Quaternion)**.** *We call the centralizer of a quaternion $q \in \mathbb{H}$ the set:*

$$Z(q) = \{p \in \mathbb{H}; qp = pq\} \tag{2.52}$$

Proposition 2.13. *For every $q \in \mathbb{H}$, $q = a_1 + b_1 i + c_1 j + d_1 k$*

$$Z(q) = \left\{p = a_2 + b_2 i + c_2 j + d_2 k\,; \frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2}\right\} \tag{2.53}$$

Proof. $qp = a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 + (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)i + (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)j + (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)k = pq = (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2) + (a_2b_1 + b_2a_1 + c_2d_1 - d_2c_1)i + (a_2c_1 - b_2d_1 + c_2a_1 + d_2b_1)j + (a_2d_1 + b_2c_1 + d_2a_1 - c_2b_1)k$
Using the definition of equality of quaternions, all this becomes:

$$\begin{cases} a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 = a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2 \\ a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2 = a_2b_1 + b_2a_1 + c_2d_1 - d_2c_1 \\ a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2 = a_2c_1 - b_2d_1 + c_2a_1 + d_2b_1 \\ a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2 = a_2d_1 + b_2c_1 + d_2a_1 - c_2b_1 \end{cases} \tag{2.54}$$

The first is trivially always verified. Manipulating the other 3, we obtain:

$$\begin{cases} d_2c_1 = c_2d_1 \\ d_1b_2 = d_2b_1 \\ b_1c_2 = b_2c_1 \end{cases} \tag{2.55}$$

From the third we get $\frac{b_1}{b_2} = \frac{c_1}{c_2}$, but from the first we get: $\frac{c_1}{c_2} = \frac{d_1}{d_2}$, and therefore $\frac{b_1}{b_2} = \frac{c_1}{c_2} = \frac{d_1}{d_2}$. □

It is a well-known fact that the centralizer of a subset of an algebra over a field (like the quaternions) is a subalgebra over the same field.
Let's now define the concept of the centralizer of a polynomial:

Definition 2.20 (Centralizer of a Quaternionic Polynomial)**.** *Let $\phi(q)$ be a left quaternionic coefficient polynomial, expressed as $\phi(q) = \sum_{i=0}^{n} a_i q^i$, with $a_n = 1$. We call the centralizer of ϕ the set:*

$$Z(\phi(q)) = \bigcap_{i=0}^{n-1} Z(a_i) \tag{2.56}$$

in other words, the intersection of the centralizers of the coefficients.

The definition is completely analogous for a right polynomial.
Moreover, as from now on we will work exclusively with polynomials with leading coefficient (i.e., the coefficient of the highest-degree monomial) equal to 1, we will give these a name; we will call them **canonical polynomials** on $\mathbb{H}$. We now introduce an operational result:

Proposition 2.14. *Let q_0 be a root of the polynomial $\phi(q) \in \mathbb{H}[X]$, then for every $\eta \in Z(\phi)$, the quaternion $\eta q_o \eta^{-1}$ is a root of ϕ.*

Proof. The proof is very simple provided that you know the properties of those quaternionic applications that we call "rotations" (for reasons we will see later), i.e., applications of the type $\rho_\eta(q) = \eta q \eta^{-1}$. For a matter of thematic organization of the chapters we will prove the facts we will use here a little further on in the text; however, the reader can already now verify them themselves with a direct calculation.

The value of ϕ in $\eta q_o \eta^{-1}$ is: $\phi(\eta q_o \eta^{-1}) = \sum_{i=0}^{n} a_i(\eta q_o \eta^{-1})^i = \sum_{i=0}^{n} a_i(\eta q_o^i \eta^{-1}) = \sum_{i=0}^{n} \eta a_i q_o^i \eta^{-1} = \eta\Big(\sum_{i=0}^{n} a_i q_o^i\Big)\eta^{-1} = \eta\phi(q_o)\eta^{-1} = 0$, which demonstrates the assertion. □

It will be profitable for us at this point to divide our investigation into various cases:

- When the dimension of the centralizer of our polynomial is 4, in other words, it is isomorphic to $\mathbb{H}$ itself. This case will correspond to polynomials with coefficients in the center of $\mathbb{H}$ (which we know to be $\mathbb{R}$), as the centralizer of the center of a non-commutative ring is the ring itself.
- When the dimension of the centralizer of our polynomial is 2 (isomorphic to $\mathbb{C}$).
- When it is equal to 1, the most general case.

Let's start with the first case:

Quaternionic Polynomials with Real Coefficients $(\dim(Z(\phi) = 4)$

A left quaternionic polynomial whose centralizer has a dimension of 4 is a real-coefficient polynomial, in the form:

$$\Phi(q) = \sum_{i=0}^{n} r_i q^i \tag{2.57}$$

with $r_i \in \mathbb{R}$, for all i. Let's now introduce a first result regarding the roots of a real-coefficient quaternionic polynomial:

Proposition 2.15. *Let* $\Phi(q) = \sum_{i=0}^{n} r_i q^i$ *be a quaternionic polynomial with real coefficients. If it has at least one non-real root,* $\zeta \in \ker(\mathrm{ev}_\Phi)$, $\zeta \notin \mathbb{R}$, *then it has infinitely many quaternionic roots.*

Proof. Let's use the result proven just now (proposition 2.14): let $\zeta \in \ker(\mathrm{ev}_\Phi)$ be our non-real root, then for every $\eta \in Z(\Phi)$, also $\eta\zeta\eta^{-1}$ is a root. But here, since ζ is non-real, we will obtain a different solution for every η, and therefore we will have infinitely many solutions. □

Proposition 2.16. *Let $\Phi(q) = \sum_{i=0}^{n} r_i q^i$ be a quaternionic polynomial with real coefficients. Then $\ker(\mathrm{ev}_\Phi)$ consists of t points, where t is the number of real roots of the polynomial Φ, and of s spheres $S^2 \subseteq \mathbb{R}^3$. Furthermore, the following inequality is satisfied:*

$$t + 2s \leq n \tag{2.58}$$

Proof. Since Φ is a polynomial with real coefficients, we know it can be factored in the following way:

$$\Phi(q) = \prod_{k=1}^{m} \varphi_k(q) \tag{2.59}$$

where $\varphi_k(q)$ is a first-degree polynomial or an irreducible second-degree polynomial.
Suppose we have t first-degree polynomials, i.e., t real roots of the polynomial. Let's denote the number of irreducible second-degree polynomials in the factorization as s, then we can't have more than $\frac{n-t}{2}$. But these present as solutions spheres; we can thus call s also the number of spheres. From this follows the inequality of the assertion: $t + 2s \leq n$ □

Corollary 2.2. *Let $\Phi(q)$ be a quaternionic polynomial with real coefficients. Then the set of its zeros ($\ker(\mathrm{ev}_\Phi)$) is finite if and only if all the roots of Φ are real.*

Proposition 2.17. *The solution to the equation $q^n = b$ where $b \in \mathbb{R}/\{0\}$ consists of:*

- *2 points and $m - 1$ spheres for $n = 2m$ and $b > 0$.*
- *m spheres for $n = 2m$ and $b < 0$.*
- *One point and $m - 1$ spheres for $n = 2m - 1$.*

Proof. Firstly, recall that we can factorize a real-coefficient polynomial of the type $q^n - b$ on $\mathbb{C}$ as $q^n - b = \prod_{k=0}^{n}(q - \zeta_k)$, where ζ_k are the complex roots of b.

- For the first case, $(n = 2m, b > 0)$, we will have $\zeta_k = \sqrt[n]{b}(\cos(\frac{2k\pi}{n}) + i\sin(\frac{2k\pi}{n}))$ with $0 \leq k \leq n-1$. Since b is a real positive number, ζ_0 will trivially be a real number, as a consequence of the restrictions imposed on b ($b \in \mathbb{R}^+$). However, the evenness of the degree of our polynomial also imposes that for $k = m$ we will have another real root. These two are the only real roots we will obtain. Due to another important property of real-coefficient polynomials, we know that complex roots come in "conjugate pairs" (if ζ is a root then so is ζ^*). Knowing that $(q-\zeta)(q-\zeta^*)$ is an irreducible real-coefficient second-degree polynomial, we learn that we can write our polynomial as a product of 2 first-degree real-coefficient polynomials, and $(m-1)$ irreducible real-coefficient second-degree polynomials. Remembering that the solutions of the former in $\mathbb{H}$ are isolated points, and the solutions of the latter are spherical solutions, it follows that we will have 2 isolated points and $m-1$ spheres.

- The second case is entirely analogous to the first, except that, having no real solutions, we can factorize our polynomial on $\mathbb{R}[x]$ writing it as a product of m irreducible real-coefficient second-degree polynomials, which represent all spherical solutions in $\mathbb{H}$.

- For the third case, finally, we recall that a real-coefficient polynomial of the type $q^n = b$ with n odd will have only one real solution, whether $b > 0$ or $b < 0$, and thus we can factorize $q^n - b$ on $\mathbb{R}[x]$ as a product of a first-degree real-coefficient polynomial and $m-1$ irreducible real-coefficient second-degree polynomials.

□

We now proceed to analyze quaternionic polynomials with a 2-dimensional centralizer.

Quaternionic Polynomials with $\dim(Z(\Phi)) = 2$

The first question that arises at this point is: what conditions must the coefficients of a left polynomial meet so that its centralizer has dimension 2? Let's introduce some preparatory results:

Lemma 2.1. *Every quaternion $p = a + bi + cj + dk$ satisfies the following equation with real coefficients:*

$$q^2 - 2aq + a^2 + b^2 + c^2 + d^2 = q^2 - 2\,\mathrm{Sc}(p) + |p|^2 = 0$$

we will call the polynomial $f_p(q) = q^2 - 2\,\mathrm{Sc}(p) + |p|^2$ the ***characteristic polynomial of the quaternion p****.*

Exercise 2.8. *Prove lemma 2.1.*

Lemma 2.2. *Let's call $\mathbb{K}_\alpha$ a set of elements of the type $m + n\alpha$ with $m, n \in \mathbb{R}$ and $\alpha \in \mathbb{H} - \mathbb{R}$, then it forms a subalgebra isomorphic to the algebra of complex numbers.*

Proof. The set is trivially closed with respect to addition inherited from $\mathbb{H}$. As for multiplication, observe that every quaternion $\alpha \in \mathbb{H}$ satisfies a quadratic equation with a negative discriminant, let

$$q^2 + sq + t = 0 \tag{2.60}$$

be the aforementioned equation. Then: $\alpha^2 = -s\alpha - t \Longrightarrow \mathbb{K}_\alpha$ is closed with respect to multiplication, i.e., $\mathbb{K}_\alpha$ is a subalgebra of dimension 2 of $\mathbb{H}$.
Thus $\mathbb{K}_\alpha$ is a $\mathbb{R}$-Algebra associative unitary of dimension 2, we have 3 possibilities:

- It is isomorphic to $\mathbb{C}$ (the complex numbers), if $\frac{s^2}{4} - t < 0$.
- It is isomorphic to $\mathbb{R}[\epsilon]$ (Clifford's dual numbers), if $\frac{s^2}{4} - t = 0$.
- It is isomorphic to $\mathbb{R}[j]$ (the split-complex numbers, also called pseudocomplex numbers), if $\frac{s^2}{4} - t > 0$.

Having a negative discriminant ($\frac{s^2}{4} - t < 0$), it is isomorphic to $\mathbb{C}$. □

We now prove a necessary and sufficient condition for the centralizer of a left quaternionic polynomial to be of dimension 2.

Proposition 2.18. *The dimension of the centralizer $Z(\Phi)$ of a polynomial $\Phi(q) = \sum_{i=0}^{n} \alpha_i q^i$ is equal to 2 $\Longleftrightarrow$ $\alpha_s\alpha_t = \alpha_t\alpha_s$ for all t, s, in other words, if the coefficients commute with each other.*

Proof. Let's first prove the direct implication ($\Longrightarrow$).
Let $\dim(Z(\Phi)) = 2$. Being $Z(\Phi) = \bigcap_{i=0}^{n-1} Z(\alpha_i)$ we have that for all $\zeta \in Z(\Phi)$, $\zeta\alpha_t = \alpha_t\zeta$, for all $t \in \mathbb{N} \cap [0, n-1]$, or in other words $\zeta \in Z(\alpha_t)$, for all $t \in \mathbb{N} \cap [0, n-1]$. Let $\zeta = a + bi + cj + dk$. Recalling now a fact proven a little earlier, we can interpret this fact as:

$$\frac{b_s}{b} = \frac{c_s}{c} = \frac{d_s}{d} \tag{2.61}$$

and

$$\frac{b_t}{b} = \frac{c_t}{c} = \frac{d_t}{d} \tag{2.62}$$

Where s and t are two non-equal indices coming from the set $\mathbb{N} \cap [0, n-1]$. Simple algebraic manipulation shows us that:

$$\frac{b_t}{b_s} = \frac{c_t}{c_s} = \frac{d_t}{d_s} \tag{2.63}$$

which means, in light of proposition 2.13, that $\alpha_t\alpha_s = \alpha_s\alpha_t$.
Let's now prove the opposite direction ($\Longleftarrow$). Saying that α_t commutes with α_s, for two non-equal indices t and s belonging to $\mathbb{N} \cap [0, n-1]$, means (again in light of proposition 2.13):

$$\frac{b_t}{b_s} = \frac{c_t}{c_s} = \frac{d_t}{d_s} = \gamma \in \mathbb{R} \tag{2.64}$$

therefore:

$$\begin{cases} b_t = \gamma b_s \\ c_t = \gamma c_s \\ d_t = \gamma d_s \end{cases} \tag{2.65}$$

From this it follows that we can write α_t as $\alpha_t = a_t + \gamma b_s i + \gamma c_s j + \gamma d_s k = (a_t - \gamma a_s) + \gamma(a_s + b_s i + c_s j + d_s k) = \rho + \gamma\alpha_s$, where here $\rho = a_t - \gamma a_s$, and $\alpha_s \in \mathbb{H} - \mathbb{R}$. It follows then that the coefficients of the polynomial belong to $\mathbb{K}_{\alpha_s}$, a subalgebra of $\mathbb{H}$ that we have proven to be isomorphic to the complex numbers. Being the complex numbers a field, they are their own center, and the centralizer of each element is the whole set $\mathbb{C}$. From this, it follows that the centralizer of a polynomial with such coefficients must be 2-dimensional. □

Corollary 2.2. *The centralizer $Z(\Phi)$ of a left quaternionic coefficient polynomial Φ is of dimension 2 if and only if it is isomorphic to the algebra of complex numbers $\mathbb{C}$.*

Proof. Directly follows from the considerations made in the proof of the previous proposition. □

From what was said in the last propositions, we deduce that we can write a left quaternionic coefficient polynomial Φ with $\dim(Z(\Phi)) = 2$ as:

$$\Phi(q) = \prod_{k=1}^{n} (q - \zeta_k) \tag{2.66}$$

with $\zeta_k \in \ker(\mathrm{ev}\,\Phi)$ for all $k \in \mathbb{N} \cap [1, n]$, i.e., where $\zeta_k \in \mathbb{C}$ are the zeros of the polynomial in question.
From this way of expressing a left polynomial belonging to such a class, the following proposition naturally follows:

Proposition 2.19 (Structure of the Zeros of a Left Quaternionic Polynomial with $\dim(Z(\Phi)) = 2$)**.** *The set of zeros $\ker(\mathrm{ev}\,\Phi)$ of a left quaternionic polynomial $\Phi(q) = \sum_{i=0}^{n} a_i q^i$ with $\dim(Z(\Phi)) = 2$ consists of t points and s spheres, satisfying the following inequality:*

$$t + 2s \leq n \tag{2.67}$$

where n is the degree of the polynomial.

Proof. Using the fact just proven, we write Φ as:

$$\Phi(q) = \prod_{k=1}^{n} (q - \zeta_k) \tag{2.68}$$

Among the roots ζ_i, we call s the number of pairs of roots $(\zeta,\ \zeta^*)$, that is, the number of roots that also present their conjugate among the roots. We can combine them to obtain an irreducible trinomial with real coefficients, the characteristic polynomial P_ζ. But such a polynomial has exactly S^2 as its set of roots. The remaining roots, which we will call t, will instead be isolated points. Thus:

$$t + 2s \leq n \tag{2.69}$$

□

Quaternionic Polynomials with $\dim(Z(\Phi)) = 1$

We now work on the most general case, when the dimension of the centralizer of our polynomial is equal to 1. Before proving some results, let's introduce some preliminary notions:

Definition 2.21 (Conjugate Polynomial)**.** *Let $\Phi(q) = \sum_{i=0}^{n} a_i q^i$ be a left quaternionic polynomial: we call $\Phi^*(q)$ its conjugate polynomial, defined as:*

$$\Phi^*(q) := \sum_{i=0}^{n} a_i^* q^i \tag{2.70}$$

Definition 2.22 (Pseudo-norm of a Quaternionic Polynomial)**.** *Let* $\Phi(q) = \sum_{i=0}^{n} a_i q^i$ *be a left quaternionic polynomial: we call* $N(\Phi(q))$ *its pseudo-norm, defined as:*

$$N(\Phi(q)) := \Phi(q)\Phi^*(q) \tag{2.71}$$

i.e., the product of the polynomial Φ *with its conjugate polynomial.*

Exercise 2.9. *Let* $\Phi(q) = (1 + j + k)q - 2jq^2 + (i - k)q^4$ *and* $\Psi(q) = k + 2kq - 2iq^2 - 3jq^3$ *be two left quaternionic polynomials. Calculate:*

1. $\Phi^*(q)$
2. $\Psi^*(q)$
3. $N(\Phi(q))$
4. $N(\Psi(q))$

Proposition 2.20. *Let* $\Phi(q) = \sum_{k=0}^{n} \alpha_k q^k$ *be a canonical left quaternionic coefficient polynomial. Then its pseudo-norm* $N(\Phi(q)) = \Phi(q)\Phi^*(q)$ *is a polynomial with real coefficients of degree* $\deg(N(\Phi(q)) = 2\deg(\Phi(q))$.

Proof. As we have defined multiplication on the ring of left quaternionic coefficient polynomials:

$$\Phi(q)\Phi^*(q) = \sum_{i=0}^{2n} \beta_i q^i$$

where

$$\beta_i = \sum_{j=0}^{i} \alpha_j \alpha^*_{m-j}$$

We need to prove that β_i is a real number $\forall i$, which we will see is quite simple. In fact, it is enough to notice that even-degree terms will have sums of terms of the type $\alpha_j\alpha_j^*$ and of the type $\alpha_j\alpha_k^* + \alpha_k\alpha_j^*$, while odd-degree terms will have sums of terms only of the second type. However, we know that the algebra of quaternions is well-normed, and thus such terms are all real numbers. Let's expand some of the coefficients in question to give the reader a better idea of the proof:

$$\beta_0 = \alpha_0\alpha_0^*$$

$$\beta_1 = \alpha_0\alpha_1^* + \alpha_1\alpha_0^*$$
$$\vdots$$
$$\beta_{2n-1} = \alpha_0\alpha_{2n-1}^* + \alpha_1\alpha_{2n-2}^* + \cdots + \alpha_{n-1}\alpha_n^* + \alpha_n\alpha_{n-1}^* + \cdots + \alpha_{2n-2}\alpha_1^* + \alpha_{2n-1}\alpha_0^*$$
$$\beta_{2n} = \alpha_0\alpha_{2n}^* + \alpha_1\alpha_{2n-1}^* + \cdots + \alpha_n\alpha_n^* + \cdots + \alpha_{2n-1}\alpha_1^* + \alpha_{2n}\alpha_0^*$$

From this, the assertion follows. □

Definition 2.23 (Similar Quaternions)**.** *Let $q_1, q_2 \in \mathbb{H}$ be two quaternions; we say that q_1 and q_2 are **similar**, and write $q_1 \sim q_2$, if there exists a non-zero quaternion $h \in \mathbb{H} \setminus \{0\}$ such that $hq_1 = q_2h$.*

Exercise 2.10. *Prove that the relation "being similar" on quaternions is an equivalence relation on $\mathbb{H}$.*

Exercise 2.11. *Prove that i and k are similar, and find a quaternion h that makes them so.*

Exercise 2.12. *Given two quaternions $q_1 = a_1 + b_1i + c_1j + d_1k$ and $q_2 = a_2 + b_2i + c_2j + d_2k$, prove that if*

$$\begin{vmatrix} a_1 - a_2 & b_2 - b_1 & c_2 - c_1 & d_2 - d_1 \\ b_1 - b_2 & a_1 - a_2 & d_2 - d_1 & -c_1 - c_2 \\ c_1 - c_2 & -d_1 - d_2 & a_1 - a_2 & b_1 + b_2 \\ d_1 - d_2 & c_1 + c_2 & -b_1 - b_2 & a_1 - a_2 \end{vmatrix} = 0 \tag{2.72}$$

then $q_1 \sim q_2$.

Corollary 2.3. *Let $q = a+bi+cj+dk$ be a quaternion, where $a, b, c, d \in \mathbb{R}\setminus\{0\}$, then $a + \sqrt{b^2 + c^2 + d^2} \sim q$.*

Proof. Follows as an immediate corollary of exercise 2.12. In fact, just notice that the determinant of the matrix:

$$\begin{vmatrix} 0 & \sqrt{b^2+c^2+d^2} - b & -c & -d \\ b - \sqrt{b^2+c^2+d^2} & 0 & -d & -c \\ c & -d & 0 & b + \sqrt{b^2+c^2+d^2} \\ d & c & -b - \sqrt{b^2+c^2+d^2} & 0 \end{vmatrix} = 0$$

□

Before continuing with the treatment started on quaternionic polynomials, let's prove a lemma that will be useful shortly:

Lemma 2.3. *Let q_1 and q_2 be two quaternions, then they are similar if and only if their scalar parts and their norms are equal, i.e:*

$$q_1 \sim q_2 \Longleftrightarrow \mathrm{Sc}(q_1) = \mathrm{Sc}(q_2)\ e\ |q_1| = |q_2|$$

Proof. We start from the direct implication; suppose that $q_1 \sim q_2$, then we will have that there exists a quaternion $h \in \mathbb{H} \setminus \{0\}$ such that:

$$hq_1 = q_2h \tag{2.73}$$

which implies that $|hq_1| = |q_2h|$. But due to the properties of the quaternion norm, we will have that $|hq_1| = |h| \cdot |q_1| = |q_2| \cdot |h| = |q_2h|$ and therefore $|q_1| = |q_2|$ since we can divide both sides by $|h|$ being h non-zero. Furthermore, always as a consequence of equation (2.73), we will have that $q_1 = h^{-1}q_2h$, from which $\mathrm{Sc}(q_1) = \mathrm{Sc}((h^{-1}q_2)h) = \mathrm{Sc}(h(h^{-1}q_2)) = \mathrm{Sc}(q_2)$. For the reverse direction, instead, suppose that q_1 and q_2 are two quaternions with equal scalar part and norm, $\mathrm{Sc}(q_1) = \mathrm{Sc}(q_2)$ and $|q_1| = |q_2|$. Let $q_1 = a_1 + b_1i + c_1j + d_1k$ and $q_2 = a_2 + b_2i + c_2j + d_2k$. By corollary 2.3, we will have that:

$$a_1 + \sqrt{b_1^2 + c_1^2 + d_1^2}i \sim q_1$$

and

$$a_2 + \sqrt{b_2^2 + c_2^2 + d_2^2}i \sim q_2$$

but since $\mathrm{Sc}(q_1) = \mathrm{Sc}(q_2) = a_1$ we will have that $b_1^2 + c_1^2 + d_1^2 = b_2^2 + c_2^2 + d_2^2$ and $a_2 = a_1$ and thus:

$$q_1 \sim a_1 + \sqrt{b_1^2 + c_1^2 + d_1^2}i = a_2 + \sqrt{b_2^2 + c_2^2 + d_2^2}i \sim q_2$$

therefore $q_1 \sim q_2$ by the transitive property of quaternion similarity. □

Lemma 2.4. *Let Φ be a left quaternionic coefficient polynomial and let $\zeta \in \mathbb{H}$. If the characteristic polynomial of ζ, P_ζ, divides Φ, then the entire similarity class of the quaternion ζ, $[\zeta]$ (i.e., the equivalence class of ζ under the relation "being similar") consists of roots of the polynomial Φ.*
If instead P_ζ does not divide Φ, i.e., there is a non-zero remainder, then if there is, there is no more than one root in $[\zeta]$.

Proof. If $P_\zeta \mid \Phi(q)$, then $\Phi(q) = Q(q)P_\zeta(q)$. Therefore, all roots of P_ζ are also roots of Φ. Remember that $P_\zeta = q^2 - 2\,\mathrm{Sc}(\zeta) + |\zeta|^2$, from which it follows that for all $z \in [\zeta]$, being $|z| = |\zeta|$ and $\mathrm{Sc}(z) = \mathrm{Sc}(\zeta)$ by the previous lemma, $P_\zeta(z) = P_\zeta(\zeta) = 0$, hence we obtain that $\Phi(z) = 0$.
Let's now see the second case:

If P_ζ does not divide Φ, then the division of Φ by P_ζ will have a non-zero remainder. We write:

$$\Phi(q) = Q(q)P_\zeta(q) + R(q) \tag{2.74}$$

Where $R(q)$ is a first-degree polynomial of the form $\alpha q + \beta$. If $\alpha = 0, \beta \neq 0$, there will be no solutions in $[\zeta]$. In the case where $\alpha, \beta \neq 0$, however:
Let $z \sim \zeta$, such that $\Phi(z) = 0$. Then we will have:

$$\Phi(z) = Q(z)P_\zeta(z) + \alpha z + \beta \tag{2.75}$$

but being z in the equivalence class of ζ under the relation of similarity of quaternions defined in this chapter, we know that it is a root of P_ζ. We will therefore have $0 = 0 + \alpha z + \beta$, that is $z = -\alpha^{-1}\beta$. Therefore, if we have a root belonging to the equivalence class of ζ in the case where P_ζ does not divide Φ, we can have a single root equal to $-\alpha^{-1}\beta$. □

From this proposition, the following fact follows immediately as a corollary:

Corollary 2.4. *The set of zeros* $\ker(\mathrm{ev}_\Phi)$ *of a left quaternionic polynomial* Φ *is infinite if and only if* $\exists \zeta \in \mathbb{H}$ *such that* $P_\zeta \mid \Phi(q)$.

Proposition 2.21. *If* ζ *is a root of the polynomial* $\Phi(q)$*, then the characteristic polynomial of the latter,* P_ζ *divides the pseudo-norm of* Φ*,* $N(\Phi(q))$*.*

Proof. Let $\zeta \in \ker(\mathrm{ev}_\Phi)$, we divide Φ by P_ζ with remainder:

$$\Phi(q) = Q(q)P_\zeta(q) + R(q) \tag{2.76}$$

its conjugate, using the properties of quaternion conjugation, is:

$$\Phi^*(q) = Q^*(q)P_\zeta(q) + R^*(q) \tag{2.77}$$

The pseudo-norm, which is given by the product of Φ with Φ^*, will be:

$$N(\Phi) = \Phi\Phi^* = Q(q)Q^*(q)P_\zeta^2(q) + (R(q)Q^*(q) + Q(q)R^*(q))P_\zeta(q) + R(q)R^*(q) \tag{2.78}$$

the assertion is trivially true if the remainder is equal to 0 (i.e. if P_ζ divides Φ). Therefore, let us suppose that it is not equal to 0, and prove that P_ζ divides it, and therefore in this case too P_ζ divides Φ.
Remembering that $R(q) = \alpha q + \beta$, we calculate its pseudo-norm:

$$N(R) = RR^* = \alpha\alpha^* q^2 + (\alpha\beta^* + \beta\alpha^*)q + \beta\beta^* \tag{2.79}$$

which we can rewrite as

$$N(R) = \alpha\alpha^*\Big(q^2 + \frac{\alpha\beta^* + \beta\alpha^*}{\alpha\alpha^*}q + \frac{\beta\beta^*}{\alpha\alpha^*}\Big) \tag{2.80}$$

We know that ζ is a zero belonging to $[\zeta]$. Knowing that the remainder, by the assumption made earlier, is not equal to 0, by lemma 2.4 we know that ζ is of the form $\zeta = -\alpha^{-1}\beta$. Calculating the characteristic polynomial of ζ, we will see that it is precisely $q^2 + \frac{\alpha\beta^*+\beta\alpha^*}{\alpha\alpha^*}q + \frac{\beta\beta^*}{\alpha\alpha^*}$; let's prove it:

$$-2Sc(\zeta) = 2Sc(\alpha^{-1}\beta) = \alpha^{-1}\beta + (\alpha^{-1}\beta)^* \tag{2.81}$$

being $\alpha^{-1}\beta = \frac{\alpha^*\beta}{\|\alpha\|^2}$, we have $-2Sc(\zeta) = \frac{\alpha^*\beta}{|\alpha|^2} + \frac{\beta^*\alpha}{|\alpha|^2} = \frac{\alpha^*\beta+\beta^*\alpha}{|\alpha|^2}$
Concerning $|\zeta|^2$, however:

$$|\zeta|^2 = (\alpha^{-1}\beta)(\alpha^{-1}\beta)^* = \alpha^{-1}\beta\beta^*(\alpha^{-1})^* \tag{2.82}$$

using the properties of quaternion conjugation. With a final algebraic manipulation, we will obtain:

$$|\zeta|^2 = \frac{\alpha^*}{|\alpha|^2}\beta\beta^*\frac{\alpha}{|\alpha|^2} = \frac{\beta\beta^*}{\alpha\alpha^*} \tag{2.83}$$

Indeed, P_ζ divides RR^*, and with this we have proven the assertion, namely we have proven that P_ζ divides the pseudo-norm of the polynomial Φ. □

We close this section with an extension of a result already proven earlier (for left polynomials with centralizers of dimensions 4 and 2 respectively) to more general left polynomials (with coefficients in $\mathbb{H}$, i.e. $\dim(Z(\Phi)) = 1$):

Theorem 2.3. *The set of zeros of a canonical left quaternionic polynomial $\Phi(q) = \sum_{i=0}^{n} a_i q^i$ consists of t points and s 2-dimensional spheres S^2; moreover, the following inequality is valid:*

$$t + 2s \leq n \tag{2.84}$$

Proof. Let $\zeta \in \mathbb{H}$. If P_ζ divides our polynomial Φ, then $[\zeta] \cong S^2 \subset \ker(\mathrm{ev}_\Phi)$ and $P_\zeta^2 \mid N(\Phi)$. However, if P_ζ does not divide Φ, then the polynomial Φ will have no more than one root from the similarity class of ζ, and if such a root exists, $N(\Phi)$ will be divisible by P_ζ.
We call t the number of isolated roots ζ (i.e. with P_ζ that does not divide

Φ) and we call s the number of roots ζ' such that $P_{\zeta'} \mid \Phi$. Applying mathematical induction to the pseudo-norm of Φ, $N(\Phi)$, we obtain:

$$2t + 4s \leq 2n \Longrightarrow t + 2s \leq n$$

□

Note: all the results previously obtained, in which for simplicity we assumed that we were dealing exclusively with left polynomials, are equally obtainable with entirely analogous proofs for right quaternionic polynomials.
With this result, we have managed to classify the types of connected components of the set of zeros of a left/right quaternionic polynomial. We have seen that they are components homeomorphic to 2-dimensional spheres S^2 or points. We have also seen that the number of the aforementioned does not exceed the degree of our polynomial. However, an adequate treatment of a class of quaternionic coefficient polynomials, those that we called ”general polynomials” at the beginning of the section, remains pending. Things get much more complicated for the latter, so we will not give an extensive treatment of their zeros and the theory concerning them in this text. Nevertheless, in 2 chapters we will extend the fundamental theorem of algebra also to linear general polynomials with quaternionic coefficients.

2.5 Matrix representations of $\mathbb{H}$

2.5.1 Complex matrix representation of quaternions

There is an injective homomorphism from $\mathbb{H}$ to the ring of matrices $M_{2x2}(\mathbb{C})$, i.e., the 2x2 complex matrices. We can therefore represent a quaternion with a complex matrix belonging to the set $M_{2x2}(\mathbb{C})$.
In particular, the quaternion $q = a + bi + cj + dk$ corresponds to the matrix $M_q = \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix} = a\sigma_0 + b\sigma_1 + c\sigma_2 + d\sigma_3$ $(a,b,c,d \in \mathbb{R})$ where here:

$$\sigma_0 = I_{2x2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sigma_1 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = i\sigma_z$$

$$\sigma_2 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = i\sigma_y,\ \sigma_3 = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = i\sigma_x$$

Where $\sigma_x, \sigma_y, \sigma_z$ are the Pauli matrices.

Proposition 2.22. *The algebra of quaternions $\mathbb{H}$ is isomorphic to the algebra of matrices representable as a linear combination of σ_i, $i = 0, 1, 2, 3$, in symbols:*

$$\mathbb{H} \cong \operatorname{span}\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\} = \operatorname{span}\{I_{2x2}, i\sigma_x, i\sigma_y, i\sigma_z\}$$

Proof. Let φ be the application that maps $1 \longrightarrow \sigma_0,\ i \longrightarrow \sigma_1,\ j \longrightarrow \sigma_2,$ $k \longrightarrow \sigma_3$, and the quaternion $q = \alpha + \beta i + \gamma j + \delta k \longrightarrow \begin{bmatrix} \alpha+\beta i & \gamma+\delta i \\ -\gamma+\delta i & \alpha-\beta i \end{bmatrix}$.
The application is well-defined and injective: in fact, suppose that $q_1, q_2 \in \mathbb{H}$, are equal, then, by the definition of canonical equivalence on quaternions, we have $x_{1j} = x_{2j}\ \forall j \in [1,4] \cap \mathbb{N}$. From this trivially follows that:

$$\begin{cases} x_{11} + x_{12}i = x_{21} + x_{22}i \\ x_{13} + x_{14}i = x_{23} + x_{24}i \\ -x_{13} + x_{14}i = -x_{23} + x_{24}i \\ x_{11} - x_{12}i = x_{21} - x_{22}i \end{cases}$$

and therefore also $\varphi(q_1) = \varphi(q_2)$. The converse $\varphi(q_1) = \varphi(q_2) \Longrightarrow q_1 = q_2$ is in turn trivial to prove: starting this time from the system of equations above, and using the definition of equality of complex numbers, we will deduce that $x_{1j} = x_{2j}\ \forall j \in [1,4] \cap \mathbb{N}$, then $q_1 = q_2$.

Let's calculate $\varphi(q_1+q_2) = \begin{bmatrix} (\alpha_1+\alpha_2)+(\beta_1+\beta_2)i & (\gamma_1+\gamma_2)+(\delta_1+\delta_2)i \\ -(\gamma_1+\gamma_2)+(\delta_1+\delta_2)i & (\alpha_1+\alpha_2)-(\beta_1+\beta_2)i \end{bmatrix} =$
$\begin{bmatrix} \alpha_1+\beta_1 i & \gamma_1+\delta_1 i \\ -\gamma_1+\delta_1 i & \alpha_1-\beta_1 i \end{bmatrix} + \begin{bmatrix} \alpha_2+\beta_2 i & \gamma_2+\delta_2 i \\ -\gamma_2+\delta_2 i & \alpha_2-\beta_2 i \end{bmatrix} = \varphi(q_1)+\varphi(q_2)$
$\varphi(q_1q_2) = \begin{bmatrix} \tilde{\alpha}+\tilde{\beta}i & \tilde{\gamma}+\tilde{\delta}i \\ -\tilde{\gamma}+\tilde{\delta}i & \tilde{\alpha}-\tilde{\beta}i \end{bmatrix}$
where $\tilde{\alpha} = (\alpha_1\alpha_2-\beta_1\beta_2-\gamma_1\gamma_2-\delta_1\delta_2)$, $\tilde{\beta} = (\alpha_1\beta_2+\beta_1\alpha_2-\delta_1\gamma_2+\gamma_1\delta_2)$, $\tilde{\gamma} = (\alpha_1\gamma_2-\beta_1\delta_2+\gamma_1\alpha_2+\delta_1\beta_2)$, $\tilde{\delta} = (\alpha_1\delta_2+\beta_1\gamma_2-\gamma_1\beta_2+\delta_1\alpha_2)$.
$\varphi(q_1)\varphi(q_2) = \begin{bmatrix} \alpha_1+\beta_1 i & \gamma_1+\delta_1 i \\ -\gamma_1+\delta_1 i & \alpha_1-\beta_1 i \end{bmatrix} \begin{bmatrix} \alpha_2+\beta_2 i & \gamma_2+\delta_2 i \\ -\gamma_2+\delta_2 i & \alpha_2-\beta_2 i \end{bmatrix} =$
$\begin{bmatrix} (\alpha_1+\beta_1 i)(\alpha_2+\beta_2 i)+(\gamma_1+\delta_1 i)(-\gamma_2+\delta_2 i) & (\alpha_1+\beta_1 i)(\gamma_2+\delta_2 i)+(\gamma_1+\delta_1 i)(\alpha_2-\beta_2 i) \\ (-\gamma_1+\delta_1 i)(\alpha_2+\beta_2 i)+(\alpha_1-\beta_1 i)(-\gamma_2+\delta_2 i) & (-\gamma_1+\delta_1 i)(\gamma_2+\delta_2 i)+(\alpha_1-\beta_1 i)(\alpha_2-\beta_2 i) \end{bmatrix}$
Performing the calculations, we will realize that it is exactly equal to $\varphi(q_1q_2)$. The surjectivity follows trivially from the restriction of the image of the injective algebra homomorphism in question, φ, from $M_{2x2}(\mathbb{C})$ to simply $\mathrm{span}\{\sigma_0,\sigma_1,\sigma_2,\sigma_3\}$ □

Remark 2.1. *The squared norm of a quaternion* $|q|^2 = a^2+b^2+c^2+d^2$ *is equal to the determinant of its corresponding matrix* $\det(M_q)$. *The verification is immediate:*

$$\det(M_q) = (a+bi)(a-bi)-(-c+di)(c+di) = a^2+b^2+c^2+d^2$$

Remark 2.2. *The conjugate quaternion* q^* *corresponds to the hermitian transpose of the matrix corresponding to* q, $M_q^\dagger = (M_q^T)^*$. *We leave the verification of this fact to the reader, as it is a simple direct calculation.*

We have thus seen that there is a subtle connection between quaternions and Pauli matrices, which for those with rudiments of quantum mechanics or quantum information theory will know that, together with the identity matrix I_{2x2}, form a basis for hermitian matrices (often called observables in contexts related to quantum mechanics) that act on 2-dimensional Hilbert spaces (for example the space of spinors for spin $\frac{1}{2}$ particles or the space of states of a qubit).

2.5.2 Writing quaternions as 4x4 real matrices

Another matrix representation of the algebra of quaternions is the following:

$$q = a+bi+cj+dk \longrightarrow^{\varphi^1} M_q^{\mathbb{R}1} = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}$$

However, the representations of quaternions as 4x4 real matrices are multiple. Here are other representations:

$$q = a + bi + cj + dk \longrightarrow^{\varphi^2} M_q^{\mathbb{R}2} = \begin{bmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{bmatrix}$$

$$q = a + bi + cj + dk \longrightarrow^{\varphi^3} M_q^{\mathbb{R}3} = \begin{bmatrix} a & d & -b & -c \\ -d & a & c & -b \\ b & -c & a & -d \\ c & b & d & a \end{bmatrix}$$

$$q = a + bi + cj + dk \longrightarrow^{\varphi^4} M_q^{\mathbb{R}4} = \begin{bmatrix} a & -b & d & -c \\ b & a & -c & -d \\ -d & c & a & -b \\ c & d & b & a \end{bmatrix}$$

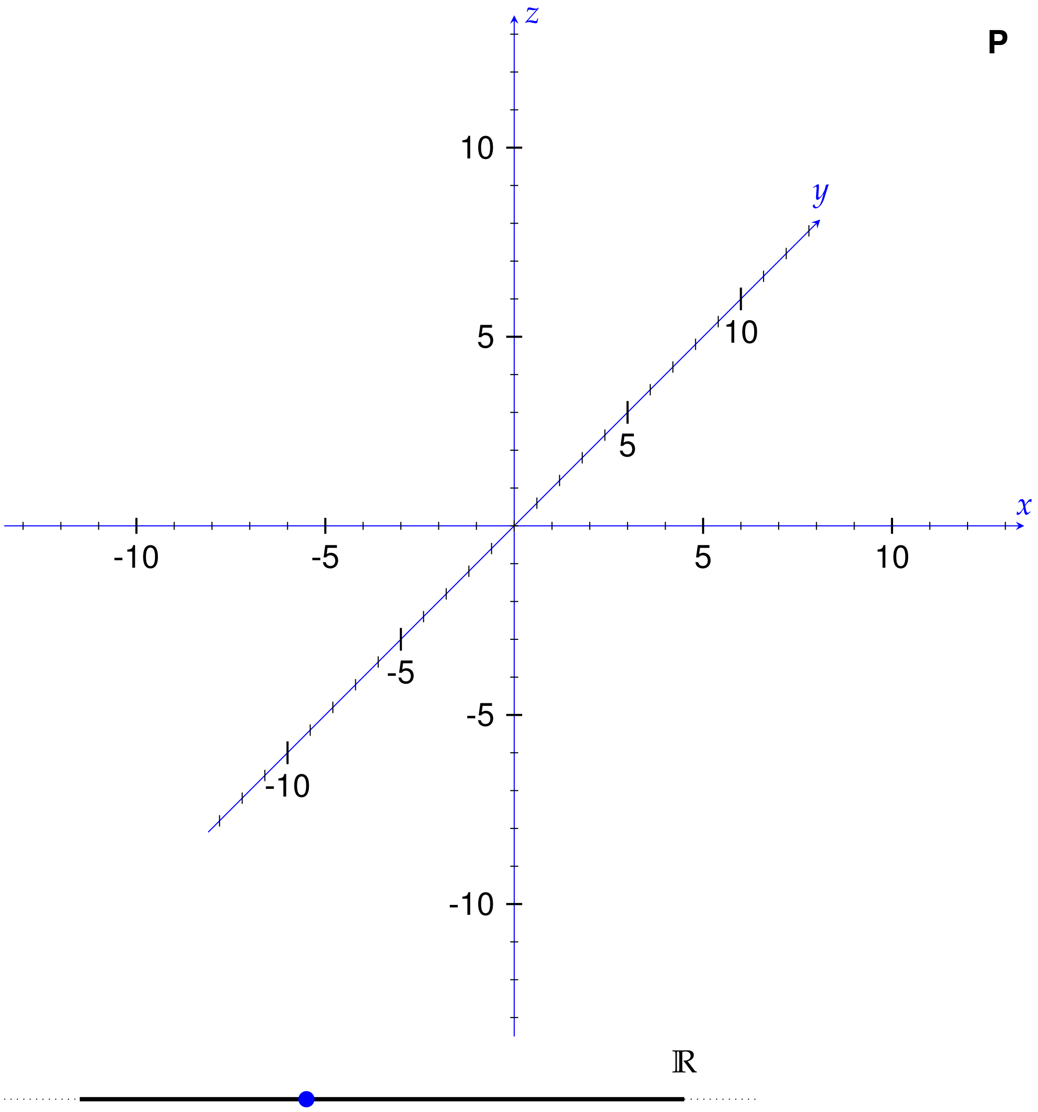
Chapter 3

Geometry of Quaternions

One of the most important applications of quaternions is undoubtedly their use in the theory of rotations in three-dimensional space. At first glance, it may be difficult for someone new to quaternions to understand how three-dimensional rotations can be generated with quaternions and what benefits they offer compared to Euler matrices. In this chapter, we will talk exactly about this, i.e., we will explain why quaternions are able to represent rotations and also cover many other important topics, such as techniques for visualizing subsets of the hyperspace $\mathbb{H}$ and how to interpolate rotations using SLERP. The topic of rotation interpolation is indeed a very important procedure easily achievable with quaternionic rotations, and it is one of the reasons why this way of representing rotations is so powerful.

3.1 Visualizing quaternions geometrically

In this section, it will be fundamentally important to have a good geometric intuition, therefore we will start by developing some ways to visualize various geometric loci living in $\mathbb{H}$. Quaternions, as seen in the previous chapter, can be written as $\mathbb{H} \cong \mathbb{R} \oplus \mathbf{P}$, where here we will call $\{xi + yj + zk \ ; \ x, y, z \in \mathbb{R}\} =: \mathbf{P} \cong \mathbb{R}^3$ the space of pure quaternions. We can interpret the vector part of the quaternion as its "spatial" part, and its real part as the "temporal" part, and therefore we can visualize quaternions as a continuous timeline in which each point is associated with a three-dimensional space, as illustrated in the following image:



Pure quaternions, therefore, live in the three-dimensional space $\mathbb{R}^3$ and are easily visualizable (in this case, it is as if we have fixed our temporal moment at $t = 0$). Normal pure quaternions live on the sphere S^2, and such quaternions are the solutions of the quaternionic equation $q^2 = -1$, as seen in the previous chapter. Normal quaternions, on the other hand, live on the sphere $S^3 \subset \mathbb{H}$ in four-dimensional space. How can we more easily visualize them and "bypass" their four-dimensionality? One way would be through a stereographic projection of the 3-sphere S^3 onto $\mathbb{R}^3$, but we can also visualize them by dividing the 3-sphere S^3 into 2 hemispheres and an equator, all 3 visualizable in $\mathbb{R}^3$. Let's build this visualization "inductively" starting from spheres of lower dimension:
Consider the 1-sphere (circle) S^1:

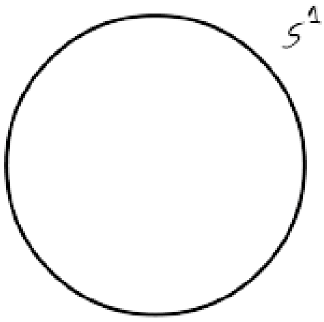


Figure 3.1: Circle S^1 visualized bi-dimensionally.

We can divide our circle into 3 regions: the northern hemisphere, the southern hemisphere, and the equator in the following way:

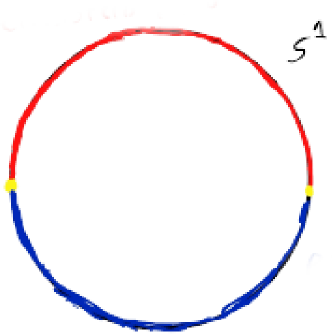


Figure 3.2: Northern hemisphere, southern hemisphere, and equator of the circle S^1.

where here we have indicated the northern hemisphere in red, the southern hemisphere in blue, and the equator in yellow. At this point, we can map the equator to 2 points, and the northern and southern hemispheres to two segments, obtaining a one-dimensional visualization of S^1 on $\mathbb{R}$:

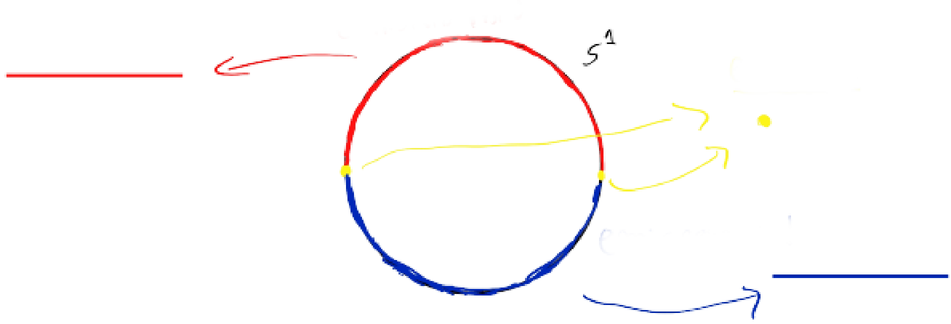


Figure 3.3: One-dimensional visualization of the circle S^1.

Similarly, this can be done with a 2-sphere S^2:

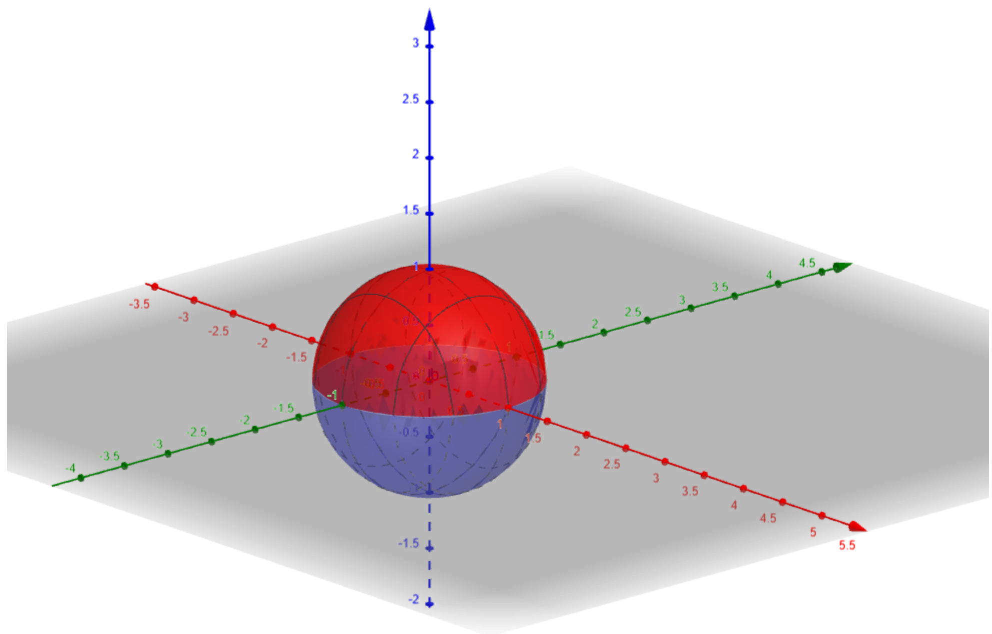


Figure 3.4: Visualization of the sphere S^2 immersed in $\mathbb{R}^3$: the northern hemisphere is indicated in red, while the southern hemisphere is indicated in blue.

where, as before, we have outlined the northern hemisphere in red and the southern hemisphere in blue, and the equator in yellow (barely visible in the image). These regions of the sphere can be transformed into two-dimensional regions that are easily visualizable: 2 filled balls (discs) and 1 empty circle.

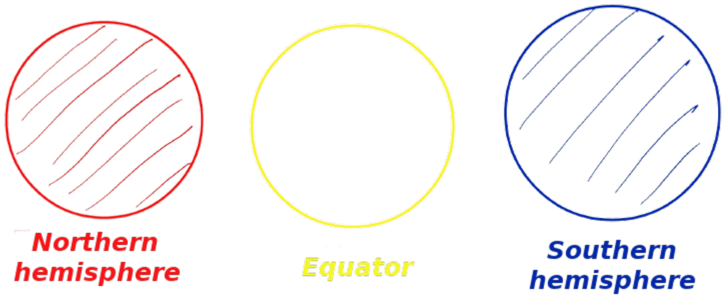


Figure 3.5: Two-dimensional visualization of the sphere S^2.

We can repeat the same procedure to visualize the 3-sphere S^3 with 3 three-dimensional regions. The result will be, as the reader might have expected, 2 filled balls and a sphere:

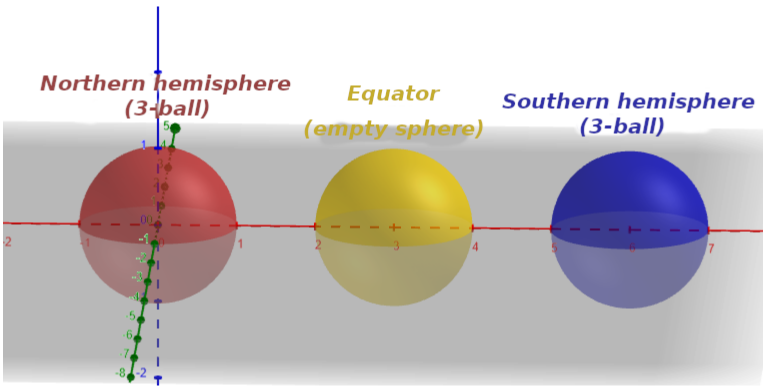


Figure 3.6: Three-dimensional visualization of the hypersphere S^3.

3.2 Quaternions and Rotations

In this section, we will see why and how normal quaternions produce three-dimensional rotations, constructing (almost) an isomorphism between them and the special orthogonal group $SO(3)$ (which is the group of 3-dimensional rotation matrices). Before proving these facts for quaternions, let's make a brief digression about some analogous facts for complex numbers. We know that multiplying a complex number by a normal complex number (i.e., with norm equal to 1) geometrically expresses a rotation; indeed, given a complex number $z = r_1 e^{i\phi}$ and a normal complex number $w = e^{i\psi}$, we have:

$$zw = r_1 e^{i(\phi+\psi)}$$

the multiplication of z by w has thus rotated z by an angle ψ. In the case of multiplying a complex number by a non-normal complex number, we simply obtain a composition of a rotation and a dilation/contraction.
Mathematically, this fact is due to the isomorphism of groups (Lie groups) constructible between the group $SO(2)$ of orthogonal matrices with determinant 1 (which we will call the special orthogonal group of degree 2) and the group of normal complex numbers $U(1) := \{e^{i\theta} \ ; \ \theta \in \mathbb{R}\}$.

Remark 3.1. *A Lie group is simply a group that also has a structure of a smooth manifold.*

The isomorphism in question is given by the following consideration: we know that we can identify a complex number with a matrix in the following way:

$$a + ib = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

As a consequence of this, we will have that a normal complex number of the type $z = e^{i\theta} = \cos\theta + i\sin\theta$ will have a matrix representation of the type:

$$z = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

but the identification just presented of a complex number with a matrix (matrix representation) is an isomorphism with the image. Therefore, we will have that the group of normal complex numbers $U(1)$ will be isomorphic to the group of matrices $\left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \ ; \ \theta \in \mathbb{R} \right\}$.
But this group is precisely $SO(2)$, and therefore we will have:

$$U(1) \cong SO(2)$$

Now we would like to demonstrate an analogous result for quaternions $\mathbb{H}$; however, we will see that the case of quaternions will be slightly different:

instead of $U(1)$, we will have $SU(2)$, instead of $SO(2)$, we will have $SO(3)$, and instead of an isomorphism, we will have a double covering, which we will define precisely later.
Let's start by defining our "candidate" for a rotation in space: given a normal quaternion $q \in S^3 = \{p \in \mathbb{H} \ ; \ |p| = 1\}$, which we will write in polar form as:

$$q = \cos\left(\frac{\theta}{2}\right) + \operatorname{sgn}(\operatorname{Vec}(q)) \sin\left(\frac{\theta}{2}\right) \tag{3.1}$$

and given a pure quaternion $v \in \mathbf{P} \cong \mathbb{R}^3$, we want to demonstrate that the application defined as:

$$\rho_q(v) = qvq^{-1} \tag{3.2}$$

produces a rotation by an angle θ around the axis of rotation given by $\operatorname{sgn}(\operatorname{Vec}(q))$. Before proving this, however, let's see some properties of this application:

Proposition 3.1 (Properties of Quaternionic Rotation)**.** *Let $q \in \mathbb{H} \setminus \{0\}$ (we will demonstrate these properties for a more general q, possibly non-normal, even if we will work only with normal quaternions for three-dimensional rotations), let $p \in \mathbb{H}$ and let $\rho_q(p)$ be defined as before by:*

$$\rho_q(p) = qpq^{-1}$$

*then this application, which we will call **rotation**, has the following properties:*

1. *The application is $\mathbb{R}$-linear, i.e:*

$$\rho_q(\lambda_1 p_1 + \lambda_2 p_2) = \lambda_1 \rho_q(p_1) + \lambda_2 \rho_q(p_2) \quad \forall\, p_1, p_2 \in \mathbb{H} \ , \ \forall \lambda_1, \lambda_2 \in \mathbb{R}$$

2. *Preservation of quaternionic multiplication:*

$$\rho_q(p_1 p_2) = \rho_q(p_1)\rho_q(p_2) \quad \forall p_1, p_2 \in \mathbb{H}$$

3. $(\rho_{q_1} \circ \rho_{q_2})(p) = \rho_{q_1 q_2}(p)$. $\forall q_1, q_2 \in \mathbb{H} \setminus \{0\}$.

4. *ρ_q is an isometric automorphism (an automorphism that preserves the norm function, i.e $|\rho_q(p)| = |p|$) of $\mathbb{H}$.*

Exercise 3.1. *Prove the properties of quaternionic rotations stated in proposition 3.1.*

Let's now prove that, for $q \in S^3$ and $p \in \mathbb{R}^3$, such an application is indeed a rotation. Let's start by proving that the special unitary group $SU(2)$, composed of matrices of the type:

$$SU(2) = \left\{ \begin{bmatrix} a+bi & c+di \\ -c+di & a-bi \end{bmatrix} \ ; \ a^2+b^2+c^2+d^2 = 1, \ ,a,b,c,d \in \mathbb{R} \right\} \tag{3.3}$$

is isomorphic to the normal quaternions living on the 3-sphere $S^3 \subset \mathbb{H}$.
Recalling the matrix representation seen in subsection 2.5.1, we notice that this group presents precisely the matrix representations of unitary quaternions, and therefore the group of unitary quaternions, geometrically visualizable as S^3, is isomorphic to the group $SU(2)$.
Unfortunately, in the quaternionic case, the Lie group $SU(2)$ will not be globally isomorphic to $SO(3)$, but only locally. This means, more specifically, that we will not have an isomorphism of Lie groups but rather a double covering, i.e., we will have a surjective continuous homomorphism $\varphi : SU(2) \to SO(3)$ that maps each unitary quaternion $q \in SU(2)$ to a rotation ρ_q ; however, this homomorphism is not injective, as it maps antipodal points of the sphere S^3, q and $-q$, to the same rotation ρ_q.

Let's now construct the homomorphism in question by proving the following theorems:

Theorem 3.1. *Let $q \in \mathbb{H} \setminus \{0\}$ be a quaternion, and let $p \in \mathbb{H}$ be another quaternion; the application ρ_q defined as:*

$$\rho_q : p \to qpq^{-1}$$

is a rotation in $SO(4)$, whose restriction to the set $\mathbf{P} \subset \mathbb{H}$ of pure quaternions is a rotation in $SO(3)$. Moreover, every rotation R in $SO(3)$ is of the form:

$$\rho_q : p \to qpq^{-1}$$

for a non-zero quaternion q and for a pure quaternion $p \in \mathbf{P}$. Moreover, if two quaternions $q_1, q_2 \neq 0$ represent the same rotation, then we have $q_1 = \lambda q_2$ for $0 \neq \lambda \in \mathbb{R}$.

Proof. Let's first prove that every rotation of $SO(3)$ is indeed of the announced form ρ_q. For the Cartan-Dieudonne theorem, we will have that every non-identity rotation is the composition of an even number of reflections. Therefore, it will be enough to show that for each reflection σ of $\mathbf{P} \cong \mathbb{R}^3$ around a plane $H \subset \mathbf{P}$, there exists a non-zero quaternion q such that $\sigma(p) = -qpq^{-1} \ \forall p \in \mathbf{P}$.

Assuming that q is a pure quaternion orthogonal to H, we know that the mathematical expression for our reflection σ around H will be of the type:

$$\sigma(p) = p - 2\frac{p \cdot q}{q \cdot q}q \quad \forall p \in \mathbf{P} \tag{3.4}$$

(where here $\cdot$ is the scalar product, to avoid confusion). But we have seen in a problem addressed in chapter 2 that, for two pure quaternions $r, s \in \mathbf{P}$, we have $2(r \cdot s) = -(rs + sr)$. Therefore, we can rewrite equation (3.4) as:

$$\sigma(p) = p + 2(p \cdot q)q^{-1} = p - (pq + qp)q^{-1} = -qpq^{-1} \tag{3.5}$$

obtaining the desired result. To prove that an application of the type $\rho_q : p \to qpq^{-1}$ is indeed a rotation of $SO(4)$ (for a generic p), we will need to explicitly derive its matrix form.
A direct calculation gives us the following expression for $\rho_q(p)$:

$$\rho_q(p) = \frac{1}{|q|^2}\{(a^2+b^2+c^2+d^2)p_0+[(a^2+b^2-c^2-d^2)p_1+2(bc-ad)p_2+2(ac+bd)p_3]i+$$
$$[2(ad + bc)p_1 + (a^2 - b^2 + c^2 - d^2)p_2 + 2(cd - ab)p_3]j+$$
$$[2(bd - ac)p_1 + 2(cd + ab)p_2 + (a^2 - b^2 - c^2 + d^2)p_3]k\}$$

Where here $p = p_0 + p_1 i + p_2 j + p_3 k$ and $q = a + bi + cj + dk$.
Identifying, through an isomorphism, $\rho_q(p)$ and p with column vectors of $\mathbb{R}^4$, we can rewrite $\rho_q(p)$ as the vector p multiplied by a matrix M_ρ that represents the operation $q\square q^{-1} : p \to qpq^{-1}$. The matrix in question (also obtainable through a direct calculation) is:

$$M_\rho = \frac{1}{||q||^2}\begin{bmatrix} ||q||^2 & 0 & 0 & 0 \\ 0 & a^2+b^2-c^2-d^2 & 2(bc-ad) & 2(ac+bd) \\ 0 & 2(ad+bc) & a^2-b^2+c^2-d^2 & 2(cd-ab) \\ 0 & 2(bd-ac) & 2(ab+cd) & a^2-b^2-c^2-d^2 \end{bmatrix}$$

Therefore:

$$\rho_q(p) = \frac{1}{||q||^2}\begin{bmatrix} ||q||^2 & 0 & 0 & 0 \\ 0 & a^2+b^2-c^2-d^2 & 2(bc-ad) & 2(ac+bd) \\ 0 & 2(ad+bc) & a^2-b^2+c^2-d^2 & 2(cd-ab) \\ 0 & 2(bd-ac) & 2(ab+cd) & a^2-b^2-c^2-d^2 \end{bmatrix}\begin{bmatrix} p_0 \\ p_1 \\ p_2 \\ p_3 \end{bmatrix}$$

The matrix M_ρ thus obtained is an orthogonal matrix belonging to the special orthogonal group $SO(4)$. If the quaternion p were to be a pure quaternion, then the matrix would be an element of $SO(3)$.

To prove the last point, instead, let's assume that given two quaternions $q_1, q_2 \in \mathbb{H} \setminus \{0\}$, $\rho_{q_1}(p) = \rho_{q_2}(p)\ \forall p \in \mathbb{H}$. In other words, we have:

$$q_1 p q_1^{-1} = q_2 p q_2^{-1}$$

Multiplying from the left by q_2^{-1} we get:

$$q_2^{-1} q_1 p = p q_2^{-1} q_1$$

i.e., $q_2^{-1} q_1 \in Z(\mathbb{H})$ (the center of quaternions). Thus, the quaternion $q_2^{-1} q_1$ commutes with $p\ \forall p \in \mathbb{H}$, hence it is an element of the center. But in chapter 2, we saw that the center of the algebra of quaternions are the real numbers, and therefore we will have $q_2^{-1} q_1 \in \mathbb{R}$. We will call this number $\frac{1}{\lambda}$. Indeed, multiplying both sides by q_2 from the left, we get:

$$q_2 = \lambda q_1$$

as we wanted to prove. □

Proposition 3.1 and theorem 3.1 allow us to construct the double covering of $SO(3)$ by $SU(2)$, as announced at the beginning of the section. Our continuous surjective homomorphism will be $\rho : SU(2) \to SO(3)$, which maps a normal quaternion $q \to \rho_q := q \square q^{-1}$. The application is obviously surjective as, as seen in theorem 3.1, every rotation of $SO(3)$ can be represented as $q \square q^{-1}$ for a pure quaternion $q \in \mathbf{P}$. The application is also a group homomorphism, as a consequence of proposition 3.1. Finally, we note that the application $\rho : SU(2) \to SO(3)$ maps antipodal points of the sphere $SU(2)$ to the same rotation $SO(3)$: indeed, for two normal quaternions $q, -q \in SU(2)$, we have $\rho_q(p) = qpq^{-1} = \rho_{-q}(p) = (-q)p(-q)^{-1}$. To verify that these points are the only cases in which two points are mapped to the same rotation, we must use the final part of theorem 3.1, which states that if two quaternions $q_1, q_2 \in \mathbb{H}$ represent the same rotation then we have $q_1 = \lambda q_2$ for $\lambda \in \mathbb{R}$. But having restricted the domain of ρ from $\mathbb{H} \setminus \{0\}$ to $SU(2)$, we will have that necessarily $|q_1| = |q_2| = 1$. But then λ can only be -1. Therefore, two different quaternions representing the same rotation must necessarily be antipodal points of $SU(2)$, i.e., one the negation of the other.

This result also allows us to inquire about the topological-geometric structure of $SO(3)$. Remember that $SU(2)$ is topologically the sphere S^3. By identifying the antipodal points of the sphere S^3 as equivalent, and constructing the quotient space under this identification, we will obtain a space homeomorphic to $SO(3)$. But the space thus constructed, i.e., the sphere

S^3 in which antipodal points are identified, is precisely the real projective space $\mathbb{RP}^3$.
The topological and differential structure of the Lie group $SO(3)$ will therefore be diffeomorphic to that of the real projective space $\mathbb{RP}^3$. Let's quickly recap the results obtained in this section:

- Let $q = \cos\frac{\theta}{2} + \mathrm{sgn}(\mathrm{Vec}(q))\sin\frac{\theta}{2}$ be a normal quaternion (i.e., a quaternion with $|q| = 1$), and let $v \in \mathbf{P}$ be a pure quaternion. Then the application $\rho_q : \mathbf{P} \to \mathbf{P}$ defined as:

$$\rho_q(v) = qvq^{-1}$$

  represents the rotation of the vector $v \in \mathbf{P} \cong \mathbb{R}^3$ by an angle θ around the axis of rotation of $\mathrm{sgn}(\mathrm{Vec}(q))$.

- Normal quaternions, geometrically $S^3 \subset \mathbb{H}$, are a Lie group isomorphic to $SU(2)$.

- The possibility of performing three-dimensional rotations with normal quaternions is explained by the existence of a continuous surjective homomorphism $\rho : SU(2) \to SO(3)$ that maps a normal quaternion q to its associated rotation function $\rho_q(p) = qpq^{-1}$. However, this homomorphism maps antipodal points of $SU(2)$ to the same rotation, and will therefore be called a **double covering** of $SO(3)$.

- $SO(3)$ is geometrically interpretable as a sphere S^3 where antipodal points have been identified with each other, or more precisely, as the quotient of the topological space S^3 under the equivalence relation that identifies antipodal points as equivalent. Such a space is $\mathbb{RP}^3$.

Exercise 3.2. *Rotate the vector* $\vec{v} = \frac{1}{\sqrt{3}}i + 2j - k$ *by an angle of* $45\,°$ *around the axis of rotation given by* $\vec{u} = \sqrt{3}i - \pi j + k$.

Exercise 3.3. *Prove that, for a quaternion* $p \in \mathbb{H}$*, if it has the same direction as* q *then the application* $\rho_q(p)$ *will not change its direction.*

3.3 Problems with Eulerian matrices

At this point, the reader may wonder why it is advantageous to use quaternions instead of Eulerian matrices to represent three-dimensional rotations. As already mentioned in the introduction, the biggest problem with

Eulerian matrices is the (in)famous **Gimbal block**. Let's start by briefly recalling the theory of Eulerian matrices; given a vector $\vec{v}$ in $\mathbb{R}^3$, it is possible to rotate it around the axes x, y or z using the following rotation matrices:

$$R_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \quad R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}$$

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

We call the angle ϕ by which we rotate our vector $\vec{v}$ around the x axis "roll" angle, the angle θ by which we rotate our vector around the y axis "pitch" angle, and finally the angle ψ "yaw" angle.
The idea is then to express every three-dimensional rotation as a composition of these matrices. However, there is a problem: you have surely seen online, perhaps in some video compilation of bugs, bizarre 3d animations where in the course of a rotation of a body there is a sudden abrupt and unrealistic movement. Also, the episode of Apollo 11 is very famous, in which an IMU ("Inertial measurement unit", i.e., inertial measurement unit) with 3 apparently independent rotating axes was used. The choice not to include a fourth one forced engineers to find other ways to avoid the problem of gimbal lock, and the solution was to include an indicator on the console that warned of the approach to "critical" angles. But what exactly is meant by critical angles? What exactly is a gimbal lock and when does it occur? Why is it such an important problem? First, let's clarify the language, as when we talk about "gimbal lock" we do not mean an actual block, but rather a "loss of a degree of freedom" (by loss of a degree of freedom here we mean the loss of a possibility to rotate around an axis).
The question to ask at this point is: how can you lose a degree of freedom? We will see that this happens when two of the rotation rings of a gimbal suspension align. Let's now confirm this fact mathematically: suppose we have a sequence xyz of rotations $R_{xyz}(\phi, \theta, \psi)$ defined as:

$$R_{xyz}(\phi, \theta, \psi) = R_x(\phi)R_y(\theta)R_z(\psi)$$

let's set the "pitch" angle θ to $\frac{\pi}{2}$, i.e., align it with the x ring. At this point, replacing the rotation matrices introduced at the beginning of the section in place of R_x, R_y and R_z, and setting $\theta = \frac{\pi}{2}$ we obtain:

$$R_{xyz}(\phi, \frac{\pi}{2}, \psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

developing the multiplication we obtain:

$$R_{xyz}(\phi, \frac{\pi}{2}, \psi) = \begin{bmatrix} 0 & 0 & 1 \\ \sin\phi\cos\psi + \cos\phi\sin\psi & -\sin\phi\sin\psi + \cos\phi\cos\psi & 0 \\ -\cos\phi\cos\psi + \sin\phi\sin\psi & \cos\phi\sin\psi + \sin\phi\cos\psi & 0 \end{bmatrix}$$

and finally, using the trigonometric sum identities we obtain:

$$R_{xyz}(\phi, \frac{\pi}{2}, \psi) = \begin{bmatrix} 0 & 0 & 1 \\ \sin(\phi+\psi) & \cos(\phi+\psi) & 0 \\ -\cos(\phi+\psi) & \sin(\phi+\psi) & 0 \end{bmatrix}$$

We observe that the algebraic form obtained confirms our geometric intuition: by aligning two rotation axes we lose a degree of freedom. Indeed now, changing ϕ and ψ, we can only rotate around the z axis. This can be visualized geometrically with the following image:

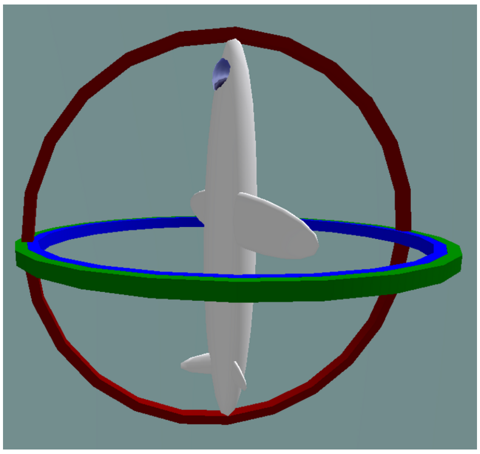


Figure 3.7: Visualization of the gimbal lock of a 3d model. [the photo in the figure is taken from Wikipedia, uploaded by user "MathsPoetry": *https : //upload.wikimedia.org/wikipedia/commons/3/38/Gimbal_lock.png* license CC BY-SA 3.0].

In the figure, in fact, two of the rings of the gimbal suspension are aligned, and therefore our object will lose a direction of possible rotation and we will say that it has lost a degree of freedom.
During the landing of the Apollo 11 spacecraft, there was a very close approach to achieving the gimbal lock of the spacecraft's IMU. From this arose the historic phrase of Michael Collins: «How about sending me a

fourth gimbal for Christmas?», referring to the possibility of having been able to avoid such an inconvenience if a fourth ring had been included. Representing three-dimensional rotations through quaternions therefore provides a barrier to the problem of gimbal lock, and this is why the latter are so popular in all those situations where one works with rotations in three-dimensional space.

3.4 Interpolation of rotations

In this section, we will investigate the following question: if we have two rotations (which, in light of the last section, we know to be represented by two normal quaternions on the sphere S^3), how can we find a continuous series of ”intermediate” rotations? This process is of great importance for computer graphics and other fields of applied science where three-dimensional rotations are used, as it will allow us to represent the rotations of entities smoothly and realistically. We will call this process **rotation interpolation**. The reader has probably already heard this word, most likely in the following instances: ”linear interpolation,” ”polynomial interpolation.” Rotation interpolation, in fact, just like linear interpolation or polynomial interpolation, aims to connect a finite number of points through segments and, respectively, through a polynomial (of degree equal to the number of points minus 1), aims to connect two rotations with each other.
Let’s now recall the formula for linear interpolation (which we will often abbreviate as ”LERP”) of two points in space x_1 and x_2:

$$x(t) = (1 - t)x_1 + tx_2 \quad 0 \leq t \leq 1 \tag{3.6}$$

This method of interpolation can be geometrically visualized as a segment connecting the two points x_1, x_2. The intermediate positions between the two live on this segment, as shown in the figure below:

Let’s now return to our initial situation: we have two normal quaternions q_1, q_2 living on the 3-sphere S^3. Ideally, we want to be able to interpolate them in a continuum of intermediate quaternions, also living on the 3-sphere S^3, that are equally angularly spaced (more suggestively, we want this interpolation to have ”constant angular velocity”).
We see that interpolating them linearly will result in the ”intermediate” quaternions not being on S^3, as also suggested by the figure below (reduced to the two-dimensional case for simplicity):

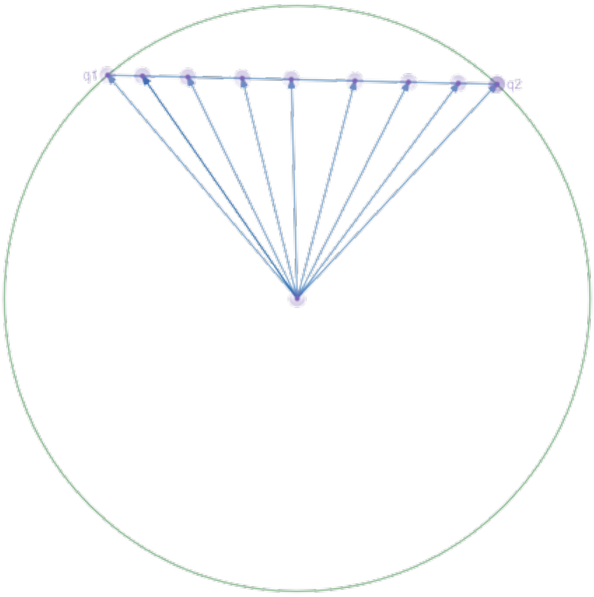


Figure 3.8: Visualization of linear interpolation of points on a sphere.

We can mathematically verify this by calculating the norm of an intermediate quaternion $q(t) = (1-t)q_1 + tq_2$ at a moment t:

$$|q(t)|^2 = q(t) \cdot q(t) = (1-t)^2 q_1 \cdot q_1 + t(1-t)q_1 \cdot q_2 + t(1-t)q_2 \cdot q_1 + t^2 q_2 \cdot q_2$$

where here $\cdot$ is the scalar product of quaternions. Since by hypothesis q_1 and q_2 are normal quaternions, we have that $|q_1| = 1 = q_1 \cdot q_1$ and $|q_2| = 1 = q_2 \cdot q_2$. Also, calling ϕ the angle between the two, such that $q_1 \cdot q_2 = \cos\phi$, we get:

$$|q(t)|^2 = 1 + 2t^2 - 2t + 2t(1-t)\cos\phi$$

which confirms our initial assertion. At this point, we might think of projecting $q(t)$ onto S^3, and although doing so will give us a normal quaternion living on the hypersphere, we will not obtain an interpolation of quaternions equally angularly spaced, as we can see in the graph below.

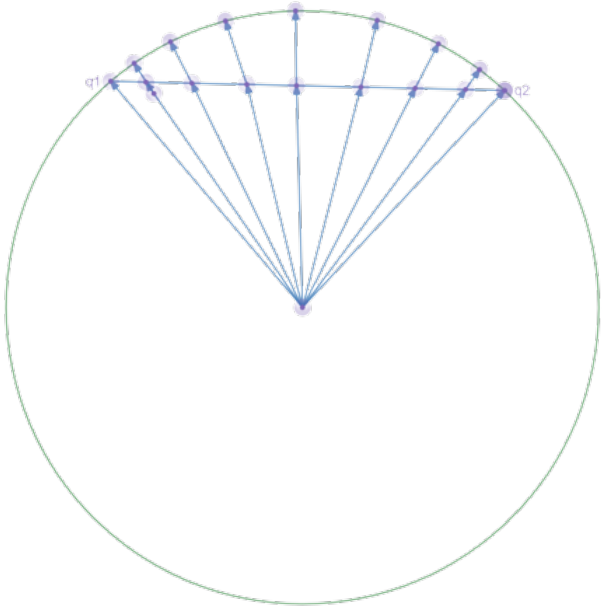


Figure 3.9: Projection of the "intermediate quaternions" between q_1 and q_2 obtained by linear interpolation of quaternions onto the sphere S^3.

To achieve our goal, we will need to use the technique of spherical interpolation (which we will often call "SLERP" by default); we will derive it using the Gram-Schmidt procedure, which the reader has probably already encountered in a linear algebra course. As before, let q_1 and q_2 be two normal quaternions; let's start by finding, with the Gram-Schmidt orthogonalization method, a normal quaternion orthogonal to q_1, which we will call q':

$$q' = \frac{q_2 - q_1(q_1 \cdot q_2)}{|q_2 - q_1(q_1 \cdot q_2)|}$$

Note that the denominator of the expression above has the following property:

$$|q_2 - q_1(q_1 \cdot q_2)|^2 = -2\cos^2\phi + \cos^2\phi = \sin^2\phi$$

where here ϕ is the angle between q_1 and q_2. From this, it follows that $|q_2 - q_1(q_1 \cdot q_2)| = |\sin\phi|$. For simplicity, we require that $0 \leq \phi < \pi$, so that we have $\sin\phi = |\sin\phi|$. The system of two orthonormal vectors $\{q_1, q'\}$ will allow us to rotate our initial orthonormal quaternion q_1 in a range of values between q_1 and q_2 while maintaining a "constant angular velocity"; we can visualize the situation through the following illustration (simplified):

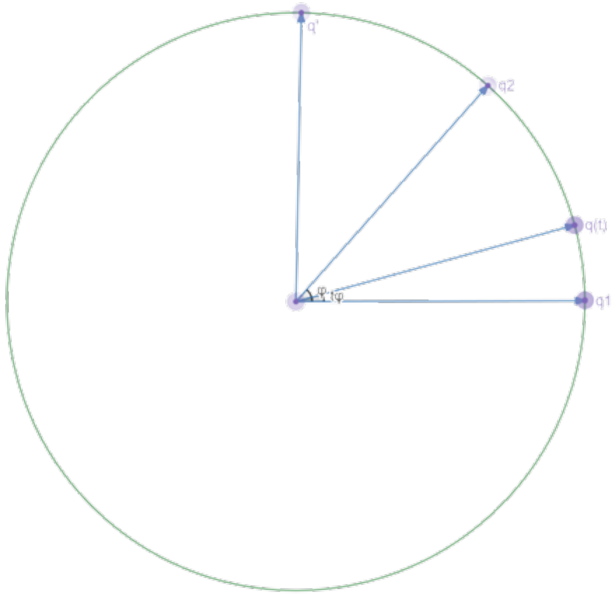


Figure 3.10: Visualization of the spherical interpolation procedure of quaternions on S^3.

Let's set our intermediate quaternion $q(t)$ as a function of t equal to:

$$q(t) = q_1 \cos(t\phi) + q' \sin(t\phi) \quad 0 \leq t \leq 1 \tag{3.7}$$

recalling that $q' = \dfrac{q_2 - q_1(q_1 \cdot q_2)}{|q_2 - q_1(q_1 \cdot q_2)|} = \dfrac{q_2 - q_1 \cos\phi}{\sin\phi}$, we can substitute this in equation (3.7), and developing the calculations we get:

$$q(t) = q_1 \frac{\cos(t\phi)\sin\phi - \sin(t\phi)\cos\phi}{\sin\phi} + q_2 \frac{\sin(t\phi)}{\sin\phi}$$

Finally, using the sum formulas of the sine, we arrive at the final form of equation (3.7):

$$q(t) = q_1 \frac{\sin(\phi - t\phi)}{\sin\phi} + q_2 \frac{\sin(t\phi)}{\sin\phi} \quad 0 \leq t \leq 1 \tag{3.8}$$

The equation above is the formula for spherical interpolation of two normal quaternions q_1, q_2.

Exercise 3.4. *Let $q(t)$ be defined as in equation* (3.8)*: verify that $|q(t)| = 1$ $\forall t \in [0, 1]$.*

Chapter 4

Topology of Quaternions

In this fourth chapter, we will study quaternions from a topological perspective. As we will see, unlike their somewhat bizarre algebraic behavior, quaternions will have a canonical topological structure (i.e., one induced by the metric naturally assignable to a well-normed algebra) with many desirable properties. The canonical topology of quaternions is a complete metric space, homeomorphic to $(\mathbb{R}^4, \mathscr{U})$, and the definition of quaternionic analogs of the limit of a sequence, limit of a function, continuous function, etc., will be easily realizable, and will obey almost all the properties found in their real and complex counterparts.

4.1 Construction of $(\mathbb{H}, \mathscr{U})$

Let us begin by constructing the canonical topology on the set of quaternions. We proved in the second chapter that the function $\delta(q_1, q_2) = |q_1 - q_2|$, $\delta : \mathbb{H} \times \mathbb{H} \longrightarrow \mathbb{R}$ was a metric on $\mathbb{H}$, i.e. it satisfied the following axioms:

1. $\delta(q_1, q_2) \geq 0$.
2. $\delta(q_1, q_2) = \delta(q_2, q_1)$.
3. $\delta(q_1, q_3) \leq \delta(q_1, q_2) + \delta(q_2, q_3)$.
4. $\delta(q_1, q_2) = 0 \Longleftrightarrow q_1 = q_2$.

$\forall q_1, q_2 \in \mathbb{H}$.
Readers with rudiments of topology will certainly remember the construction of a topology on a set given a metric, but we will repeat it here anyway. Let's begin by defining the notion of open ball and closed ball in $\mathbb{H}$ with respect to the canonical metric δ.

Definition 4.1 (Open and Closed Ball in $\mathbb{H}$)**.** *Let $q \in \mathbb{H}$ be a quaternion and $r > 0$ be a positive real number. Then we will call the following sets open and closed ball of center q and radius r:*

$$B(q,r) = \{p \in \mathbb{H} \mid \delta(q,p) < r\} = \{p \in \mathbb{H} \mid (t_1 - t_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 < r^2\}$$

$$\bar{B}(q,r) = \{p \in \mathbb{H} \mid \delta(q,p) \leq r\} = \{p \in \mathbb{H} \mid (t_1 - t_2)^2 + (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 \leq r^2\}$$

Geometrically, as also made explicit by the condition imposed on the t, x, y, z components of a quaternion belonging to an open/closed ball, they represent respectively the interior of a hypersphere (3-sphere), and the interior with border of a hypersphere.
Let's now prove some properties of open balls in $\mathbb{H}$, however generally valid also for open balls in more general sets equipped with a metric.

Proposition 4.1. *Let $B(p,\epsilon)$ be an open ball with center p and radius ϵ. Then $\forall q \in B(p,\epsilon)$, there exists an open ball $C(q,\rho)$ with center in q contained in $B(p,\epsilon)$, $C(q,\rho) \subset B(p,\epsilon)$.*

Proof. Let $q \in B(p,\epsilon)$, then by the definition we gave of open ball, we will have $\delta(q,p) < \epsilon \Longrightarrow \epsilon - \delta(q,p) > 0$.
Let's call this positive real number ρ, $\rho = \epsilon - \delta(q,p)$. Let's now prove that $C(q,\rho)$ is a subset of $B(p,\epsilon)$. Let $x \in C(q,\rho)$, then we will require that $\delta(x,q) < \rho = \epsilon - \delta(q,p)$.
Using the triangle inequality we have:

$$\delta(x,p) \leq \delta(x,q) + \delta(q,p) < (\epsilon - \delta(q,p)) + \delta(q,p) = \epsilon$$

so $x \in B(p,\epsilon)$, i.e. $C(q,\rho) \subset B(p,\epsilon)$. □

Lemma 4.1. *Let $B_1(q,\epsilon)$ and $B_2(q,\rho)$ be two open quaternionic balls centered at the same point $q \in \mathbb{H}$. Then either $B_1 \subset B_2$ or $B_2 \subset B_1$.*

Proof. Since ϵ and ρ are two positive real numbers, we will have that $\epsilon \leq \rho$ or $\rho \leq \epsilon$, given that $\mathbb{R}$ is a totally ordered set (with respect to the canonical ordering $\leq$). From this it immediately follows that one of the two balls must be contained in the other. □

Lemma 4.2. *Let B_1, B_2 be two open quaternionic balls, and let $q \in B_1 \cap B_2$. Then there exists an open ball with center q, $B_3(q,\rho)$, such that $q \in B_3 \subset B_1 \cap B_2$.*

Proof. Let $q \in B_1 \cap B_2$, then $q \in B_1$ and $q \in B_2$. By proposition 4.1, we know that in both cases there are balls centered at q P_1, P_2 contained in B_1 and B_2 respectively. However, having the same center, as demonstrated previously, we will have that $P_1 \subset P_2 \subset B_2$ or $P_2 \subset P_1 \subset B_1$. In the first case we will have $B_3 = P_1 \subset B_1 \cap B_2$, while in the second $B_3 = P_2 \subset B_1 \cap B_2$. □

Graphically, we can visualize the lemma just demonstrated in the following way:

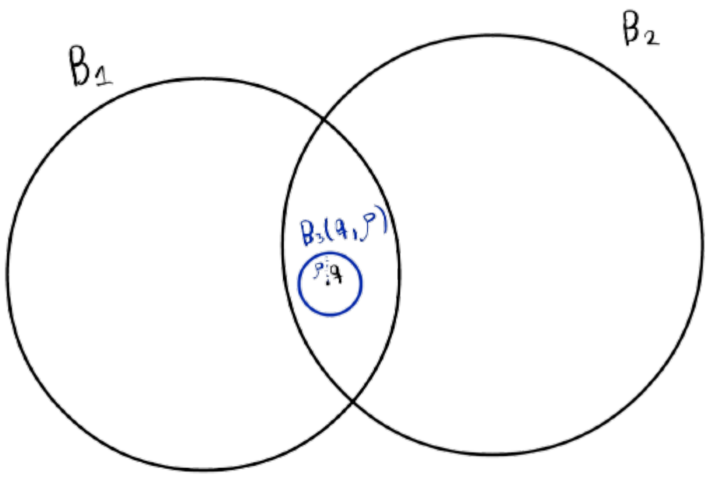


Figure 4.1: Visualization of lemma 4.2.

These two preparatory propositions allow us to very quickly prove the following fact:

Theorem 4.1. *The set of open balls in* $\mathbb{H}$ $\Sigma_{\mathbb{H}} = \{B(q, \epsilon) | \ q \in \mathbb{H}, \epsilon > 0\}$ *with respect to the metric* δ *forms a basis for a topology on* $\mathbb{H}$*. Such topology is the metric topology of* $\mathbb{H}$*, which we will call the canonical topology of quaternions, and denote* $(\mathbb{H}, \mathscr{U})$*.*

The topology whose open sets are given by unions of open balls with respect to a metric is generally inducible on all sets equipped with a metric δ; we will call these spaces metric spaces, and denote them as (X, τ, δ). Being a metric space is the "ideal" property for setting up analysis on a space, as it allows us to define the notions of analysis very similarly to the real and complex counterparts.

Metric spaces are also considered at the apex of the "hierarchy" of spaces, if we order them in a hierarchy according to the various separation axioms (Kolmogorov Classification).

Explanatory Note: The separation axioms are axioms that "give structure" to a topological space, i.e. require that some extra properties be satisfied that make the structure more interesting. Some important examples are the T_2 axiom; spaces that respect it are called Hausdorff spaces, of great importance for the study of smooth manifolds as we saw in the introduction. There are many others, for example the T_0, T_1, T_3 axioms, etc. Why are they named using an increasing natural number? The answer lies in propositions generally demonstrated in an introductory course in general topology; every T_6 space is T_5, and in turn every T_5 space is T_4, and in turn every T_4 space is T_3 and so on. Spaces are then "hierarchized" based on the axioms respected in light of this "chain of relations", according to what is called "Kolmogorov Classification", illustrated below:

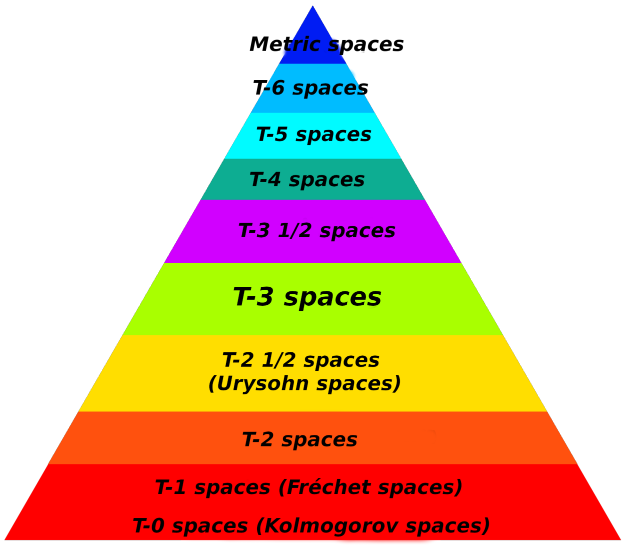


Figure 4.2: "Hierarchy" of topological spaces visualized through a pyramid.

At the apex as we see there are metric spaces. They indeed respect all the underlying properties, plus those extra exclusive to their structure.

From now on we will denote the topology just constructed on $\mathbb{H}$, $(\mathbb{H}, \mathscr{U}, \delta)$ as simply $\mathbb{H}$, and when we refer to open/closed in $\mathbb{H}$, we will always mean with respect to the metric topology in question, unless previously specified. Let's now prove another fact about the just constructed canonical topology of quaternions. Although very simple to demonstrate, we will include a proof here.

Theorem 4.2. *Let* $(\mathbb{H}, \mathscr{U})$ *be the canonical metric topology of quaternions and* $(\mathbb{R}^4, \mathscr{U})$ *be the metric topology induced by the 4-Euclidean metric on* $\mathbb{R}^4$*. Then the function* $\varphi : \mathbb{H} \longrightarrow \mathbb{R}^4$ *defined in the following way:*

$$\varphi(a + bi + cj + dk) = [a, b, c, d] \in \mathbb{R}^4$$

is an isometric homeomorphism of metric spaces. Thus $(\mathbb{H}, \mathscr{U}) \cong (\mathbb{R}^4, \mathscr{U})$*.*

Proof. Let's begin the proof by showing that φ is bijective: suppose we have two equal quaternions $q_1 = q_2$, then their components are equal $a_1 = a_2, b_1 = b_2$ and so on; we have $\varphi(q_1) = [a_1, b_1, c_1, d_1]$ and $\varphi(q_2) = [a_2, b_2, c_2, d_2]$, and since the components of the quaternions are equal, we will have that the elements of the image quaternions of $\mathbb{R}^4$ will have their components orderly equal, i.e. $\varphi(q_1) = \varphi(q_2)$. The inverse is equally true: $\varphi(q_1) = \varphi(q_2) \Longrightarrow a_1 = a_2$, $b_1 = b_2$, $c_1 = c_2, d_1 = d_2$, and therefore $q_1 = q_2$. We have proved that φ is well defined (necessary condition for it to be a function) and injective. Suppose now we have a point of $\mathbb{R}^4$, $[a, b, c, d]$, $a, b, c, d \in \mathbb{R}$. There is then a $q \in \mathbb{H}$, expandable in components as $q = a + bi + cj + dk$, such that $\varphi(q) = [a, b, c, d]$, therefore φ is also surjective.

It remains to show that φ is bicontinuous, i.e. that φ and φ^{-1} are both continuous functions, or alternatively that φ is continuous and open.

Let $B(q, \epsilon)$ be an open quaternionic ball (an element of the basis of $(\mathbb{H}, \mathscr{U})$). Its image $\varphi(B)$ is an open ball of $\mathbb{R}^4$, $P(\varphi(q), \epsilon)$, thus an element of the basis of $(\mathbb{R}^4, \mathscr{U})$. Similarly, given an open ball of $\mathbb{R}^4$, its image under φ^{-1} (i.e. its pre-image under φ) is an open quaternionic ball, i.e. an element of the basis of $\mathbb{H}$, and thus φ is a homeomorphism. We leave the simple task of verifying that it is also an isometry, i.e. an application that preserves distances between 2 metric spaces, to the reader. $\square$

4.2 Quaternion Sequences

Now, let's study quaternion sequences, which we will define in a manner entirely analogous to complex and real sequences.

Definition 4.2. *A* ***quaternion sequence*** *is a function* $\varphi : \mathbb{N} \longrightarrow \mathbb{H}$*; for* $n \in \mathbb{N}$*, we'll call* $\varphi(n) = q_n$ *its value calculated at* n*, and we'll denote sequences simply by their image,* $\mathrm{Im}(\varphi) = \{q_n\}_{n=1}^{\infty}$.

Example 4.1. *Here are some examples of quaternion sequences:*

- $\{j^n\}_{n=1}^{\infty} = \{j, -1, -j, 1, j, -1, \dots\}$
- $\{\frac{n+1}{2}(i+j)^n\}_{n=1}^{\infty} = \{i+j, -3, -4i-4j, \dots\}$
- $\{n+ni-nj\}_{n=1}^{\infty} = \{1+i-j, 2+2i-2j, \dots\}$
- $\{\mathrm{Vec}((2+j+k)^n)\}_{n=1}^{\infty} = \{j+k, 4j+4k, 10j+10k, \dots\}$

Notation: From now on, we will write a quaternion sequence $\{q_n\}_{n=1}^{\infty}$ as simply $\{q_n\}$ to simplify the notation.
Note that we can write a quaternion sequence $\{q_n\}$ as $\{q_n\} = \{a_n\} + \{b_n\}i + \{c_n\}j + \{d_n\}k$, where $\{a_n\}, \{b_n\}, \{c_n\}, \{d_n\}$ are real sequences.

Example 4.2. *In the example presented earlier,* $\{j^n\}$ *can be written as* $\{j^n\} = \{\alpha_n\}+\{\beta_n\}j$*, with* $\alpha_n = \begin{cases} 0 & \textit{if } n \equiv 1(\textit{mod } 2)\,, \\ 1 & \textit{if } n \equiv 0(\textit{mod } 4)\,, \\ -1 & \textit{otherwise} \end{cases}$ *and* $\beta_n = \begin{cases} 0 & \textit{if } n \equiv 0(\textit{mod } 2)\,, \\ -1 & \textit{if } n \equiv 0(\textit{mod } 3)\,, \\ 1 & \textit{otherwise} \end{cases}$

Exercise 4.1. *Write the sequence* $\{q_n\} = \{(2nk+1)(i+nj-nk)(\mathrm{Sc}((n-j)(n+k)))\}$ *in the form* $\{q_n\} = \{a_n\} + \{b_n\}i + \{c_n\}j + \{d_n\}k$

Exercise 4.2. *Write the first 5 terms of the sequence* $\{(nj+nk)\times\mathrm{Vec}((i+j)^n)\}$

4.2.1 Unary and Binary Operations on Quaternionic Sequences

We define the following unary and binary operations for quaternionic sequences:

- Addition and subtraction:
  $\{q_n\} \pm \{r_n\} = \{q_n \pm r_n\}$

- Multiplication:
  $\{q_n\}\{r_n\} = \{q_n r_n\}$
- Conjugation:
  $\{q_n\}^* = \{q_n^*\}$
- Norm:
  $|\{q_n\}| = \{|q_n|\}$

Exercise 4.3. *Prove that, under these operations, the set of quaternionic sequences with non zero entries (except the all zero sequence, which will serve as the additive identitity) forms a skew-field.*

Exercise 4.4. *Let* $\{q_n\} = \{(n^2 j - j + k) \times (n^2 k - i)\}$, $\{r_n\} = \{(ni - nk)((i - nj) \times (k + ni))\}$, *calculate:*

- $\{q_n\}^* + \{r_n\} - \{1 + nj\}$
- $\{r_n\}^* + \{q_n\}$
- $|\{q_n\}|$
- $|\{r_n\}|$
- $\{q_n\}\{r_n\}$

4.2.2 Convergence of Quaternionic Sequences

We now introduce a notion of convergence for quaternionic sequences.

Definition 4.3. *Let* $\{q_n\}$ *be a quaternionic sequence; we say that this sequence converges to the quaternion* $q \in \mathbb{H}$*, writing:*

$$\lim_{n \to +\infty} q_n = q$$

if for every $\epsilon > 0$*, there exists an integer* N *such that* $|q_n - q| < \epsilon$ *for all* $n \geq N$*.*

This definition is analogous to its real and complex counterparts; in fact, they are all special cases of the notion of the limit of a sequence in a metric space.
Having introduced the notion of the limit of a convergent sequence in $\mathbb{H}$, we proceed to prove some algebraic properties regarding limits:

Proposition 4.2. *Every convergent quaternionic sequence is bounded.*

Proof. Require that q_n converges, and that it converges to q; then for every $\epsilon > 0$, there exists an integer N such that for all $n \geq N$ $|q_n - q| < \epsilon$.
Since ϵ is an arbitrary positive real number, we choose $\epsilon = 1$, there will exist an integer $\bar{N}$ such that:

$$|\ |q_n| - |q|\ | \leq |q_n - q| < 1\ ,\ \forall n \geq \bar{N}$$

Manipulating the inequality using properties of the absolute value, we get:

$$|q| - 1 < |q_n| < 1 + |q|$$

We have shown that the norm of each member of the sequence with n greater than or equal to N is less than the real number $|q| + 1$. For the indices between 1 and N, however:

$$|q_n| \leq \max\{|q_1|, |q_2|, |q_3|, \ldots, |q_N|, |q| + 1\}$$

Therefore, q_n is bounded. □

Theorem 4.3. *Let $\{q_n\}$ and $\{r_n\}$ be two quaternionic sequences with $\lim\limits_{n\to+\infty} q_n = q$ and $\lim\limits_{n\to+\infty} r_n = r$. Then the following algebraic properties hold:*

- $\lim\limits_{n\to+\infty} q_n \pm r_n = q \pm r$.
- $\lim\limits_{n\to+\infty} q_n r_n = qr$.
- $\lim\limits_{n\to+\infty} r_n^* = r^*$.
- $\lim\limits_{n\to+\infty} |r_n| = |r|$.
- $\lim\limits_{n\to+\infty} r_n^{-1} = r^{-1}$.

Proof.
- We assume that q_n and r_n converge. Let $\epsilon > 0$ be a positive real number; since q_n and r_n converge, we can make $|q_n - q|$ and $|r_n - r|$ less than $\frac{\epsilon}{2}$, for all $n \geq N_1$ and $n \geq N_2$ respectively, by choosing two integers called N_1 and N_2.
  Let $\bar{N} = \max\{N_1, N_2\}$, thus for all $n \geq \bar{N}$:

$$|(q_n \pm r_n) - (q \pm r)| = |(q_n - q) \pm (r_n - r)| \leq |q_n - q| + |r_n - r| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

- Assuming that r_n and q_n are convergent sequences, we can select 3 integers N_1, N_2, N_3 such that:

$$|q_n - q| < \frac{\epsilon}{2(1+|r|)} \quad \forall n \geq N_1$$

$$|r_n - r| < \frac{\epsilon}{2(1+|q|)} \quad \forall n \geq N_2$$

$$|r_n - r| < 1 \quad \forall n \geq N_3$$

where here $\epsilon > 0$. For the last of the three equations, we have that:

$$|r_n| = |r_n - r + r| \leq |r_n - r| + |r| < 1 + |r| \quad \forall n \geq N_3$$

Let $\overline{N} := \max\{N_1, N_2, N_3\}$, observe that, for all $n \geq \overline{N}$:

$$|q_n r_n - qr| = |q_n r_n - qr_n + qr_n - qr| \leq |q_n r_n - qr_n| + |qr_n - qr|$$

applying properties of the quaternion norm, we get:

$$|q_n r_n - qr_n| + |qr_n - qr| = |q_n - q||r_n| + |q||r_n - r|$$
$$< (1+|r|)\frac{\epsilon}{2(1+|r|)} + (1+|q|)\frac{\epsilon}{2(1+|q|)} = \epsilon$$

- Write r as $r = a + bi + cj + dk$ and $r_n = a_n + b_n i + c_n j + d_n k$.
  Suppose that $\lim_{n\to\infty} r_n = r$, then for every $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that for all $n \geq N$:

$$\sqrt{(a_n - a)^2 + (b_n - b)^2 + (c_n - c)^2 + (d_n - d)^2} = |r_n - r| < \epsilon$$

  Notice that $|r_n^* - r^*| = \sqrt{(a_n - a)^2 + (-b_n + b)^2 + (-c_n + c)^2 + (-d_n + d)^2} = |r_n - r|$, and is therefore less than ϵ for every n greater than the N selected previously thanks to the assumed convergence of r_n.

- For every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that for all $n \geq N$, $|r_n - r| < \epsilon$. But since $||r_n| - |r|| \leq |r_n - r|$, we have that $\lim_{n\to\infty} |r_n| = |r|$

- Remember that for every $q \in \mathbb{H} - \{0\}$, $q^{-1} = \frac{q^*}{||q||^2}$.
  Suppose that $\lim_{n\to\infty} r_n = r$. Then both r_n^* and $|r_n|$ will be convergent sequences, precisely converging to r^* and $|r|$. Then:

$$\lim_{n\to\infty} r_n^{-1} = \lim_{n\to\infty} \frac{r_n^*}{|r_n|} = \left(\lim_{n\to\infty} \frac{1}{|r_n|}\right)(\lim_{n\to\infty} r_n^*) = \frac{r^*}{|r|} = r^{-1} \tag{4.1}$$

□

If a quaternionic sequence does not converge, then we say that it is divergent.

Exercise 4.5. *Determine whether the following quaternionic sequences converge or not:*

- $q_n = \frac{n}{\sqrt[3]{n+1}}(j-k) + \frac{i}{n^2}$
- $q_n = n - n^3 i - n^6 j - n^9 k$

Theorem 4.4. *Let* $q_n = a_n + b_n i + c_n j + d_n k$ *be a quaternionic sequence and* $q = a + bi + cj + dk$ *a quaternion. Then,*

$$\lim_{n\to\infty} q_n = q \tag{4.2}$$

if and only if

$$\lim_{n\to\infty} a_n = a \ , \ \lim_{n\to\infty} b_n = b \ , \ \lim_{n\to\infty} c_n = c \ , \ \lim_{n\to\infty} d_n = d$$

Proof. Start with the direct implication ($\Longrightarrow$); suppose that $\lim\limits_{n\to\infty} q_n = q$, then for every $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that:

$$\sqrt{(a_n-a)^2 + (b_n-b)^2 + (c_n-c)^2 + (d_n-d)^2} < \epsilon$$

Observe that:

$$\begin{cases} |a_n - a| \leq \sqrt{(a_n-a)^2 + (b_n-b)^2 + (c_n-c)^2 + (d_n-d)^2} < \epsilon \\ |b_n - b| \leq \sqrt{(a_n-a)^2 + (b_n-b)^2 + (c_n-c)^2 + (d_n-d)^2} < \epsilon \\ |c_n - c| \leq \sqrt{(a_n-a)^2 + (b_n-b)^2 + (c_n-c)^2 + (d_n-d)^2} < \epsilon \\ |d_n - d| \leq \sqrt{(a_n-a)^2 + (b_n-b)^2 + (c_n-c)^2 + (d_n-d)^2} < \epsilon \end{cases}$$

Thus, the sequences a_n, b_n, c_n, d_n are convergent real sequences in a, b, c, d respectively.
Now prove the converse ($\Longleftarrow$). Let $\epsilon > 0$; having assumed the convergence of the 4 real sequences a_n, b_n, c_n, d_n in a, b, c, d, we can find 4 integers N_1, N_2, N_3, N_4 such that:

$$|a_n - a| < \frac{\epsilon}{4} \ \ \forall n \geq N_1 \ \ ; \ \ |b_n - b| < \frac{\epsilon}{4} \ \ \forall n \geq N_2$$

$$|c_n - c| < \frac{\epsilon}{4} \ \ \forall n \geq N_3 \ \ ; \ \ |d_n - d| < \frac{\epsilon}{4} \ \ \forall n \geq N_4$$

Set $N = \max\{N_1, N_2, N_3, N_4\}$. For all $n \geq N$ we have:

$$\sqrt{(a_n-a)^2} + \sqrt{(b_n-b)^2} + \sqrt{(c_n-c)^2} + \sqrt{(d_n-d)^2} = |a_n-a| + |b_n-b| + |c_n-c| + |d_n-d| < \epsilon$$

But

$$\sqrt{(a_n-a)^2+(b_n-b)^2+(c_n-c)^2+(d_n-d)^2} \leq \sqrt{(a_n-a)^2}+\sqrt{(b_n-b)^2}+\sqrt{(c_n-c)^2}+\sqrt{(d_n-d)^2}$$

Therefore, $\sqrt{(a_n-a)^2+(b_n-b)^2+(c_n-c)^2+(d_n-d)^2} < \epsilon$, that is $\lim_{n\to\infty} q_n = q$. □

Corollary 4.1. *Let q_n be a quaternionic sequence, then it converges to $q = a + bi + cj + dk$ if and only if $\mathrm{Sc}(q_n)$ converges to $\mathrm{Sc}(q) = a$ and $\mathrm{Vec}(q_n)$ converges to $\mathrm{Vec}(q) = v = bi + cj + dk$.*

We now have a tool to more easily determine the limits of quaternionic sequences, reducing it to a problem in terms of real sequences.

Example 4.3. *We have the sequence $q_n = (1+\frac{1}{n})^n[(\frac{j}{n}+\frac{k}{n^2})\times(\frac{i}{n^3}-\frac{k}{n})]$. First, let's reduce it to the more manageable form $q_n = a_n + b_n i + c_n j + d_n k$.*

$$q_n = (1+\frac{1}{n})^n(\frac{j}{n^5}-\frac{i}{n^2}-\frac{k}{n^4})$$

From here we can trivially deduce that it converges to 0, $\lim_{n\to\infty} q_n = 0$.

Here is a slightly less trivial example.

Example 4.4. *We have the sequence $q_n = \frac{n}{\sqrt[n]{n!}}(i+j+\frac{k}{n})(i+\frac{j}{n}+k)$. Expanding the product, we get:*

$$q_n = \frac{n}{\sqrt[n]{n!}}\Big(-1-\frac{2}{n}+i\Big(1-\frac{1}{n^2}\Big)+j\Big(\frac{1}{n}-1\Big)+k\Big(\frac{1}{n}-1\Big)\Big)$$

Remembering that $\lim_{n\to\infty}\frac{n}{\sqrt[n]{n!}} = e$, we have that:

$$\lim_{n\to\infty} q_n = -e + ei - ej - ek$$

Exercise 4.6. *Determine whether the following sequences converge, and if they do, find their limit*

1. $q_n = \arctan(n)i - \frac{\arctan(n)^2}{2}k$
2. $q_n = (1+\frac{1}{n})^n(i-j)(\frac{i}{n}+k)$
3. $q_n = (i+j+k)^n$

4.2.3 Subsequences of a Quaternion Sequence

The definition is completely analogous to the real and complex counterparts.

Definition 4.4. *Let $\{q_n\}$ be a quaternion sequence. We say that $\{q_{m(n)}\}$ is a subsequence of $\{q_n\}$ if $\{m(n)\}_{n=1}^{\infty}$ is a strictly increasing sequence of natural numbers.*

Exercise 4.7. *Prove that, given a quaternion sequence $\{q_n\}$ converging to q ($\lim\limits_{n\to\infty} q_n = q$), every subsequence of it also converges to q.*

4.3 Quaternionic cauchy sequences

Definition 4.5 (Cauchy Sequence of Quaternions)**.** *Let q_n be a quaternion sequence, we say that it is a Cauchy sequence if $\forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that:*

$$|q_n - q_m| < \epsilon \ \ \forall n, m \geq N$$

Proposition 4.3. *Every convergent quaternion sequence is a Cauchy sequence.*

Proof. Let $\epsilon > 0$; we know that there exists a natural number N, such that $\forall n \geq N$, $|q_n - q| < \frac{\epsilon}{2}$.
Let $m, n \in \mathbb{N}$, $m, n \geq N$. We have

$$|q_m - q| < \frac{\epsilon}{2} \ \ and \ \ |q_n - q| < \frac{\epsilon}{2}$$

But by the triangle inequality:

$$|q_m - q_n| = |q_n - q - q_m + q| \leq |q_n - q| + |q_m - q| < \epsilon$$

Thus q_n is a Cauchy sequence. □

The proposition just proved is generally valid for metric spaces. However, we will see that the quaternion metric space is a special type of metric space, called a complete metric space, where the converse of the proposition is also true, that is, every Cauchy sequence is convergent.

Definition 4.6 (Complete Metric Space)**.** *Let (X, τ, δ) be a metric space. It is called a complete metric space if all Cauchy sequences of said space are convergent.*

We want to prove that quaternions are a complete metric space. To do this, however, we first need a preparatory lemma:

Lemma 4.3. *Let $q_n = a_n + b_n i + c_n j + d_n k$ be a Cauchy quaternion sequence, then a_n, b_n, c_n, d_n are Cauchy real sequences.*

Proof. Knowing that q_n is a Cauchy sequence, we know that $\forall \epsilon > 0$, $\exists N$ natural such that $\forall n, m \geq N$:

$$|q_n - q_m|^2 = |a_n - a_m|^2 + |b_n - b_m|^2 + |c_n - c_m|^2 + |d_n - d_m|^2 < \epsilon^2$$

And so, $\forall n, m \geq N$:

$$\begin{cases} |a_n - a_m|^2 < \epsilon^2 \\ |b_n - b_m|^2 < \epsilon^2 \\ |c_n - c_m|^2 < \epsilon^2 \\ |d_n - d_m|^2 < \epsilon^2 \end{cases}$$

thus a_n, b_n, c_n, d_n are all Cauchy real sequences. □

Theorem 4.5 (Quaternions are a Complete Metric Space)**.** *Quaternions, equipped with their canonical topology, are a complete metric space.*

Proof. Let q_n be a Cauchy quaternion sequence. As we saw earlier, it can be expressed as $q_n = a_n + b_n i + c_n j + d_n k$ where a_n, b_n, c_n, d_n are real sequences. As assured by the proposition proven earlier, they are all Cauchy real sequences, therefore they converge as $\mathbb{R}$ is a complete metric space, and we have seen that a sequence q_n converges if and only if all its "real component sequences" converge in turn. Hence we deduce that q_n converges. □

That is, for quaternion sequences, being a Cauchy sequence is a sufficient and necessary condition for them to be convergent.

Corollary 4.2. *Let (q_n) be a quaternion sequence. Then it is a convergent sequence if and only if it is a Cauchy sequence.*

Quaternions, therefore, like reals and complex numbers, are a complete metric space. This brings several good additional properties to our structure that will be helpful for the study of analysis in this space. For example, the Heine-Borel theorem is also valid for quaternions, in other words, boundedness and total boundedness coincide as notions. Our current goal is indeed to prove precisely this result, and then to continue with our dissertation introducing new entities of interest.

4.4 The Heine-Borel Theorem for Quaternions

We could introduce the Heine-Borel theorem directly, taking for granted some results typically proved in an introductory course in General Topology. However, to make the text more accessible and less dispersive, we will report below the proofs of all preparatory propositions and definitions that we will use in the demonstration of the Heine-Borel theorem.

Definition 4.7 (Diameter of a Metric Space)**.** *Let (X, τ, δ) be a metric space. We call the diameter of X:*

$$\operatorname{diam}(X) = \sup_{a,b \in X} \delta(a, b)$$

Remark 4.1. *The word diameter etymologically derives from Greek, dia (across) + metron (measure), which literally translates to "measure across". The quantity just defined is exactly a generalization to abstract spaces (with a geometric visualization in many cases not possible) of the classic concept of the diameter of a circle or sphere, which is exactly the longest chord; the diameter of a circle or sphere, in fact, measures the distance between the 2 farthest points on its surface with respect to the Euclidean metric. Moreover, in the canonical topology of a Euclidean space it coincides precisely with the classic geometric definition. By inducing the property of the diameter of a circle being the measure of the distance between the two farthest points on its surface, and choosing as a distance not the Euclidean metric, but a general metric, we have managed to give a definition for more general sets equipped with more general metrics.*

It is easily verifiable that the diameter of the canonical topology on $\mathbb{H}$ is ∞.

Definition 4.8 (Bounded Space/Set)**.** *Let (X, τ, δ), then we will say that X is a bounded metric space if its diameter is a finite number:*

$$\operatorname{diam}(X) < \infty$$

Remark 4.2. *This also has a rather simple intuitive meaning:*
We see that the set of distances between two points of spaces like $\mathbb{R}^2$ or $\mathbb{R}^3$ has no supremum, and is therefore unbounded from above. Our geometric intuition suggests that spaces like $\mathbb{R}^2$ or $\mathbb{R}^3$ should not be bounded, but spaces like the ball $B(0, 1) \in \mathbb{R}^3$ should be. We define as bounded spaces those spaces whose set of distances between 2 points has a supremum, i.e., is bounded.

Definition 4.9 (Covering and Subcovering of a Space). *Let X be a topological space, and let $\mathscr{A} = \{R_i\}_{i\in I}$ be a set of subsets of X.*
We will say that $\mathscr{A}$ is a covering of a subset $Y \subset X$ if:

$$Y \subset \cup_{i\in I} R_i$$

We will say that such a covering is open if the sets R_i are all open $\forall i$ and we will say that it is reducible to a finite subcovering if there exists a subclass of $\mathscr{A}$, $\mathscr{A}^* = \{R_{i_k}\}_{k=1}^m$, finite, such that it is still a covering of Y, i.e.:

$$Y \subset \cup_{k=1}^m R_{i_k}$$

Definition 4.10 (Compact Space). *Let X be a topological space. We will then say that X is **compact** if we can reduce every open covering of the latter to a finite open subcovering.*

Remark 4.3. *The intuitive motivation behind the concept of a compact space might instead seem less immediate at first glance. The word "compact" was introduced for the first time by Fréchet, in the context of the study of metric spaces. The definition we have given above, however, differs from that given by Fréchet (which is now called sequential compactness), and was provided for the first time by Alexandrov and Urysohn, two important Russian mathematicians, among the "leaders" of the Russian school of topology (in fact, many results or mathematical entities studied in topology take their name). As the etymology of the adjective suggests, "being compact" means being cohesive/restricted, and derives from the Latin "compactus". The reason for the choice of such a word is due to the fact that the concept of "compact space" is nothing but a generalization of the concept of finite space: in fact, compact spaces, in addition to respecting some properties also common to finite spaces, are defined precisely in the following perspective: "covering" a potentially infinite set of opens, and reducing the covering to a finite number.*

Theorem 4.6 (Heine-Borel Theorem). *Let $K \subset \mathbb{H}$, then K is compact if and only if it is closed and bounded.*

Proof. We start from the direct direction ($\Longrightarrow$).
Let $K \subset \mathbb{H}$ be a compact, we know that is, every open covering of the latter admits a finite subcovering. We choose the following specific covering of open balls $\mathscr{A} = \{B(0,n) : n \in \mathbb{N}\}$. Being K by hypothesis a compact set, we know that there exists a finite subcovering $\mathscr{A}^*$, and being finite (being the natural numbers totally ordered) there exists a ball with the maximum radius $B(0,m)$, inside which, according to a theorem proved earlier, are

contained all the other balls of the covering (because they have the same center).
Therefore, $K \subset B(0, m)$, and thus $\mathrm{diam}(K) \leq m$, which is the definition of a bounded set. Furthermore, we know that compact sets are closed in Hausdorff spaces, but as explained earlier $\mathbb{H}$ is Hausdorff because being a metric space it inherits the separation axioms in a "cascading" fashion. Thus K is closed and bounded.
Conversely, suppose that K is closed and bounded, that is, closed and with a finite diameter. Then we will have that K is contained in a closed quaternion ball $K \subset \bar{B}(0, \rho)$. But the closed ball $\bar{B}(0, \rho)$ is a compact set, and thus K is a closed subset of a compact set, hence K is compact. □

4.5 Limits of Functions of a Quaternionic Variable

We now introduce the notions of continuity and other topological notions on functions of a quaternionic variable.

Definition 4.11 (Limit of a Quaternionic Function)**.** *Let $f : \mathbb{H} \longrightarrow \mathbb{H}$ be a function of a quaternionic variable, which we will denote by $f(q)$. We write that the limit of $f(q)$ as $q \to \alpha$ equals λ, in symbols:*

$$\lim_{q \to \alpha} f(q) = \lambda$$

if $\forall \epsilon > 0$, there exists a $\delta > 0$ such that $0 < |q - \alpha| < \delta \Longrightarrow |f(q) - \lambda| < \epsilon$.

We know that the limit of a function between metric spaces is unique, and therefore this definition makes sense. Below we provide a proof for the particular case of the quaternion space.

Proposition 4.4. *Let $f(q)$ be a quaternionic variable function. If $\lim_{q\to\alpha} f(q) = \lambda$, then λ is unique.*

Proof. Let $\epsilon > 0$ be a positive real number; suppose that $f(q)$ has two limits as q tends to α: $\lim_{q\to\alpha} f(q) = \lambda$ and $\lim_{q\to\alpha} f(q) = \tilde{\lambda}$, $\lambda \neq \tilde{\lambda}$.
Then there exists a $\delta > 0$ such that $|q - \alpha| < \delta \Longrightarrow |f(q) - \lambda| < \frac{\epsilon}{2}$ and there exists a $\tilde{\delta} > 0$ such that $|q - \alpha| < \tilde{\delta} \Longrightarrow |f(q) - \tilde{\lambda}| < \frac{\epsilon}{2}$.
Let $\Delta := \min\{\delta, \tilde{\delta}\}$, then:

$$|q - \alpha| < \Delta \Longrightarrow |f(q) - \lambda| < \frac{\epsilon}{2} \;\; and \;\; |q - \alpha| < \Delta \Longrightarrow |f(q) - \tilde{\lambda}| < \frac{\epsilon}{2}$$

Thus:

$$|q-\alpha| < \Delta \Longrightarrow |\lambda - f(q)| + |f(q) - \tilde{\lambda}| < \epsilon$$

But by the triangle inequality, we have $|\lambda - \tilde{\lambda}| \leq |\lambda - f(q)| + |f(q) - \tilde{\lambda}| < \epsilon$, but since ϵ is arbitrary, we are saying that $|\lambda - \tilde{\lambda}|$ is 0.
But we know that $|\lambda - \tilde{\lambda}| = 0$ if and only if $\lambda = \tilde{\lambda}$ according to the axioms that a metric must respect, and therefore we have reached the result we wanted to prove, i.e., the limit is unique. □

Theorem 4.7 (Algebraic Properties of Limits)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{H}$ be two functions of a quaternionic variable and $\lambda_1, \lambda_2 \in \mathbb{H}$ two quaternions, then if:*

$$\lim_{q\to\alpha} f(q) = \lambda_1 \;\; and \;\; \lim_{q\to\alpha} g(q) = \lambda_2$$

we will have that:

$$\lim_{q\to\alpha} f(q) \pm g(q) = \lambda_1 \pm \lambda_2$$

$$\lim_{q\to\alpha} f(q)g(q) = \lambda_1\lambda_2$$

$$\lim_{q\to\alpha} f(q)^* = \lambda_1^*$$

$$\lim_{q\to\alpha} |f(q)| = |\lambda_1|$$

$$\lim_{q\to\alpha} f(q)^{-1} = \lambda_1^{-1} \;\; if \;\; \lambda_1 \neq 0$$

Proof. The proof is very similar to the one given for the same properties valid for quaternionic sequences.

- Starting with the first one, being by assumption the limit for q tending to α of f and g equal to λ_1 and λ_2, respectively, we know that given a real number $\epsilon > 0$, there exists a $\delta_1 > 0$ and a $\delta_2 > 0$ such that:

$$0 < |q-\alpha| < \delta_1 \Longrightarrow |f(q) - \lambda_1| < \frac{\epsilon}{2}$$

$$0 < |q-\alpha| < \delta_2 \Longrightarrow |g(q) - \lambda_2| < \frac{\epsilon}{2}$$

  Let $\Delta := \min\{\delta_1, \delta_2\}$, and using the triangle inequality:

$$0 < |q-\alpha| < \Delta \Longrightarrow |f(q) + g(q) - (\lambda_1 + \lambda_2)| \leq |f(q) - \lambda_1| + |g(q) - \lambda_2| < \epsilon$$

  The proof for the $-$ is entirely analogous.

- Let $\epsilon > 0$ be a positive real number; since $\lim_{q\to\alpha} f(q) = \lambda_1$ and $\lim_{q\to\alpha} g(q) = \lambda_2$, we know that there are 3 positive real numbers, δ_1, δ_2 and δ_3 such that:

$$0 < |q-\alpha| < \delta_1 \Longrightarrow |f(q) - \lambda_1| < \frac{\epsilon}{2(1+|\lambda_2|)}$$

$$0 < |q-\alpha| < \delta_2 \Longrightarrow |g(q) - \lambda_2| < \frac{\epsilon}{2(1+|\lambda_1|)}$$

$$0 < |q-\alpha| < \delta_3 \Longrightarrow |g(q) - \lambda_2| < 1$$

  Using the triangle inequality, we see that:

$$|g(q)| \leq |g(q) - \lambda_2| + |\lambda_2| < 1 + |\lambda_2|$$

  Now, observing that:

$$|f(q)g(q) - \lambda_1\lambda_2| = |f(q)g(q) - \lambda_1 g(q) + \lambda_1 g(q) - \lambda_1\lambda_2|$$

  by the triangle inequality

$$\leq |f(q)g(q) - \lambda_1 g(q)| + |\lambda_1 g(q) - \lambda_1\lambda_2| = |g(q)||f(q) - \lambda_1| + |\lambda_1||g(q) - \lambda_2|$$

$$< (1+|\lambda_2|)\frac{\epsilon}{2(1+|\lambda_2|)} + (1+|\lambda_1|)\frac{\epsilon}{2(1+|\lambda_1|)} = \epsilon$$

  and thus $\lim_{q\to\alpha} f(q)g(q) = \lambda_1\lambda_2$.

- Let $f(q) = x + yi + zj + wk$, then $f(q)^* = x - yi - zj - wk$, similarly let $\lambda_1 = a + bi + cj + dk$ then $\lambda_1^* = a - bi - cj - dk$. Observing that:

$$|f(q)^* - \lambda_1^*| = \sqrt{(x-a)^2 + (b-y)^2 + (c-z)^2 + (d-w)^2} =$$

$$\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2 + (w-d)^2} = |f(q) - \lambda_1|$$

  Being $\lim_{q\to\alpha} f(q) = \lambda_1$, we will have that $\forall\epsilon > 0$, $\exists\delta > 0$ such that:

$$0 < |q-\alpha| < \delta \Longrightarrow |f(q) - \lambda_1| = |f(q)^* - \lambda_1^*| < \epsilon$$

  As we wanted to prove.

- $\forall\epsilon > 0$, $\exists\delta > 0$ such that:

$$0 < |q-\alpha| < \delta \Longrightarrow |f(q) - \lambda_1| < \epsilon$$

  but $||f(q)| - |\lambda_1|| \leq |f(q) - \lambda_1|$, and thus it follows that $\lim_{q\to\alpha} |f(q)| = |\lambda_1|$.

- Follows as a corollary of the 3 parts of the theorem we have just proven; just remember that $q^{-1} = \frac{q^*}{|q|^2}\ \forall q \in \mathbb{H} \setminus \{0\}$, and at this point, we can apply the properties 2,3 and 4 demonstrated before.

$\square$

Theorem 4.8 (Necessary and Sufficient Condition for the Convergence of a Limit)**.** *Let $f(q) = t(q) + x(q)i + y(q)j + z(q)k$ be a function of a quaternionic variable and $\lambda = a + bi + cj + dk$ a quaternion, then:*

$$\lim_{q\to\alpha} f(q) = \lambda \iff \begin{cases} \lim_{q\to\alpha} t(q) = a \\ \lim_{q\to\alpha} x(q) = b \\ \lim_{q\to\alpha} y(q) = c \\ \lim_{q\to\alpha} z(q) = d \end{cases}$$

Proof. Starting from the direct direction ($\Longrightarrow$):
Let $\lim_{q\to\alpha} f(q) = \lambda$, then $\forall\epsilon > 0$, $\exists\delta > 0$, such that $0 < |q - \alpha| < \delta \Longrightarrow |f(q) - \lambda| < \epsilon$, i.e.:

$$\sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} < \epsilon$$

trivially

$$\begin{cases} |t-a| \le \sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} < \epsilon \\ |x-b| \le \sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} < \epsilon \\ |y-c| \le \sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} < \epsilon \\ |z-d| \le \sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} < \epsilon \end{cases}$$

Now let's prove the reverse implication ($\Longleftarrow$):
Let $\epsilon > 0$ be a positive real number, then we can find four positive real numbers $\delta_1, \delta_2, \delta_3, \delta_4$ such that:

$$|q-\alpha| < \delta_1 \Longrightarrow |t-a| < \frac{\epsilon}{4}\ \ ;\ \ |q-\alpha| < \delta_2 \Longrightarrow |x-b| < \frac{\epsilon}{4}$$

$$|q-\alpha| < \delta_3 \Longrightarrow |y-c| < \frac{\epsilon}{4}\ \ ;\ \ |q-\alpha| < \delta_4 \Longrightarrow |z-d| < \frac{\epsilon}{4}$$

Let $\Delta = \min\{\delta_1, \delta_2, \delta_3, \delta_4\}$, then:

$$|q-\alpha| < \Delta \Longrightarrow |t-a| + |x-b| + |y-c| + |z-d| < \epsilon$$

But $\sqrt{(t-a)^2 + (x-b)^2 + (y-c)^2 + (z-d)^2} \le |t-a| + |x-b| + |y-c| + |z-d| < \epsilon$, and thus $\lim_{q\to\alpha} f(q) = \lambda$. $\square$

Corollary 4.3. *Let $f(q) = t(q) + x(q)i + y(q)j + z(q)k$ be a function of a quaternionic variable and $\lambda = a + bi + cj + dk$ a quaternion, then:*

$$\lim_{q\to\alpha} f(q) = \lambda \iff \begin{cases} \lim_{q\to\alpha} \mathrm{Sc}(f(q)) = \mathrm{Sc}(\lambda) = a \\ \lim_{q\to\alpha} \mathrm{Vec}(f(q)) = \mathrm{Vec}(\lambda) = bi + cj + dk \end{cases}$$

We now tackle some operational examples of calculating limits of quaternionic variable functions to increase our familiarity with the definitions and theorems just given.

Example 4.5. *Consider the function $f(q) = q^2$, we want to calculate its limit as $q \to k$. Write q as $q = t + xi + yj + zk$, and substituting it into our expression we obtain:*

$$f(q) = (t + xi + yj + zk)^2 = t^2 - x^2 - y^2 - z^2 + 2txi + 2tyj + 2tzk$$

We have "decomposed" the function into its various components, the real and imaginary ones (i, j, k); we will call them $f_1(t, x, y, z) = t^2 - x^2 - y^2 - z^2$, $f_2(t, x, y, z) = 2tx$, $f_3(t, x, y, z) = 2ty$, $f_4(t, x, y, z) = 2tz$. These "component" functions are functions $f_i : \mathbb{R}^4 \to \mathbb{R}$, $i = 1, 2, 3, 4$, and we know well how to calculate their limits:

$$\begin{cases} \lim_{q\to k} t^2 - x^2 - y^2 - z^2 = -1 \\ \lim_{q\to k} 2tx = 0 \\ \lim_{q\to k} 2ty = 0 \\ \lim_{q\to k} 2tz = 0 \end{cases}$$

Therefore: $\lim_{q\to k} q^2 = -1$.
We notice that it coincides with the value of f at k, $f(k) = \lim_{q\to k} f(q)$, and we will call this property for quaternionic functions (just like in the real and complex case) continuity.

Example 4.6. *In light of the result just shown, let's find the limit for q tending to $1 + i - j$ of the function $f(q) = 2kqiq$. Write q as $q = t + xi + yj + zk$, then $f(q)$ becomes:*

$$f(q) = 4xz + 4xyi + (2t^2 - 2x^2 + 2y^2)j - 4txk$$

Let's consider separately the limits of the components of the function: $f_1(t, x, y, z) = 4xz$, $f_2(t, x, y, z) = 4xy$, $f_3(t, x, y, z) = 2t^2-2x^2+2y^2$, $f_4(t, x, y, z) =$

$-4tx$. *However, noting that these are continuous functions from* $\mathbb{R}^4$ *to* $\mathbb{R}$, *we have:*

$$\begin{cases} \lim\limits_{q\to(1+i-j)} 4xz = 0 \\ \lim\limits_{q\to(1+i-j)} 4xy = -4 \\ \lim\limits_{q\to(1+i-j)} 2t^2 - 2x^2 + 2y^2 = 2 \\ \lim\limits_{q\to(1+i-j)} -4tx = -4 \end{cases}$$

Therefore:

$$\lim_{q\to(1+i-j)} f(q) = -4i + 2j - 4k$$

4.6 Continuous functions

Definition 4.12 (Continuous function). *Let* $f : \mathbb{H} \to \mathbb{H}$ *be a function of a quaternionic variable, we say that it is continuous at a point* $p \in \mathbb{H}$ *if:*

$$f(p) = \lim_{q\to p} f(q)$$

we will say that it is continuous on a set $U \subset \mathbb{H}$ *if it is continuous at every point of* U, *i.e. if* $\forall p \in U$:

$$f(p) = \lim_{q\to p} f(q)$$

if a function is continuous on the whole $\mathbb{H}$, *then we will simply say that it is continuous.*

The intuitive meaning behind the definition is always the same: it is a generalization of the continuity of real functions, which intuitively is the property of a function's graph to "have no holes", and at the same time a particularization of the topological definition of continuity, where continuous functions are seen as morphisms of topological spaces, i.e., applications between the aforementioned that preserve some structural characteristics.
According to the second interpretation, continuous functions are functions between two topological spaces such that the pre-image of every open set of the second is open in the first, as seen in the topology section of the introduction.

Proposition 4.5. *Let* $f(q) = f_1(t,x,y,z)+f_2(t,x,y,z)i+f_3(t,x,y,z)j+f_4(t,x,y,z)k$ *be a quaternionic variable function, then* $f(q)$ *is continuous at* $\alpha \in \mathbb{H}$, $\alpha := a + bi + cj + dk$ *if and only if its component functions* $f_i(t,x,y,z)$, $i = 1,2,3,4$ *are continuous at* $[a,b,c,d] \in \mathbb{R}^4$.

Proof. Follows as an immediate corollary of the definition of continuity and theorem 4.8 previously demonstrated. □

Proposition 4.6. *Let $f(q)$ and $g(q)$ be two continuous quaternionic variable functions, then $f(q) \pm g(q)$, $f(q)g(q)$, and $g(q)f(q)$ are also continuous. The function $f(q)^{-1}$ is instead continuous at every point $r \in \mathbb{H}$ such that $f(r) \neq 0$.*

Proof. Trivial, follows from the properties of limits and the definition of continuity. □

Proposition 4.7. *Let $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{H}$ be two continuous quaternionic functions, then $g \circ f = g(f(q))$ is continuous.*

Proof. This fact is generally valid for two continuous functions between topological spaces, and can be proven starting from the following simple observation:
Let $f : X \to Y$ and $g : Y \to Z$ be continuous functions between topological spaces: by the definition of continuity, we will have that if A is an open set in Z, then $g^{-1}[A]$ is an open set in Y, and being f also continuous, we will have that $(g \circ f)^{-1} = f^{-1}[g^{-1}[A]]$ is open in X, therefore $g \circ f : X \to Z$ is continuous. □

Proposition 4.8. *Quaternionic constant functions, of the type $f(q) = \gamma$, $\gamma \in \mathbb{H}$ are continuous.*

Proof. This is also a fact generally valid for constant functions between topological spaces. Let $f : X \to Y$ be a function between topological spaces, defined as $f(x) = k \in Y$, $\forall x \in X$, i.e., a constant function. Let $A \subset Y$ be an open set of Y, then its pre-image under f will be:

$$f^{-1}[A] = \begin{cases} \emptyset \ \ if \ \ k \notin A \\ X \ \ if \ \ k \in A \end{cases}$$

in both cases, they are open sets in X, and therefore f is continuous. □

We have introduced an operational notion of continuity in the space of quaternions; now let's prepare to test it through a series of explanatory examples.

Example 4.7. *We have the identity function $f(q) = q$; writing $q = t + xi + yj + zk$, we can write it as $f(q) = t + xi + yj + zk = f_1(x, y, z, t) + f_2(x, y, z, t)i + f_3(x, y, z, t)j + f_4(x, y, z, t)k$, with $f_1(x, y, z, t) = t$, $f_2(x, y, z, t) = x$, $f_3(x, y, z, t) = y$, $f_4(x, y, z, t) = z$. Being the component functions continuous, we conclude that also $f(q) = q$ is continuous.*

Example 4.8. *The function analyzed in the last section, $f(q) = 2kqiq$ is also continuous, as mentioned in the example where we mentioned it, its component functions are continuous.*

Example 4.9. *Let's now look at the function $f(q) = q^2$. We can rewrite it, specifying the component functions, as:*

$$f(q) = t^2 - x^2 - y^2 - z^2 + 2txi + 2tyj + 2tzk$$

here again, the component functions are all continuous functions $f_i : \mathbb{R}^4 \to \mathbb{R}$, and therefore $f(q) = q^2$ is continuous.

Proposition 4.9. *The quaternionic variable function $f : \mathbb{H} \to \mathbb{H}$ defined as $f(q) = q^n$ with $n \in \mathbb{N}$ is continuous.*

Proof. We prove this fact by induction: we have already covered the cases $n = 1$ and $n = 2$ in the examples, thus we have already covered the base case.
For the inductive step, we assume that for $n = k$ the assertion is satisfied, that is, $f(q) = q^k$ is continuous, we now show that it is true also for $n = k+1$: $f(q) = q^{k+1} = q^k q$, but q is continuous (example 4.7) and so is q^k by the inductive step, but we know that the product of two continuous functions is continuous, and therefore $f(q) = q^{k+1}$ is continuous. □

Corollary 4.4. *Functions associated with right and left quaternionic polynomials, i.e., functions of the type $f(q) = \sum_{m=0}^{n} \alpha_m q^m$ and of the type $f(q) = \sum_{m=0}^{n} q^m \alpha_m$ are continuous.*

Proof. Follows as an immediate corollary of the previous proposition and the fact that sums of continuous functions are continuous. □

Corollary 4.5. *General linear quaternionic polynomials, i.e., functions of the type $\Phi(q) = a_0 q a_1 q \ldots q a_n q + \phi(q)$ where $\phi(q)$ is a finite sum of similar monomials $a_0 q a_1 q \ldots q a_k$ with $k < n$ are continuous functions.*

Proof. Follows as a corollary of the law of product and sum of limits of quaternionic functions and the fact that $f(q) = q$ is a continuous function. Note that they are the sum of products of constant functions with the identity function $f(q) = q$. But products of continuous functions are continuous, and the sum of continuous functions is a continuous function, so general linear quaternionic polynomials are continuous. □

Corollary 4.6. *General quaternionic polynomials (whose general form is difficult to express) are continuous functions.*

Proof. The proof is also here an immediate corollary of the results presented previously by a very similar reasoning; they are sums of products of continuous functions, and therefore continuous functions. □

Proposition 4.10. *The function $f : \mathbb{H} \to \mathbb{H}$ defined as $f(q) := q^*$, i.e., that associates each quaternion with its conjugate, is a continuous function.*

Proof. The identity function $f(q) = q$ is continuous, therefore $\lim_{q\to\alpha} q = \alpha$ $\forall\alpha \in \mathbb{H}$, and by a property of limits demonstrated in theorem 4.7, we can say that $\lim_{q\to\alpha} q^* = \alpha^*$ $\forall\alpha \in \mathbb{H}$, therefore $f(q) = q^*$ is continuous. □

Exercise 4.8. *Prove that the norm function $f : \mathbb{H} \to \mathbb{R}$ $f(q) = |q|$ of a quaternion is a continuous function.*

Exercise 4.9. *Prove that the functions $f : \mathbb{H} \to \mathbb{R}$ and $g : \mathbb{H} \to$ **P** defined as $f(q) := \mathrm{Sc}(q)$ and $g(q) := \mathrm{Vec}(q)$ are continuous functions.*

Exercise 4.10. *Prove that if $f(q)$ and $g(q)$ are two continuous quaternionic variable functions, then their scalar product $\sigma(q) = \langle f(q)|g(q)\rangle$ is a continuous function.*

Having not yet introduced the quaternionic analogues of some important elementary functions such as the exponential, trigonometric, or hyperbolic functions, the subject of our dissertation is currently somewhat limited. We will address the values for which these functions are continuous in due time when we will have defined them.

Exercise 4.11. *Determine the subsets of $\mathbb{H}$ within which the following functions are continuous:*

- $f(q) = (q^3 + 1)^{-1}$
- $g(q) = q^{-1}|q|^2(q - 3)^{-3}$
- $h(q) = \mathrm{Vec}(q) \times (j - \mathrm{Sc}(q)^2 k)$
- $\phi(q) = (q^8 + 1)^{-1}$
- $\mathrm{sgn}(q) = \frac{q}{|q|}$

4.7 Alexandroff Compactification of $\mathbb{H}$

In real and complex analysis, we often work with compactifications of the spaces $\mathbb{R}$ and $\mathbb{C}$, such as the Alexandroff compactification of complex numbers, better known as the Riemann sphere, or a compactification of the reals known as the extended real line, obtained by adding two new points $+\infty$ and $-\infty$ to $\mathbb{R}$.
Let's start by recalling the compactification of the complex numbers, and the stereographic projection of a sphere onto a plane (the inverse of the homeomorphism that maps the complex plane into the sphere S^2 without the north pole).

4.7.1 Stereographic Projection and Riemann Sphere

How can a sphere be "flattened"? The question just posed has been an area of great interest for a long period of time for cartographers but also for mathematicians, and is partly also one of the reasons why differential geometry was developed as we now know it.
Despite numerous predecessors, historically the development of differential geometry as an area of study as we know it now is due to Gauss.
What motivated Gauss? Exactly the matter we talked about before; in fact, in the early 1800s, Gauss was very interested in problems of topography and cartography: he carried out a land survey for the kingdom of Hanover in 1818, and also received numerous assignments from the kingdom of Denmark for cartographic/topographic issues.
In particular, the Hanover land survey significatively increased his interest in the study of surfaces and their "flattening", leading him to focus in the following years on this area of study, concluding with an important conclusion from a "practical" point of view: it's not possible to represent the Earth's surface on a flat surface (like a map) without distortion (this fact is a corollary of the so-called "Theorema Egregium").
What is the stereographic projection of a sphere? (we will focus on the case $S^2 \subset \mathbb{R}^3$ for now) It is simply a way of flattening a sphere, not the preferred one by cartographers however, who instead prefer the Mercator projection (which is still very interesting from a mathematical point of view, especially for its relation to the Gudermannian function).
The question that arises spontaneously is, however, how to realize such flattening? Our goal is to identify every point of the plane with a unique point on the sphere; the strategy we will use is the following: select a point on our sphere $S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\}$, $N = (0, 0, 1) \in S^2$, let $p \in \mathbb{R}^2$ be

a point on the plane, let $\overline{Np}$ be the line that joins p and N; then we identify each point of the plane $p \in \mathbb{R}^2$ with the intersection of the line $\overline{Np}$ with the sphere S^2: such intersection is unique. Let's visualize it graphically:

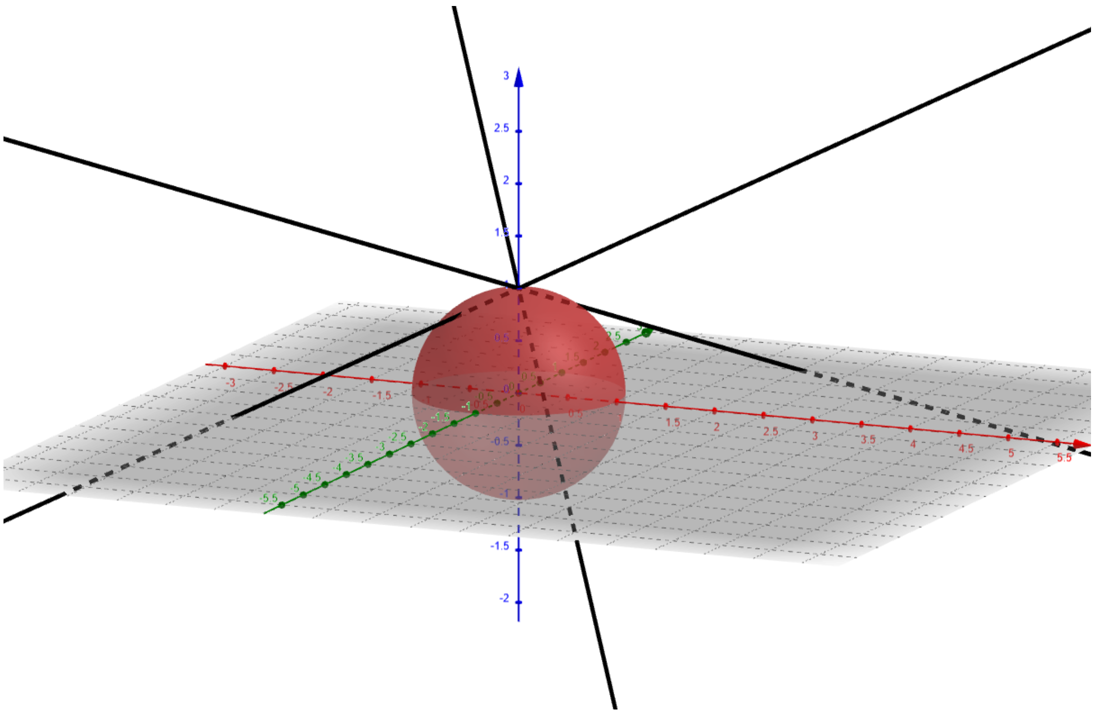


Figure 4.3: Visualization of the intersection between the line $\overline{Np}$ and the sphere S^2.

The sphere, intersected with the plane, forms a circumference. The points of this circumference, under the stereographic projection, remain unchanged; moreover, the points outside the circumference will be mapped on the northern hemisphere, while those inside on the southern hemisphere, as we can also see from the image above.

Let's now derive an analytical expression for this application:

Let $P = (x, y)$ be a point of the plane, $P^* = (x_1, x_2, x_3)$ its image on the sphere and N the north pole (or infinity point) just defined, the equation of the line between the two will be given, in its vector form, by:

$$\vec{r}(t) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} x \\ y \\ -1 \end{bmatrix}$$

expressible with parametric equations as:

$$\begin{cases} x_1 = tx \\ x_2 = ty \\ x_3 = 1 - t \end{cases} \tag{4.3}$$

However, we know that $r(t)$, in addition to belonging to the line, also belongs to the sphere S^2, and thus:

$$x_1^2 + y_1^2 + z_1^2 = 1 = t^2x^2 + t^2y^2 + (1-t)^2 = t^2(x^2 + y^2 + 1) + 1 - 2t$$

this equation has two roots, $t = 0$ and $t = \frac{2}{x^2+y^2+1}$. Discarding the first, as we want to map a point of the plane into a point of the sphere $S^2 \setminus \{N\}$, we take the second, and substituting it in the system of equations derived earlier, we obtain:

$$\begin{cases} x_1 = \dfrac{2x}{x^2 + y^2 + 1} \\ x_2 = \dfrac{2y}{x^2 + y^2 + 1} \\ x_3 = 1 - \dfrac{2}{x^2 + y^2 + 1} = \dfrac{x^2 + y^2 - 1}{x^2 + y^2 + 1} \end{cases}$$

Thus given $P = (x, y)$, we will have that $P^* = \left(\dfrac{2x}{x^2 + y^2 + 1}, \dfrac{2y}{x^2 + y^2 + 1}, \dfrac{x^2 + y^2 - 1}{x^2 + y^2 + 1}\right) \in S^2$.

For the inverse of this application, deriving x and y from equation (4.3), we will have that given a point $P = (x_1, x_2, x_3) \in S^2$, its stereographic projection onto the plane will be equal to $P^* = (\frac{x_1}{1-x_3}, \frac{x_2}{1-x_3})$.

We thus now have a way to map every point of a plane into a point of a sphere minus the "point at infinity". Before continuing, let's recall some topological definitions:

Definition 4.13 (Compactification). *Let (X, τ) and $(Y, \tilde{\tau})$ be two topological spaces, then if X is homeomorphic to a subset of Y and Y is a compact space, we will say that Y is a compactification of X.*

Definition 4.14 (Alexandroff Compactification). *Let (X, τ) be a topological space, we define the topological space $(X_{\mathscr{A}}, \tau_{\mathscr{A}})$ as follows:*

1. *We define the set $X_{\mathscr{A}}$ as X plus a new point that we will call the point at infinity; $X_{\mathscr{A}} = X \cup \{\infty\}$.*

2. *The topology $\tau_{\mathscr{A}}$ under which such a set will form a compactification, instead, will be given by all the open sets of X plus all the complements in $X_{\mathscr{A}}$ of the closed and compact sets of X.*

It is shown in introductory courses of General Topology that such a topological space is a compactification of X, and therefore we call this topological space Alexandroff compactification, or one-point compactification.
The sphere S^2 is a compactification of the plane $\mathbb{R}^2$; we can prove this by showing that the inverse stereographic projection $\varphi : \mathbb{R}^2 \to S^2 \setminus \{N\}$ is a homeomorphism. Let's start by observing that the function inverse stereographic projection defined as:

$$\varphi(x,y) = \left(\frac{2x}{x^2+y^2+1}, \frac{2y}{x^2+y^2+1}, \frac{x^2+y^2-1}{x^2+y^2+1}\right) \subset \mathbb{R}^3$$

is first of all injective if we consider $\mathbb{R}^3 \setminus \{(x,y,z) \ : z = -1\}$ as its codomain, as it has an inverse $\forall x_3 \neq -1$:

$$\varphi^{-1}(x_1,x_2,x_3) = \left(\frac{x_1}{1-x_3}, \frac{x_2}{1-x_3}\right) \in \mathbb{R}^2$$

Our function φ is also bicontinuous, as both φ and φ^{-1} are continuous functions with respect to their respective sets of departure and arrival (as the components of both are continuous functions). Finally, composing φ with the inclusion $\imath : \mathbb{R}^3 \setminus \{(x,y,z) \ : z = -1\} \to S^2 \setminus \{N\}$, and observing that $\varphi(x,y) \in S^2 \setminus \{N\} \ \forall x,y \in \mathbb{R}^2$ (i.e φ is surjective) we deduce that φ is a homeomorphism between $\mathbb{R}^2$ and $S^2 \setminus \{N\}$. We can formalize the result just obtained through the following proposition:

Proposition 4.11. *S^2 under the subspace topology inherited from $\mathbb{R}^3$ is a compactification of $\mathbb{R}^2$.*

We could have alternatively proved the bicontinuity of φ in a more geometric way: it was enough to observe that circles in $\mathbb{R}^2$ are mapped into circles of S^2 not containing the infinity point, while lines of $\mathbb{R}^2$ are mapped into circles of S^2 containing the infinity point. Conversely, every circle on S^2 is the image under the inverse stereographic projection of a line or a circle in $\mathbb{R}^2$. Excluding the point at infinity, we will have that circles on S^2 are mapped into circles on $\mathbb{R}^2$ and vice versa: similarly, open balls of S^2 are mapped into open balls of $\mathbb{R}^2$ and vice versa; from this follows the bicontinuity of $\varphi : \mathbb{R}^2 \to S^2 \setminus \{N\}$. As with the plane $\mathbb{R}^2$, also $\mathbb{C}$, being homeomorphic to the 2-Euclidean space, can be compactified in the same way: we call the compactification of complex numbers **Riemann Sphere**.

Proposition 4.12. *The Riemann Sphere (the compactification of* $\mathbb{C}$ *just shown) is homeomorphic to the Alexandroff compactification of* $\mathbb{C}$*.*

For this reason, we often call the north pole of the Riemann Sphere the infinity point, indicated with ∞.

4.7.2 Alexandroff Compactification of Quaternions

We now wish to realize a compactification of quaternions similar to that of the complex plane, adding an "infinity point".
Doing this is not difficult, as we will see, since it is possible to extend the procedure done earlier with $\mathbb{R}^2$ for more general Euclidean spaces $\mathbb{R}^n$ in the following way:

Definition 4.15. *Let* $N \in S^n$ *be a point of the circle* $S^n \subset \mathbb{R}^{n+1}$ *which we call the "point at infinity" and let* $E \subset \mathbb{R}^n$ *be the hyperplane identified by setting the* $(n+1)$*-th coordinate of* $\mathbb{R}^{n+1}$ *equal to 0. The stereographic projection of* $P \in S^n \setminus \{N\}$ *is the point* $\tilde{P}$ *of intersection between the line* $\overline{QP}$ *and the hyperplane* E*.*
The coordinates of $\tilde{P} = (X_i)_{i=1}^n$*, the stereographic projection of* $P = (x_i)_{i=0}^n$ *are given by:*

$$X_i = \frac{x_i}{1-x_0} \quad i = 1, 2, 3, \ldots, n \tag{4.4}$$

We call the function $\varphi : S^n \setminus \{N\} \to \mathbb{R}^n$ *stereographic projection, defined as:*

$$\varphi(P) = \tilde{P} = \Big(\frac{x_i}{1-x_0}\Big)_{i=1}^n = \big(\frac{x_1}{1-x_0}, \frac{x_2}{1-x_0}, \ldots, \frac{x_n}{1-x_0}\big) \tag{4.5}$$

The idea behind it is the same: we want to identify every point of the plane $\mathbb{R}^n$ with a point on the hypersphere S^n.
The expression for the inverse stereographic projection φ^{-1} is given by:

$$x_0 = \frac{\sum_{i=1}^{n} X_i^2 - 1}{\sum_{i=1}^{n} X_i^2 + 1} \quad x_j = \frac{2X_j}{\sum_{i=1}^{n} X_i^2 + 1} \tag{4.6}$$

for $j = 1, 2, 3, \ldots n$.

$$\varphi^{-1}(\tilde{P}) = (x_j)_{j=0}^n \in S^n \setminus \{N\}$$

The existence of an inverse confirms that the function is bijective: it remains to show bicontinuity, whose proof, however, proceeds in a very similar way to the previous one.

Corollary 4.7. *The hypersphere S^n is a compactification of $\mathbb{R}^n$.*

Proposition 4.13. *The hypersphere S^n is homeomorphic to the Alexandroff compactification of $\mathbb{R}^n$, $(\mathbb{R}^n_{\mathscr{A}}, \tau_{\mathscr{A}})$.*

This result, like the one seen earlier for the complex case, follows from the uniqueness with respect to homeomorphisms of a one-point compactification of a Hausdorff space (such as $\mathbb{R}^n$).
From this, it follows that we can realize a one-point compactification of the quaternions $\mathbb{H}$ in a way analogous to that realized for complex numbers, that is, by "enclosing them" in a sphere S^4 and adding an infinity point. What has just been said was exactly what we wanted; we were indeed looking for an "infinity point" irrespective of the sign, given the lack of a total order on Quaternions (and in general for all $\mathbb{R}$-Algebras of the "Cayley-Dickson chain" from complex numbers onwards). Such a point allows us to express some concepts of interest for the theory of functions, such as that of a ball centered at ∞:

$$B(\infty, r) = \left\{ q \in \mathbb{H} \; : \; \frac{1}{|q|} < r \right\}$$

where here $r > 0$. Again, now having well defined an infinity point for quaternions, we can analyze the behavior "at infinity" of quaternion functions with the introduction of the following definitions:

$$\lim_{q \to \alpha} f(q) = \infty \iff \lim_{q \to \alpha} \frac{1}{f(q)} = 0$$

$$\lim_{q \to \infty} f(q) = \lambda \iff \lim_{q \to 0} f(q^{-1}) = \lambda$$

$$\lim_{q \to \infty} f(q) = \infty \iff \lim_{q \to 0} \frac{1}{f(q^{-1})} = 0$$

We call $\mathbb{H}_\infty = \mathbb{H} \cup \{\infty\}$ the space of "extended quaternions", and we will call the Alexandroff compactification of quaternions the Riemann hypersphere.

4.8 Eilenberg-Niven Theorem

In this section, we will discuss the extension of the fundamental theorem of algebra to general linear quaternions (which we already mentioned in chapter 2), better known as the Eilenberg-Niven theorem. Let's state it here:

Theorem 4.9 (Eilenberg-Niven Theorem, 1944)**.** *Let f be a general linear quaternionic polynomial of degree n, expressible as:*

$$f(q) = a_0 q a_1 q \dots q a_n + \phi(q)$$

where here $a_i \in \mathbb{H} \setminus \{0\}$ and $\phi(q)$ is a sum of monomials $b_0 q b_1 q \dots q b_k$, where $k < n$. Then, the equation $f(q) = 0$ has at least one solution in $\mathbb{H}$.

To prove this theorem, we will need the concept of the **Brouwer degree of a continuous function between spheres**. Let's define it here:

Definition 4.16 (Brouwer degree of a continuous function)**.** *Given a continuous function $f : S^n \to S^n$; such a function induces a homomorphism of groups $f_* : H_n(S^n) \to H_n(S^n)$ from the n-th homology group of S^n, $H_n(S^n)$ to $H_n(S^n)$. It is known that $H_n(S^n) \cong \mathbb{Z}$, and therefore the induced homomorphism becomes a function from $\mathbb{Z}$ to $\mathbb{Z}$. We call the* ***Brouwer degree of f*** *the value:*

$$\deg(f) = f_*(1) \tag{4.7}$$

Within this chapter, we will sometimes simply call the Brouwer degree "degree". It will be clear from the context when we are talking about the topological degree of a function and when we are instead talking about the algebraic degree of a polynomial.
One of the first results that is proven about this notion in a course of topology is that homotopic functions have the same Brouwer degree.
We are now ready to prove the Eilenberg-Niven theorem in a topological language by proving the following proposition from which the aforementioned theorem follows.

Proposition 4.14. *Let S^4 be the 4-sphere, i.e., the Alexandroff compactification of quaternions. Extend the general linear polynomial of the form of theorem 4.9 to the aforementioned by setting $f(\infty) = \infty$. The quaternionic general linear polynomial $f : S^4 \to S^4$ so extended is a function of Brouwer degree equal to n (its degree as a polynomial).*

Proof. Define a function $g : S^4 \to S^4$ as:

$$g(q) = \begin{cases} q^n & if \quad q \in \mathbb{H} \\ \infty & if \quad q = \infty \end{cases}$$

The strategy to prove this theorem is the following: we will first show that f is homotopic to g, and then prove that $g(q)$ has degree n: from this, it will follow that f will also have degree n (i.e., the assertion) since homotopic

functions have the same Brouwer degree.
Let's start by constructing the homotopy between f and g in two steps: observe that the function defined as:

$$F(q,t) := \begin{cases} a_0 q a_1 q a_2 \dots q a_n + (1-t)\phi(q) & if \quad q \in \mathbb{H} \\ \infty \quad if \quad q = \infty \end{cases}$$

is a continuous function such that $F(q,0) = f(q)$ and

$$F(q,1) = \begin{cases} a_0 q a_1 q a_2 \dots q a_n & if \quad q \in \mathbb{H} \\ \infty \quad if \quad q = \infty \end{cases}$$

Now choose for each $i \in \mathbb{N} \cap [0,n]$ a continuous path $a_i(t)$ in $\mathbb{H} \setminus \{0\}$, with $1 \leq t \leq 2$, such that $a_i(1) = a_i$ and $a_i(2) = 1$. For such values of t ($1 \leq t \leq 2$) define a function:

$$G(q,t) := \begin{cases} a_0(t) q a_1(t) q a_2(t) \dots q a_n(t) & if \quad q \in \mathbb{H} \\ \infty \quad if \quad q = \infty \end{cases} \tag{4.8}$$

The function G thus defined is a continuous function such that $G(q,1) = F(q,1)$ and $G(q,2) = g(q)$; thus, we have obtained the desired homotopy between f and g. Now let's prove that g has a Brouwer degree equal to n. In light of what was seen in chapter 2 (in section 2.4 to be exact), we know that the equation

$$g(q) = q^n = i$$

has n solutions, and its n solutions are the complex roots of i. We will prove that g has a Brouwer degree equal to n by showing that the determinant of its Jacobian is positive at every root of i [Note 4.1]. Write our quaternion $q = t + xi + yj + zk$ in its polar form:

$$q = |q|(\cos\theta + \operatorname{sgn}(\operatorname{Vec}(q))\sin\theta)$$

from the quaternionic De Moivre theorem, it follows that $q^n = |q|^n(\cos(n\theta) + \operatorname{sgn}(\operatorname{Vec}(q))\sin(n\theta))$.
Writing

$$g(q) = q^n = g_1(t,x,y,z) + ig_2(t,x,y,z) + jg_3(t,x,y,z) + kg_4(t,x,y,z)$$

we obtain that, setting it equal to the polar form of q^n:

$$\begin{cases} g_1(q) = |q|^n \cos(n\theta) \\ g_2(q) = \dfrac{x|q|^n \sin(n\theta)}{\pm|\operatorname{Vec}(q)|} \\ g_3(q) = \dfrac{y|q|^n \sin(n\theta)}{\pm|\operatorname{Vec}(q)|} \\ g_4(q) = \dfrac{z|q|^n \sin(n\theta)}{\pm|\operatorname{Vec}(q)|} \end{cases}$$

Calculating the partial derivatives of each component of the function g, g_i, with respect to t, x, y, and z, we obtain the following expression for the Jacobian matrix of $g(q)$:

$$J = \begin{bmatrix} nx & -nt & 0 & 0 \\ nt & nx & 0 & 0 \\ 0 & 0 & \frac{1}{x} & 0 \\ 0 & 0 & 0\frac{1}{x} & \end{bmatrix}$$

from which, through a simple direct calculation, we have that $\det(J) = n^2\left(\frac{t^2}{x^2} + 1\right)$. But since $t^2 + x^2 = 1$ and $x \neq 0$, we have that the determinant will be equal to $\dfrac{n^2}{x^2}$, a positive number. This completes the proof. □

4.9 Quaternionic Series

We conclude the chapter with the last important topic remaining to be addressed before moving on to elementary functions and various formulations of differential calculus on $\mathbb{H}$: infinite quaternionic series.

Definition 4.17 (Infinite quaternionic series)**.** *Let $\{q_n\}$ be a sequence with values in $\mathbb{H}$. We call $S_n = q_1 + q_2 + \cdots + q_n$ the n-th partial sum. We will call the sequence $\{S_n\}$ the series associated with $\{q_n\}$. Notationally, we indicate:*

$$\lim_{n\to\infty} S_n = \sum_{n=1}^{\infty} q_n \tag{4.9}$$

If such a limit converges to a quaternion $S \in \mathbb{H}$, then we will say that the infinite series converges to S, writing $\sum_{n=1}^{\infty} q_n = S$; otherwise, we will say that it diverges.

Let's now prove some properties of infinite quaternionic series:

Proposition 4.15. *Let* $\sum_{n=1}^{\infty} q_n$ *and* $\sum_{n=1}^{\infty} r_n$ *be two convergent quaternionic series, and* $\lambda \in \mathbb{H}$ *a quaternion, then the following properties hold:*

- $\sum_{n=1}^{\infty}(q_n + r_n) = \sum_{n=1}^{\infty} q_n + \sum_{n=1}^{\infty} r_n.$
- $\sum_{n=1}^{\infty} \lambda q_n = \lambda \sum_{n=1}^{\infty} q_n.$

Proof.

- $\sum_{n=1}^{\infty}(q_n + r_n) = \lim_{n\to\infty} S_{q_n+r_n} = \lim_{n\to\infty} S_{q_n} + S_{r_n} = \lim_{n\to\infty} S_{q_n} + \lim_{n\to\infty} S_{r_n} = \sum_{n=1}^{\infty} q_n + \sum_{n=1}^{\infty} r_n.$
- $\sum_{n=1}^{\infty} \lambda q_n = \lim_{n\to\infty} S_{\lambda q_n} = \lim_{n\to\infty} \lambda S_{q_n} = \lambda \lim_{n\to\infty} S_{q_n} = \lambda \sum_{n=1}^{\infty} q_n.$

□

Example 4.10. *We have the sequence* $q_n = 1 + \frac{i}{n} + \frac{j}{n} + k$. *We want to investigate the convergence of* $\sum_{n=1}^{\infty} q_n$.
We see that the m-th partial sum will be given by:

$$S_m = \sum_{n=1}^{m}(1 + \frac{i}{n} + \frac{j}{n} + k) = m + mk + \sum_{n=1}^{m} \frac{i}{n} + \frac{j}{n}$$

letting $m \to \infty$*, we see that both the coefficients of 1 and* k*, as well as those of* i *and* j *diverge, and hence the series is divergent.*

Example 4.11. *Let* $\sum_{n=0}^{\infty} q^n$ *be the quaternionic geometric series. The m-th partial sum is equal to:*

$$S_m = 1 + q + \cdots + q^{m-1}$$

multiplying both sides of the equation by q *on the right, we obtain:*

$$S_m q = q + q^2 + \cdots + q^m$$

Now subtract $S_m q$ from S_m:

$$S_m - S_m q = 1 - q^m$$

from which $S_m(1-q) = (1-q^m)$, i.e $S_m = (1-q^m)(1-q)^{-1}$. Taking the limit of this sequence as $m \to \infty$, assuming that $|q| < 1$, we get:

$$\sum_{n=0}^{\infty} q^n = (1-q)^{-1}$$

therefore, as in the real and complex case, such a series converges to $(1-q)^{-1}$ for $|q| < 1$.

Now imagine we want to investigate the convergence of a series like $\sum_{n=1}^{\infty}(i+j)^n$:

Example 4.12. *We have $q_n = (i+j)^n$. Observe that:*

$$(i+j)^n = \begin{cases} i+j & n=1 \\ -2 & n=2 \\ -2i-2j & n=3 \\ 4 & n=4 \\ 4i+4j & n=5 \\ -8 & n=6 \\ -8i-8j & n=7 \\ 16 & n=8 \\ \dots \end{cases}$$

Divide the sequence into 2 subsequences, $q_{n1} = \{i+j, -2i-2j, 4i+4j, \dots\}$ and $q_{n2} = \{-2, 4, -8, 16, \dots\}$; they can be written in a closed form as $q_{n1} = (-1)^{n+1}2^{n-1}(i+j)$ and $q_{n2} = (-1)^n 2^n$.

Observe then that the m-th partial sum of q_n, $\Sigma_m = \sum_{n=1}^{m} q_n$ can be written as $\Sigma_m = \sum_{n=1}^{m} q_{n1} + \sum_{n=1}^{m} q_{n2} = S_{m1} + S_{m2}$.

To determine the convergence of Σ_m, consider separately S_{m1} and S_{m2}:

$$S_{m1} = \sum_{n=1}^{m}(-1)^{n+1}2^{n-1}(i+j) \quad ; \quad S_{m2} = \sum_{n=1}^{m}(-1)^n 2^n$$

We can write the first in a more suggestive form, explicating the real component sequences of the associated sequence of the series as:

$$S_{m1} = \sum_{n=1}^{m}(-1)^{n+1}2^{n-1}i + \sum_{n=1}^{m}(-1)^{n+1}2^{n-1}j = \sum_{n=1}^{m}\sigma_1(n)i + \sum_{n=1}^{m}\sigma_2(n)j$$

for the second one, it is not necessary as it is a real series.

$$\Sigma_m = \sum_{n=1}^{m}(-1)^n 2^n + \sum_{n=1}^{m}\sigma_1(n)i + \sum_{n=1}^{m}\sigma_2(n)j$$

We have explicated the real series components of our quaternionic series Σ_m, and thus reduced a "quaternionic" problem to a real one, and now using known results from real analysis, we can easily determine that Σ_m is divergent, as its component sequences $\sum_{n=1}^{m}(-1)^n 2^n$, $\sum_{n=1}^{m}\sigma_1(n)$ and $\sum_{n=1}^{m}\sigma_2(n)$ are divergent (by a theorem proven in the section on quaternionic sequences).

Note quickly that verifying it directly becomes tedious, and we had to transform the problem into a real problem to resort to results regarding real series that allow us to determine convergence or divergence immediately. We have 2 problems: this is not always possible, and even if it is possible, it can be a lengthy process, as in the example above.
A question naturally arises: can we extend to quaternions some "tests" for the convergence of series? The answer is affirmative; let's start with a proposition about a necessary but not sufficient condition for a series to converge.

Proposition 4.16. *Let q_n be a sequence, then if $\sum_{n=1}^{\infty} q_n$ converges, necessarily $\lim_{n\to\infty} q_n = 0$.*

Proof. We can prove this fact in two ways; either noting that $\sum_{n=1}^{\infty} q_n$ can be written as $\sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n i + \sum_{n=1}^{\infty} c_n j + \sum_{n=1}^{\infty} d_n k$, and that, as a corollary of the same fact for real sequences, it follows that $a_n, b_n, c_n, d_n \to 0$ for $n \to \infty$ and therefore $q_n \to 0$, or alternatively noting that $S_n - S_{n-1} = q_n$, but for $n \to \infty$, $S_n - S_{n-1} \to 0$ and therefore $q_n \to 0$. □

With this theorem, we could have immediately determined the behavior of the series in the previous example, as we already knew that the sequence $(i + j)^n$ was divergent.
However, what we have just derived is a necessary but not sufficient condition, so we can only deduce that if q_n diverges or does not converge to 0, the series associated with it is divergent, but we cannot conclude anything if we know that $\lim_{n\to\infty} q_n = 0$. We want to obtain stronger results.

Definition 4.18. *Let $\sum_{n=1}^{\infty} q_n$ be a quaternionic series, then we will say that it is absolutely convergent if the real series $\sum_{n=1}^{\infty} |q_n|$ converges.*
Conversely, we will say that it is conditionally convergent if the series $\sum_{n=1}^{\infty} q_n$ converges, but the series associated with the norm of the sequence $\sum_{n=1}^{\infty} |q_n|$ does not.

Theorem 4.10. *Let $\sum_{n=1}^{\infty} q_n$ be a quaternionic series. Then if the series is absolutely convergent, it is convergent.*

Proof. Assume that the series is absolutely convergent, to prove that $\sum_{n=1}^{\infty} q_n$ also converges, we need to show that the sequence of partial sums converges. First of all, observe that, since $\sum_{n=1}^{\infty} |q_n|$ is a convergent series, it will follow that the sequence of partial sums associated with it will be a Cauchy sequence.
$\sum_{n=1}^{\infty} |q_n| = L \iff \forall \epsilon > 0$, $\exists N \in \mathbb{N}$ such that $\forall m, l \geq N$,
$\sum_{n=l+1}^{m} |q_n| = \left| \sum_{n=l+1}^{m} |q_n| \right| = |\Sigma_m - \Sigma_l| < \epsilon$
Let $\{\tilde{\Sigma}_m\}$ be the sequence of partial sums of the series $\sum_{n=1}^{\infty} q_n$. Let $m, l \geq N$, where N is the natural number existing by hypothesis chosen earlier; by

the triangle inequality, we have:

$$|\tilde{\Sigma}_m - \tilde{\Sigma}_l| = \left| \sum_{n=l+1}^{m} q_n \right| \leq \sum_{n=l+1}^{m} |q_n| < \epsilon$$

Therefore, the sequence $\{\tilde{\Sigma}_m\}$ is a Cauchy sequence, that is, a convergent sequence. Thus, the series $\sum_{n=1}^{\infty} q_n$ converges. □

This is, on the contrary, a sufficient but not necessary condition, and thus will only be fruitful in the case where we can show that the series of the norm of the associated sequence is convergent.

Example 4.13. *We have the series $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$. It is well known that this series converges, however, the series associated with its norm: $\sum_{n=1}^{\infty} \frac{1}{n}$ is divergent (it is the famous harmonic series). The converse of the proposition just proven is therefore generally false.*

Proposition 4.17 (Cauchy's Root Test)**.** *Let $\sum_{n=1}^{\infty} q_n$ be a quaternionic series; we define:*

$$C = \lim_{n\to\infty} \sup \sqrt[n]{|q_n|} \tag{4.10}$$

Then:

1. *If $C < 1$, the series converges absolutely.*
2. *If $C > 1$, the series diverges.*
3. *If $C = 1$, the test is inconclusive.*

Proof. Starting from the first, suppose that $C < 1$, then for every $n \geq N$, $\exists k$ such that $\sqrt[n]{|q_n|} \leq k < 1 \implies |q_n| \leq k^n < 1$. Since the real geometric series $\sum_{n=1}^{\infty} k^n$ converges for $|k| < 1$, we will have that, by the comparison test for series, $\sum_{n=1}^{\infty} |q_n|$ also converges, and hence $\sum_{n=1}^{\infty} q_n$ converges absolutely.
For the second point, instead, supposing we have $C > 1$, we would have $\forall \epsilon > 0$, $|q_n| > C - \epsilon$ for an infinite number of indices. Being ϵ arbitrary,

let's choose it in such a way as to make $C - \epsilon > 1$, and hence $|q_n| > 1$ for an infinite number of indices, from which it follows that the series is divergent. $\square$

Proposition 4.18 (D'Alembert's Ratio Test)**.** *Let* $\sum_{n=1}^{\infty} q_n$ *be a quaternionic series of non-zero terms; we define:*

$$C = \lim_{n\to\infty} \sup \frac{|q_{n+1}|}{|q_n|} \tag{4.11}$$

Then:

1. *If* $C < 1$*, the series converges absolutely.*
2. *If* $C > 1$*, the series diverges.*
3. *If* $C = 1$*, the test is inconclusive.*

Proof. Starting from the first, choose a real number r, $C < r < 1$; there exists an $N \in \mathbb{N}$ such that $\forall n \geq N$ we have:

$$|q_{n+1}| < r|q_n|$$

more generally, $\forall i > 0$:

$$|q_{n+i}| < r^i |q_n|$$

and hence:

$$\sum_{i=N+1}^{\infty} |q_i| = \sum_{i=1}^{\infty} |q_{N+i}| < |q_n| \sum_{i=1}^{\infty} r^i$$

and by the comparison criterion, we will have that the series associated with q_n converges absolutely.
For the second case, instead, we will have that for an infinite number of indices $|q_{n+1}| > |q_n|$, and hence the series diverges. $\square$

We conclude this section with a generalization of the Cauchy-Hadamard theorem for quaternions, which will be useful for finding the radius of convergence of quaternionic power series that we will analyze.
But first, let's introduce the concept of series of functions and quaternionic power series.

Definition 4.19 (Convergent sequence of quaternionic functions)**.** *We will say that the sequence of functions* $\{f_n\}_{n=1}^{\infty}$*, where* $\forall n \in \mathbb{N}$*,* $f_n : \mathbb{H} \to \mathbb{H}$*,* *converges to* $f : \mathbb{H} \to \mathbb{H}$ *if* $\forall q \in \mathbb{H}$ *and* $\forall \epsilon > 0$ *there exists a natural number* $N(\epsilon, q)$ *such that* $\forall n \geq \mathbb{N}$

$$|f_n(q) - f(q)| < \epsilon$$

if this is the case, we will write $\lim_{n\to\infty} f_n(q) = f(q)$.

Definition 4.20 (Uniformly Convergent Sequence of Quaternionic Functions)**.** *We say that a sequence of quaternionic functions* $\{f_n\}_{n=1}^{\infty}$ *converges uniformly to* $f : \mathbb{H} \to \mathbb{H}$ *if* $\forall \epsilon > 0$*, there exists a natural number* $N(\epsilon)$ *such that:*

$$|f_n(q) - f(q)| < \epsilon \quad \forall n \geq N \ \ and \ \ \forall q \in \mathbb{H}$$

Once we have defined the concept of sequences of quaternionic functions, we can define the concept of series of quaternionic functions.

Definition 4.21 (Series of Quaternionic Functions)**.** *Let* $\{f_n\}_{n=1}^{\infty}$ *be a sequence of quaternionic functions, i.e., a sequence where* $\forall n$*,* $f_n : \mathbb{H} \to \mathbb{H}$*. We call the pointwise sum of functions* $S_m = f_1(q) + f_2(q) + \cdots + f_m(q)$ *the* m*-th partial sum of* f_n*. We call the sequence* $\{S_m\}_{m=1}^{\infty}$ ***the series of functions associated*** *with the sequence of functions* $\{f_n\}_{n=1}^{\infty}$*. If such series of functions converges to a function* $\phi : \mathbb{H} \to \mathbb{H}$*, we will formally write:*

$$\phi(q) = \sum_{n=1}^{\infty} f_n(q) \tag{4.12}$$

If the sequence of partial sums of a sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ *converges uniformly, then we will say that the associated series* $\sum_{n=1}^{\infty} f_n(q)$ *is a* ***uniformly convergent series****.*
We can define, as a particular case of quaternionic function series, the concept of ***left/right quaternionic power series****.*

Definition 4.22 (Left/Right Quaternionic Power Series)**.** *We call quaternionic function series of the type* $\displaystyle\sum_{n=0}^{\infty} a_n(q - q_0)^n$ *left quaternionic power series; moreover, we call the quaternion* q_0 *the center of the series.*
In a completely analogous manner, we call series of functions of the type $\displaystyle\sum_{n=0}^{\infty}(q-q_0)^n a_n$ *right quaternionic power series, and in both cases we call the quaternion* q_0 *the center of the series.*

Theorem 4.11 (Cauchy-Hadamard Theorem)**.** *Let* $\sum_{n=0}^{\infty} a_n(q-q_0)^n$ *be a left quaternionic power series, then there exists a positive extended real number* $R \in \mathbb{R}^+ \cup \{+\infty\}$ *called the radius of convergence of the series, for which the series converges absolutely for all values of* q *such that* $|q-q_0| < R$ *and diverges for all values such that* $|q-q_0| > R$. *We can express the radius of convergence R with the following formula in terms of the sequence of coefficients of the power series:*

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}} \tag{4.13}$$

Proof. Assume that $q_0 = 0$ without loss of generality. In the case where $R = 0$ ($\frac{1}{R} = \infty$) we trivially have that the series diverges everywhere.
So let's assume that $0 \leq \frac{1}{R} < \infty$ and that $|q| < R$; then $\exists r \in \mathbb{R}$ such that $|q| < r < R$.
Since $\frac{1}{r} > \frac{1}{R}$, we know from the definition of $\limsup$ that there exists a natural number N such that $\forall n \geq N$ $|a_n|^{\frac{1}{n}} < \frac{1}{r}$. From this last inequality, it follows that:

$$|a_n| < \frac{1}{r^n} \Longrightarrow |a_n q^n| < \left(\frac{|q|}{r}\right)^n$$

But we know that $|q| < r$, and hence $\frac{|q|}{r} < 1$. From this, it follows that the series $\sum_{n=0}^{\infty} \left(\frac{|q|}{r}\right)^n$ converges, therefore the series $\sum_{n=0}^{\infty} a_n q^n$ converges absolutely.
Conversely, suppose that $|q| > R$. Similarly, we will find a real number $r \in \mathbb{R}$, such that $|q| > r > R$. Since $\frac{1}{r} < \frac{1}{R}$ here, on the contrary, always by the definition of $\limsup$, we will have that $|a_n|^{\frac{1}{n}} > \frac{1}{r}$ for infinite values of n. From which follows, with an algebraic manipulation entirely analogous to that done earlier, that $|a_n q^n| > \left(\frac{|q|}{r}\right)^n$ for infinite values of n. Since here we have that $|q| > r$, we deduce that $\frac{|q|}{r} > 1$, i.e., the terms of the series of functions are greater than 1 for infinite values of n, and therefore the series diverges. □

Weierstrass's M-test is also valid for quaternions, which is used to verify whether an infinite series of functions converges uniformly. The proof is entirely analogous (with a few adjustments) to that provided in a classic function theory course, let's report it here below.

Theorem 4.12 (Weierstrass M-test)**.** *Let* $\{f_n\}$ *be a sequence of quaternionic functions* $f_n : U \to \mathbb{H}$. *If there exists a sequence of positive real* *numbers $\{M_n\}$ such that:*

$$\sum_{n=1}^{\infty} M_n = M \in \mathbb{R} \quad (i.e \sum_{n=1}^{\infty} M_n \; converges)$$

$$|f_n(q)| \leq M_n \quad \forall n \geq 1 \quad \forall q \in U$$

then the series

$$\sum_{n=1}^{\infty} f_n(q)$$

converges uniformly over U.

Proof. Let Σ_m be the m-th partial sum defined as:

$$\Sigma_m = \sum_{k=1}^{m} f_k(q)$$

Since the series associated with the real sequence $\{M_n\}$ converges, we will have that, by the Cauchy criterion, $\forall \epsilon > 0$, $\exists N$ such that $\forall m > n > N$:

$$\sum_{k=n+1}^{m} M_k < \epsilon$$

Choosing an N, we will have that $\forall q \in U$, $\forall m > n > N$

$$|\Sigma_m - \Sigma_n| = \left| \sum_{k=n+1}^{m} f_k(q) \right| \leq \sum_{k=n+1}^{m} |f_k(q)| \leq \sum_{k=n+1}^{m} M_k < \epsilon \tag{4.14}$$

The sequence of partial sums $\{\Sigma_n\}$ is therefore a sequence of Cauchy quaternions, and this implies, due to the completeness of the metric space of quaternions, that it is a convergent sequence. Let $\Sigma(q)$ be the function to which such a sequence converges: taking the limit $m \to \infty$ of equation (4.14) we will have that:

$$|\Sigma(q) - \Sigma_n(q)| = \lim_{m \to \infty} |\Sigma_m(q) - \Sigma_n(q)| \leq \epsilon$$

hence, the function converges uniformly in U. □

The theorem just proved allows us to demonstrate the following equally important result.

Proposition 4.19. *Let $\sum_{n=0}^{\infty} a_n(q-q_0)^n$ be a left quaternionic power series, and let R be its convergence radius as defined earlier. Then, given a positive real number $r > 0$ such that $0 < r < R$, the series converges absolutely in the closed ball $\overline{B}(q_0, r)$.*

Proof. Here too, we will set $q_0 = 0$; let $r \in \mathbb{R}^+$ be such that $0 < r < R$, where $R \neq 0$ is the radius of convergence of our series. Let ρ be a positive real number such that $0 < r < \rho < R$. From this inequality, it follows that $\frac{1}{\rho} > \frac{1}{R} = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$.
From the just obtained relation, we will have the guarantee of the existence of an N such that $\forall n \geq N$

$$\sqrt[n]{|a_n|} < \frac{1}{\rho} \Longrightarrow |a_n| < \frac{1}{\rho^n}$$

Let now $q \in \overline{B}(0, r)$, from the above inequality it follows that $\forall n \geq N$:

$$|a_n q^n| < \left(\frac{|q|}{\rho}\right)^n \leq \left(\frac{r}{\rho}\right)^n$$

This condition, along with the fact that $\sum_n^{\infty} \left(\frac{r}{\rho}\right)^n$ converges, forms precisely the set of prerequisites for the Weierstrass M-test; therefore our series $\sum_{n=0}^{\infty} a_n q^n$ is a uniformly convergent series in $\overline{B}(0, r)$. □

Theorem 4.13 (Ratio Test)**.** *Let $\displaystyle\sum_{n=0}^{\infty} a_n(q-q_0)^n$ be a left quaternionic power series with radius of convergence equal to R. Then we will have that:*

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} \tag{4.15}$$

if such limit exists.

Proof. Let's also set $q_0 = 0$ without loss of generality. Suppose that the sequence $\frac{|a_n|}{|a_{n+1}|}$ is convergent, and let $r \in \mathbb{R}$ be a positive real number such that:

$$|q| < r < \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$$

Being $\frac{|a_n|}{|a_{n+1}|}$ convergent, we will be able to find a natural number N such that, $\forall n \geq N$:

$$\frac{|a_n|}{|a_{n+1}|} > r$$

Let $A = |a_N|r^N$. We observe that, as a consequence of the above inequality:

$$|a_{N+k}|r^{N+k} = |a_{N+k-1}|r^{N+k-1}\frac{|a_{N+k}|}{|a_{N+k-1}|}r < |a_{N+k-1}|r^{N+k-1}$$

where here $k \in \mathbb{N}$. From this, it follows that $|a_n|r^n \leq A \ \forall n \geq N$. Manipulating the just obtained inequality we obtain:

$$|a_n q^n| = |a_n|r^n\left(\frac{|q|}{r}\right)^n \leq A\left(\frac{|q|}{r}\right)^n \quad \forall n \geq N$$

But since $|q| < r$, then $\dfrac{|q|}{r} < 1$ and the geometric series $\displaystyle\sum_{n=0}^{\infty}\left(\frac{|q|}{r}\right)^n$ converges, and as a consequence of this, by the comparison criterion, we will have that also the series $\displaystyle\sum_{n=0}^{\infty}|a_n q^n|$ converges, i.e. the series $\displaystyle\sum_{n=0}^{\infty} a_n q^n$ converges absolutely for $|q| < r$. However, since r is an arbitrary real number, and being this true for every $|q| < r < \lim_{n\to\infty}\dfrac{|a_n|}{|a_{n+1}|}$, we must necessarily have that the radius of convergence R of our series is greater than the limit of the sequence $\dfrac{|a_n|}{|a_{n+1}|}$, $R \geq \lim_{n\to\infty}\dfrac{|a_n|}{|a_{n+1}|}$.
Let now $r \in \mathbb{R}^+$ be such that:

$$|q| > r > \lim_{n\to\infty}\frac{|a_n|}{|a_{n+1}|}$$

(always assuming the convergence of the sequence $\dfrac{|a_n|}{|a_{n+1}|}$). For the convergence of the aforementioned, we will have that there exists a natural number N such that, $\forall n \geq N$, $\dfrac{|a_n|}{|a_{n+1}|} < r$. We again call $A = |a_N|r^N$, and with a procedure analogous to the previous one, we will obtain that:

$$|a_{N+k}|r^{N+k} > |a_{N+k-1}|r^{N+k-1} \quad k \in \mathbb{N}$$

From this, it follows that $|a_n|r^n \geq A \ \forall n \geq N$. As a consequence of this inequality, we obtain that:

$$|a_n q^n| = |a_n|r^n\left(\frac{|q|}{r}\right)^n \geq A\left(\frac{|q|}{r}\right)^n \quad \forall n \geq N$$

But since this time $|q| > r$, we will have that $\dfrac{|q|}{r} > 1$ and hence the geometric series $\displaystyle\sum_{n=0}^{\infty}\left(\frac{|q|}{r}\right)^n$ diverges. From this, it follows, by the comparison criterion,

that also the series $\sum_{n=0}^{\infty} a_n q^n$ diverges for $|q| > r$. Therefore, for considerations entirely analogous to those made earlier, we must necessarily have $R \leq \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$.
In these two "proof steps" of the theorem, we have therefore obtained the following inequalities:

$$R \leq \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} \quad \textit{and} \quad R \geq \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$$

from which it follows that $R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}$, proving the assertion.

□

Exercise 4.12. *Demonstrate that, given $\{f_n\}_{n\in\mathbb{N}}$ a sequence of continuous quaternionic functions $f_n : \mathbb{H} \to \mathbb{H}\ \forall n$, if it converges uniformly to $f : \mathbb{H} \to \mathbb{H}$, then f is continuous.*

Chapter 5

Elementary Functions

In this chapter, we will discuss the extensions to the quaternion skew-field of all the most important elementary functions on the reals (exponential, logarithm, trigonometric functions, hyperbolic functions), deriving operational expressions for the aforementioned.

5.1 Quaternionic Exponential

We begin this section by extending the exponential function $\exp(x)$ to the quaternions, of which we already know its real and complex version. Similarly to the complex case, to derive an expression for the latter we will use its Taylor series expansion, which in the real and complex case is:

$$\exp(x) = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

Within the chapter we will interchangeably use e^x and $\exp(x)$, the two most standard notations for the exponential function, depending on which of the two seems more suggestive judging by the circumstances.

Now consider the same infinite series on the quaternion skew-field, $\displaystyle\sum_{n=0}^{\infty} \frac{q^n}{n!}$.

Observing that $|q^n| \leq |q|^n \ \forall q \in \mathbb{H}$, we will have that, by the comparison criterion, the real sequence

$$\sum_{n=0}^{\infty} \frac{|q^n|}{n!}$$

converges, as each term is less than that of the other real sequence (which is convergent) $\displaystyle\sum_{n=0}^{\infty} \frac{|q|^n}{n!}$, which is precisely equal to $e^{|q|}$. Alternatively, we could have observed that the series thus defined is a quaternionic series that converges absolutely everywhere using the ratio criterion (see theorem 4.13 from chapter 4) by calculating its radius of convergence:

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = \lim_{n\to\infty} \frac{(n+1!)}{n!} = \lim_{n\to\infty}(n+1) = \infty$$

With this, we have proved that the series $\sum_{n=0}^{\infty} \frac{q^n}{n!}$ is absolutely convergent $\forall q \in \mathbb{H}$.
At this point we know that in the quaternions a function defined as

$$\exp(q) = e^q = \sum_{n=0}^{\infty} \frac{q^n}{n!} \tag{5.1}$$

is well defined and assumes a finite value $\forall q \in \mathbb{H}$. Now all that remains is to obtain an algebraic form of the latter that is easier to use.
Let $q \in \mathbb{H}$ be a quaternion, let $\mathrm{Sc}(q) = q_0$ be its scalar part and $\mathrm{Vec}(q)$ its vector part. Consider the two quaternionic series:

$$e^{q_0} = \sum_{n=0}^{\infty} \frac{q_0^n}{n!} \quad and \;\; e^{\mathrm{Vec}(q)} = \sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^n}{n!}$$

then both converge, the first as it is precisely the real function e^{q_0}, and the second by an argument similar to the one mentioned earlier.
Taking the Cauchy product of the two series we obtain:

$$e^{q_0} e^{\mathrm{Vec}(q)} = \sum_{l=0}^{\infty} \Big[\sum_{n=0}^{l} \frac{q_0^n}{n!} \frac{\mathrm{Vec}(q)^{l-n}}{(l-n)!} \Big] = \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{n=0}^{l} \binom{l}{n} q_0^n \,\mathrm{Vec}(q)^{l-n}$$

$$\Longrightarrow e^{q_0} e^{\mathrm{Vec}(q)} = \sum_{l=0}^{\infty} \frac{(q_0 + \mathrm{Vec}(q))^l}{l!} = e^{q_0 + \mathrm{Vec}(q)} \tag{5.2}$$

Finally, to obtain an expression for the quaternionic function e^q, we prove and use the following result:

Lemma 5.1. *Let* $q \in \mathbb{H}$ *be a quaternion and* $\mathrm{Vec}(q)$ *its vector part. Then:*

$$\sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^n}{n!} = \cos|\mathrm{Vec}(q)| + \mathrm{sgn}(\mathrm{Vec}(q)) \sin|\mathrm{Vec}(q)|$$

where here $\mathrm{sgn}(\mathrm{Vec}(q)) = \frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|}$.

Proof. To prove the fact, let's start by showing that for a pure quaternion $p \in \mathbf{P}$, $p^{2n} = (-1)^n |p|^{2n}$.
We proceed by induction, considering the base case $n = 2$. We observe that, by expanding the coordinates of $p = xi + yj + zk$, with $x, y, z \in \mathbb{R}$:

$$p^2 = (xi + yj + zk)^2 = -x^2 - y^2 - z^2 = -|p|^2$$

for the inductive step, assume that it is true for $n = k$, i.e:

$$p^{2k} = (-1)^k |p|^{2k}$$

then for $n = k + 1$ we will have:

$$p^{2k+2} = p^2(p^{2k}) = (-x^2 - y^2 - z^2)(-1)^k |p|^{2k} = (-1)^{k+1} |p|^{2k+2}$$

Now, separate the series in the statement into a part with even powers and one with odd powers:

$$\sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^n}{n!} = \sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^{2n+1}}{(2n+1)!}$$

for the first, we can directly substitute $\mathrm{Vec}(q)^{2n}$ with $(-1)^n |\mathrm{Vec}(q)|^{2n}$, while for the second:

$$\sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^{2n+1}}{(2n+1)!} = \mathrm{Vec}(q) \sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^{2n}}{(2n+1)!} = \mathrm{Vec}(q) \sum_{n=0}^{\infty} \frac{(-1)^n |\mathrm{Vec}(q)|^{2n}}{(2n+1)!}$$

$$= \frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|} \sum_{n=0}^{\infty} \frac{(-1)^n |\mathrm{Vec}(q)|^{2n+1}}{(2n+1)!}$$

from which:

$$\sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^n}{n!} = \sum_{n=0}^{\infty} (-1)^n \frac{|\mathrm{Vec}(q)|^{2n}}{(2n)!} + \frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|} \sum_{n=0}^{\infty} \frac{(-1)^n |\mathrm{Vec}(q)|^{2n+1}}{(2n+1)!}$$

but the first series on the right side of the equation above is precisely $\cos(|\mathrm{Vec}(q)|)$, while the second is $\sin(|\mathrm{Vec}(q)|)$, multiplied by $\frac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|} = \mathrm{sgn}(\mathrm{Vec}(q))$, from which follows the assertion:

$$\sum_{n=0}^{\infty} \frac{\mathrm{Vec}(q)^n}{n!} = \cos|\mathrm{Vec}(q)| + \mathrm{sgn}(\mathrm{Vec}(q)) \sin|\mathrm{Vec}(q)|$$

$\square$

Combining the result just obtained with equation (5.2), we obtain the following expression for the quaternionic exponential $\exp(q) = e^q$:

Definition 5.1 (Quaternionic Exponential)**.** *We will call the following function the quaternionic exponential function, quaternionic exponential, or natural quaternionic exponential:*

$$\exp(q) = e^q = e^{\mathrm{Sc}(q)}(\cos|\mathrm{Vec}(q)| + \mathrm{sgn}(\mathrm{Vec}(q))\sin|\mathrm{Vec}(q)|)$$

where here, as usual, $\mathrm{Sc}(q)$ *and* $\mathrm{Vec}(q)$ *indicate the scalar and vector part of the quaternion* $q \in \mathbb{H}$*, and* $\mathrm{sgn}(\mathrm{Vec}(q)) = \dfrac{\mathrm{Vec}(q)}{|\mathrm{Vec}(q)|}$*. If* $|\mathrm{Vec}(q)| = 0$ *(i.e.* $q \in \mathbb{R}$*), then* $\mathrm{sgn}(\mathrm{Vec}(q))$ *is not well-defined, and in that case we define the exponential function as the classical real exponential function.*
The introduced function is defined $\forall q \in \mathbb{H}$*.*

The relationship just obtained is often called the "quaternionic Euler formula", given its great similarity to the Euler formula valid for complex numbers:

$$e^{i\theta} = \cos\theta + i\sin\theta$$

Due to the non-commutativity of the quaternions, however, we will have that the quaternionic exponential will lose some properties, such as, generally:

$$e^{q_1+q_2} \neq e^{q_1}e^{q_2} \quad ; \quad q_1, q_2 \in \mathbb{H}$$

The reader can convince themselves of this fact with the following exercise:

Exercise 5.1. *Calculate:*

- e^{i+j+k}
- $e^{\pi i}e^{\pi j}$
- $e^{\pi i+\pi j}$
- $e^{\sqrt{2}j-5k}$
- $e^{\frac{\pi}{2}j}$

Aside from the latter, the exponential still has many "desirable" properties, common to the real and complex exponential. Let's state and prove them:

Proposition 5.1 (Properties of Quaternionic Exponential)**.** *Let* $e^q : \mathbb{H} \to \mathbb{H}$ *be the quaternionic exponential just defined. Then the following properties are valid:*

- *If* $q_1, q_2 \in \mathbb{H}$ *are two quaternions that commute with each other, i.e.* $q_1 q_2 = q_2 q_1$*, then* $e^{q_1+q_2} = e^{q_1}e^{q_2}$.
- $e^{\mathrm{sgn}(q)\pi} = -1\ \forall q \in \mathbb{H}$.
- $e^{-q}e^{q} = e^{q}e^{-q} = 1\ \forall q \in \mathbb{H}$.
- $(e^q)^n = e^{nq}$ *for* $n \in \mathbb{N}$, $\forall q \in \mathbb{H}$.
- $\overline{e^q} = e^{\bar{q}}$.

Proof: Let's prove all the statements in order:

- Let's start by proving the first part of the proposition: suppose that $q_1, q_2 \in \mathbb{H}$ are two quaternions that commute with each other.
  The infinite series $e^{q_1} = \sum_{n=0}^{\infty} \frac{q_1^n}{n!}$ and $e^{q_2} = \sum_{n=0}^{\infty} \frac{q_2^n}{n!}$ are uniformly convergent, and their Cauchy product is:
  $$e^{q_1}e^{q_2} = \sum_{n=0}^{\infty} \frac{q_1^n}{n!} \sum_{m=0}^{\infty} \frac{q_2^m}{m!} = \sum_{l=0}^{\infty} \Big[\sum_{n=0}^{l} \frac{q_1^n}{n!} \frac{q_2^{l-n}}{(l-n)!} \Big] = \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{n=0}^{l} \binom{l}{n} q_1^n q_2^{l-n}$$
  but since $q_1 q_2 = q_2 q_1$, then $(q_1+q_2)^l = \sum_{n=0}^{l} \binom{l}{n} q_1^n q_2^{l-n}$ and therefore:
  $$\sum_{l=0}^{\infty} \frac{1}{l!} \sum_{n=0}^{l} \binom{l}{n} q_1^n q_2^{l-n} = \sum_{l=0}^{\infty} \frac{(q_1+q_2)^l}{l!} = e^{q_1+q_2}$$
- $e^{\mathrm{sgn}(q)\pi} = \cos(\pi) + \mathrm{sgn}(q)\sin(\pi) = \cos(\pi) = -1$
- This follows as an immediate corollary of point 1. Indeed, we observe that q and $-q$ commute with each other:
  $$q(-q) = -q^2 = (-q)q$$
  and thus $e^{-q}e^{q} = e^{-q+q} = e^0 = 1$, $\forall q \in \mathbb{H}$.
- This is also an immediate corollary of the first point. However, in this case, we proceed by induction. For the base case ($n = 2$), we observe that q trivially commutes with itself and thus:
  $$(e^q)^2 = e^q e^q = e^{q+q} = e^{2q}$$

For the inductive step, assume that this identity is true for any $n = k$:

$$(e^q)^k = e^{kq}$$

then, for $n = k + 1$:

$$(e^q)^{k+1} = e^q (e^q)^k = e^q e^{kq}$$

but observing that q and kq commute with each other (being $k \in \mathbb{N}$) then:

$$(e^q)^{k+1} = e^q e^{kq} = e^{(k+1)q}$$

- $\overline{e^q} = e^{q_0}(\cos|\operatorname{Vec}(q)| - \operatorname{sgn}(\operatorname{Vec}(q))\sin|\operatorname{Vec}(q)|) = e^{q_0}(\cos|-\operatorname{Vec}(q)| - \operatorname{sgn}(\operatorname{Vec}(q))\sin|-\operatorname{Vec}(q)|)$.
  But $\operatorname{sgn}(-\operatorname{Vec}(q)) = -\operatorname{sgn}(\operatorname{Vec}(q))$ and therefore:

$$\overline{e^q} = e^{q_0}(\cos|-\operatorname{Vec}(q)| + \operatorname{sgn}(-\operatorname{Vec}(q))\sin|-\operatorname{Vec}(q)|) =$$

$$e^{\operatorname{Sc}(\overline{q})}(\cos|\operatorname{Vec}(\overline{q})| + \operatorname{sgn}(\operatorname{Vec}(\overline{q}))\sin|\operatorname{Vec}(\overline{q})|) = e^{\overline{q}}$$

□

Furthermore, in accordance with the real and complex exponential function, the quaternionic exponential function is also never equal to 0 when calculated at a point in its domain, i.e., $e^q \neq 0$, $\forall q \in \mathbb{H}$.
The following definition of the exponential function for $\mathbb{H}$ is therefore certainly convenient for these properties, but we want to verify others: for example, for the reals and complexes we know the following identity to be valid:

$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n$$

We wonder if this identity is also valid for the quaternionic exponential; the answer is affirmative, let's prove it in the following proposition:

Proposition 5.2. *Let $q \in \mathbb{H}$, and denote e^q, as usual, the natural quaternionic exponential function. Then:*

$$e^q = \lim_{n\to\infty}\left(1 + \frac{q}{n}\right)^n \tag{5.3}$$

Proof. To prove the proposition, we must first obtain an expression for the quaternionic sequence $\left(1 + \frac{q}{n}\right)^n$. Let $q := a + bi + cj + dk$, then the sequence $1 + \frac{q}{n}$ is equal to $1 + \frac{a}{n} + \frac{b}{n}i + \frac{c}{n}j + \frac{d}{n}k$.

Now express the quaternion just obtained in polar form, and by exploiting the De Moivre formula (chapter 2 section 3) we will have:

$$\left(1+\frac{q}{n}\right)^n = \left(1+\frac{2a}{n}+\frac{a^2+b^2+c^2+d^2}{n^2}\right)^{\frac{n}{2}}\Bigg[\cos\left(n\arccos\left(\frac{1+\frac{a}{n}}{\sqrt{1+\frac{2a}{n}+\frac{a^2+b^2+c^2+d^2}{n^2}}}\right)\right)$$

$$+\frac{bi+cj+dk}{\sqrt{b^2+c^2+d^2}}\sin\left(n\arccos\left(\frac{1+\frac{a}{n}}{\sqrt{1+\frac{2a}{n}+\frac{a^2+b^2+c^2+d^2}{n^2}}}\right)\right)$$

now, observing that:

$$\lim_{n\to\infty} n\arccos\left(\frac{1+\frac{a}{n}}{\sqrt{1+\frac{2a}{n}+\frac{a^2+b^2+c^2+d^2}{n^2}}}\right) = \sqrt{b^2+c^2+d^2} = |\operatorname{Vec}(q)|$$

and

$$\lim_{n\to\infty}\left(1+\frac{2a}{n}+\frac{a^2+b^2+c^2+d^2}{n^2}\right)^{\frac{n}{2}} = e^a = e^{\operatorname{Sc}(q)}$$

we will have

$$\lim_{n\to\infty}\left(1+\frac{q}{n}\right)^n = e^{\operatorname{Sc}(q)}(\cos|\operatorname{Vec}(q)| + \operatorname{sgn}(\operatorname{Vec}(q))\sin|\operatorname{Vec}(q)|) = e^q$$

□

The quaternionic exponential also has other properties in common with the complex one. The reader may recall the famous identity:

$$e^{i\pi} = -1$$

We observe that, through a direct calculation, we obtain:

$$\begin{cases} e^{\pi i} = -1 \\ e^{\pi j} = -1 \\ e^{\pi k} = -1 \end{cases}$$

from which follows that we can write a quaternionic analogue of the famous identity $e^{i\pi}+1=0$, often called the ”most beautiful formula in mathematics” as:

$$e^{\pi i}+e^{\pi j}+e^{\pi k}+3=0 \tag{5.4}$$

This identity is an immediate corollary of the second relationship demonstrated in proposition 5.1, which states that e raised to any pure normal quaternion is equal to -1.

Exercise 5.2. *Prove that the quaternionic exponential function e^q is a continuous function on all $\mathbb{H}$.*

5.2 Quaternionic Natural Logarithm

We will now derive an expression for the quaternionic natural logarithm, defining it based on the following relationship:

$$w = \log(q) \quad if \quad q = e^w$$

from which, writing q in polar form, we obtain:

$$|q|(\cos\theta + \operatorname{sgn}(\operatorname{Vec}(q))\sin\theta) = e^{w_0}(\cos|\operatorname{Vec}(w)| + \operatorname{sgn}(\operatorname{Vec}(w))\sin|\operatorname{Vec}(w)|)$$

from which, equating modulus, argument, and sign, we obtain:

$$\begin{cases} e^{w_0} = |q| \implies w_0 = \ln|q| \\ \operatorname{sgn}(\operatorname{Vec}(q)) = \operatorname{sgn}(\operatorname{Vec}(w)) \\ |\operatorname{Vec}(w)| = \theta + 2\pi n = \arg(q) \quad n \in \mathbb{Z} \end{cases}$$

from which we obtain the following expression for the quaternionic natural logarithm:

$$\log(q) = \ln|q| + \operatorname{sgn}(\operatorname{Vec}(q))\arg(q) \tag{5.5}$$

As in the case of the complex logarithm, it is possible to define the principal value of $\log(q)$ by restricting the argument $\arg(q) = \theta + 2\pi n$ to the principal argument $\operatorname{Arg}(q) = \theta$, with $\theta \in (-\pi, \pi]$.

Definition 5.2 (Principal Value of the Quaternionic Logarithm)**.** *We call the **principal value of the quaternionic logarithm**, denoted by* $\operatorname{Log}(q)$*, the following expression:*

$$\operatorname{Log}(q) = \ln|q| + \operatorname{Arg}(q)\operatorname{sgn}(\operatorname{Vec}(q)) = \ln|q| + \theta\operatorname{sgn}(\operatorname{Vec}(q)) \tag{5.6}$$

$\theta \in (-\pi, \pi]$.

As in the complex case, it is possible to select branch lines and consider the branches of $\log(q)$ whose domain is the quaternions excluding the said branch line; the branches thus obtained are effectively functions.

Example 5.1. *Solve the following quaternionic equation:*

$$e^q = i \implies q = \log(i) = \ln|i| + \operatorname{sgn}(i)\arg(i)$$

but since $|i| = 1$, $\ln|i| = 0$, $\operatorname{sgn}(i) = i$ *and* $\arg(i) = \frac{\pi}{2} + 2\pi n \; n \in \mathbb{Z}$*, we have:*

$$q = \left(\frac{\pi}{2} + 2\pi n\right)i = \frac{(4n+1)\pi}{2}i \quad n \in \mathbb{Z}$$

The solutions are exactly the solutions of the same equation in the complex field.

Example 5.2. *Now solve*

$$e^q = j$$

we see that also for the other imaginary units of $\mathbb{H}$ *we will have results entirely analogous to those obtained with the unit i. As before:*

$$q = \log(j) = \ln|j| + \operatorname{sgn}(j)\arg(j)$$

and as before, we have $\ln|j| = 0$ *and* $\operatorname{sgn}(j) = j$*. Moreover,* $\arg(j) = \frac{\pi}{2} + 2\pi n$ *and therefore:*

$$q = \Big(\frac{\pi}{2} + 2\pi n\Big)j = \frac{(4n+1)\pi}{2}j \quad n \in \mathbb{Z}$$

In an entirely analogous way, also $e^q = k \implies q = \frac{(4n+1)\pi}{2}k$. We can rewrite these identities more compactly as:

$$e^{\frac{(4n+1)}{2}\pi e_i} = e_i \quad i = 1, 2, 3$$

Exercise 5.3. *Calculate* $\log(q)$ *for the following values of* q*:*

- $\sqrt{\pi} - \sqrt{\pi}i + ej$.
- $i - j$.

5.3 Powers of Quaternions

Using the notions of exponential and logarithm for quaternions, it is possible to define what it means to "raise q_1 to the power of q_2", where $q_1, q_2 \in \mathbb{H}$ are two quaternions.
To define the "function" (it assumes multiple values for each quaternion, so it's not strictly a function but a multivalued function) of power for the quaternions, we proceed exactly as in the complex case, i.e., we define $q_1^{q_2}$ to be:

Definition 5.3 (q_1 raised to the power of q_2)**.** *Let* q_1, q_2 *be two quaternions. Then, we define* $q_1^{q_2}$ *through the following equation:*

$$q_1^{q_2} := e^{\log(q_1)q_2} \tag{5.7}$$

Proposition 5.3 (Properties of the Quaternionic Power Function)**.** *Let* $q_1, q_2, q_3 \in \mathbb{H}$*, then:*

- $(q_1^{q_2})^n = q_1^{nq_2}$.
- *If* $\log(q_1)q_2$ *and* $\log(q_1)q_3$ *commute, then* $q_1^{q_2}q_1^{q_3} = q_1^{q_2+q_3}$.

Proof. Follows as an immediate corollary of the properties of the exponential e^q demonstrated in section 1 of this chapter. □

In this case too, we can select a principal value of the multivalued power function, let's formally define it here:

Definition 5.4 (Principal Value of the Quaternionic Power Function)**.** *Let* q_1, q_2 *be two quaternions; we call the principal value of the quaternionic power function the quaternion:*

$$q_1^{q_2} = e^{\mathrm{Log}(q_1)q_2}$$

that is, the value obtained by restricting the quaternionic logarithm to its principal value $\mathrm{Log}(q) = \ln|q| + \mathrm{sgn}(q)\,\mathrm{Arg}(q)$.

Example 5.3. *Consider quaternionic powers of the type* $e_i^{e_j}$, *where* $i, j = 1, 2, 3$ *and, as usual,* e_i, e_j *are the quaternionic imaginary units.*
Let's start with the first case, $i^{\,i}$. *We observe that, according to the definition of one quaternion raised to the power of another quaternion:*

$$i^{\,i} = e^{\log(i)i}$$

but, as seen in the previous section, $\log(i) = \frac{(4n+1)}{2}\pi i$, *from which :*

$$i^{\,i} = e^{-\frac{(4n+1)}{2}\pi} \qquad n \in \mathbb{Z}$$

Its principal value is

$$i^{\,i} = e^{\mathrm{Log}(i)i} = e^{-\frac{\pi}{2}}$$

The result is congruent with the same result often demonstrated in complex analysis courses. We see that, with almost identical calculations, this result also applies to the other imaginary units, i.e.:

$$j^j = e^{-\frac{\pi}{2}}$$

$$k^k = e^{-\frac{\pi}{2}}$$

Summarizing these 3 results in a single proposition, we can say that the quaternionic imaginary units (i, j, k) *raised to the power of themselves result in a real number, precisely* $e^{-\frac{\pi}{2}}$.

Now we just have to check the cases of $e_i^{e_j}$ where $i \neq j$. Starting from i^j: by the definition of quaternionic powers given earlier, we have:

$$i^j = e^{\log(i)j} = e^{\frac{(4n+1)\pi}{2}ij} = e^{\frac{(4n+1)\pi}{2}k} = k$$

Similarly for j^i, i^k, k^i, j^k, k^j we will have:

$$j^i = e^{\log(j)i} = e^{\frac{(4n+1)\pi}{2}ji} = -k$$

$$i^k = -j$$

$$k^i = j$$

$$k^j = -i$$

$$j^k = i$$

We can combine all the results just obtained for powers of quaternionic imaginary units to other imaginary units not equal to themselves with the equation:

$$e_i^{e_j} = e_i e_j \quad i, j \in \{1,2,3\} \;\; i \neq j \tag{5.8}$$

Including also the cases where i and j are equal, we can write the results of this example in the following single identity:

$$e_i^{e_j} = \begin{cases} e^{-\frac{(4n+1)\pi}{2}} & if \;\; i = j \\ e_i e_j & if \;\; i \neq j \end{cases}$$

Alternatively, it can be rewritten in a more compact form using the Kronecker delta:

$$e_i^{e_j} = e^{-\frac{(4n+1)\pi}{2}}\delta_{ij} + (1-\delta_{ij})e_i e_j \tag{5.9}$$

Corollary 5.1. *Let $i, j \in \{1,2,3\}$, and let e_i, e_j be quaternionic imaginary units ($e_1 = i, e_2 = j, e_3 = k$), then the following identity is verified:*

$$\sum_{i \neq j} e_i^{e_j} = i^j + j^k + i^k + k^j + k^i + j^i = 0 \tag{5.10}$$

Example 5.4. *Let's work on the following example, which is slightly more laborious:*

$$(i+j)^k = e^{\log(i+j)k}$$

With a direct calculation, we see that (using the definition of quaternionic logarithm):

$$\log(i+j) = \frac{1}{2}\ln\, 2 + \Big(\frac{i}{\sqrt{2}} + \frac{j}{\sqrt{2}}\Big)\frac{(4n+1)\pi}{2}$$

from which

$$(i+j)^k = e^{(\frac{ln\ 2}{2}+\frac{(4n+1)\pi i}{2\sqrt{2}}+\frac{(4n+1)\pi j}{2\sqrt{2}})k}$$

Developing the multiplication in the exponent, and using the definition of quaternionic exponential we finally obtain:

$$(i+j)^k =$$

$$\cos\left(\sqrt{\frac{ln^2\ 2}{4}+\frac{(4n+1)^2\pi^2}{4}}\right)+\frac{\frac{kln\ 2}{2}-\frac{(4n+1)\pi j}{2\sqrt{2}}+\frac{(4n+1)\pi i}{2\sqrt{2}}}{\sqrt{\frac{ln^2\ 2}{4}+\frac{(4n+1)^2\pi^2}{4}}}\sin\left(\sqrt{\frac{ln^2\ 2}{4}+\frac{(4n+1)^2\pi^2}{4}}\right)$$

where $n \in \mathbb{Z}$ is an integer.
The principal value of $(i+j)^k$, however, is:

$$(i+j)^k = \cos\left(\sqrt{\frac{ln^2\ 2}{4}+\frac{\pi^2}{4}}\right)+\frac{\frac{kln\ 2}{2}-\frac{\pi j}{2\sqrt{2}}+\frac{\pi i}{2\sqrt{2}}}{\sqrt{\frac{ln^2\ 2}{4}+\frac{\pi^2}{4}}}\sin\left(\sqrt{\frac{ln^2\ 2}{4}+\frac{\pi^2}{4}}\right)$$

5.4 Sine and Cosine of Quaternions

In a complete discussion of extending elementary functions to the skew-field of quaternions, it is essential to include a treatment of the trigonometric functions on $\mathbb{H}$.
To extend these functions, we will proceed similarly to how we did for the quaternionic exponential, i.e., we will define them starting from their infinite series, in the following way:

$$\sin(q) := \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}q^{2n+1} \quad q \in \mathbb{H} \tag{5.11}$$

$$\cos(q) := \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}q^{2n} \quad q \in \mathbb{H} \tag{5.12}$$

However, to ensure that this definition makes sense, we first need to verify that these series are convergent for all $q \in \mathbb{H}$. The strategy is not much different from that used for the exponential function: it's enough to notice that, since for all $q \in \mathbb{H}$, $|q^n| \leq |q|^n$, and the series

$$\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}|q|^{2n+1} = \sin(|q|) \in \mathbb{R} \qquad \textit{and} \qquad \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}|q|^{2n} = \cos(|q|) \in \mathbb{R}$$

converge (precisely converging to the real values $\sin(|q|)$ and $\cos(|q|)$ respectively), we have that (by the comparison test) the series (5.11) and (5.12) absolutely converge for all $q \in \mathbb{H}$.
The next step is to derive an operational form of the just-defined trigonometric functions, which makes calculations simpler and possibly explicates the scalar and vectorial part of the function.
Let's start by writing the infinite series for $e^{\pm q \ \mathrm{sgn}(\mathrm{Vec}(q))}$. To simplify the notation, given a quaternion $q \in \mathbb{H}$, let's call $\vec{v} \in$ **P** its vector part and q_0 its scalar part. Using the definition of quaternionic exponential given in section 1, we have:

$$e^{q \ \mathrm{sgn}(\vec{v})} = \sum_{n=0}^{\infty} \frac{(q \ \mathrm{sgn}(\vec{v}))^n}{n!}$$

$$e^{-q \ \mathrm{sgn}(\vec{v})} = \sum_{n=0}^{\infty} \frac{(-q \ \mathrm{sgn}(\vec{v}))^n}{n!}$$

Remembering that, as seen in chapter 2, pure normal quaternions are square roots of -1, i.e $\forall \vec{v} \in \{q \in \mathbb{H} \ ; \ |q| = 1, \ Sc(q) = 0\} \cong S_2, \ \vec{v}^2 = -1$, we have:

$$e^{q \ \mathrm{sgn}(\vec{v})} = \sum_{n=0}^{\infty} \frac{(-1)^n q^{2n}}{(2n)!} + \mathrm{sgn}(\vec{v}) \sum_{n=0}^{\infty} \frac{(-1)^n q^{2n+1}}{(2n+1)!} = \cos(q) + \mathrm{sgn}(\vec{v}) \sin(q)$$

$$e^{-q \ \mathrm{sgn}(\vec{v})} = \sum_{n=0}^{\infty} \frac{(-1)^n q^{2n}}{(2n)!} - \mathrm{sgn}(\vec{v}) \sum_{n=0}^{\infty} \frac{(-1)^n q^{2n+1}}{(2n+1)!} = \cos(q) - \mathrm{sgn}(\vec{v}) \sin(q)$$

From this, we get:

$$\cos(q) = \frac{1}{2}(e^{qsgn(\vec{v})} + e^{-qsgn(\vec{v})})$$

$$\sin(q) = \frac{-sgn(\vec{v})}{2}(e^{qsgn(\vec{v})} - e^{-qsgn(\vec{v})})$$

This result allows us to give an equivalent definition to the one given earlier for sine and cosine of quaternions that is more useful and suggestive:

Definition 5.5 (Sine and Cosine of Quaternions)**.** *We call sine and cosine of quaternions the following functions of a quaternion variable:*

$$\sin(q) = \begin{cases} \dfrac{-\mathrm{sgn}(\vec{v})}{2}(e^{q\,\mathrm{sgn}(\vec{v})} - e^{-q\,\mathrm{sgn}(\vec{v})}) & if \ \ |\vec{v}| \neq 0 \\ \sin(q_0) \quad if \ |\vec{v}| = 0 \end{cases} \tag{5.13}$$

$$\cos(q) = \begin{cases} \frac{1}{2}(e^{q\,\mathrm{sgn}(\vec{v})} + e^{-q\,\mathrm{sgn}(\vec{v})}) & if \ |\vec{v}| \neq 0 \\ \cos(q_0) \quad if \ |\vec{v}| = 0 \end{cases} \tag{5.14}$$

Proposition 5.4 (Properties of Sine and Cosine of Quaternions)**.** *Let $q \in \mathbb{H}$ be a quaternion, let t be its real part and $\vec{v} := xi + yj + zk \in$ **P** its vector part, then:*

- $\sin(-q) = -\sin(q)$.
- $\cos(-q) = \cos(q)$.
- $\sin(q) = \sin(t)\cos(\vec{v}) + \cos(t)\sin(\vec{v})$.
- $\cos(q) = \cos(t)\cos(\vec{v}) - \sin(t)\sin(\vec{v})$.
- $\cos^2(q) + \sin^2(q) = 1$.

Proof. Let's prove the facts in order:

- To calculate $\sin(-q)$ directly:

$$\sin(-q) = -\frac{\operatorname{sgn}(-\vec{v})}{2}(e^{-q\ \operatorname{sgn}(-\vec{v})} - e^{q\ \operatorname{sgn}(-\vec{v}))}) = \frac{sgn(\vec{v})}{2}(e^{q\operatorname{sgn}(\vec{v})} - e^{-q\operatorname{sgn}(\vec{v})}) = -\sin(q)$$

  In the case where q is a real number, this identity is already well known, and therefore it is necessary to prove it only for a quaternion with a non-zero vector part.

- Similarly, for $\cos(-q)$:

$$\cos(-q) = \frac{1}{2}(e^{-q\ \operatorname{sgn}(-\vec{v})} + e^{-(-q)\ \operatorname{sgn}(-\vec{v})}) = \cos(q)$$

- To prove this fact, we must first derive some preliminary relations; first, observe that:

$$q\ \operatorname{sgn}(\vec{v}) = q_0\ \operatorname{sgn}(\vec{v}) + \vec{v}\ \operatorname{sgn}(\vec{v}) = q_0\ \operatorname{sgn}(\vec{v}) + \frac{\vec{v}^2}{|\vec{v}|} = q_0\ \operatorname{sgn}(\vec{v}) - |\vec{v}|$$

  since, as proven earlier, $\vec{v}^2 = -|\vec{v}|^2$. Moreover:

$$\operatorname{sgn}(q_0\vec{v}) = \frac{q_0\operatorname{sgn}(\vec{v})}{|q_0|} = \operatorname{sgn}(q_0)\operatorname{sgn}(\vec{v})$$

  Thanks to these facts, and remembering that, since $-|q|$ (a real number) and $q_0\ \operatorname{sgn}(\vec{v})$ commute, we have:

$$e^{q\ \operatorname{sgn}(\vec{v})} = e^{-|\vec{v}|}e^{q_0\ \operatorname{sgn}(\vec{v})} = e^{-|\vec{v}|}(\cos(q_0) + \operatorname{sgn}(\vec{v})\sin(q_0))$$

The proof of the identity now becomes very simple; from the definition of quaternionic sine we have:

$$\sin(q) = \frac{-\operatorname{sgn}(\vec{v})}{2}(e^{q\ \operatorname{sgn}(\vec{v})} - e^{-q\ \operatorname{sgn}(\vec{v})})$$

substituting the obtained relations for $e^{\pm q\ \operatorname{sgn}(\vec{v})}$:

$$\sin(q) = \frac{-\operatorname{sgn}(\vec{v})}{2}(e^{-|\vec{v}|}\cos(q_0) + e^{-|\vec{v}|}\sin(q_0)\ \operatorname{sgn}(\vec{v}) - e^{|\vec{v}|}\cos(q_0)$$

$$+e^{-|\vec{v}|}\ \operatorname{sgn}(\vec{v})\sin(q_0)) = \frac{-\operatorname{sgn}(\vec{v})}{2}(\cos(q_0)(e^{-|\vec{v}|} - e^{|\vec{v}|}) + \sin(q_0)\ \operatorname{sgn}(\vec{v})(e^{-|\vec{v}|} + e^{|\vec{v}|}))$$

Notice that, substituting the obtained relations for $e^{\pm q\ \operatorname{sgn}(\vec{v})}$ into the expressions for sine and cosine, and setting $q_0 = 0$, we get:

$$\sin(\vec{v}) = -\frac{\operatorname{sgn}(\vec{v})}{2}(e^{-|\vec{v}|} - e^{|\vec{v}|}) \tag{5.15}$$

and

$$\cos(\vec{v}) = \frac{1}{2}(e^{-|\vec{v}|} + e^{|\vec{v}|}) \tag{5.16}$$

distributing the term $-\frac{\operatorname{sgn}(\vec{v})}{2}$ in the expression for $\sin(q)$ we get:

$$\sin(q) = \frac{1}{2}\sin(q_0)(e^{|\vec{v}|} + e^{-|\vec{v}|}) - \frac{\operatorname{sgn}(\vec{v})}{2}\cos(q_0)(e^{-|\vec{v}|} - e^{|\vec{v}|})$$

which is precisely equal to, considering equations (5.15) and (5.16):

$$\sin(q) = \sin(q_0)\cos(\vec{v}) + \cos(q_0)\sin(\vec{v}) \tag{5.17}$$

- The process is almost the same; we start from the equation:

$$\cos(q) = \frac{1}{2}(e^{q\ \operatorname{sgn}(\vec{v})} + e^{-q\ \operatorname{sgn}(\vec{v})})$$

then we substitute the identities obtained in the previous point for $e^{\pm q\ \operatorname{sgn}(\vec{v})}$ obtaining:

$$\cos(q) = \frac{1}{2}(e^{-|\vec{v}|}\cos(q_0) + e^{-|\vec{v}|}\ \operatorname{sgn}(\vec{v})\sin(q_0) + e^{|\vec{v}|}\cos(q_0)$$

$$-e^{-|\vec{v}|}\ \operatorname{sgn}(\vec{v})\sin(q_0)) = \frac{1}{2}\cos(q_0)(e^{|\vec{v}|} + e^{-|\vec{v}|}) + \frac{\operatorname{sgn}(\vec{v})}{2}\sin(q_0)(e^{-|\vec{v}|} - e^{|\vec{v}|})$$

$$\implies \cos(q) = \cos(q_0)\cos(\vec{v}) - \sin(q_0)\sin(\vec{v}) \tag{5.18}$$

- First of all, let's see that, in light of equations (5.16) and (5.15) we have:
$$\sin^2(\vec{v}) + \cos^2(\vec{v}) = -\frac{1}{4}(e^{-|\vec{v}|} - e^{|\vec{v}|})^2 + \frac{1}{4}(e^{-|\vec{v}|} + e^{|\vec{v}|})^2$$
expanding the squares we obtain:
$$\sin^2(\vec{v}) + \cos^2(\vec{v}) = -\frac{1}{4}e^{-2|\vec{v}|} - \frac{1}{4}e^{2|\vec{v}|} + \frac{1}{2} + \frac{1}{4}e^{-2|\vec{v}|} + \frac{1}{4}e^{2|\vec{v}|} + \frac{1}{2} = 1$$
This fact will be useful in a moment: now, let's derive an explicit expression for $\sin^2(q)$ and $\cos^2(q)$, using equations (5.17) and (5.18)
$$\sin^2(q) = (\sin(q_0)\cos(\vec{v}) + \cos(q_0)\sin(\vec{v}))^2 = \sin^2(q_0)\cos^2(\vec{v})$$
$$+\sin(q_0)\cos(q_0)\cos(\vec{v})\sin(\vec{v}) + \cos^2(q_0)\sin^2(\vec{v}) + \cos(q_0)\sin(q_0)\sin(\vec{v})\cos(\vec{v})$$
$$\cos^2(q) = (\cos(q_0)\cos(\vec{v}) - \sin(q_0)\sin(\vec{v}))^2 = \cos^2(q_0)\cos^2(\vec{v}) + \sin^2(q_0)\sin^2(\vec{v})$$
$$-\sin(q_0)\cos(q_0)\sin(\vec{v})\cos(\vec{v}) - \cos(q_0)\sin(q_0)\cos(\vec{v})\sin(\vec{v})$$
From which, with a simple calculation, we will obtain that:
$$\sin^2(q)+\cos^2(q) = \sin^2(q_0)\cos^2(\vec{v}) + \sin^2(q_0)\sin^2(\vec{v}) + \cos^2(q_0)\sin^2(\vec{v}) + \cos^2(q_0)\cos^2(\vec{v})$$
$$= \sin^2(q_0)(\cos^2(\vec{v}) + \sin^2(\vec{v})) + \cos^2(q_0)(\sin^2(\vec{v}) + \cos^2(\vec{v})) = 1$$
therefore:
$$\sin^2(q) + \cos^2(q) = 1 \qquad \forall q \in \mathbb{H} \tag{5.19}$$

□

The expressions derived at the beginning of the section for the sine and cosine of a quaternion are certainly more useful than their expansions in series if we need to use the aforementioned functions, and they also allowed us to prove some important properties of the sine and cosine of quaternions; however, if our purpose is purely to calculate the value of the functions in question in a quaternion $q \in \mathbb{H}$, or alternatively, to explicate its real component and the coefficients of i, j, k, in most cases, a better strategy will be to use the (equivalent) expression that we will derive shortly.
First of all, let's recall a result already proven in the first section, and state it in a lemma:

Lemma 5.2. *Let $\vec{q} \in \mathbb{H}$ be a pure quaternion, i.e., a quaternion with a scalar part equal to 0,* $\mathrm{Sc}(\vec{q}) = 0$. *Then:*
$$\begin{cases} \vec{q}^{\,2n} = (-1)^n|\vec{q}|^{2n} \\ \vec{q}^{\,2n+1} = (-1)^n|\vec{q}|^{2n}\vec{q} \end{cases} \tag{5.20}$$

Now, as before, let $q \in \mathbb{H}$ be a quaternion with a scalar part $q_0 \in \mathbb{R}$ and a vector part $\vec{v} \in \mathbf{P}$. In light of Lemma 5.2, and using the definition of quaternionic sine (in terms of its infinite series expansion):

$$\sin(\vec{v}) = \sum_{n=0}^{\infty} \frac{(-1)^n \vec{v}^{\,2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{|\vec{v}|^{2n}\vec{v}}{(2n+1)!} = \sum_{n=0}^{\infty} \frac{|\vec{v}|^{2n+1}}{(2n+1)!}\ \mathrm{sgn}(\vec{v}) = \sinh(|\vec{v}|)\ \mathrm{sgn}(\vec{v})$$

similarly, for the cosine function:

$$\cos(\vec{v}) = \sum_{n=0}^{\infty} \frac{(-1)^n \vec{v}^{\,2n}}{(2n)!} = \sum_{n=0}^{\infty} \frac{|\vec{v}|^{\,2n}}{(2n)!} = \cosh(|\vec{v}|)$$

Finally, using the sum formulas derived earlier, we will have that for a quaternion $q := q_0 + \vec{v} \in \mathbb{H}$:

$$\sin(q) = \sin(q_0)\cosh(|\vec{v}|) + \cos(q_0)\sinh(|\vec{v}|)\,\mathrm{sgn}(\vec{v}) \tag{5.21}$$

$$\cos(q) = \cos(q_0)\cosh(|\vec{v}|) - \sin(q_0)\sinh(|\vec{v}|)\,\mathrm{sgn}(\vec{v}) \tag{5.22}$$

Proposition 5.5. *Let $q \in \mathbb{H}$ be a quaternion, then:*

- $\overline{\sin(q)} = \sin(\overline{q})$
- $\overline{\cos(q)} = -\cos(\overline{q})$

Proof.

- For the first point:

$$\overline{\sin(q)} = \sin(q_0)\cosh(|\vec{v}|) - \cos(q_0)\sinh(|\vec{v}|)\,\mathrm{sgn}(\vec{v}) =$$

$$\sin(q_0)\cosh(|-\vec{v}|) + \cos(q_0)\sinh(|-\vec{v}|)\ \mathrm{sgn}(-\vec{v}) = \sin(\overline{q})$$

- Similarly, for the second:

$$\overline{\cos(q)} = \cos(q_0)\cosh(|\vec{v}|) + \sin(q_0)\sinh(|\vec{v}|)\ \mathrm{sgn}(\vec{v}) =$$

$$-(-\cos(q_0)\cosh(|-\vec{v}|) + \sin(q_0)\sinh(|-\vec{v}|)\,\mathrm{sgn}(-\vec{v})) = -\cos(\overline{q})$$

$\square$

Exercise 5.4. *Using the expressions* (5.21) *and* (5.22) *for the sine and cosine function, prove that the two commute with each other, i.e*

$$\sin(q)\cos(q) = \cos(q)\sin(q) \quad \forall q \in \mathbb{H}$$

Exercise 5.5. *Prove that, for a non-zero pure quaternion, i.e. a quaternion* $\vec{q} \in \boldsymbol{P}$*,* $\vec{q} \neq 0$*,* $\mathrm{Sc}(\vec{q}) = 0$*, we have:*

$$\sin(\vec{q}) = \mathrm{sgn}(\vec{q}) \sinh(|\vec{q}|)$$

$$\cos(\vec{q}) = \cosh(|\vec{q}|)$$

Exercise 5.6. *Find all zeros of the quaternionic sine and cosine functions,* $\sin(q)$ *and* $\cos(q)$*, and prove that they are the same zeros known on the real and complex numbers, i.e:*

$$\sin(q) = 0 \Longleftrightarrow q = n\pi \quad n \in \mathbb{Z}$$

$$\cos(q) = 0 \Longleftrightarrow q = \frac{\pi}{2} + n\pi \quad n \in \mathbb{Z}$$

Exercise 5.7. *Calculate the sine and cosine function for the following quaternionic values:*

- $q_1 = -\sqrt{2}i + j \quad ; \quad q_2 = i \quad ; \quad q_3 = j \quad ; \quad q_4 = k$
- $q_5 = -3j - 2k \quad ; \quad q_6 = \sqrt{2}(1 - k)$

5.5 Tangent and Cotangent

As we saw in the exercises of the previous section, the sine and cosine functions commute, from which it follows that we can unambiguously define a quaternionic analogue of the tangent and cotangent functions.
We start by defining the quaternionic analogue of the secant and cosecant functions:

Definition 5.6 (Quaternionic Secant Function)**.** *Let* $q \in \mathbb{H}$ *be a quaternion; we define the function* $\sec(q)$ *as:*

$$\sec(q) := (\cos(q))^{-1} = \frac{1}{\cos(q)} \tag{5.23}$$

Definition 5.7 (Quaternionic Cosecant Function)**.** *Let* $q \in \mathbb{H}$ *be a quaternion; we define the function* $\csc(q)$ *as:*

$$\csc(q) := (\sin(q))^{-1} = \frac{1}{\sin(q)} \tag{5.24}$$

Now we define the tangent and cotangent functions:

Definition 5.8 (Quaternionic Tangent Function)**.** *Let $q \in \mathbb{H}$ be a quaternion; we define the function* $\tan(q)$ *as:*

$$\tan(q) := \sin(q)(\cos(q))^{-1} = (\cos(q))^{-1} \sin(q) \tag{5.25}$$

Definition 5.9 (Quaternionic Cotangent Function)**.** *Let $q \in \mathbb{H}$ be a quaternion; we define the function* $\cot(q)$ *as:*

$$\cot(q) := \cos(q)(\sin(q))^{-1} = (\sin(q))^{-1} \cos(q) \tag{5.26}$$

5.6 Hyperbolic Functions

We can also extend hyperbolic functions, $\sinh$ and $\cosh$, to quaternions, defining them based on the famous relationships with the exponential function derived in the real case:

Definition 5.10 (Quaternionic hyperbolic sine)**.** *Let $q \in \mathbb{H}$ be a quaternion; we call hyperbolic sine the following function:*

$$\sinh(q) := \frac{e^q - e^{-q}}{2} \tag{5.27}$$

Definition 5.11 (Quaternionic hyperbolic cosine)**.** *Let $q \in \mathbb{H}$ be a quaternion; we call hyperbolic cosine the following function:*

$$\cosh(q) := \frac{e^q + e^{-q}}{2} \tag{5.28}$$

Alternatively, these functions can be defined based on their well-known series expansions:

$$\sinh(q) = \sum_{n=0}^{\infty} \frac{q^{2n+1}}{(2n+1)!} \quad ; \quad \cosh(q) = \sum_{n=0}^{\infty} \frac{q^{2n}}{(2n)!}$$

It is easy to verify, however, that the two definitions are equivalent; we leave the verification of this simple fact to the reader, which can be further simplified by keeping in mind that it is an immediate corollary of the agreement of the quaternionic exponential e^q with its infinite series $\sum_{n=0}^{\infty} \frac{q^n}{n!}$ for quaternions.

Proposition 5.6 (Properties of Hyperbolic Sine and Cosine)**.** *Let $q \in \mathbb{H}$ be a quaternion, then the following identities are valid:*

- $\sinh(-q) = -\sinh(q)$
- $\cosh(-q) = \cosh(q)$
- $\cosh^2(q) - \sinh^2(q) = 1$

Proof. The proof of the first two facts is elementary, and the result is obtained by direct substitution of $-q$ into the expressions for $\sinh(q)$ and $\cosh(q)$.
For the second fact, it suffices to explicitly calculate $\cosh^2(q)$ and $\sinh^2(q)$, respectively equal to:

$$\cosh^2(q) = \frac{e^{2q} + 2 + e^{-2q}}{2}$$

$$\sinh^2(q) = \frac{e^{2q} - 2 + e^{-2q}}{2}$$

from which:

$$\cosh^2(q) - \sinh^2(q) = 1 \tag{5.29}$$

□

Exercise 5.8. *Calculate the functions* $\sinh(q)$ *and* $\cosh(q)$ *for the following quaternionic values:*

- $q_1 = i \quad ; \quad q_2 = j \quad ; \quad q_3 = k \quad ; \quad q_4 = -i + \sqrt{2}j - \pi k$
- $q_5 = \pi - i$

Exercise 5.9. *Prove that, inspired by similar theorems proved in previous sections, the formulas for sum and subtraction for the quaternionic hyperbolic sine and cosine hold only for 2 quaternions that commute with each other, i.e:*

$$\cosh(q_1 + q_2) = \cosh(q_1)\cosh(q_2) + \sinh(q_1)\sinh(q_2) \quad \Longleftrightarrow q_1q_2 = q_2q_1$$

$$\sinh(q_1 + q_2) = \sinh(q_1)\cosh(q_2) + \sinh(q_2)\cosh(q_1) \quad \Longleftrightarrow q_1q_2 = q_2q_1$$

Now let's derive a new expression for the two functions just introduced, linking them to the classic trigonometric functions and allowing us to perform calculations and demonstrate facts with these functions more easily. First, we see that

$$e^{\pm\vec{v}} = \cos|\vec{v}| \pm \operatorname{sgn}(\vec{v})\sin|\vec{v}|$$

where here $\vec{v}$ is a pure quaternion. Using the definitions of hyperbolic sine and cosine, we finally obtain the following expressions for the hyperbolic sine/cosine of a pure quaternion:

$$\cosh(\vec{v}) = \cos|\vec{v}|$$

$$\sinh(\vec{v}) = \operatorname{sgn}(\vec{v})\sin|\vec{v}|$$

Finally, combining this result with the formulas from exercise 5.9, we obtain that for a general quaternion $q \in \mathbb{H}$, with real part q_0 and vector part $\vec{v}$, i.e $q = q_0 + \vec{v}$ we have:

$$\sinh(q) = \sinh(q_0)\cos|\vec{v}| + \cosh(q_0)\operatorname{sgn}(\vec{v})\sin|\vec{v}| \tag{5.30}$$

$$\cosh(q) = \cosh(q_0)\cos|\vec{v}| + \sinh(q_0)\operatorname{sgn}(\vec{v})\sin|\vec{v}| \tag{5.31}$$

Now we state other properties of hyperbolic functions that will be useful for the rest of this chapter:

Proposition 5.7 (Further Properties of Hyperbolic Sine and Cosine)**.** *Let* $q \in \mathbb{H}$*, then the following identities are valid:*

- $\overline{\cosh(q)} = \cosh(\overline{q})$.

- $\overline{\sinh(q)} = \sinh(\overline{q})$.

Proof. For the first point:

- $\overline{\cosh(q)} = \overline{\frac{e^q+e^{-q}}{2}} = \frac{e^{\overline{q}}+e^{-\overline{q}}}{2} = \cosh(\overline{q})$.
- Similarly, for the second point:

$$\overline{\sinh(q)} = \overline{\frac{e^q - e^{-q}}{2}} = \sinh(\overline{q})$$

□

Finally, as with trigonometric functions, we can unambiguously define a quaternionic analogue of hyperbolic tangent and cotangent, tanh and coth. First, however, we define a quaternionic analogue of hyperbolic secant and cosecant:

Definition 5.12 (Hyperbolic Secant). *Let $q \in \mathbb{H}$ be a quaternion; we call the function $\operatorname{sech}(q)$ hyperbolic secant, defined as:*

$$\operatorname{sech}(q) := (\cosh(q))^{-1} = \frac{1}{\cosh(q)} \tag{5.32}$$

Definition 5.13 (Hyperbolic Cosecant). *Let $q \in \mathbb{H}$ be a quaternion; we call the function $\operatorname{csch}(q)$ hyperbolic cosecant, defined as:*

$$\operatorname{csch}(q) := (\sinh(q))^{-1} = \frac{1}{\sinh(q)} \tag{5.33}$$

Based on the identity on the multiplicative inverse of a quaternion, derived in chapter 2, we can write the two functions just introduced also as:

$$\operatorname{sech}(q) = \frac{\cosh(\overline{q})}{|\cosh(q)|^2}$$

$$\operatorname{csch}(q) = \frac{\sinh(\overline{q})}{|\sinh(q)|^2}$$

To be able to define unambiguously a concept of a quaternionic hyperbolic tangent, ideally we would want these two functions to commute with hyperbolic sine and cosine respectively, i.e:

$$\operatorname{sech}(q)\sinh(q) = \sinh(q)\operatorname{sech}(q) \quad e \quad \operatorname{csch}(q)\cosh(q) = \cosh(q)\operatorname{csch}(q)$$

This fact follows as a corollary of the following lemma:

Lemma 5.3. *The hyperbolic sine and hyperbolic cosine functions commute with each other, i.e:*

$$\sinh(q)\cosh(q) = \cosh(q)\sinh(q)$$

Proof. It suffices to observe that $(e^q - e^{-q})(e^q + e^{-q}) = e^{2q} - e^{-2q} = (e^q + e^{-q})(e^q - e^{-q})$ thus:

$$\sinh(q)\cosh(q) = \cosh(q)\sinh(q) = \frac{\sinh(2q)}{2}$$

□

We are now ready to give a definition of a quaternionic hyperbolic tangent and a quaternionic hyperbolic cotangent.

Definition 5.14 (Hyperbolic Tangent)**.** *Let $q \in \mathbb{H}$ be a quaternion; we call hyperbolic tangent the function $\tanh(q)$, defined as:*

$$\tanh(q) := \sinh(q)\operatorname{sech}(q) = \operatorname{sech}(q)\sinh(q) = \frac{\sinh(q)}{\cosh(q)} \tag{5.34}$$

Definition 5.15 (Hyperbolic Cotangent)**.** *Let $q \in \mathbb{H}$ be a quaternion; we call hyperbolic cotangent the function $\coth(q)$, defined as:*

$$\coth(q) := \cosh(q)\operatorname{csch}(q) = \operatorname{csch}(q)\sinh(q) = \frac{\cosh(q)}{\sinh(q)} \tag{5.35}$$

Exercise 5.10. *Calculate the hyperbolic tangent and cotangent functions for the following quaternionic values:*

- $q_1 = i \quad ; \quad q_2 = j \quad ; \quad q_3 = k \quad ; \quad q_4 = -\ln\ 2 + \sqrt{3}i - k$

Exercise 5.11. *Prove the following hyperbolic identities:*

- $\tanh^2(q) = 1 - \operatorname{sech}^2(q)$
- $\coth^2(q) = 1 + \operatorname{csch}^2(q)$
- $\tanh(-q) = -\tanh(q)$
- $\coth(-q) = -\coth(q)$
- $\operatorname{sech}(-q) = \operatorname{sech}(q)$
- $\operatorname{csch}(-q) = -\operatorname{csch}(q)$

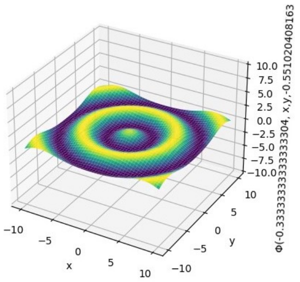
5.7 Visualizing Quaternionic Functions

A significant challenge in quaternionic analysis is the following: the four-dimensionality of quaternions makes it very difficult to visualize functions of a quaternionic variable $f : \mathbb{H} \to \mathbb{H}$. We are dealing with functions that map quadruples of real numbers to quadruples of real numbers, i.e., functions that would ideally require eight dimensions for easy visualization. However, we should not give up: it is possible to ”circumvent” this problem in various ways, and in this section, we will propose one. As we have already seen in the previous parts of the monograph, we can write a function of a quaternionic variable $f : \mathbb{H} \to \mathbb{H}$ by explicating its components as follows:

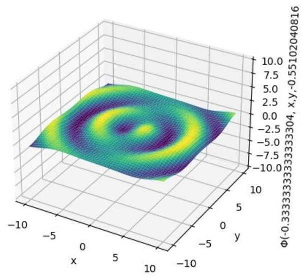
$$f(q) = f_1(q) + f_2(q)i + f_3(q)j + f_4(q)k$$

where f_i is a function from quaternions to reals, $f_i : \mathbb{H} \to \mathbb{R}\ i = 1, 2, 3, 4$.
Deciding to visualize the four component functions of a quaternionic function instead of the function itself already simplifies the problem considerably: we are now looking at visualizing functions that map quadruples of real numbers to real numbers, and thus only require five dimensions instead of eight. However, we still face the inability to directly visualize five dimensions, given our innate three-dimensionality.
The problem, however, is only apparent and can be circumvented; to visualize the remaining two dimensions we will use time, i.e., we will fix two parameters of the function $f_i(t, x, y, z)$ and let them evolve over time, thus visualizing our function through three spatial coordinates in the strict sense and two ”spatialized temporal” coordinates.
The process is easily achievable with Python’s ”matplotlib” and ”numpy” libraries, using the PillowWriter class from matplotlib. Below are some

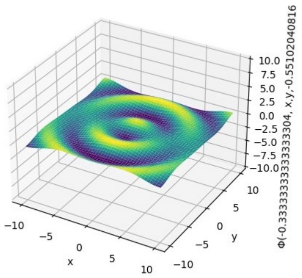
”temporal snapshots” of the evolutions of the graphs of some quaternionic functions. To view the entire evolution of the graphs, I will leave links in the notes section to videos I created with matplotlib [Note 5.1]. Let’s start by visualizing the exponential function $\exp(q) = f_1(q)+f_2(q)i+f_3(q)j+f_4(q)k$; we will report the graphs of its four component functions visualized in ”snapshots”.



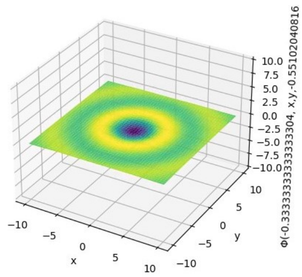
(a) Real part.



(b) i part.



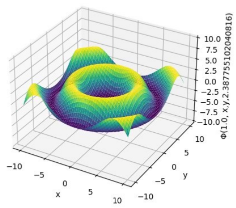
(c) j part.



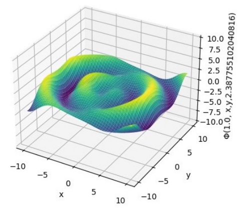
(d) k part.

Figure 5.1: Quaternionic exponential function $\exp(q)$ visualized at the time instant $(-0.33, x, y, -0.551)$. On the x axis, the x is reported, on the y axis, the y is reported, and on the z axis, the value of $f_i(t, x, y, z)$ is reported.

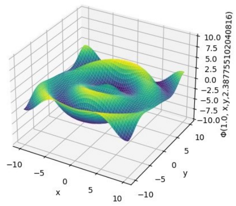
Here by ”real part” we mean $f_1(q)$, by ”i part” $f_2(q)$, by ”j part” $f_3(q)$, and by ”k part” $f_4(q)$, i.e., the coefficient functions of the respective imaginary units.
The sub-region of the domain chosen for visualization is the quaternionic hyper-cuboidal region $[-3,3]\times[-10,10]\times[-10,10]\times[-3,3]$.



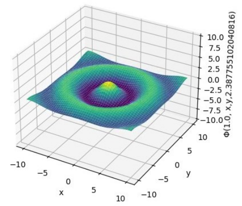
(a) Real part.



(b) i part.

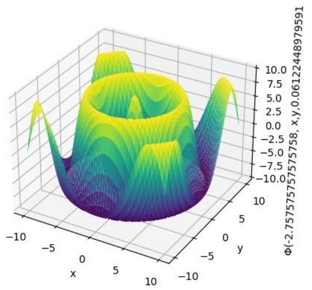


(c) j part.

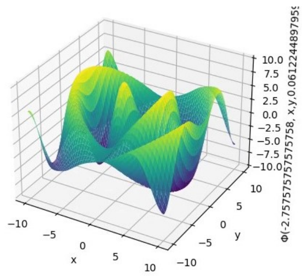


(d) k part.

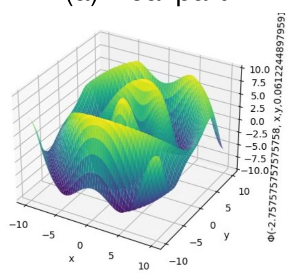
Figure 5.2: Quaternionic exponential function $\exp(q)$ visualized at the time instant $(1, x, y, 2.38)$.



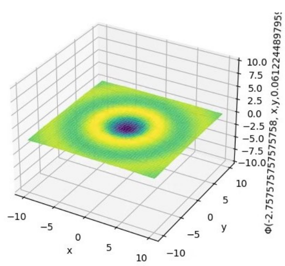
(a) Real part.



(b) i part.

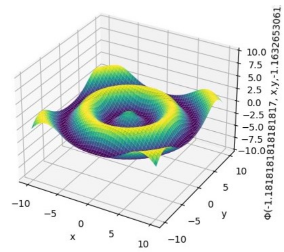


(c) j part.

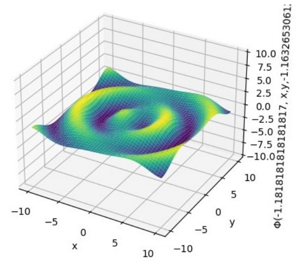


(d) k part.

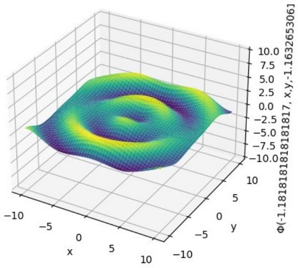
Figure 5.3: Quaternionic hyperbolic cosine function $\cosh(q)$ visualized at the time instant $(-2.7575, x, y, 0.06122)$.



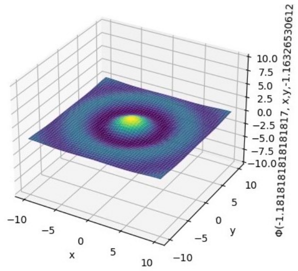
(a) Real part.



(b) i part.

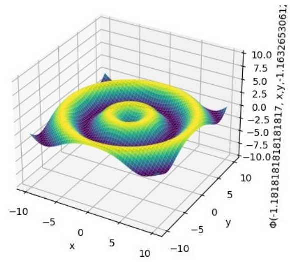


(c) j part.

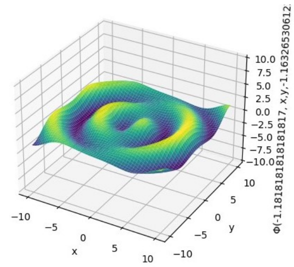


(d) k part.

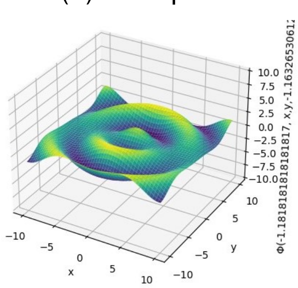
Figure 5.4: Quaternionic hyperbolic cosine function $\cosh(q)$ visualized at the time instant $(-1.181818, x, y, -1.1632653)$.



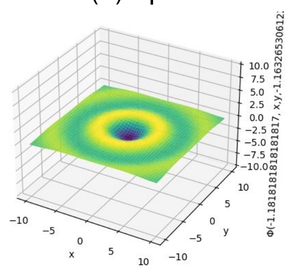
(a) Real part.



(b) i part.

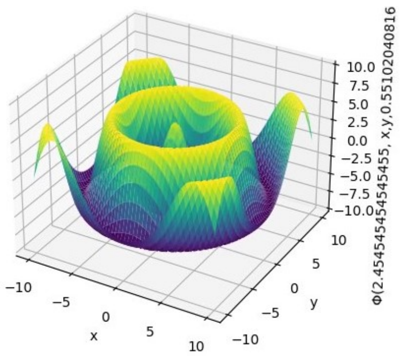


(c) j part.

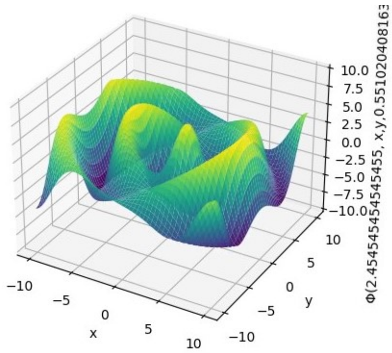


(d) k part.

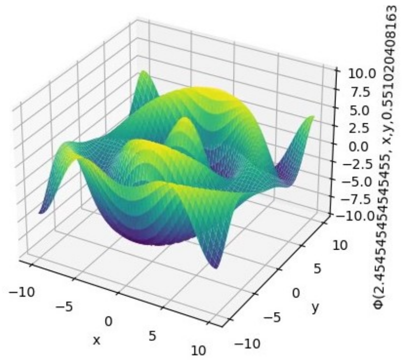
Figure 5.5: Quaternionic hyperbolic sine function visualized at the time instant $(-1.181818, x, y, -1.1632653)$.



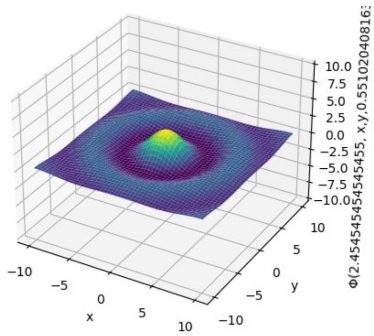
(a) Real part.



(b) i part.



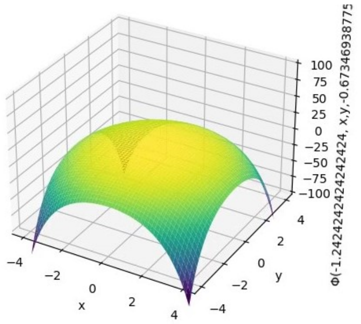
(c) j part.



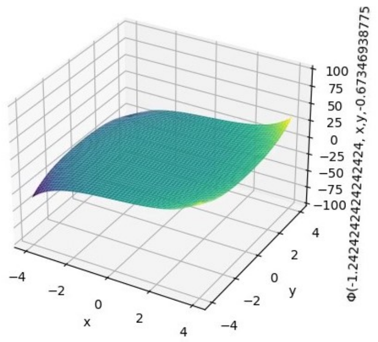
(d) k part.

Figure 5.6: Quaternionic hyperbolic sine function $\sinh(q)$ visualized at time $(2.4545, x, y, 0.5510)$.

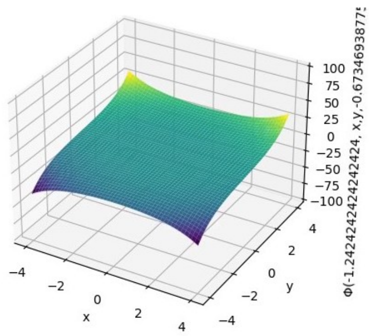
Also for the visualization of the hyperbolic quaternionic functions, we have taken into consideration the hyper-domain $[-3, 3] \times [-10, 10] \times [-10, 10] \times [-3, 3]$.
For the visualization of the sine function $\sin(q)$, we will opt for a change in the hypercuboidal region of visualization to make this graphical representation more suggestive: in particular, we will restrict the possible values of x and y to the set $[-4, 4]$ instead of $[-10, 10]$. In addition, we will graphically depict on the z axis the values of the component functions in a range from -100 to 100. The ranges over which t and z are allowed to vary will remain the same.
As always, for the ”moving” visualization of these graphs (i.e., the complete visualization), I refer the reader to the link in the notes, specifically the link in [Note 5.1].



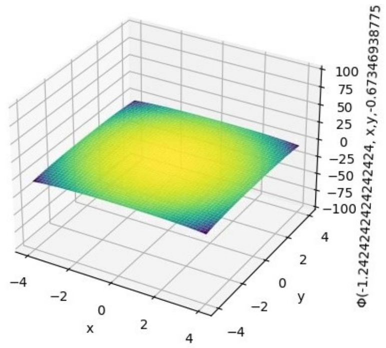
(a) Real part.



(b) i part.



(c) j part.



(d) k part.

Figure 5.7: :
Quaternionic sine function $\sin(q)$ visualized at the instant in time $(-1.2424, x, y, -0.6734)$.

Chapter 6

Differential and Integral Calculus on $\mathbb{H}$

Having proved some fundamental results on the algebraic and topological structure of quaternions, an important issue remains open: how can we develop a theory of regular functions of a quaternionic variable and calculate derivatives of quaternionic functions, similarly to the real and complex cases?

The answer is more difficult than expected; as already mentioned in the introduction of the text, most of the problems are caused by the non-commutativity of the structure, and therefore a ”standard” approach analogous to the real and complex case is not possible. In the first sections, we will see, by trying to develop a notion of differentiability in a manner similar to the real and complex case, what will ”go wrong.”

That said, as anticipated in the introduction, various mathematicians have managed to circumvent this serious problem by providing alternative definitions of regularity, and in this chapter, we will see one of the various approaches that have been used (we will see alternative approaches in the next chapter).

6.1 Some preliminary definitions

Before starting a discussion on the differential calculus for functions of a quaternionic variable, let's make some notes here about conventions that will be in force for the duration of this chapter.

Finding ourselves in a position where we will have to use some tools of differential geometry on smooth manifolds, we first specify with respect to which differentiable structure we are considering the 4-manifold of quater-

nions.
In Chapter 4, we defined the canonical topology $\mathscr{U}$ on $\mathbb{H}$, induced by its metric $\delta(q_1, q_2) = |q_1 - q_2|$, obtaining the topological space $(\mathbb{H}, \mathscr{U})$.
As an immediate corollary of what was said in Chapter 4, we conclude that this space is Hausdorff and respects the second countability axiom.
Also recalling the homeomorphism from theorem 4.2, we note that it will be possible to equip $(\mathbb{H}, \mathscr{U})$ with a smooth atlas composed of a single chart, $\mathscr{A} = \{(\varphi, \mathbb{H})\}$ (where $\varphi : \mathbb{H} \to \mathbb{R}^4$ is the homeomorphism in question) thus obtaining a 4-dimensional real manifold.
We observe that such a differentiable structure provides $\mathbb{H}$ with global coordinates. From now on, unless otherwise specified, we will consider quaternions as the smooth manifold $(\mathbb{H}, \mathscr{U}, \mathscr{A})$ with respect to this differentiable structure.
We now define an $\mathbb{R}$-linear application, $\Gamma : \mathbb{H}^* \to \mathbb{H}$, from $\mathbb{H}^*$, the dual space of quaternions, to $\mathbb{H}$ as:

$$\langle \Gamma(\alpha), q \rangle = \alpha(q) \tag{6.1}$$

with $\alpha \in \mathbb{H}^*$ and $q = t + xi + yj + zk \in \mathbb{H}$, and where $\langle\ ,\ \rangle$ denotes the quaternionic scalar product introduced in Chapter 2. Manipulating expression (6.1) exploiting the linearity of functionals of $\mathbb{H}^*$, we obtain:

$$\Gamma(\alpha)_t t + \Gamma(\alpha)_x x + \Gamma(\alpha)_y y + \Gamma(\alpha)_z z = t\alpha(1) + x\alpha(i) + y\alpha(j) + z\alpha(k) \tag{6.2}$$

from which:

$$\Gamma(\alpha) = \alpha(1) + i\alpha(i) + j\alpha(j) + k\alpha(k)$$

We can extend this application to $\mathbb{R}$-linear applications from $\mathbb{H}$ to $\mathbb{H}$, $\hom_{\mathbb{R}}(\mathbb{H}, \mathbb{H})$, however, being careful about the non-commutativity of quaternions. Such a set of applications will form an $\mathbb{H}$-module, which we will call F, linearly spanned by linear functionals $\lambda \in \mathbb{H}^*$.
We define the right $\mathbb{H}$-linear extension, Γ_r of Γ as:

$$\Gamma_r(\alpha) = \alpha(1) + i\alpha(i) + j\alpha(j) + k\alpha(k) \tag{6.3}$$

and its left $\mathbb{H}$-linear extension Γ_l:

$$\Gamma_l(\alpha) = \alpha(1) + \alpha(i)i + \alpha(j)j + \alpha(k)k \tag{6.4}$$

where here α is an $\mathbb{R}$-linear application from $\mathbb{H}$ to $\mathbb{H}$, $\alpha \in \hom_{\mathbb{R}}(\mathbb{H}, \mathbb{H}) = F$.

6.2 Quaternionic Differential Forms

We now define the differential forms on the canonical manifold structure defined on $\mathbb{H}$, which we will use in our development of differential calculus on $\mathbb{H}$.
Let $p \in \mathbb{H}$, then we know from a well-known result of differential geometry that the tangent space of $\mathbb{H}$ at p, $T_p\mathbb{H}$, is of dimension 4, as is its dual space, the cotangent space $T^*_p\mathbb{H}$. By constructing an isomorphism, we can therefore treat tangent spaces at points of the quaternionic hyperspace as points in $\mathbb{H}$. We classically define the tangent bundle and cotangent bundle of $\mathbb{H}$:

$$T\mathbb{H} = \coprod_{p\in\mathbb{H}} T_p\mathbb{H} \tag{6.5}$$

$$T^*\mathbb{H} = \coprod_{p\in\mathbb{H}} T^*_p\mathbb{H} \tag{6.6}$$

and naturally make them into 8-dimensional real manifolds. Having done this, we can talk about vector fields and n-forms on quaternions, trivially adapting the notions of vector field and differential n-forms of manifolds developed in the final part of the introduction. We also define the operations of exterior differentiation and exterior product in the usual way.
After this brief digression, we are ready to introduce some important objects for our discussion.
The differential of a quaternionic function f will be indicated by the following 1-form:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

or, more compactly, using the Einstein convention:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x_i}dx_i$$

Of particular importance is the differential of the identity function, the 1-form:

$$dq = dt + idx + jdy + kdz = dt + e_i dx_i$$

Also important is the exterior product of dq with itself:

$$dq \wedge dq = 2(idy \wedge dz + jdz \wedge dx + kdx \wedge dy)$$

On quaternions, moreover, being an oriented manifold, we will have a non-null 4-form of volume, which we will call v:

$$v = dt \wedge dx \wedge dy \wedge dz$$

We finally define the 3-form Dq as follows:

$$Dq = dx \wedge dy \wedge dz - idt \wedge dy \wedge dz - jdt \wedge dz \wedge dx - kdt \wedge dx \wedge dy \tag{6.7}$$

We can rewrite the above in a more compact way using the Einstein convention and the Levi-Civita symbol as:

$$Dq = dx \wedge dy \wedge dz - \frac{1}{2}\epsilon_{ijk}e_i dt \wedge dx_j \wedge dx_k$$

We observe that for the definition just given the following equality will be valid $\forall h_1, h_2, h_3, h_4 \in \mathbb{H}$:

$$\langle h_1, Dq(h_2, h_3, h_4)\rangle = v(h_1, h_2, h_3, h_4)$$

where here $\langle\ ,\ \rangle$ indicates the scalar product of quaternions introduced in Chapter 2. We can easily verify this identity noting that, starting from the right side of the equation:

$$v(h_1, h_2, h_3, h_4) = \begin{vmatrix} t_1 & x_1 & y_1 & z_1 \\ t_2 & x_2 & y_2 & z_2 \\ t_3 & x_3 & y_3 & z_3 \\ t_4 & x_4 & y_4 & z_4 \end{vmatrix}$$

By a cofactor expansion (better known as Laplace's development), we obtain:

$$\begin{vmatrix} t_1 & x_1 & y_1 & z_1 \\ t_2 & x_2 & y_2 & z_2 \\ t_3 & x_3 & y_3 & z_3 \\ t_4 & x_4 & y_4 & z_4 \end{vmatrix} = t_1 \begin{vmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_4 & y_4 & z_4 \end{vmatrix} - x_1 \begin{vmatrix} t_2 & y_2 & z_2 \\ t_3 & y_3 & z_3 \\ t_4 & y_4 & z_4 \end{vmatrix} + y_1 \begin{vmatrix} t_2 & x_2 & z_2 \\ t_3 & x_3 & z_3 \\ t_4 & x_4 & z_4 \end{vmatrix} - z_1 \begin{vmatrix} t_2 & x_2 & y_2 \\ t_3 & x_3 & y_3 \\ t_4 & x_4 & y_4 \end{vmatrix}$$

But observing that:

$$\begin{cases} dx \wedge dy \wedge dz = Dq_t(h_2, h_3, h_4) = \begin{vmatrix} x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_4 & y_4 & z_4 \end{vmatrix} \\ -dt \wedge dy \wedge dz = Dq_x(h_2, h_3, h_4) = -\begin{vmatrix} t_2 & y_2 & z_2 \\ t_3 & y_3 & z_3 \\ t_4 & y_4 & z_4 \end{vmatrix} \\ -dt \wedge dz \wedge dx = dt \wedge dx \wedge dz = Dq_y(h_2, h_3, h_4) = \begin{vmatrix} t_2 & x_2 & z_2 \\ t_3 & x_3 & z_3 \\ t_4 & x_4 & z_4 \end{vmatrix} \\ -dt \wedge dx \wedge dy = Dq_z(h_2, h_3, h_4) = -\begin{vmatrix} t_2 & x_2 & y_2 \\ t_3 & x_3 & y_3 \\ t_4 & x_4 & y_4 \end{vmatrix} \end{cases}$$

we will have that the cofactor expansion of the expression for $v(h_1, h_2, h_3, h_4)$ will simply be the scalar product of h_1 with $Dq(h_2, h_3, h_4)$, i.e:

$$\langle h_1, Dq(h_2, h_3, h_4)\rangle = v(h_1, h_2, h_3, h_4)$$

Geometrically, given 3 quaternions q_1, q_2, q_3, the 3-form just introduced calculated in this triplet, $Dq(q_1, q_2, q_3)$ will be a quaternion perpendicular to q_1, q_2 and q_3, with a norm equal to the volume of the parallelepiped defined by the 3 quaternions q_1, q_2 and q_3.

Exercise 6.1. *Determine, through direct calculation, the value of:*

1. $Dq(i, j, k)$
2. $Dq(1, i, j)$
3. $Dq(1, j, k)$
4. $Dq(1, i, k)$
5. $Dq(1 - \sqrt{2}i + 5j - k, i - j + k, \sqrt{5} + \sqrt{3}i - j - k)$

Exercise 6.2. *Prove, motivated by the previous exercise, that:*

$$Dq(1, e_i, e_j) = \frac{1}{2}(e_j e_i - e_i e_j)$$

Lemma 6.1. *Let $q, p \in \mathbb{H}$ be two quaternions, then:*

$$Dq(1, q, p) = \frac{1}{2}(pq - qp)$$

Proof. Write q as $q = t_1 + x_1 i + y_1 j + z_1 k = t_1 + x_{1i}e_i$ and p as $p = t_2 + x_2 i + y_2 j + z_2 k = t_2 + x_{2i}e_i$. Exploiting the $\mathbb{R}$-trilinearity of Dq:

$$Dq(1, q, p) = Dq(1, t_1 + x_{1i}e_i, t_2 + x_{2j}e_j) = Dq(1, t_1, t_2 + x_{2j}e_j) + Dq(1, x_{1i}e_i, t_2 + x_{2j}e_j)$$

$$= Dq(1, t_1, t_2) + Dq(1, t_1, x_{2j}e_j) + Dq(1, x_{1i}e_i, t_2) + Dq(1, x_{1i}e_i, x_{2j}e_j)$$

$$= t_1 t_2 Dq(1, 1, 1) + t_1 x_{2j} Dq(1, 1, e_j) + x_{1i} t_2 Dq(1, e_i, 1) + x_{1i} x_{2j} Dq(1, e_i, e_j)$$

But we know that $Dq(1, e_i, e_j) = \frac{1}{2}(e_j e_i - e_i e_j)$, $Dq(1, 1, 1) = 0$, $Dq(1, 1, e_j) = Dq(1, e_j, 1) = 0$ and so the expression becomes simply:

$$Dq(1, q, p) = \sum_{i=1}^{3}\sum_{j=1}^{3} \frac{x_{1i}x_{2j}}{2}(e_j e_i - e_i e_j) = \sum_{i \neq j \;\; i,j=1,2,3} \frac{x_{1i}x_{2j}}{2}(e_j e_i - e_i e_j)$$

For simplicity of notation let's put $x_{11} = x_1, x_{12} = y_1, x_{13} = z_1, x_{21} = x_2, x_{22} = y_2, x_{23} = z_2$.
Expanding the sum on the right:

$$\frac{x_1y_2}{2}(ji-ij)+\frac{x_1z_2}{2}(ki-ik)+\frac{y_1z_2}{2}(kj-jk)+\frac{y_1x_2}{2}(ij-ji)+\frac{z_1x_2}{2}(ik-ki)+\frac{z_1y_2}{2}(jk-kj)$$

Simplifying, we obtain:

$$(z_1y_2 - y_1z_2)i + (x_1z_2 - z_1x_2)j + (y_1x_2 - x_1y_2)k = Dq(1, p, q)$$

This expression should not be new to the reader; it is nothing but half of the commutator of p and q, $[p, q] = pq - qp = (2z_1y_2 - 2y_1z_2)i + (2x_1z_2 - 2z_1x_2)j + (2y_1x_2 - 2x_1y_2)k$, a formula derived in the exercises of the second chapter. With this observation, we arrive at the desired fact, that is:

$$Dq(1, q, p) = \frac{1}{2}[p, q] = \frac{1}{2}(pq - qp)$$

□

We now prove a proposition that will allow us to calculate the value of our 3-form Dq in an alternative way.

Proposition 6.1. *Let q_1, q_2, q_3 be 3 quaternions, then:*

$$Dq(q_1, q_2, q_3) = \frac{1}{2}(q_3q_1^*q_2 - q_2q_1^*q_3)$$

Proof. Let $u \in S^3 \subset \mathbb{H}$ be a unit quaternion (quaternion with norm 1), then the application $l : q \to uq$ is an orthogonal linear application with determinant equal to 1, and therefore we have:

$$Dq(uq_1, uq_2, uq_3) = uDq(q_1, q_2, q_3)$$

More specifically, we select u to be $u = |q_1|q_1^{-1}$, then:

$$Dq(|q_1|, |q_1|q_1^{-1}q_2, |q_1|q_1^{-1}q_3) = |q_1|q_1^{-1}Dq(q_1, q_2, q_3)$$

From which it follows that, using the $\mathbb{R}$-trilinearity of the form Dq:

$$Dq(q_1, q_2, q_3) = \frac{|q_1|^3}{|q_1|}q_1Dq(1, q_1^{-1}q_2, q_1^{-1}q_3) = |q_1|^2q_1Dq(1, q_1^{-1}q_2, q_1^{-1}q_3)$$

Now, for the previous lemma, we know that $\forall q, p \in \mathbb{H}$, $Dq(1, q, p) = \frac{1}{2}(pq - qp)$. Since $Dq(1, q_1^{-1}q_2, q_1^{-1}q_3)$ is precisely an expression of this form, we rewrite the equation above as:

$$Dq(q_1, q_2, q_3) = \frac{1}{2}|q_1|^2q_1(q_1^{-1}q_3q_1^{-1}q_2 - q_1^{-1}q_2q_1^{-1}q_3) = \frac{1}{2}(q_3q_1^*q_2 - q_2q_1^*q_3)$$

□

Proposition 6.2. *Let $q_1, q_2, q_3, a, b \in \mathbb{H}$ be 5 quaternions, then:*

$$Dq(aq_1b, aq_2b, aq_3b) = |a|^2|b|^2 aDq(q_1, q_2, q_3)b$$

Proof. Let $u, v \in S^3 \subset \mathbb{H}$ be two unit quaternions. We observe that the linear application $q \to uqv$ is a rotation, and therefore an orthogonal application with determinant 1. Therefore:

$$Dq(uq_1v, uq_2v, uq_3v) = uDq(q_1, q_2, q_3)v$$

Choosing $u = a|a|^{-1}$ and $v = b|b|^{-1}$ we obtain:

$$Dq(a|a|^{-1}q_1b|b|^{-1}, a|a|^{-1}q_2b|b|^{-1}, a|a|^{-1}q_3b|b|^{-1}) = |a|^{-1}|b|^{-1}aDq(q_1, q_2, q_3)b$$

Using the $\mathbb{R}$-trilinearity of Dq on the left side of the equation, we get:

$$|a|^{-3}|b|^{-3}Dq(aq_1b, aq_2b, aq_3b) = |a|^{-1}|b|^{-1}aDq(q_1, q_2, q_3)b$$

From which follows the assertion:

$$Dq(aq_1b, aq_2b, aq_3b) = |a|^2|b|^2 aDq(q_1, q_2, q_3)b$$

□

6.3 Quaternionic Differential Operators

We now introduce some important quaternionic operators that we will use later in our discussion. First, we note that the differential of a quaternionic function $f : \mathbb{H} \to \mathbb{H}$ is an $\mathbb{R}$-linear application from $\mathbb{H}$ to $\mathbb{H}$, i.e $df \in \hom_{\mathbb{R}}(\mathbb{H}, \mathbb{H}) = F$. Recall the Γ_r and Γ_l applications, respectively right and left linear from F to $\mathbb{H}$:

$$\Gamma_r(\alpha) = \alpha(1) + i\alpha(i) + j\alpha(j) + k\alpha(k)$$

$$\Gamma_l(\alpha) = \alpha(1) + \alpha(i)i + \alpha(j)j + \alpha(k)k$$

$$\Gamma_r : F = \hom_{\mathbb{R}}(\mathbb{H}, \mathbb{H}) \to \mathbb{H} \;\; ; \;\; \Gamma_l : F = \hom_{\mathbb{R}}(\mathbb{H}, \mathbb{H}) \to \mathbb{H}$$

Now we calculate the image under Γ_r and Γ_l of the differential of a quaternionic function:

$$\Gamma_r(df) = \frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}$$

$$\Gamma_l(df) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k$$

In light of these developments, we now define the following new entities:

Definition 6.1. *Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternionic function (differentiable in the classical sense of differential geometry), then we define the following operators acting on it with the following symbols:*

$$\begin{cases} \bar{\partial}_l f = \frac{1}{2}\Gamma_r(df) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}\Big) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} + e_i\frac{\partial f}{\partial x_i}\Big) \\ \partial_l f = \frac{1}{2}\Gamma_r^*(df) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} - i\frac{\partial f}{\partial x} - j\frac{\partial f}{\partial y} - k\frac{\partial f}{\partial z}\Big) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} - e_i\frac{\partial f}{\partial x_i}\Big) \\ \bar{\partial}_r f = \frac{1}{2}\Gamma_l(df) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k\Big) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x_i}e_i\Big) \\ \partial_r f = \frac{1}{2}\Gamma_l^*(df) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} - \frac{\partial f}{\partial x}i - \frac{\partial f}{\partial y}j - \frac{\partial f}{\partial z}k\Big) = \frac{1}{2}\Big(\frac{\partial f}{\partial t} - \frac{\partial f}{\partial x_i}e_i\Big) \end{cases} \tag{6.8}$$

Definition 6.2. *We define the Laplacian of a quaternionic function (differentiable in the classical sense of differential geometry) as:*

$$\triangle f = \nabla^2 f = \frac{\partial^2 f}{\partial t^2} + \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

We observe that the Laplacian operator can be written as:

$$\triangle = 4\partial_l\bar{\partial}_l = 4\bar{\partial}_l\partial_l = 4\partial_r\bar{\partial}_r = 4\bar{\partial}_r\partial_r \tag{6.9}$$

Once the notational conventions that we will use in the course of the chapter are established, we now go on to operationally define a concept of derivative of a quaternionic function.

6.4 The Classic Approach

We have now reached the point of defining what it means for a function of a quaternionic variable to be differentiable, and to give meaning to the term "derivative of f". The first approach we will use is the "classic" one (in the sense that we will define it in a way entirely analogous to the complex and real case). We therefore want to define the derivative of a quaternionic function as the limit as a quaternion $h \to 0$ of $f(x + h) - f(x)$ divided by h. The only precaution we must take is the following: since quaternions are not a field, we will necessarily have to define a right and a left derivative, since we can divide by h on the left and on the right.

Definition 6.3 (Left Derivative/Left Differentiability)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternionic function. We say that f is left-differentiable or differentiable on the left at the point $q_0 \in \mathbb{H}$, if the following limit exists:*

$$\frac{\overleftarrow{d} f}{\overleftarrow{d} q} = \lim_{h \to 0} h^{-1}(f(q_0 + h) - f(q_0)) \tag{6.10}$$

if it exists, we will call the latter "left derivative of f calculated at the point q_0".

Definition 6.4 (Right Derivative/Right Differentiability)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternionic function. We say that f is right-differentiable or differentiable on the right at the point $q_0 \in \mathbb{H}$, if the following limit exists:*

$$\frac{\overrightarrow{d} f}{\overrightarrow{d} q} = \lim_{h \to 0} (f(q_0 + h) - f(q_0))h^{-1} \tag{6.11}$$

if it exists, we will call the latter "right derivative of f calculated at the point q_0".

We now prove some elementary properties of these newly introduced derivatives:

Proposition 6.3 (Properties of the Left Derivative)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{H}$ be functions left-differentiable at $q_0 \in \mathbb{H}$, and let $\lambda \in \mathbb{H}$ be a constant, then:*

- *$f + g$ is left-differentiable at q_0, and its left derivative is:*

$$\frac{\overleftarrow{d}(f + g)}{\overleftarrow{d} q} = \frac{\overleftarrow{d} f}{\overleftarrow{d} q} + \frac{\overleftarrow{d} g}{\overleftarrow{d} q}$$

- *$f\lambda$ is left-differentiable at q_0 and its left derivative is:*

$$\frac{\overleftarrow{d}(f\lambda)}{\overleftarrow{d} q} = \frac{\overleftarrow{d} f}{\overleftarrow{d} q}\lambda$$

Proof. By hypothesis, the limits $\frac{\overleftarrow{d} f}{\overleftarrow{d} q} = \lim_{h \to 0} h^{-1}(f(q_0 + h) - f(q_0))$ and $\frac{\overleftarrow{d} g}{\overleftarrow{d} q} = \lim_{h \to 0} h^{-1}(g(q_0 + h) - g(q_0))$ exist. We write the left derivative of $f + g$:

$$\frac{\overleftarrow{d}(f + g)}{\overleftarrow{d} q} = \lim_{h \to 0} h^{-1}((f(q_0 + h) + g(q_0 + h)) - (f(q_0) + g(q_0))$$

We observe that, due to the algebraic properties and properties of quaternionic limits demonstrated in Chapter 4:

$$\frac{\overleftarrow{d}(f+g)}{\overleftarrow{d}q} = \lim_{h\to 0} h^{-1}(f(q_0+h) - f(q_0)) + \lim_{h\to 0} h^{-1}(g(q_0+h) - g(q_0))$$

and thus:

$$\frac{\overleftarrow{d}(f+g)}{\overleftarrow{d}q} = \frac{\overleftarrow{d}f}{\overleftarrow{d}q} + \frac{\overleftarrow{d}g}{\overleftarrow{d}q}$$

For the second point, remembering the properties of quaternionic limits:

$$\frac{\overleftarrow{d}(f\lambda)}{\overleftarrow{d}q} = \lim_{h\to 0} h^{-1}(f(q_0+h)\lambda - f(q_0)\lambda) = \lim_{h\to 0} h^{-1}(f(q_0+h) - f(q_0))\lambda = \frac{\overleftarrow{d}f}{\overleftarrow{d}q}\lambda$$

□

Right derivatives of quaternionic functions also have the same properties:

Proposition 6.4 (Linearity of the Right Derivative)**.** *Let* $f : \mathbb{H} \to \mathbb{H}$ *and* $g : \mathbb{H} \to \mathbb{H}$ *be functions right-differentiable at* $q_0 \in \mathbb{H}$*, and let* $\lambda \in \mathbb{H}$ *be a constant, then:*

- $f + g$ *is right-differentiable at* q_0*, and its right derivative is:*

$$\frac{\overrightarrow{d}(f+g)}{\overrightarrow{d}q} = \frac{\overrightarrow{d}f}{\overrightarrow{d}q} + \frac{\overrightarrow{d}g}{\overrightarrow{d}q}$$

- λf *is right-differentiable at* q_0 *and its right derivative is:*

$$\frac{\overrightarrow{d}(\lambda f)}{\overrightarrow{d}q} = \lambda\frac{\overrightarrow{d}f}{\overrightarrow{d}q}$$

We leave the easy task of proving this proposition to the reader, in light of the fact that it is an almost identical proof to the one presented earlier.

Exercise 6.3. *Let* $f(q) = \gamma$*, with* $\gamma \in \mathbb{H}$ *a constant function. Prove that the function is both left and right differentiable, and that both derivatives are equal to 0, i.e:*

$$\frac{\overleftarrow{d}f}{\overleftarrow{d}q} = \frac{\overrightarrow{d}f}{\overrightarrow{d}q} = 0$$

Exercise 6.4. *Let $f(q) = q$ be the identity function. Prove that the function is both left and right differentiable, and prove that both derivatives are equal to 1:*

$$\frac{\overleftarrow{d} q}{\overleftarrow{d} q} = \frac{\overrightarrow{d} q}{\overrightarrow{d} q} = 1$$

As in the real and complex case, we will instead say that a function is differentiable (right/left) in a domain (which is customary to define as an open connected set of $\mathbb{H}$) if it is differentiable (right/left) at every point of the latter.
Despite this definition of the derivative having some desirable properties (such as additive linearity, and a multiplicative linearity for "one side only"), there are just as many rather unpleasant ones.
Firstly, we note that the classic product rule of real and complex analysis is not generally valid:

$$(fg)' = f'g + fg'$$

Thus given two right/left differentiable quaternionic functions, we cannot conclude anything about the differentiability of their product, and we do not have a quick formula to compute it.
But if that were not enough, this definition of the derivative has an even worse property, already mentioned earlier in the introduction of the text: being a left/right differentiable function implies being a left/right linear function. This result was already proved in the late 1930s by V.C.A Ferraro (see [12]), however, we will provide here the proof given by A.Sudbery [43].

Theorem 6.1 (Sudbery, 1979)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternionic function left-differentiable in an open and connected domain $U \subset \mathbb{H}$. Then, in U, it will have the following form:*

$$f(q) = a + qb$$

with $a, b \in \mathbb{H}$.

Proof. First, we observe that left quaternionic differentiability at a point q is a condition stronger than the canonical differentiability of f, intended in the classical sense (reduced to the notion of differentiability of functions from $\mathbb{R}^4$ to $\mathbb{R}^4$). Moreover, the differential of f at q is given by:

$$df_q = dq\frac{\overleftarrow{d} f}{\overleftarrow{d} q}$$

Expanding the left and right sides of the equivalence, we obtain:

$$\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = \frac{\overleftarrow{d} f}{\overleftarrow{d} q}dt + idx\frac{\overleftarrow{d} f}{\overleftarrow{d} q} + jdy\frac{\overleftarrow{d} f}{\overleftarrow{d} q} + kdz\frac{\overleftarrow{d} f}{\overleftarrow{d} q}$$

Equating the coefficients of dt, dx, dy, and dz we obtain the following system:

$$\begin{cases} \frac{\not{d} f}{\not{d} q} = \frac{\partial f}{\partial t} \\ \frac{\not{d} f}{\not{d} q} = -i\frac{\partial f}{\partial x} \\ \frac{\not{d} f}{\not{d} q} = -j\frac{\partial f}{\partial y} \\ \frac{\not{d} f}{\not{d} q} = -k\frac{\partial f}{\partial z} \end{cases} \tag{6.12}$$

However, we recall that we can generally write a quaternion $q = t + xi + yj + zk$ as $q = v + jw$, where $v, w \in \mathbb{C}$ are complex numbers, $v = t + ix, w = y - iz$. We also write our function $f(q)$ as $f(q) = g(v, w) + jh(v, w)$, where g and h are functions in the two complex variables v and w. We can then convert the equations of (6.12) into two systems of complex equations:

$$\frac{\partial g}{\partial t} = -i\frac{\partial g}{\partial x} = \frac{\partial h}{\partial y} = i\frac{\partial h}{\partial z}$$

$$\frac{\partial h}{\partial t} = i\frac{\partial h}{\partial x} = -\frac{\partial g}{\partial y} = i\frac{\partial g}{\partial z}$$

We can rewrite the above in terms of complex partial differential equations as:

$$\frac{\partial g}{\partial v} = \frac{\partial h}{\partial w} \tag{6.13}$$

$$\frac{\partial h}{\partial \overline{v}} = -\frac{\partial g}{\partial \overline{w}} \tag{6.14}$$

$$\frac{\partial g}{\partial \bar{v}} = \frac{\partial h}{\partial \bar{w}} = \frac{\partial h}{\partial v} = \frac{\partial g}{\partial w} = 0$$

From this last equation, in particular, we deduce that g is an analytic function of v and $\overline{w}$, while h is an analytic function of $\overline{v}$ and w. According to Hartogs' theorem [Note 6.1], g and h are $\mathbb{C}$-continuous functions and have continuous partial derivatives of every order, and thus:

$$\frac{\partial^2 g}{\partial v^2} = \frac{\partial}{\partial v}\left(\frac{\partial h}{\partial w}\right) = \frac{\partial}{\partial w}\left(\frac{\partial h}{\partial v}\right) = 0$$

Let us assume for now that U is convex (we will see later how to apply to open connected sets the arguments that we will present below); then g is linear in $\bar{w}$ and v, while h is linear in w and $\bar{v}$:

$$g(v, w) = \alpha + \beta v + \gamma \bar{w} + \delta v \bar{w}$$

$$h(v,w) = \epsilon + \zeta\bar{v} + \eta w + \theta\bar{v}w$$

with $\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta \in \mathbb{C}$. The equations (6.13) and (6.14), calculated explicitly, become:

$$\frac{\partial g}{\partial v} = \beta + \delta\bar{w} = \frac{\partial h}{\partial w} = \eta + \theta\bar{v}$$
$$\frac{\partial h}{\partial \bar{v}} = \zeta + \theta w = -\frac{\partial g}{\partial \bar{w}} = -\gamma - \delta v$$

Substituting v and w with $v = t + ix$ and $w = y - iz$:

$$\beta + \delta y + \delta zi = \eta + \theta t - \theta xi$$

$$\zeta + \theta y - \theta zi = -\gamma - \delta t - \delta xi$$

From which, by the definition of equality of complex numbers, we obtain the following system of equalities:

$$\begin{cases} \beta + \delta y = \eta + \theta t \\ \delta z = -\theta x \\ \zeta + \theta y = -\gamma - \delta t \\ \theta z = \delta x \end{cases}$$

$\forall x, y, z, t \in \mathbb{R}$. Being the value of these 4 real constants arbitrary, we can assume that they are all equal to 1. By doing so, we obtain:

$$\begin{cases} \beta + \delta = \eta + \theta \\ \delta = -\theta \\ \zeta + \theta = -\gamma - \delta \\ \theta = \delta \end{cases}$$

Manipulating the second and fourth equations of the system, we obtain that:

$$\delta = -\theta = -\delta \Longrightarrow 2\delta = 0 \Longrightarrow \theta = \delta = 0$$

Substituting the values of θ and δ just obtained into the first system, we obtain the other relationships between the constants:

$$\beta = \eta \qquad \zeta = -\gamma$$

Hence:

$$g(v,w) = \alpha + \beta v + \gamma\bar{w}$$
$$h(v,w) = \epsilon - \gamma\bar{v} + \beta w$$

And since $f = g + jh$, we obtain an expression for f on a convex set U:

$$f(q) = \alpha + j\epsilon + (v + jw)(\beta - j\gamma) = a + qb \tag{6.15}$$

where $a = \alpha + j\epsilon$ and $b = \beta - j\gamma$ To extend this argument to open and connected domains, it suffices to note that the aforementioned can be covered by convex sets, any two of which can be connected by a chain of convex sets that overlap in pairs; from this it follows that the form of f on the entire domain U will be $f(q) = a + qb$ □

An entirely analogous theorem is valid for right-differentiable functions.

Theorem 6.2 (Sudbery, 1979)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a right-differentiable quaternionic function in an open and connected domain $U \subset \mathbb{H}$. Then, in U, it will have the following form:*

$$f(q) = a + bq$$

with $a, b \in \mathbb{H}$.

The proof of this fact is entirely analogous to the one provided previously, with some minor adjustments.
The theorems just presented convince us that developing a theory of differential calculus on functions of a quaternionic variable using the definitions of derivatives given earlier cannot be fruitful: the class of differentiable functions is in fact too small to the point of only comprising right and left affine functions. Consequently, we must find an alternative way to define the concept of differentiability and derivative on quaternions.

6.5 Regular functions and the Cauchy-Riemann-Fueter equation

Definition 6.5 (Left regular function of a quaternionic variable)**.** *We say that a function of a quaternionic variable $f : \mathbb{H} \to \mathbb{H}$ is left regular at $q \in \mathbb{H}$ if it is differentiable in the classical sense of differential geometry (which reduces to the notion of differentiability between functions from $\mathbb{R}^4$ to $\mathbb{R}^4$) at q and if there exists a quaternion $f_l'(q)$, which we will call the left derivative of f at q, such that:*

$$d(dq \wedge dqf) = Dqf_l'(q) \tag{6.16}$$

Definition 6.6 (Right regular function of a quaternionic variable)**.** *We say that a function of a quaternionic variable $f : \mathbb{H} \to \mathbb{H}$ is right regular at $q \in \mathbb{H}$ if it is differentiable in the classical sense of differential geometry (which reduces to the notion of differentiability between functions from $\mathbb{R}^4$ to $\mathbb{R}^4$) at q and if there exists a quaternion $f_r'(q)$, which we will call the right derivative of f at q, such that:*

$$d(f dq \wedge dq) = f_r'(q) Dq \tag{6.17}$$

The two theories will be entirely analogous, and hence, as we did before, we will derive results only for left regular functions, making a point later for right ones, if necessary.
In light of what has just been said, to lighten the notation, we will call $f_l'(q)$ simply $f'(q)$.
Let's now prove an important theorem, the Cauchy-Riemann-Fueter equation, a quaternionic equivalent of the Cauchy-Riemann equations in complex analysis.

Theorem 6.3 (Left Cauchy-Riemann-Fueter equation)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a function "differentiable in the classical sense" (i.e., $\mathbb{R}$-differentiable). Then f is left regular at $q \in \mathbb{H}$ if and only if $\bar{\partial}_l f = 0$, that is if:*

$$\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} = 0 \tag{6.18}$$

Proof. First, we prove the direct direction ($\Longrightarrow$). Suppose that f is a left regular function at q, then, by the definition given just before:

$$d(dq \wedge dqf) = Dqf_l'(q)$$

Observing that $d(dq \wedge dqf) = dq \wedge dq \wedge df$ we can rewrite the above equation as:

$$dq \wedge dq \wedge df = Dqf_l'(q)$$

This equation is valid for every triplet of quaternions in which we calculate these 3-forms. From the arbitrariness of the latter, it follows that we can choose to calculate both sides in 2 triplets: $(i, j, k) \in \mathbb{H}^3$ and $(1, i, j) \in \mathbb{H}^3$. Starting with the first: to calculate the left side, we apply the following formula to calculate the value of the $r+s$-form given by the exterior product of an r-form and an s-form in an $r+s$-tuple:

$$\omega \wedge \alpha(h_1, h_2, \dots, h_{r+s}) = \frac{1}{r!s!} \sum_{\sigma \in S_{r+s}} sgn(\sigma)\omega(h_{\sigma(1)}, h_{\sigma(2)}, \dots, h_{\sigma(r)})\alpha(h_{\sigma(r+1)}, \dots, h_{\sigma(r+s)})$$

Applying the above formula, $dq \wedge dq \wedge df$ becomes:

$$dq \wedge dq \wedge df = \frac{1}{2}(dq \wedge dq(h_1, h_2)df(h_3) - dq \wedge dq(h_1, h_3)df(h_2) + dq \wedge dq(h_3, h_1)df(h_2) -$$

$$dq \wedge dq(h_3, h_2)df(h_1) + dq \wedge dq(h_2, h_3)df(h_1) - dq \wedge dq(h_2, h_1)df(h_3))$$

Putting $h_1 = i, h_2 = j, h_3 = k$, we get: $dq \wedge dq \wedge df(i, j, k) = \frac{1}{2}\Big(2k\frac{\partial f}{\partial z} + 2j\frac{\partial f}{\partial y} + 2j\frac{\partial f}{\partial y} + 2i\frac{\partial f}{\partial x} + 2i\frac{\partial f}{\partial x} + 2k\frac{\partial f}{\partial z}\Big) = 2i\frac{\partial f}{\partial x} + 2j\frac{\partial f}{\partial y} + 2k\frac{\partial f}{\partial z}$
For the right side, instead, knowing that $Dq(i, j, k) = 1$, we simply have f_l'. Thus, we arrive at the following equality:

$$f_l'(q) = 2i\frac{\partial f}{\partial x} + 2j\frac{\partial f}{\partial y} + 2k\frac{\partial f}{\partial z} \tag{6.19}$$

Now we calculate the 3-forms on both sides in the triplet $(1, i, j)$; on the left, using the same formula, we get:

$$dq \wedge dq \wedge df(1, i, j) = \frac{1}{2}(2k\frac{\partial f}{\partial t} + 2k\frac{\partial f}{\partial t}) = 2k\frac{\partial f}{\partial t}$$

On the right, however, remembering the identity $Dq(1, e_i, e_j) = \frac{1}{2}(e_je_i - e_ie_j)$ demonstrated just now, we get $Dq(1, i, j) = \frac{1}{2}(ji - ij) = -k$ and thus we arrive at the equality:

$$-kf_l'(q) = 2k\frac{\partial f}{\partial t}$$

From which:

$$f_l'(q) = -2\frac{\partial f}{\partial t} \tag{6.20}$$

Now, recalling the expression (6.19) just derived for the same quantity (i.e. f_l') and subtracting f_l' from itself, we get:

$$f_l'(q) - f_l'(q) = 0 = 2\frac{\partial f}{\partial t} + 2i\frac{\partial f}{\partial x} + 2j\frac{\partial f}{\partial y} + 2k\frac{\partial f}{\partial z}$$

from which the left Cauchy-Riemann-Fueter equation:

$$\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} = 0 \tag{6.21}$$

For the reverse direction, however, we will have a constructive proof that will also allow us to determine an operational expression for the left quaternionic derivative $f_l'(q)$.
Suppose that the Cauchy-Riemann-Fueter equations are true at q, that is:

$$\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} = 0$$

We want to prove the existence of a left quaternionic derivative, $f_l'(q)$ such that:

$$dq \wedge dq \wedge df = Dqf_l'(q)$$

We expand the left side of the equation, calculating in the usual way the exterior product:

$$dq \wedge dq \wedge df = 2(idy \wedge dz + jdz \wedge dx + kdx \wedge dy) \wedge \Big(\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\Big)$$

$$= 2\Big(idy \wedge dz \wedge dt\frac{\partial f}{\partial t} + idy \wedge dz \wedge dx\frac{\partial f}{\partial x} + jdz \wedge dx \wedge dt\frac{\partial f}{\partial t} + jdz \wedge dx \wedge dy\frac{\partial f}{\partial y} +$$

$$kdx \wedge dy \wedge dt\frac{\partial f}{\partial t} + kdx \wedge dy \wedge dz\frac{\partial f}{\partial z}\Big) = 2\Big(\Big[i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}\Big]dx \wedge dy \wedge dz + i\frac{\partial f}{\partial t}dt \wedge dy \wedge dz$$

$$+ j\frac{\partial f}{\partial t}dt \wedge dz \wedge dx + k\frac{\partial f}{\partial t}dt \wedge dx \wedge dy)$$

But, by hypothesis $\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} = 0$ and thus:

$$-\frac{\partial f}{\partial t} = i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}$$

and thus $dq \wedge dq \wedge df$ becomes:

$$2\Big(-\frac{\partial f}{\partial t}dx \wedge dy \wedge dz + i\frac{\partial f}{\partial t}dt \wedge dy \wedge dz + j\frac{\partial f}{\partial t}dt \wedge dz \wedge dx + k\frac{\partial f}{\partial t}dt \wedge dx \wedge dy\Big)$$

But remembering that $Dq = dx \wedge dy \wedge dz - idt \wedge dy \wedge dz - jdt \wedge dz \wedge dx - kdt \wedge dx \wedge dy$, we notice that the equation just obtained is nothing but Dq multiplied on the right by $-2\frac{\partial f}{\partial t} = f_l'(q)$. So, we have found a quaternion for which, $\forall(h_1, h_2, h_3) \in \mathbb{H}^3$:

$$dq \wedge dq \wedge df = Dqf_l'(q) \tag{6.22}$$

Therefore, f is regular at q, as we wanted to prove. □

The Cauchy-Riemann-Fueter equations are an important result as they allow us to verify much more quickly whether a function of a quaternionic variable is regular or not. Furthermore, they establish a desirable property of this newly defined class of functions, as they are a direct generalization of the Cauchy-Riemann equations for quaternionic functions. To convince ourselves of this fact, let's rewrite the equations as a system of two complex partial differential equations; we write the argument quaternion of the function as $q = u + vj$ and the function as $f(q) = g(u,v) + jh(u,v)$ where $g : \mathbb{C}^2 \to \mathbb{C}^2$ and $h : \mathbb{C}^2 \to \mathbb{C}^2$, then the CRF (Cauchy-Riemann Fueter) equations become:

$$\frac{\partial g}{\partial \bar{u}} = \frac{\partial h}{\partial \bar{v}}$$

and

$$\frac{\partial g}{\partial v} = -\frac{\partial h}{\partial u}$$

Let's now prove a theorem that will allow us to actually calculate the derivatives of regular functions at a point $q_0 \in \mathbb{H}$.

Corollary 6.1. *If $f : \mathbb{H} \to \mathbb{H}$ is a regular function at q, then its left quaternionic derivative, $f_l'(q)$, is given by:*

$$f_l'(q) = -2\partial_l f = -\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} \tag{6.23}$$

Proof. In the previous theorem, we found the identities:

$$f_l'(q) = -2\frac{\partial f}{\partial t}$$

and

$$f_l'(q) = 2i\frac{\partial f}{\partial x} + 2j\frac{\partial f}{\partial y} + 2k\frac{\partial f}{\partial z}$$

from which:

$$2f_l'(q) = -2\frac{\partial f}{\partial t}+2i\frac{\partial f}{\partial x}+2j\frac{\partial f}{\partial y}+2k\frac{\partial f}{\partial z} \implies f_l'(q) = -2\partial_l f = -\frac{\partial f}{\partial t}+i\frac{\partial f}{\partial x}+j\frac{\partial f}{\partial y}+k\frac{\partial f}{\partial z}$$

□

As already mentioned above, a very similar result applies to right-regular functions, which we will call the right Cauchy-Riemann-Fueter equations (CRF-dx for short). Let's state it formally:

Theorem 6.4 (Right Cauchy-Riemann-Fueter Equation)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a function "differentiable in the classical sense" (i.e., $\mathbb{R}$-differentiable). Then f is right-regular at $q \in \mathbb{H}$ if and only if $\overline{\partial}_r f = 0$, that is, if:*

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k = 0 \tag{6.24}$$

Proof. Let's start by proving the direct implication ($\Longrightarrow$). By hypothesis, being f right-regular, we have that $d(fdq \wedge dq) = f_r'(q)Dq$ for a certain quaternion $f_r'(q)$. Due to the properties of the exterior differential, we know that $d(fdq \wedge dq) = df \wedge dq \wedge dq$. We now have the following equality of 3-forms:

$$df \wedge dq \wedge dq(h_1, h_2, h_3) = f_r'(q)Dq(h_1, h_2, h_3)$$

valid $\forall h_1, h_2, h_3 \in \mathbb{H}$. Being valid for every triplet of quaternions, it will also be valid for the triplets (i, j, k) and $(1, i, j)$, in symbols:

$$df \wedge dq \wedge dq(i, j, k) = f_r'(q)Dq(i, j, k) \quad ; \quad df \wedge dq \wedge dq(1, i, j) = f_r'(q)Dq(1, i, j)$$

Starting from the first: remembering that $df \wedge dq \wedge dq(h_1, h_2, h_3) = \frac{1}{2}[df(h_1)dq \wedge dq(h_2, h_3) - df(h_1)dq \wedge dq(h_3, h_2) - df(h_2)dq \wedge dq(h_1, h_3) + df(h_2)dq \wedge dq(h_3, h_1) + df(h_3)dq \wedge dq(h_1, h_2) - df(h_3)dq \wedge dq(h_2, h_1)]$, and substituting $h_1 = i, h_2 = j, h_3 = k$, we obtain the following expression for $df \wedge dq \wedge dq(i, j, k)$:
$df \wedge dq \wedge dq(i, j, k) = \frac{1}{2}(2\frac{\partial f}{\partial x}i + 2\frac{\partial f}{\partial x}i + 2\frac{\partial f}{\partial y}j + 2\frac{\partial f}{\partial y}j + 2\frac{\partial f}{\partial z}k + 2\frac{\partial f}{\partial z}k) = 2\frac{\partial f}{\partial x}i + 2\frac{\partial f}{\partial y}j + 2\frac{\partial f}{\partial z}k$. The right side of the equality, instead, will simply be equal to $f_r'(q)$ as $Dq(i, j, k) = 1$ (as demonstrated in the second section of this chapter). From this, we obtain the following equality:

$$f_r'(q) = 2\frac{\partial f}{\partial x}i + 2\frac{\partial f}{\partial y}j + 2\frac{\partial f}{\partial z}k$$

Let's now compute the two 3-forms in the triplet $(1, i, j)$. For the left side, we will have, substituting $h_1 = 1, h_2 = i, h_3 = j$ in the same expression as before for $df \wedge dq \wedge dq(h_1, h_2, h_3)$:

$$df \wedge dq \wedge dq(1, i, j) = 2\frac{\partial f}{\partial t}k$$

Moreover, remembering that $Dq(1, i, j) = -k$, we obtain the following equality:

$$2\frac{\partial f}{\partial t}k = -f_r'(q)k \Longrightarrow f_r'(q) = -2\frac{\partial f}{\partial t}$$

Finally, we observe that with the two identities just derived we can rewrite $f_r'(q) - f_r'(q) = 0$ as:

$$2\frac{\partial f}{\partial t} + 2\frac{\partial f}{\partial x}i + 2\frac{\partial f}{\partial y}j + 2\frac{\partial f}{\partial z}k = 0 \Longrightarrow \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k = 0$$

Let's now prove the reverse implication ($\Longleftarrow$). Suppose that the right Cauchy-Riemann-Fueter equation is satisfied at q, i.e:

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k = 0 \tag{6.25}$$

We want to prove the existence of a right quaternionic derivative $f_r'(q)$ for which the following identity will hold:

$$df \wedge dq \wedge dq = f_r'(q)Dq$$

The right side of the equality is equal to $f_r'(q)dx \wedge dy \wedge dz - f_r'(q)idt \wedge dy \wedge dz - f_r'(q)jdt \wedge dz \wedge dx - f_r'(q)kdt \wedge dx \wedge dy$.
The left side of the equality, $df \wedge dq \wedge dq$, is equal to:
$df \wedge dq \wedge dq = (\frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz) \wedge (2idy \wedge dz + 2jdz \wedge dx + 2kdx \wedge dy) = 2[(\frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k)dx \wedge dy \wedge dz + \frac{\partial f}{\partial t}idt \wedge dy \wedge dz + \frac{\partial f}{\partial t}jdt \wedge dz \wedge dx + \frac{\partial f}{\partial t}kdt \wedge dx \wedge dy]$. From the right Cauchy-Riemann-Fueter equation (6.25), however, we deduce that $\frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k = -\frac{\partial f}{\partial t}$ and therefore:

$$df \wedge dq \wedge dq = -2\frac{\partial f}{\partial t}(dx \wedge dy \wedge dz - idt \wedge dy \wedge dz - jdt \wedge dz \wedge dx - kdt \wedge dx \wedge dy) \quad (6.26)$$

But the equation above is nothing but our 3-form Dq multiplied on the right by a quaternion $-2\frac{\partial f}{\partial t}$; we have thus proved the existence of a quaternion $f_r'(q) = -2\frac{\partial f}{\partial t}$ and thus the function f is right regular at q. □

Corollary 6.2. *If a function $f : \mathbb{H} \to \mathbb{H}$ is right-regular at $q \in \mathbb{H}$, then its right derivative is given by:*

$$f_r'(q) = -2\partial_r f = -\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x}i + \frac{\partial f}{\partial y}j + \frac{\partial f}{\partial z}k \quad (6.27)$$

Another property respected by left/right $\mathbb{H}$-regular functions that we would expect is that they are harmonic functions, i.e., solutions of the Laplace equation. We see that this is undoubtedly the case in the following proposition:

Proposition 6.5. *Let $f : \mathbb{H} \to \mathbb{H}$ be a left/right-regular and classically (in the classical sense) differentiable at least twice function. Then f is a harmonic function, i.e. $\triangle f = 0$.*

Proof. Recalling the formula derived in section 3 of this chapter:

$$\triangle f = \partial_l \bar{\partial}_l f = \partial_r \bar{\partial}_r f$$

But we see that being left regular $\iff \bar{\partial}_l f = 0$ and being right regular $\iff \bar{\partial}_r f = 0$, which immediately follows the assertion. □

For the proof of the proposition above, however, we had to postulate that the function was also of class C^∞ in the classical sense. We will show later in this chapter that being a left/right regular quaternionic function implies being a C^∞ class function.
We see that it is possible to obtain new regular functions by summing regular functions or multiplying a regular function by a quaternionic constant $\gamma \in \mathbb{H}$. We group these properties in the following proposition

Proposition 6.6. *Let $\gamma \in \mathbb{H}$ be a quaternionic constant and let $f, g : \mathbb{H} \to \mathbb{H}$ be left-regular quaternionic functions of a variable. Then:*

- *$f + g$ is left regular and its left derivative is given by:*

$$(f+g)'_l(q) = f'_l(q) + g'_l(q)$$

- *$f\gamma$ is left regular and its left derivative is given by:*

$$(f\gamma)'_l(q) = f'_l(q)\gamma$$

Proof. By hypothesis, f and g are left-regular, therefore:

$$\bar{\partial}_l f = 0 \quad and \quad \bar{\partial}_l g = 0$$

Adding them:

$$0 = \bar{\partial}_l f + \bar{\partial}_l g = \frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} + \frac{\partial g}{\partial t} + i\frac{\partial g}{\partial x} + j\frac{\partial g}{\partial y} + k\frac{\partial g}{\partial z}$$

Using the linearity of partial derivatives and the algebraic properties of quaternions, we obtain:

$$\frac{\partial(f+g)}{\partial t} + i\frac{\partial(f+g)}{\partial x} + j\frac{\partial(f+g)}{\partial y} + k\frac{\partial(f+g)}{\partial z} = 0$$

thus $f + g$ is a left-regular function. Now let's calculate its left derivative:

$$(f+g)'_l(q) = -2\partial_l(f+g) = -\frac{\partial(f+g)}{\partial t} + i\frac{\partial(f+g)}{\partial x} + j\frac{\partial(f+g)}{\partial y} + k\frac{\partial(f+g)}{\partial z}$$

$$= -\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z} - \frac{\partial g}{\partial t} + i\frac{\partial g}{\partial x} + j\frac{\partial g}{\partial y} + k\frac{\partial g}{\partial z} = f'_l(q) + g'_l(q)$$

For the second point, instead, it is enough to observe that:

$$\frac{\partial f\gamma}{\partial t} + i\frac{\partial f\gamma}{\partial x} + j\frac{\partial f\gamma}{\partial y} + k\frac{\partial f\gamma}{\partial z} = \bar{\partial}_l f\gamma = 0$$

And similarly, for the calculation of its derivative:

$$(f\gamma)'_l(q) = -2\partial_l(f\gamma) = \Big(-\frac{\partial f}{\partial t} + i\frac{\partial f}{\partial x} + j\frac{\partial f}{\partial y} + k\frac{\partial f}{\partial z}\Big)\gamma = f'_l(q)\gamma$$

□

As the reader might have expected, a similar result applies to right-regular quaternionic functions and their right derivatives:

Proposition 6.7. *Let $\lambda \in \mathbb{H}$ be a quaternionic constant and let $f, g : \mathbb{H} \to \mathbb{H}$ be right-regular quaternionic functions of a variable. Then:*

- *$f + g$ is right-regular and its right derivative is given by:*

$$(f+g)'_r(q) = f'_r(q) + g'_r(q)$$

- *λf is right-regular and its right derivative is given by:*

$$(\lambda f)'_r(q) = \lambda f'_r(q)$$

We have thus defined and demonstrated some basic results about left and right $\mathbb{H}$-regular functions, let's now show some examples of left and right regular functions.

Example 6.1. *Let $f(q) = \gamma$, with $\gamma \in \mathbb{H}$ being a constant quaternionic function. Trivially, since $\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x_i} = 0$, $i = 1, 2, 3$, we have that $\bar{\partial}_l f = \bar{\partial}_r f = 0$ and thus the function is left and right regular $\forall q \in \mathbb{H}$. Both its left and right derivatives are equal to 0.*

Example 6.2. *Now let's consider a less trivial example: let $\Phi(q) = \phi_1(q) + i\phi_2(q) + j\phi_3(q) + k\phi_4(q)$ be a function of a quaternionic variable with $q = t + xi + yj + zk$ and the component functions ϕ_i, $i = 1, 2, 3, 4$ defined as:*

$$\phi_1(q) = t^2 - \frac{x^2}{2} - \frac{y^2}{2} \quad ; \quad \phi_2(q) = tx$$

$$\phi_3(q) = ty \quad ; \quad \phi_4(q) = 5$$

We observe that $\bar{\partial}_l\Phi = 2t + xi + yj + i(-x + ti) + j(-y + tj) = 0$ and therefore Φ is left-regular $\forall q \in \mathbb{H}$. Its left derivative is given by:

$$\Phi'_l(q) = -2\partial_l\Phi(q) = -4t - 2xi - 2yj$$

The function Φ, in this case, is also right-regular, since $\bar{\partial}_r\Phi = 0$, and its right derivative is given by:

$$\Phi'_r(q) = -2\partial_r\Phi(q) = -4t - 2xi - 2yj$$

Example 6.3. *The identity function $f(q) = q = t + xi + yj + zk$ is neither left nor right regular. Indeed:*

$$\bar{\partial}_l f = -2 \neq 0 \quad \forall q \in \mathbb{H}$$

$$\bar{\partial}_r f = -2 \neq 0 \quad \forall q \in \mathbb{H}$$

Example 6.4. *Consider the function $f(q) = q^2 = t^2 - x^2 - y^2 - z^2 + 2xti + 2ytj + 2tzk$:*

$$\bar{\partial}_l f = 2t + 2xi + 2yj + 2zk - 2xi - 2t - 2yj - 2t - 2zk - 2t = -4t$$

The function is left-regular only for quaternions $q \in \mathbb{H}$ with $t = 0$, that is pure quaternions $\vec{q} \in P \cong \mathbb{R}^3$.
Its left derivative at a point $q \in \mathbb{H}$ with $\mathrm{Sc}(q) = 0$ is given by $f_l'(q) = -2xi - 2yj - 2zk = -2q$.
Furthermore, the function is also right-regular in the same subset of quaternions, since $\bar{\partial}_r = -4t$, and its right derivative calculated in a pure quaternion is given by $f_r'(q) = -2xi - 2yj - 2zk = -2q = f_l'(q)$.

In these examples, we have seen some of the problems with this definition of quaternionic differentiability. Indeed, although it is a much larger class of functions than the one seen with the definition given in section 4 (which we had to immediately discard for this reason), it is still quite limited compared to its real and complex counterparts. Indeed, some functions of great importance such as the identity q are not differentiable at any point (neither to the right nor to the left, and others like q^2 only in some subsets of $\mathbb{H}$).

Exercise 6.5. *Determine the points in $\mathbb{H}$ where the function $f(q) = q^3$ is left-differentiable, and the points where it is right-differentiable. Also determine an expression for the left and right derivative of f in the sets of points in question.*

6.6 Vector Form of the Cauchy-Riemann-Fueter Equation

In the previous section, we discussed the Cauchy-Riemann-Fueter equation and also defined the concepts of regularity and left/right quaternionic derivative. We extensively talked about the deep connection between quaternions and three-dimensional vectors in chapter 2, both geometrically and algebraically, and we also saw how it was possible to write quaternionic operations in terms of vector operations (like scalar or vector products).

In this section, we want to transform the conditions imposed by the Cauchy-Riemann-Fueter equation (left and right) into a pair of vector equations, and also introduce some new notation that will be convenient later on.

Firstly, let's recall, as seen in section 2 of Chapter 2, that $\mathbb{H} \cong \mathbb{R} \oplus \mathbf{P}$ and that every quaternion can be written as the sum of a scalar part and a vector part; $\forall q \in \mathbb{H}$, $q = t + \vec{v}$. Let's first introduce the following formal quaternionic operators (this is a notation abuse and hence it is important to understand that we are not literally talking about a quaternion):

$$\Box := \frac{\partial}{\partial t} + \vec{\nabla} = \frac{\partial}{\partial t} + i\frac{\partial}{\partial x} + j\frac{\partial}{\partial y} + k\frac{\partial}{\partial z} \tag{6.28}$$

$$\overline{\Box} := \frac{\partial}{\partial t} - \vec{\nabla} = \frac{\partial}{\partial t} - i\frac{\partial}{\partial x} - j\frac{\partial}{\partial y} - k\frac{\partial}{\partial z} \tag{6.29}$$

which we will call respectively "gradient quaternion" and the "conjugate gradient quaternion".

We observe, through a simple direct calculation, that:

$$\begin{cases} \Box f = \left(\dfrac{\partial}{\partial t} + \vec{\nabla}\right) f = 2\bar{\partial}_l f \\ \overline{\Box} f = \left(\dfrac{\partial}{\partial t} - \vec{\nabla}\right) f = 2\partial_l f \\ f\Box = f\left(\dfrac{\partial}{\partial t} + \vec{\nabla}\right) = 2\overline{\partial}_r f \\ f\overline{\Box} = f\left(\dfrac{\partial}{\partial t} - \vec{\nabla}\right) = 2\partial_r f \end{cases}$$

From which it follows that we can rewrite the left Cauchy-Riemann-Fueter equation as $\Box f = 0$ and the right Cauchy-Riemann-Fueter equation as $f\Box = 0$.

Let $f : \mathbb{H} \to \mathbb{H}$ be a function of a quaternionic variable. Then it can be written as a sum of a function $\phi : \mathbb{H} \to \mathbb{R}$, its scalar part, and a function $\vec{\psi} : \mathbb{H} \to \mathbb{R}^3$, its vector part. In symbols, we write:

$$f = \phi + \vec{\psi}$$

Remember from Chapter 2 that we can write the product of two quaternions, $q_1 = t_1 + \vec{r}_1$ and $q_2 = t_2 + \vec{r}_2$ as:

$$q_1 q_2 = t_1 t_2 - \vec{r}_1 \cdot \vec{r}_2 + t_1 \vec{r}_2 + t_2 \vec{r}_1 + \vec{r}_1 \times \vec{r}_2$$

where $\cdot$ denotes the scalar product of two vectors, while $\times$ denotes their vector product.
From this it follows that we can write $\Box f$ as:

$$\Box f = \frac{\partial \phi}{\partial t} - \vec{\nabla} \cdot \vec{\psi} + \frac{\partial \vec{\psi}}{\partial t} + \vec{\nabla}\phi + \vec{\nabla} \times \vec{\psi} = 0$$

By equating the real and vector part of the following expression to 0, we can rewrite the left Cauchy-Riemann-Fueter equation as a system of two vector partial differential equations. Let's write it in more detail below:

Proposition 6.8 (Vector Form of the Left Cauchy-Riemann-Fueter Equation)**.** *The function $f = \phi + \vec{\psi}$ is left-regular in $U \subset \mathbb{H}$ if and only if $\forall q \in U \subset \mathbb{H}$:*

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = \vec{\nabla} \cdot \vec{\psi} \\ \vec{\nabla}\phi = -\dfrac{\partial \vec{\psi}}{\partial t} - \vec{\nabla} \times \psi \end{cases} \tag{6.30}$$

Proof. As already mentioned earlier, the identity is immediately derived from $\Box f = 0$ (valid in q if and only if the function f in question is left-regular at that point), by equating the scalar and vector part of the expression to 0. ◻

Proposition 6.9 (Vector Form of the Right Cauchy-Riemann-Fueter Equation)**.** *The function $f = \phi + \vec{\psi}$ is right-regular in $U \subset \mathbb{H}$ if and only if $\forall q \in \mathbb{H}$:*

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = \vec{\nabla} \cdot \vec{\psi} \\ \vec{\nabla}\phi = -\dfrac{\partial \vec{\psi}}{\partial t} + \vec{\nabla} \times \psi \end{cases} \tag{6.31}$$

Exercise 6.6. *Prove that for a quaternionic function $f = \phi + \vec{\psi}$ regular both to the right and left it is true that:*

$$\vec{\nabla} \times \vec{\psi} = 0$$

6.7 Quaternionic Integral Theorems

There are quaternionic analogs to many important integral theorems of complex analysis, such as Cauchy's Integral Theorem, Morera's Theorem, Cauchy's Integral Formula and others. Obviously, in the quaternionic case, instead of having integrals of 1-forms on contours in the complex plane, we will have integrals of 3-forms on 3-chains in quaternionic hyperspace. In this section, we will present these results and provide proofs, but first, we need some preliminary definitions.

Definition 6.7 (k-parallelepiped in quaternionic space)**.** *An oriented k-parallelepiped in quaternionic space is an application $C : I^k \to \mathbb{H}$ (where here $\mathbb{R}^k \supset I^k = \prod_{i=1}^k [0,1]$ indicates the k-unit cube) of the form:*

$$C(t_1, t_2, \dots, t_k) = q_0 + t_1 h_1 + \cdots + t_k h_k = q_0 + \sum_{i=1}^{k} t_i h_i \tag{6.32}$$

We often denote the image of the application C, $\mathrm{Im}(C) = C(I^k)$ as simply $C \subset \mathbb{H}$. Furthermore, we will call the quaternion q_0 the vertex of C, while we will call the quaternions $\{h_i\}_{i=1}^k$ edges of C. A k-parallelepiped will be called non-degenerate if its edges are $\mathbb{R}$-linearly independent.
Finally, for a non-degenerate 4-parallelepiped, we will say that it is positively oriented if $dt \wedge dx \wedge dy \wedge dz(h_1, h_2, h_3, h_4) > 0$, (i.e., if the volume form calculated on its 4 edges is positive) while we will say that it is negatively oriented if $dt \wedge dx \wedge dy \wedge dz(h_1, h_2, h_3, h_4) < 0$ (i.e., if the volume form calculated on its 4 edges is negative).

Exercise 6.7. *Given two quaternionic functions f, g $\mathbb{R}$-differentiable (i.e., differentiable in the classical sense), prove that:*

$$d(gDqf) = ((\bar{\partial}_r g)f + g(\bar{\partial}_l f))dt \wedge dx \wedge dy \wedge dz \tag{6.33}$$

Lemma 6.2. *A quaternionic function $\mathbb{R}$-differentiable f is left-regular at $q \in \mathbb{H}$ if and only if:*

$$Dq \wedge df_q = 0$$

Proof. Suppose that f is left-regular at q, then at q $\bar{\partial}_l f = 0$. Using the formula above, and substituting $g(q) = 1$, we have:

$$-Dq \wedge df_q = (\bar{\partial}_r 1) f dt \wedge dx \wedge dy \wedge dz$$

but $\bar{\partial}_r 1 = 0$ and therefore $Dq \wedge df_q = 0$.
Conversely, suppose that $Dq \wedge df_q = 0$. For the identity of the previous exercise, we have

$$Dq \wedge df_q = -\bar{\partial}_l f dt \wedge dx \wedge dy \wedge dz$$

from which necessarily $\bar{\partial}_l f = 0$, i.e., f is left-regular. □

We are now ready to prove the quaternionic analog of Cauchy's theorem for a 4-parallelepiped.

Theorem 6.5 (Cauchy-Fueter Theorem for a 4-parallelepiped)**.** *If f is left-regular at every point of the 4-parallelepiped C, then:*

$$\oiiint_{\partial C} Dqf = 0 \tag{6.34}$$

Proof. The proof proceeds in a way very similar to Goursat's theorem in complex analysis, however, here we will be dissecting a 4-parallelepiped instead of a triangle.
We write our 4-parallelepiped as $C(t_1, t_2, t_3, t_4) = q_0 + \sum_{i=1}^{4} t_i h_i$, i.e., a 4-parallelepiped with vertex q_0 and edges h_1, h_2, h_3, h_4. We now dissect our 4-parallelepiped into 16 sub-4-parallelepipeds defined as:

$$C_S(t_1, t_2, t_3, t_4) = q_0 + \frac{1}{2}\sum_{i \in S} h_i + \frac{1}{2}\sum_{j=1}^{4} t_i h_i$$

where here $S \in \mathscr{P}(\{1, 2, 3, 4\})$ is an element of the power set of $\{1, 2, 3, 4\}$, i.e., S is a subset of the set $\{1, 2, 3, 4\}$ (indeed, there are exactly 16 of them, as the cardinality of $\mathscr{P}(\{1, 2, 3, 4\})$ is $2^4 = 16$). We see that

$$\oiiint_{\partial C} Dqf(q) = \sum_{S \in \mathscr{P}(\{1,2,3,4\})} \oiiint_{\partial C_s} Dqf(q)$$

Taking the absolute value of both sides of the equation and then applying the triangle inequality we get:

$$\left|\oiiint_{\partial C} Dqf(q)\right| = \left|\sum_{S \in \mathscr{P}(\{1,2,3,4\})} \oiiint_{\partial C_s} Dqf(q)\right| \leq \sum_{S \in \mathscr{P}(\{1,2,3,4\})} \left|\oiiint_{\partial C_s} Dqf(q)\right|$$

However, among the sub-parallelepipeds obtained, there will be a "maximal" sub-parallelepiped, i.e., a sub-parallelepiped that we will call $C^{(1)}$ such that:

$$\sum_{S \in \mathscr{P}(\{1,2,3,4\})} \left|\oiiint_{\partial C_s} Dqf(q)\right| \leq \sum_{Sin \mathscr{P}(\{1,2,3,4\})} \left|\oiiint_{\partial C^{(1)}} Dqf(q)\right| = 16\left|\oiiint_{\partial C^{(1)}} Dqf(q)\right|$$

From this follows, through a sequence of $\leq$ that:

$$\left|\oiiint_{\partial C} Dqf(q)\right| \leq 16\left|\oiiint_{\partial C^{(1)}} Dqf(q)\right|$$

We now repeat on $C^{(1)}$ the dissection process just completed on C n times, obtaining a sequence of maximal sub-parallelepipeds $\{C^{(n)}\}_n$; for this sequence we will have $C \supset C_1 \supset C_2 \supset \dots C_{n-1} \supset C_n \supset \dots$. For the n-th sub-parallelepiped thus obtained we will have:

$$\left| \oiiint_{\partial C} Dqf(q) \right| \leq 16^n \left| \oiiint_{\partial C^{(n)}} Dqf(q) \right| \tag{6.35}$$

Observe that the intersection over $\mathbb{N}$ of these parallelepipeds will converge to a point, which we will call q_0, $q_0 \in \bigcap_{n\in\mathbb{N}} C^{(n)}$, and there will exist a quaternionic sequence $\{q_n\}_{n\in\mathbb{N}}$ with $q_n \in C^{(n)}\ \forall n$ such that $\lim_{n\to\infty} q_n = q_0$. By hypothesis f is regular at q_0 and thus:

$$f(q) = f(q_0) + df_{q_0}(q - q_0) + r(q)(q - q_0) \tag{6.36}$$

with $r(q) \to 0$ as $q \to q_0$. Defining the value of r at q_0 as $r(q_0) = 0$, we will have that r will be a continuous function of a quaternionic variable on our parallelepipeds, and in particular its norm $|r(q)|$ will have a maximum value $\rho_n \in C^{(n)}\ \forall n$. But observing that our sequence of maximal sub-parallelepipeds $\{C^{(n)}\}_n$, in light of what was said earlier, converges to the point q_0, we will have that $\lim_{n\to\infty} \rho_n = 0$.
As a consequence of the considerations just made, and in particular of equation (6.36), we can write $\oiiint_{C^{(n)}} Dqf(q)$ as:

$$\oiiint_{C^{(n)}} Dqf(q) = \oiiint_{C^{(n)}} Dq(f(q_0) + df_{q_0}(q - q_0) + r(q)(q - q_0))$$

Observing now that $\oiiint_{\partial C^{(n)}} Dqf(q_0) = 0$ and that $\oiiint_{\partial C^{(n)}} Dqdf_{q_0}(q - q_0) = \iiiint_{C^{(n)}} Dq \wedge df_{q_0} = 0$ (by the previous lemma, since by hypothesis f is regular at q_0 and thus $Dq \wedge df_{q_0} = 0$). We finally arrive at the following equation:

$$\oiiint_{\partial C^{(n)}} Dqf(q) = \oiiint_{\partial C^{(n)}} Dq(q - q_0)r(q)$$

Let $F : I^3 \to \mathbb{H}$ now be one of the 3-parallelepipeds that forms the face of C_n: then $F \subset \partial C^{(n)}$, and the edges of F are 3 of the edges of $C^{(n)}$, $2^{-n}h_a, 2^{-n}h_b, 2^{-n}h_c$. $\forall q \in F$ we will have $|r(q)| \leq \rho_n$ and $|q - q_0| \leq 2^{-n}(|h_1| + |h_2| + |h_3| + |h_4|)$. From this follows that:

$$\oiiint_{F} Dq(q - q_0)r(q) \leq 8^{-n}|Dq(h_a, h_b, h_c)|2^{-n}(|h_1| + |h_2| + |h_3| + |h_4|)\rho_n$$

But observing that the integral $\oiiint_{\partial C^{(n)}} Dq(q-q_0)r(q)$ is given by the sum of the quaternionic integrals of the same integrand form on the 8 faces F of C^n:

$$\oiiint_{\partial C^{(n)}} Dq(q-q_0)r(q) = \sum_{F\ face} \oiiint_F Dq(q-q_0)r(q)$$

we conclude that:

$$\left|\oiiint_{\partial C^{(n)}} Dq(q-q_0)r(q)\right| \leq 8(16^{-n}\max_{a,b,c}\{|Dq(h_a,h_b,h_c)|\}(|h_1|+|h_2|+|h_3|+|h_4|)\rho_n) \tag{6.37}$$

But remembering that $\oiiint_{\partial C^{(n)}} Dqf(q) = \oiiint_{\partial C^{(n)}} Dq(q-q_0)r(q)$ and, combining the equation (6.37) just obtained with the equation (6.35) we get:

$$\left|\oiiint_{\partial C} Dqf(q)\right| \leq 8\max_{a,b,c}\{|Dq(h_a,h_b,h_c)|\}(h_1+h_2+h_3+h_4)\rho_n \tag{6.38}$$

But $\lim_{n\to\infty}\rho_n = 0$, and thus $\oiiint_{\partial C} Dqf(q) = 0$, concluding the proof of the theorem. □

With a similar strategy, we will now prove a quaternionic analog of Cauchy's integral formula for 4-parallelepipeds, conventionally called in the literature "Cauchy-Fueter Integral Formula". The theorem in question in quaternionic analysis assumes the same importance as its complex counterpart, as it can be deduced as a corollary of the latter that left-regular functions are also of class C^∞ in the classical sense of differential geometry (i.e., infinitely $\mathbb{R}$-differentiable).

Theorem 6.6 (Cauchy-Fueter Integral Formula on a 4-parallelepiped)**.** *Let f be a left-regular function at every point of a positively oriented 4-parallelepiped C, and let $q_0 \in \mathrm{int}(C)$ be a point inside C, then:*

$$f(q_0) = \frac{1}{2\pi^2}\oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q) \tag{6.39}$$

Proof. Recall the result of an exercise presented a short while ago:

$$d(gDqf) = [(\overline{\partial}_r g)f + g(\overline{\partial}_l f)]dt \wedge dx \wedge dy \wedge dz$$

Let $f(q)$ be a left-regular function, and let $g(q)$ be the function defined as:

$$g(q) = \frac{(q-q_0)^{-1}}{|q-q_0|^2} = -\partial_r\left(\frac{1}{|q-q_0|^2}\right)$$

Moreover, let our 4-parallelepiped C be explicitly defined as $C(t_1, t_2, t_3, t_4) = q_0 + \sum_i h_i t_i$. The function g is differentiable everywhere except at the point q_0, and in the same set of points it is right-regular, since $\forall q \in \mathbb{H} \setminus \{q_0\}$, $\bar{\partial}_r g(q) = 0$. Hence, by the lemma, also mentioned above, $d\left(\frac{(q-q_0)^{-1}}{\|q-q_0\|^2} Dqf(q)\right) = 0$ everywhere except at q_0. We observe that, following a reasoning entirely analogous to that of theorem 6.5, we obtain that for every sub-4-parallelepiped C' of C that does not contain the point q_0, the integral

$$\oiiint_{\partial C'} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q) = 0 \tag{6.40}$$

is equal to 0.

We dissect the 4-parallelepiped into 81 sub-parallelepipeds with edges parallel to the edges of C, $\{C_j\}_j$. From equation (6.40) we can deduce that

$$\oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q) = \oiiint_{\partial C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q)$$

where C_0 is a member of the aforementioned partition (i.e., a 4-parallelepiped with edges parallel to C) such that $q_0 \in \mathrm{int}(C_0)$.

Let δ be a positive real number: we explicitly denote our 4-parallelepiped C_0 as

$C_0(t_1, t_2, t_3, t_4) = q_0 - \frac{1}{2}\sum_{j=1}^4 \delta h_j + \sum_{k=1}^4 \delta t_k h_k$, where q_0 is the vertex of C, as defined earlier, and h_i $i = 1, 2, 3, 4$ are the edges of C; then:

$$\min_{q \in \partial C_0} |q - q_0| = \min_{1 \le i_1, i_2, i_3 \le 4} \left| \frac{dt \wedge dx \wedge dy \wedge dz(\delta h_1, \delta h_2, \delta h_3, \delta h_4)}{Dq(\delta h_{i_1}, \delta h_{i_2}, \delta h_{i_3})} \right| = \delta W(h_1, h_2, h_3, h_4) \tag{6.41}$$

where here $W(h_1, h_2, h_3, h_4) = \min_{1 \le i_1, i_2, i_3 \le 4} \left| \frac{dt \wedge dx \wedge dy \wedge dz(h_1, h_2, h_3, h_4)}{Dq(h_{i_1}, h_{i_2}, h_{i_3})} \right|$. Since $f(q)$ by hypothesis is continuous at q_0, we can choose a $\delta > 0$ such that $q \in C_0 \Longrightarrow |f(q) - f(q_0)| \le \epsilon \ \forall \epsilon > 0$. Therefore:

$$\left| \oiiint_{\partial C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq[f(q) - f(q_0)] \right| \le \frac{\max_{1 \le i_1, i_2, i_3 \le 4} \{|Dq(h_{i_1}, h_{i_2}, h_{i_3})|\}}{W^3} \epsilon \tag{6.42}$$

Let S^3, as before, be the 3-sphere with volume element $dS = r^3 \sin^2\theta \sin\phi d\theta \wedge d\phi \wedge d\psi$. Noting that on S $Dq = (q - q_0)dS$, we will have:

$$\oiiint_{S^3} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq = \oiiint_{S^3} dS = 2\pi^2$$

But we observe that the 3-form $\frac{q-q_0}{\|q-q_0\|^2}Dq$ is closed (i.e., its exterior derivative is 0) and is continuously differentiable in $\mathbb{H} \setminus \{q_0\}$, we will have that for proposition 1.17 of the introduction:

$$\oiiint_{\partial C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq = \oiiint_{S^3} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq = \oiiint_{S^3} dS = 2\pi^2$$

substituting this fact in equation (6.42), we obtain the inequality:

$$\left| \oiiint_{\partial C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) - 2\pi^2 f(q_0) \right| \leq \frac{\max\limits_{1 \leq i_1, i_2, i_3 \leq 4} \{|Dq(h_{i_1}, h_{i_2}, h_{i_3})|\}}{W^3} \epsilon$$

From which, for the arbitrariness of ϵ, we will obtain:

$$\oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = \oiiint_{\partial C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = 2\pi^2 f(q_0)$$

and dividing both sides by $2\pi^2$ we arrive at the proof of the assertion:

$$f(q_0) = \frac{1}{2\pi^2} \oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) \tag{6.43}$$

□

We recall that the function $G(q-q_0) = \frac{(q-q_0)^{-1}}{|q-q_0|^2}$ present in the integral is an $\mathbb{R}$-analytic function except at the point $q = q_0$. From this follows that the function to be integrated on the right side of the equality of the theorem just demonstrated is a continuous function of $(q, q_0) \in \partial C \times \mathrm{int}(C)$ and, for a fixed $q \in \partial C$, an $\mathbb{R}$-analytic function (i.e C^∞ in the classical sense) of q_0 in $\mathrm{int}(C)$. From this follows that the integral of such function is an $\mathbb{R}$-analytic function of q_0 in $\mathrm{int}(C)$ [Note 6.2].
We enunciate this fact more exhaustively in the following corollary.

Corollary 6.3. *Let $f : \mathbb{H} \to \mathbb{H}$ be a left-regular function in a region $\Omega \subset \mathbb{H}$. Then we will have that f will also be $\mathbb{R}$-analytic in Ω.*

All integral theorems for left-regular quaternionic functions also hold for right-regular functions, obviously multiplying f by Dq from the right rather than the left. It will therefore follow that the corollary just enunciated is also valid for right-regular functions, i.e.:

Corollary 6.4. *Let $f : \mathbb{H} \to \mathbb{H}$ be a right-regular function in a region $\Omega \subset \mathbb{H}$. Then we will have that f will also be $\mathbb{R}$-analytic in Ω.*

This fact has 2 very important consequences: the first is that every right/left regular function is a harmonic function, as the necessary condition of being differentiable "in the classical sense" at least twice can be encompassed in the notion of a regular function due to the fact just shown.

Corollary 6.5. *Every function $f : \mathbb{H} \to \mathbb{H}$ right/left regular is harmonic.*

The second, instead, is that we can extend the Cauchy-Fueter theorem, so far proved only on borders of 4-parallelepipeds (theorem 6.5) to borders of smooth 4-chains.

Theorem 6.7 (Cauchy-Fueter Theorem)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a left-regular function in an open set $U \subset \mathbb{H}$, and let C be a smooth 3-chain in U homologous to 0 in the smooth singular homology of U, i.e $C = \partial C'$ is the border of a smooth singular 4-chain in U; then:*

$$\oiiint_C Dqf(q) = \oiiint_{\partial C'} Dqf(q) = 0 \tag{6.44}$$

Proof. We know that, by lemma 6.2, $d(Dqf(q)) = 0$. Moreover, by the corollary just mentioned we will have that the 3-form $Dqf(q)$ is a smooth form on U, therefore we can apply Stokes' theorem on smooth chains (theorem 1.9 of the introduction), obtaining:

$$\oiiint_C Dqf(q) = \oiiint_{\partial C'} Dqf(q) = \iiiint_{C'} d(Dqf(q)) = 0$$

□

Similarly, using Stokes' theorem for smooth chains, we can also extend the Cauchy-Fueter integral formula to more general contours. However, we will need to introduce the following preliminary notion:

Definition 6.8 (Winding number)**.** *Let $q \in \mathbb{H}$ and let C be a smooth 3-cycle in $\mathbb{H} \setminus \{q\}$. Then C is homologous to a 3-chain $C' : \partial I^4 \to S^3_q$, where here S^3_q is the sphere of radius 1 centered at q. The* ***winding number*** *of C around q is the Brouwer degree of the application C' (for the definition of Brouwer degree see section 8 of chapter 4).*

Theorem 6.8 (Cauchy-Fueter Integral Formula for Smooth Chains)**.** *Let f be a left-regular function in an open set $U \subset \mathbb{H}$. Let $q_0 \in U$, and let C be a smooth 3-chain in $U \setminus \{q_0\}$ which is homologous, in the smooth singular* *homology of $U \setminus \{q_0\}$, to a 3-chain whose image is ∂B for a quaternionic ball $B \subset U$; then:*

$$n f(q_0) = \frac{1}{2\pi^2} \oint\!\!\!\oint\!\!\!\oint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) \tag{6.45}$$

where here n is the winding number of C around q_0.

Proof. For the case $n = 0$ we have that C will be homologous to 0 in $U \setminus \{q_0\}$, i.e. it will be the boundary of a smooth 4-chain C_0, $C = \partial C_0$. Being the 3-form $\frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q)$ closed and smooth in $U \setminus \{q_0\}$, as a consequence of Stokes' theorem we have:

$$\frac{1}{2\pi^2} \oint\!\!\!\oint\!\!\!\oint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = \frac{1}{2\pi^2} \iiiint_{C_0} d\Big(\frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q)\Big) = 0$$

For the case $n = 1$, C will be homologous to a smooth 3-chain $C' : \partial I^4 \to \partial B$, with $q_0 \in B \subset U$. The map C' also has a Brouwer degree of 1; considering these facts, we have that C is homologous to the boundary of a 4-parallelepiped C_0, $C \simeq \partial C_0$, such that $q_0 \in \operatorname{int} C_0$ and $C_0 \subset U$. Using again the fact that the 3-form $\frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q)$ is closed and smooth in $U \setminus \{q_0\}$, we have:

$$\frac{1}{2\pi^2} \oint\!\!\!\oint\!\!\!\oint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = \frac{1}{2\pi^2} \oint\!\!\!\oint\!\!\!\oint_{C_0} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = f(q_0)$$

for the Cauchy-Fueter integral formula for 4-parallelepipeds.
For a general $n \in \mathbb{N}$ we have that C will be homologous to a 3-chain $C'' = \rho \circ C'$, where $C' : \partial I^4 \to \partial B$ has the same characteristics as the map C' earlier, and the application $\rho : \partial B \to \partial B$ is a mapping of Brouwer degree equal to n in the form:

$$\rho(q_0 + r(v + jw)) = q_0 + r(v^n + jw)$$

Now we dissect our chain C' as:

$$C' = \sum_{l=1}^{n} C'_l$$

where the image of the sub-chain C'_l is the set:

$$\operatorname{Im} C'_l = \left\{ q = q_0 + r(v + jw) \in \partial B \,;\, \frac{2\pi(l-1)}{n} \le \arg(v) \le \frac{2\pi l}{n} \right\}$$

Every composition $\rho \circ C'_l$ has as image $\operatorname{Im}(\rho \circ C'_l) = \partial B$ and winding index equal to 1 around q_0: therefore, for the case above we have:

$$\frac{1}{2\pi^2} \oiiint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = \frac{1}{2\pi^2} \sum_{l=1}^{n} \oiiint_{\rho \circ C'_l} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = n f(q_0)$$

For the case $n = -1$, the chain C is homologous to a 3-chain C'' of the form $C'' = C' \circ K$, where here $C' : \partial I^4 \to \partial B$ has a Brouwer degree of 1, and $K : \partial I^4 \to \partial I^4$ has a degree of -1. Then, for arguments similar to the above, and taking the reflection $(t_1, t_2, t_3, t_4) \to (1 - t_1, t_2, t_3, t_4)$ as our K we have:

$$\frac{1}{2\pi^2} \oiiint_C \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = \frac{1}{2\pi^2} \oiiint_{C''} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q)$$

$$= -\frac{1}{2\pi^2} \oiiint_{C'} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dq f(q) = -f(q_0)$$

Finally, for a generic $n \in \mathbb{Z}^-$ (a negative integer), it will suffice to compose the argument above for $n = -1$ with the dissection carried out for the case of a generic $n \in \mathbb{N}$. □

As a corollary of the Cauchy-Fueter integral formula, we have the quaternionic Liouville theorem, which we prove here below:

Theorem 6.9 (Liouville's Theorem)**.** *Let f be a function $f : \mathbb{H} \to \mathbb{H}$ that is left-regular for all $q \in \mathbb{H}$; then if f is bounded (i.e., there exists a positive real number M such that $|f(q)| \leq M$ for all $q \in \mathbb{H}$), f is constant.*

Proof. By hypothesis, there exists an $M \in \mathbb{R}^+$ such that $|f(q)| \leq M$ for all $q \in \mathbb{H}$. Let C be the 3-sphere centered at the origin with radius R such that $R > \max\{|q_1|, |q_2|\}$, and let q_1, q_2 be two arbitrary quaternions; then by the Cauchy-Fueter integral formula for smooth chains, we have:

$$f(q_1) - f(q_2) = \frac{1}{2\pi^2} \oiiint_C \left[\frac{(q-q_1)^{-1}}{n|q-q_1|^2} - \frac{(q-q_2)^{-1}}{m|q-q_2|^2} \right] Dq f(q)$$

where n and m indicate the winding numbers of C around q_1 and q_2, respectively. Now consider $|f(q_1) - f(q_2)|$:

$$|f(q_1) - f(q_2)| = \left| \frac{1}{2\pi^2} \oiiint_C \left[\frac{(q-q_1)^{-1}}{n|q-q_1|^2} - \frac{(q-q_2)^{-1}}{m|q-q_2|^2} \right] Dq f(q) \right|$$

$$\leq \gamma V(C)M\left|\frac{(q-q_1)^{-1}}{n|R-q_1|^2}-\frac{(q-q_2)^{-1}}{m|R-q_2|^2}\right|$$

where $\gamma \in \mathbb{R}$ and $V(C)$ is the hyper-volume of C in $\mathbb{H}$. As $R \to \infty$, we observe that the last expression in the above chain of inequalities tends to 0 for q in the ball B centered at 0 with radius R (i.e., for $|q| < R$), and therefore we have that $f(q_1) = f(q_2)$ for all $q_1, q_2 \in \mathbb{H}$. This is equivalent to saying that the function $f(q)$ is constant, thus proving the assertion. □

We conclude this section with a proof of a quaternionic analogue of Morera's theorem, another quaternionic analogue of an important integral theorem of complex analysis.

Theorem 6.10 (Morera's Theorem for Quaternions, A.Sudbery, 1979)**.** *Let* $f : \mathbb{H} \to \mathbb{H}$ *be a continuous function in an open set* $U \subset \mathbb{H}$ *such that:*

$$\oiiint_{\partial C} Dqf(q) = 0$$

for every 4-parallelepiped C *contained in* U*. Then* f *is left-regular in* U*.*

Proof. The proof strategy will be as follows: we will first demonstrate that the assumptions made are sufficient to validate the Cauchy-Fueter integral formula for f, and then we will demonstrate that such a function is indeed left-regular.

First, let's demonstrate that $\oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q) = 0$ if $q_0 \notin C$, using the same argument used for the proof of the Cauchy-Fueter integral formula. As in theorem 6.6, we dissect the 4-parallelepiped into a sequence of sub-parallelepipeds $\{C_n\}_n$ converging to a point we will call q_∞ in such a way as to respect, as before, the following inequality:

$$\left|\oiiint_{\partial C} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q)\right| \leq 16^n \left|\oiiint_{\partial C^n} \frac{(q-q_0)^{-1}}{|q-q_0|^2} Dqf(q)\right| \tag{6.46}$$

From now on, for simplicity, let's call the function $\frac{(q-q_0)^{-1}}{|q-q_0|^2} = G(q-q_0)$.
Since we have assumed that $q_0 \notin C$, $G(q-q_0)$ will be a right-regular function of q in that region, and therefore we have (by an analogue of lemma 6.2 for right-regular quaternionic functions) that:

$$dG_{q-q_0} \wedge Dq = 0$$

Furthermore, the following expression will also be valid:

$$G(q-q_0) = G(q_\infty - q_0) + dG_{q_\infty - q_0}(q-q_\infty) + (q-q_\infty)r(q) \tag{6.47}$$

where $r(q)$ tends to 0 as $q \to q_\infty$. We will also write:

$$f(q) = f(q_\infty) + s(q) \tag{6.48}$$

where $s(q) \to 0$ as $q \to q_\infty$. We can now write the integral $\oiiint_{\partial C^n} G(q - q_0)Dqf(q)$, using equations (6.47) and (6.48), as:

$$\oiiint_{\partial C^n} G(q-q_0)Dqf(q) = G(q_\infty-q_0) \oiiint_{\partial C^n} Dqf(q) + \oiiint_{\partial C^n} dG_{q_\infty-q_0}(q-q_\infty)Dqf(q_0)$$

$$+ \oiiint_{\partial C^n} dG_{q_\infty-q_0}(q - q_\infty)Dqs(q) + \oiiint_{\partial C^n} (q - q_\infty)r(q)Dqf(q)$$

The first of the four integrals is null since, by hypothesis, we have assumed that for every 4-parallelepiped C, $\oiiint_{\partial C} Dqf(q) = 0$. The second one, in turn, is null because $dG_{q_\infty-q_0} \wedge Dq = 0$, due to the right-regularity of G. Now, for $q \in \partial C^n$, we have $|q - q_\infty| \leq \frac{1}{2^n}L$, where L is the sum of the lengths of the edges of C. Since $dG_{q_\infty-q_0}$ is linear, there will be a real number M such that:

$$|dG_{q_\infty-q_0}(q - q_\infty)| \leq M|q - q_\infty|$$

The volume of each face of the parallelepiped C^n is at most $8^{-n}V$, where V is the volume of the largest face of C. Therefore:

$$\left|\oiiint_{\partial C^n} G(q - q_0)Dqf(q)\right| \leq \frac{1}{16^n}[LMV\sigma_n + L\rho_n V(|f(q_\infty)| + \sigma_n)]$$

where ρ_n and σ_n are the maximum values of the functions $r(q)$ and $s(q)$ respectively on ∂C^n. Since both tend to 0 as $n \to \infty$, $\rho_n \to 0$ and $\sigma_n \to 0$, we have that:

$$\left|\oiiint_{\partial C^n} G(q - q_0)Dqf(q)\right| \to 0 \ \ as \ \ n \to \infty \Longrightarrow$$

$$16^n \left|\oiiint_{\partial C^n} G(q - q_0)Dqf(q)\right| \to 0 \ \ as \ \ n \to \infty$$

and therefore, by the inequality (6.46), we have that $\oiiint_{\partial C} \frac{(q - q_0)^{-1}}{|q - q_0|^2} Dqf(q) = 0$. Now consider the case where $q_0 \in C$: for what we have just demonstrated, we will be able to replace the parallelepiped C with a smaller one that also contains q; from this point on, the arguments made for the proof of the Cauchy-Fueter integral formula can be adapted in the same way to the current hypotheses as they depend only on the continuity of f. The following expression is therefore valid:

$$f(q) = \frac{1}{2\pi^2} \oiiint_{\partial C} G(q' - q) Dq' f(q')$$

for every 4-parallelepiped C with $q_0 \in \operatorname{int} C \subset U$. For arguments similar to those of corollary 6.3, we will have that f is a differentiable function (in the classical sense). Directly calculating $\bar{\partial}_l f(q)$, we obtain:

$$\bar{\partial}_l f(q) = \frac{1}{2\pi^2} \oiiint_{\partial C} \bar{\partial}_l [G(q' - q)] Dq' f(q') = 0 \tag{6.49}$$

which is zero since the function G is regular. From this it follows that f is a left-regular function. □

It is precisely from Morera's quaternionic theorem that C.A Deavours [”The Quaternion Calculus”, Amer. Math. Monthly 80 (1973), pages 995-1008] [6] defines the concept of a regular function of a quaternionic variable. Given the impossibility of defining the derivative of a quaternionic function as the limit of an incremental ratio (right or left), as seen in section 4 of this chapter, one can choose to define the concept of a regular function precisely as that class of functions that respects a hypercomplex analogue of Cauchy's theorem.
From this starting point, in the paper in question, many results are obtained about the Fueter-regular functions that we have presented in this chapter.

6.8 Constructing Regular Functions from Harmonic Functions

We have previously seen that we can write the Laplacian operator in terms of the operators defined in section 3 as:

$$\triangle f = 4\bar{\partial}_l \partial_l f = 4\partial_l f \bar{\partial}_l f \tag{6.50}$$

Suppose we have a harmonic function f; being $\triangle f = 0$, we will have $4\bar{\partial}_l \partial_l f = 0$, i.e., $\partial_l f$ satisfies the Cauchy-Riemann-Fueter equation and hence is left-regular. From this observation, it follows that we can construct quaternionic regular functions starting from solutions of the Laplace equation in 4 dimensions:

$$\triangle f = \frac{\partial^2 f}{\partial^2 t} + \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} + \frac{\partial^2 f}{\partial^2 z} = 0 \tag{6.51}$$

An integral expression for (almost) all the solutions of the following equation comes from Whittaker (1903) [50]:

$$f(t,x,y,z) = \int_0^{2\pi}\int_0^{\pi} \Phi(t\cos(u_1)\cos(u_2)+x\cos(u_1)\sin(u_2)+y\sin(u_1)+iz, u_1, u_2)du_1du_2 \tag{6.52}$$

A simplification of the latter to a single integral is due to Bateman (Solution of Partial Differential Equations, 1904, pp.457) [3]:

$$f(t,x,y,z) = \int_0^{2\pi} \Phi(t\cos(u)+x\sin(u)+iy, t\sin(u)-y\cos(u)+iz, u)du \tag{6.53}$$

where here $\Phi : \mathbb{C}^3 \to \mathbb{C}$ is a function in 3 complex variables.
There are many clarifications to be made on this matter, but as this text is a monograph focused strictly on quaternions and their properties, we will limit ourselves to directing the curious reader to the various resources present in the bibliography (particularly [50], [3] and [10]).
Having said that, using the expressions just introduced and the fact presented at the beginning of the section, we will now provide some practical examples of constructing regular functions from harmonic functions:

Example 6.5. *Consider the function $\Phi : \mathbb{C}^3 \to \mathbb{C}$ defined as $\Phi(z_1, z_2, z_3) = z_1^2$. The integral* (6.53) *reduces to:*

$$f(t,x,y,z) = \int_0^{2\pi} (t\cos(u) + x\sin(u) + iy)^2 du$$

expanding the square, and using simple rules of real integral calculus we obtain:

$$f(t,x,y,z) = \pi t^2 + \pi x^2 - 2\pi y^2$$

Now calculate $\partial_l f = \frac{1}{2}(\frac{\partial f}{\partial t} - e_i \frac{\partial f}{\partial x_i})$:

$$\partial_l f = \frac{1}{2}(2\pi t - 2\pi x i + 4\pi y j) = \pi t - \pi x i + 2\pi y j$$

we call the function just obtained $\tilde{f} : \mathbb{H} \to \mathbb{H}$. Check that $\tilde{f}$ is actually a regular function:

$$\bar{\partial}_l \tilde{f} = \pi + i(-\pi i) + j(2\pi j) = 0$$

Example 6.6. *Let $\Phi(z_1, z_2, z_3) = z_2 z_1 + z_2^2 + z_3$. The integral* (6.53) *becomes:*

$$f = \int_0^{2\pi} [(t\sin(u) - y\cos(u) + iz)(t\cos(u) + x\sin(u) + iy) + (t\sin(u) - y\cos(u) + iz)^2 + u]du$$

Expanding everything, we obtain that the integrand is equal to $t^2\sin(u)\cos(u)+xt\sin^2(u)+yti\sin(u)-ty\cos^2(u)-xy\cos(u)\sin(u)-y^2i\cos(u)+tiz\cos(u)+zxi\sin(u)-yz+u+t^2\sin^2(u)+y^2\cos^2(u)-z^2+2zit\sin(u)-2ziy\cos(u)-2yt\cos(u)\sin(u)$. *Performing the integral, we obtain an expression for the function* $f(t,x,y,z)$*:*

$$f=\pi xt-\pi yt-2\pi yz+2\pi^2+\pi t^2-2\pi z^2+\pi y^2$$

We call as before $\tilde{f}=\partial_l f=\frac{1}{2}(\pi x-\pi y+2\pi t-\pi ti-2\pi yj+2\pi zj+\pi tj+2\pi yk+4\pi zk)$*. The obtained function is left-regular. Its left derivative is:*

$$\tilde{f}_l'(q)=-2\pi j-4\pi+2\pi i$$

Exercise 6.8. *Calculate the left derivative* $\tilde{f}_l'(q)$ *of the function in example 6.5.*

There are other ways to construct regular functions from harmonic functions, let's see the following:

Theorem 6.11. *Let* $u:\mathbb{R}\to\mathbb{R}$ *be a real function defined on a star-shaped subset of the quaternions,* $U\subset\mathbb{H}$*. If u is harmonic and has continuous second partial derivatives, then there exists a regular function* f *defined on* U *such that its scalar part is equal to u, i.e.:*

$$\mathrm{Sc}(f)=u$$

Proof. Assuming, without loss of generality, that the set U is star-shaped with respect to the origin; if not, in fact, we will simply have to apply a translation, slightly changing the demonstration that we are about to present.
We will proceed by setting up the following constructive demonstration: we want to show that the function f defined as:

$$f(q)=u(q)+2\,\mathrm{Vec}\int_0^1 s^2\partial_l u(sq)q\,ds \tag{6.54}$$

is a regular function.
Let's start manipulating the scalar part of the integral $\int_0^1 s^2\partial_l u(sq)q\,ds$:
First, note that $\mathrm{Sc}\int_0^1 s^2\partial_l u(sq)q\,ds=\mathrm{Sc}\int_0^1 s^2\Big(\frac{\partial u}{\partial t}(sq)-e_i\frac{\partial u}{\partial x_i}(sq)\Big)q\,ds$, and performing a quaternionic multiplication we have $\mathrm{Sc}\int_0^1 s^2\partial_l u(sq)q=\frac{1}{2}\int_0^1 s^2\Big(t\frac{\partial u}{\partial t}(sq)+x_i\frac{\partial u}{\partial x_i}(sq)\Big)ds$. Having $t\frac{\partial u}{\partial t}(sq)+x_i\frac{\partial u}{\partial x_i}(sq)=\frac{du(sq)}{ds}$, we obtain $\mathrm{Sc}\int_0^1 s^2\partial_l u(sq)q=$

$\frac{1}{2}\int_0^1 s^2 \frac{du(sq)}{ds} ds$. Integrating by parts the integral just obtained, finally, we have:

$$\operatorname{Sc}\int_0^1 s^2 \partial_l u(sq) q = \frac{1}{2}u(q) - \int_0^1 su(sq)ds$$

from which we derive $u(q) = 2\operatorname{Sc}\int_0^1 s^2 \partial_l u(sq) q ds + 2\int_0^1 su(sq)ds$. Substituting this expression in equation (6.54), we obtain

$$f(q) = 2\int_0^1 s^2 \partial_l u(sq) q ds + 2\int_0^1 su(sq)ds$$

Let's verify that f is regular by calculating $\bar{\partial}_f$. Applying the operator $\bar{\partial}_l$ to f and differentiating under the integral sign (we can do this because the integrand functions are differentiable and have continuous partial derivatives) we obtain:

$$\bar{\partial}_l f = 2\int_0^1 s^2 \bar{\partial}_l [\partial_l u(sq)] q ds + \int_0^1 s^2 (\partial_l u(sq) + e_i \partial_l u(sq) e_i) ds + 2\int_0^1 s^2 \overline{\partial}_l u(sq) ds$$

Being u harmonic, we will have that $\bar{\partial}_l \partial_l u(sq) = \triangle u(sq) = 0$, and also, manipulating the integrand of the second integral:

$$\partial_l u(sq) + e_i \partial_l u(sq) e_i = -2\overline{(\partial_l u(sq))} = -2\bar{\partial}_l u(sq)$$

i.e the second integral is equal but opposite in sign to the third, and we will therefore obtain the hoped-for result:

$$\bar{\partial}_l f = 0$$

thus f is regular. □

In the case where the star-shaped region under consideration should not be star-shaped with respect to the origin, but rather with respect to another point q_0, the expression (6.54) for f derived in the theorem just proved becomes:

$$f(q) = u(q) + 2\operatorname{Vec}\int_0^1 s^2 \partial_l u((1-s)q_0 + sq)(q - q_0) ds \tag{6.55}$$

Example 6.7. *Let U be the quaternions minus the origin and the negative real line, $U = \mathbb{H} \setminus \{x \in \mathbb{R} \ ; \ x \leq 0\}$; this set is a star-shaped set with respect to $1 \in \mathbb{H}$. Consider on this set the function $u : \mathbb{H} \to \mathbb{R}$ defined* *as* $u(q) := \frac{1}{|q|^2}$*. The function is harmonic on* U*, and thus we can apply the result just derived to find a regular function* $f : \mathbb{H} \to \mathbb{H}$ *whose real part is* $\frac{1}{|q|^2}$*. Let's calculate the partial derivatives of* u *with respect to* t, x, y, z*:*

$$\begin{cases} \frac{\partial u}{\partial t} = \frac{-2t}{(t^2+x^2+y^2+z^2)^2} \\ \frac{\partial u}{\partial x} = \frac{-2x}{(t^2+x^2+y^2+z^2)^2} \\ \frac{\partial u}{\partial y} = \frac{-2y}{(t^2+x^2+y^2+z^2)^2} \\ \frac{\partial u}{\partial z} = \frac{-2z}{(t^2+x^2+y^2+z^2)^2} \end{cases}$$

From which, substituting in $\partial_l u(q) = \frac{1}{2}\left(\frac{\partial u}{\partial t} - i\frac{\partial u}{\partial x} - i\frac{\partial u}{\partial y} - i\frac{\partial u}{\partial z}\right)$ *and remembering that* $q^{-1} = \frac{q^*}{|q|^2}$ *we will have that* $\partial_l u(q) = -\frac{q^{-1}}{|q|^2}$*, and substituting all this in the formula* (6.55) *we obtain an expression for* f*:*

$$f(q) = \begin{cases} -(q\operatorname{Vec}(q))^{-1} - \frac{\operatorname{Vec}(q)}{|\operatorname{Vec}(q)|^3}\arctan\left(\frac{|\operatorname{Vec}(q)|}{\operatorname{Sc}(q)}\right) & if\ \operatorname{Vec}(q) \neq 0 \\ \frac{1}{|q|^2} & if\ q \in \mathbb{R}^+ \end{cases}$$

We will call the function f *reported above* $-2\mathscr{L}(q)$*, from which:*

$$\mathscr{L}(q) = \begin{cases} \frac{1}{2}\left((q\operatorname{Vec}(q))^{-1} + \frac{\operatorname{Vec}(q)}{|\operatorname{Vec}(q)|^3}\arctan\left(\frac{|\operatorname{Vec}(q)|}{\operatorname{Sc}(q)}\right)\right) & if\ \operatorname{Vec}(q) \neq 0 \\ \frac{-1}{2|q|^2} & if\ q \in \mathbb{R}^+ \end{cases}$$

We can write $\mathscr{L}(q)$ *more explicitly as:*

$$\mathscr{L}(q) = \begin{cases} -\frac{|\operatorname{Vec}(q)|^2 + te_ix_i}{2|\operatorname{Vec}(q)|^2(|\operatorname{Vec}(q)|^2+t^2)} + \frac{e_ix_i}{2|\operatorname{Vec}(q)|^3}\arctan\left(\frac{|\operatorname{Vec}(q)|}{t}\right) & if\ \operatorname{Vec}(q) \neq 0 \\ \frac{-1}{2|q|^2} & if\ q \in \mathbb{R}^+ \end{cases}$$

Here $t = \operatorname{Sc}(q)$*.*
We have now rewritten $\mathscr{L}$ *in order to compute its left derivative more easily; let's calculate it using the formula derived earlier now well known to the reader:*

$$\mathscr{L}_l'(q) = -2\partial_l\mathscr{L} = \frac{q^{-1}}{|q|^2} = \mathrm{G}(q)$$

We can notice an interesting parallel with the logarithm function in complex analysis: in quaternionic analysis, the logarithm function derived in chapter 5 is not regular in the sense presented in this chapter (in the sense of Fueter), but nevertheless there will exist a regular function that serves as a primitive to the function that appears in the Cauchy-Fueter integral formula, $G(q-q_0)$, just as in $\mathbb{C}$ $log(z)$ serves as a primitive to the function $\frac{1}{z}$ that appears in the complex Cauchy integral formula.

Having concluded this discussion, we now review other methods for constructing regular functions of a quaternionic variable.
In the next section, in particular, we will see how to construct regular functions starting from holomorphic functions in the complex plane, and we will also see other properties that unite the regular function $\mathscr{L}(q)$ with the complex logarithm function.

6.9 Using $\mathbb{C}$-Holomorphic Functions to construct regular functions

Before giving a method to construct regular functions starting from $\mathbb{C}$-holomorphic functions, let's give some preliminary definitions:

Definition 6.9 (Inclusion of Complex Numbers in Quaternions)**.** $\forall q \in \mathbb{H}$ *we will call $\eta_q : \mathbb{C} \to \mathbb{H}$ the inclusion of complex numbers in quaternions such that q is the image of a complex number $\zeta(q)$ in the upper half of the complex plane; in symbols:*

$$\eta_q(x+iy) = x + \mathrm{sgn}(\mathrm{Vec}(q))y \tag{6.56}$$

$$\zeta(q) = \mathrm{Sc}(q) + i|\mathrm{Vec}(q)| \tag{6.57}$$

Now we demonstrate the result that will allow us to construct regular functions from holomorphic functions:

Theorem 6.12. *Let $f : \mathbb{C} \to \mathbb{C}$, $f = u(x,y) + iv(x,y)$ be a holomorphic function in a domain $U \subset \mathbb{C}$ and define $\tilde{f}(q)$ to be:*

$$\tilde{f}(q) = \eta_q \circ f \circ \zeta(q) \tag{6.58}$$

Then the Laplacian of $\tilde{f}$, $\triangle\tilde{f}$ is a regular function in the set $\zeta^{-1}(U) \subset \mathbb{H}$.

Proof. Let $f(z) = u(x, y)+iv(x, y)$ be our holomorphic function of a complex variable $z = x+iy$ in U. To simplify the notation of (6.56) and (6.57), we will call $|\operatorname{Vec}(q)| = r$, $\operatorname{Sc}(q) = t$ and $\operatorname{sgn}(\operatorname{Vec}(q)) = \vec{r}$. The quaternionic function $\tilde{f}(q)$ induced by $f(z)$ will then be equal to:

$$\tilde{f}(q) = u(t, r) + \vec{r}v(t, r)$$

The function just induced will not generally be a regular function; let us convince ourselves of this by calculating its Laplacian:

$$\triangle\tilde{f} = \triangle u + \triangle(\vec{r}v) = \frac{\partial^2 u}{\partial^2 t} + \frac{2}{r}\frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial^2 r} + 2\Big(\frac{1}{r}\frac{\partial v}{\partial r} - \frac{1}{r^2}v\Big)\vec{r} + \frac{\partial^2 v}{\partial t^2}\vec{r} + \frac{\partial^2 v}{\partial r^2}\vec{r}$$

Being the function f holomorphic, as an immediate consequence of the Cauchy-Riemann equations we have:

$$\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 u}{\partial r^2} = \frac{\partial^2 v}{\partial t^2} + \frac{\partial^2 v}{\partial r^2} = 0$$

It follows from this that we can reduce the Laplacian of $\tilde{f}$ to simply:

$$\triangle\tilde{f} = \frac{2}{r}\frac{\partial u}{\partial r} + 2\Big(\frac{1}{r}\frac{\partial v}{\partial r} - \frac{1}{r^2}v\Big)\vec{r} \qquad (6.59)$$

Exploiting again the assumption that f is holomorphic, we have that from the Cauchy-Riemann equations $\frac{\partial u}{\partial r} = -\frac{\partial v}{\partial t}$, and substituting this in (6.59) we obtain:

$$\triangle\tilde{f} = 2\Big[-\frac{\partial}{\partial t}\Big(\frac{v}{r}\Big) + \vec{r}\frac{\partial}{\partial r}\Big(\frac{v}{r}\Big)\Big] \qquad (6.60)$$

The expression just obtained is nothing but the negation of the operator ∂_l applied to $\frac{v}{r}$, all multiplied by 4.
Remembering what was said in the section where we faced a vectorial formulation of the Cauchy-Riemann-Fueter equations, in this case it will be convenient to write the equation (6.60) using the formal quaternions defined as $\square = \frac{\partial}{\partial t} + \vec{\nabla}$ and $\bar{\square} = \frac{\partial}{\partial t} - \vec{\nabla}$. In this case we will have:

$$\triangle\tilde{f} = -2\bar{\square}\Big(\frac{v}{r}\Big)$$

But the Laplacian, as seen in section 3, is equal to $\square\bar{\square} = \bar{\square}\square$, and thus we have $-2\bar{\square}\Big(\frac{v}{r}\Big) = \square\bar{\square}\tilde{f}$ from which:

$$\square\tilde{f} = -2\frac{v}{r} \qquad (6.61)$$

and

$$\overline{\square}\tilde{f} = 2\frac{v}{r} \tag{6.62}$$

We have just seen that, in general, functions of the type $\tilde{f}$ generated by holomorphic functions $f = u(x,y) + iv(x,y)$ are not regular: however, observing that $\triangle\left(\frac{v}{r}\right) = \frac{1}{r}\left(\frac{\partial^2 u}{\partial t^2} + \frac{\partial^2 v}{\partial r^2}\right) = 0$, and applying the Laplacian operator to both sides of equation (6.61) we will have that $\square(\triangle\tilde{f}) = 0 \Longleftrightarrow \bar{\partial}_l(\triangle\tilde{f}) = 0$, i.e., the function $\triangle\tilde{f}$ is regular $\forall q \in \zeta^{-1}(U)$.

□

This result is of great importance for a number of reasons. Before discussing them, however, let's introduce this immediate corollary.

Corollary 6.6. *Quaternionic functions induced by holomorphic functions satisfy the biharmonic equation, in symbols:*

$$\triangle\triangle\tilde{f} = 0$$

Returning to the earlier point; the reader may remember as we already mentioned in the section on the Cauchy-Riemann-Fueter equation that some important functions, like $f(q) = q$ and $f(q) = q^2$ are generally not regular (one is not at any point, the other only in a subregion of $\mathbb{H}$). As a corollary, generally functions expressible as convergent quaternionic power series of the type:

$$\phi(q) = \sum_{n=0}^{\infty} a_n(q - q_0)^n$$

(where a_n is a quaternionic constant) are not regular. However, it is possible to show that series of the type

$$\tilde{\phi}(q) = \sum_{n=0}^{\infty} a_n \triangle (q - q_0)^n$$

are regular. And that is precisely what we are going to do now:

Proposition 6.10. *Let $\phi : \mathbb{H} \to \mathbb{H}$ be a quaternionic function expressible as a series of Δ-powers:*

$$\phi(q) = \sum_{n=0}^{\infty} a_n \triangle (q - q_0)^n$$

then $\phi(q)$ is regular.

Proof. The strategy is as follows: we will first show that functions of the type $(q-q_0)^n$ are induced by the holomorphic functions $(z-z_0)^n$ in $\mathbb{C}$, and then as a corollary of theorem 6.12 we will have that the Laplacians of these functions will be regular functions. Finally, exploiting the fact that quaternionic regular functions are closed with respect to addition, it follows that these series are regular functions.
Without loss of generality, let's assume that the series in question is centered at the origin, i.e. $q_0 = 0$. We now prove by induction that functions of the type $f(z) = z^n$ induce quaternionic monomials of the type $\tilde{f} = q^n$.
For the base case, we select $n = 1$, i.e. $f(z) = z = x + iy$. Then the induced function will be equal to $\tilde{f}(q) = t + \frac{xi + yj + zk}{|xi + yj + zk|}(|xi+yj+zk|) = t+xi+yj+zk = q$.
Suppose that this relation is also valid for some $n = k$, we now want to show that it is valid for $n = k+1$. The function $f(z) = z^{k+1}$ can be written as $f(z) = zz^k$; decompose z^k into its imaginary and real parts, which we will call $z^k = \phi_1(z) + i\phi_2(z)$. Then we can write $f(z)$ in terms of its real and imaginary parts as:

$$f(z) = x\phi_1 - y\phi_2 + i(\phi_1 y + x\phi_2)$$

Its induced function, therefore, will be equal to $\tilde{f}(q) = t\phi_1(t,r) - r\phi_2(t,r) + \vec{r}\Big(r\phi_1(t,r) + t\phi_2(t,r)\Big) = t\phi_1(t,r) - r\phi_2(t,r) + \vec{r}r\phi_1(t,r) + \vec{r}t\phi_2(t,r)$.
However, remembering that for the inductive step we assumed that q^k was equal to $\phi_1(t,r) + \vec{r}\phi_2(t,r)$ (the function induced by z^k) and writing $q = t + xi + yj + zk$ as $q = t + r\vec{r}$ (since $r = |xi+yj+zk|$ and $\vec{r} = sgn(Vec(q)) = \frac{xi + yj + zk}{|xi + yj + zk|}$) we will have:

$$q^{k+1} = (t + r\vec{r})(\phi_1(t,r) + \vec{r}\phi_2(t,r)) = t\phi_1 + t\vec{r}\phi_2 + r\vec{r}\phi_1 - r\phi_2 = \tilde{f}(q)$$

thus we have proved by induction that holomorphic functions of the type $f(z) = z^n$ induce quaternionic monomials $\tilde{f}(q) = q^n$; it follows from this that their Laplacians are regular and the sums of the aforementioned also. □

We will call power series of Laplacians of quaternionic monomials of the type $(q-q_0)^n$ Δ-power series, and we will say that a function $\phi : \mathbb{H} \to \mathbb{H}$ expressible as a Δ-power series is Δ-analytic.
We observe that for the theorem just demonstrated, we will have that all Δ-analytic functions are also regular functions.
Finally, we will call weighted sums of Laplacians of monomials of the type $(q-q_0)$ "Δ-polynomials", $\eta(q) = \sum_{m=0}^{n} a_m \triangle (q-q_0)^m$.

Lemma 6.3. *Let $\tilde{f}$ be a quaternionic function induced by a holomorphic function of $\mathbb{C}$, f. Then:*

$$\bar{\Box}\tilde{f} = 2\Big(\frac{\partial \tilde{f}}{\partial t} + \frac{v}{r}\Big) \tag{6.63}$$

Exercise 6.9. *Prove lemma 6.3.*

We now prove a result that will allow us to more easily compute the Laplacians of functions induced by holomorphic functions:

Proposition 6.11. *Let $f : \mathbb{C} \to \mathbb{C}$ be a holomorphic function; let $\tilde{f} : \mathbb{H} \to \mathbb{H}$ be its induced function, then we will have:*

$$\triangle \tilde{f} = \frac{2}{r}\vec{r}\Big(\frac{\partial \tilde{f}}{\partial t} - \frac{v}{r}\Big)$$

Proof. We derived in the proof of theorem 6.12 the equation (6.61), $\Box\tilde{f} = -2\frac{v}{r}$. Applying the conjugate gradient operator $\bar{\Box}$ to both sides of this identity we get:

$$\bar{\Box}\Box\tilde{f} = \triangle\tilde{f} = -2\bar{\Box}\Big(\frac{v}{r}\Big) = -\frac{2}{r}\Big(\bar{\Box}v + \frac{v}{r}\vec{r}\Big)$$

But v is the scalar part of the induced function $\tilde{f}$, so we will have:

$$v = \frac{1}{2}(\vec{r}\tilde{f}^{*} - \vec{r}\tilde{f})$$

Applying the conjugate gradient operator $\bar{\Box}$ to both sides of the equation we get:

$$\bar{\Box}v = \frac{1}{2}(\bar{\Box}(\vec{r}\tilde{f}^{*}) - \bar{\Box}(\vec{r}\tilde{f}))$$

It is quickly shown that the functions $\vec{r}\tilde{f}^{*}$ and $\vec{r}\tilde{f}$ are the functions induced by the complex functions $i\bar{f}$ and if. Then, using equations (6.62) and (6.61), we will have:

$$\bar{\Box}v = \frac{1}{2}\Big(2\frac{u}{r}\Big) - \frac{1}{2}\Big(2\vec{r}\frac{\partial \tilde{f}}{\partial t} + \frac{u}{r}\Big) = -\vec{r}\frac{\partial \tilde{f}}{\partial t}$$

From which, substituting $\bar{\Box}v$ in the initial expression for $\triangle\tilde{f}$, we get the assertion. $\Box$

The identity just demonstrated, applied in the particular case of Laplacians of Δ-polynomials, allows us to derive the following equations:

$$\begin{cases} \triangle(q^n) = -4((n-1)q^{n-2} + (n-2)q^{n-3}\bar{q} + (n-3)q^{n-4}\bar{q}^2 + \dots) \\ \triangle(q^{-n}) = -4(nq^{-n-1}\bar{q}^{-1} + (n-1)q^{-n}\bar{q}^{-2} + \dots + q^{-2}\bar{q}^{-n}) \end{cases} \tag{6.64}$$

Exercise 6.10. *Let $\tilde{f}$ be the quaternionic function induced by the holomorphic function $f : \mathbb{C} \to \mathbb{C}$. Prove that the left derivative of $\triangle\tilde{f}$ is equal to the Laplacian of the function induced by the complex derivative of f, in symbols:*

$$(\triangle\tilde{f})'_l = \triangle\tilde{f'} \tag{6.65}$$

The following example is quite interesting and reinforces the intuition already built previously on the relationship between $\mathscr{L}(q), G(q)$ and $\log(z), \frac{1}{z}$. In fact, for $f(z) = \log(z)$ we will have $\triangle\tilde{f} = -4\mathscr{L}(q)$, while for $g(z) = \frac{1}{z}$ we will have $\triangle\tilde{g} = -4G(q)$; in other words, the function $\mathscr{L}$ primitive of $G(q)$ (the function that appears in the Cauchy-Fueter integral formula) is induced by the complex natural logarithm $\log(z)$, primitive of the function z^{-1} that appears in the complex Cauchy integral formula.

Exercise 6.11. *Calculate the induced quaternionic function $\tilde{f}$ for the following holomorphic functions in $\mathbb{C}$:*

- $f(z) = \sin(z)$
- $f(z) = e^z$
- $f(z) = z^4 - \cos(z)\sin(z)$

For each of these, calculate the Laplacian of the induced function to obtain a regular function, and calculate for each of these its left derivative.

6.10 Other Methods for Constructing Regular Functions

In this section, we introduce some additional methods for constructing regular functions. In particular, in the following proposition, we will see how to construct regular functions at a point $q \in \mathbb{H} \setminus \{0\}$ given a regular function at its multiplicative inverse q^{-1}.

Proposition 6.12. *Let $f : \mathbb{H} \to \mathbb{H}$ be a quaternionic function. Define $If : \mathbb{H} \setminus \{0\} \to \mathbb{H}$ as:*

$$If(q) = \frac{q^{-1}}{|q|^2} f(q^{-1})$$

Then, if f is regular at q^{-1}, If is regular at q.

Proof. It suffices to prove that $Dq \wedge d(If(q)) = 0$. Our function $If(q)$ is given by the product of the function $G(q) = \frac{q^{-1}}{||q||^2}$, defined earlier, with the function evaluated at the multiplicative inverse of $q \in \mathbb{H}$, $q \neq 0$. Let $\mathfrak{I} : \mathbb{H} \setminus \{0\} \to \mathbb{H}$ be the inversion map defined as $\mathfrak{I}(q) = q^{-1}$, then we can write If as:

$$If(q) = G(q)(f \circ \mathfrak{I})$$

First, we calculate the exterior differential of the latter: applying the product rule for exterior differentiation, we get:

$$d(If) = dG_q f(q^{-1}) + G(q)d(f \circ \mathfrak{I})$$

but we see that $d(f \circ \mathfrak{I})$, by a well-known result in differential geometry (see introduction), is simply the pullback of the differential of f with respect to $\mathfrak{I}$ at q^{-1}, $d(f \circ \mathfrak{I}) = \mathfrak{I}^* df_{q^{-1}}$.
From the considerations just made, it follows that we can write $Dq \wedge d(If)$ as:

$$Dq \wedge d(If) = Dq \wedge dG_q f(q^{-1}) + Dq \wedge G(q)\mathfrak{I}^* df_{q^{-1}} = Dq \wedge G(q)\mathfrak{I}^* df_{q^{-1}}$$

since G is a closed form and therefore $dG_q = 0$. We now observe that, by the properties of the 3-form Dq demonstrated in section 2 of this chapter, we have:

$$\mathfrak{I}^* Dq(h_1, h_2, h_3) = Dq(-q^{-1}h_1q^{-1}, -q^{-1}h_2q^{-1}, -q^{-1}h_3q^{-1}) = -\frac{q^{-1}}{|q|^4} Dq(h_1, h_2, h_3)q^{-1}$$

From which it follows that $DqG(q) = -|q|^2 q\mathfrak{I}^* Dq$. But pullbacks maintain the exterior product, and inserting this identity just found into our equation above, we obtain:

$$Dq \wedge d(If) = -|q|^2 q\mathfrak{I}^*(Dq \wedge df_{q^{-1}})$$

We observe that if f is regular at q^{-1}, then by lemma 6.2 we have that $Dq \wedge df_{q^{-1}} = 0$, and thus $Dq \wedge d(If) = 0$, i.e., If is regular at q. □

Proposition 6.13. *Let $f : \mathbb{H} \to \mathbb{H}$ be a function of a quaternionic variable and let $a, b \in \mathbb{H}$ be two quaternions. Define $M(a, b)f(q)$ as:*

$$M(a, b)f(q) := bf(aqb)$$

Then, if f is regular at aqb, $M(a, b)f$ is regular at q.

Proof. Let $\mu : \mathbb{H} \to \mathbb{H}$ be defined as $\mu(q) := aqb$ where a, b are two constant quaternions. Then the pullback of Dq with respect to μ is:

$$\mu^* Dq = Dq(ah_1b, ah_2b, ah_3b) = |a|^2|b|^2 aDq(h_1, h_2, h_3)b$$

From which we obtain an expression for $Dq(h_1, h_2, h_3)$ in terms of its pullback with respect to μ, which will be useful to us later:

$$Dq(h_1, h_2, h_3) = |a|^{-2}|b|^{-2}a^{-1}\mu^* Dq(h_1, h_2, h_3)b^{-1}$$

To now prove that $M(a, b)f$ is regular at q if f is regular at $\mu(q)$, we only need to compute the exterior product of Dq with the exterior differential of $M(a, b)f$, and verify that it is equal to 0.
We see that $d(M(a, b)f)_q = bd(f \circ \mu)_q = b\mu_q^* df_{\mu(q)}$ (by applying the same result used earlier), which implies that:

$$Dq \wedge d(M(a, b)f)_q = |a|^{-2}|b|^{-2}a^{-1}\mu_q^* Dq(h_1, h_2, h_3)b^{-1} \wedge b\mu_q^* df_{\mu(q)}$$

Now, using the properties of the exterior product $\wedge$ and remembering that pullbacks preserve exterior products, we have:

$$Dq \wedge d(M(a, b)f)_q = |a|^{-2}|b|^{-2}a^{-1}\mu^*(Dq \wedge df_{\mu(q)}) = 0 \tag{6.66}$$

if f is regular at $\mu(q) = aqb$, as $Dq \wedge df_{\mu(q)} = 0$

□

Finally, it is possible to construct regular functions also starting from quaternionic homographies. First, we prove the following result, which will provide us with the general algebraic form of a quaternionic homography.

Proposition 6.14. *Let $\mathbb{H}^* = \mathbb{H} \cup \{\infty\} \cong S^4$ be the Alexandroff compactification of the quaternions, seen in chapter 4. Then a function $f : \mathbb{H}^* \to \mathbb{H}^*$ is a homography (i.e., a conformal function that preserves orientation) if and only if f is of the form:*

$$f(q) = (aq + b)(cq + d)^{-1} \tag{6.67}$$

for $a, b, c, d \in \mathbb{H}$.

Proof. We start by proving the direction $\Longleftarrow$: suppose that f is of the form $f(q) = (aq + b)(cq + d)^{-1}$. Then, f will have a differential of:

$$df_q = (ac^{-1}d - b)(cq + d)^{-1}cdq(cq + d)^{-1}$$

we observe that the differential is of the form $\alpha dq\beta$, where $\alpha, \beta \in \mathbb{H}$, which is a combination of rotations and dilations, and therefore f is a conformal application that preserves orientation.
Now suppose instead that f is a conformal application that preserves orientation; the set of such applications is generated by applications of the type $q \to \alpha q\beta$, $q \to q + \gamma$ and $q \to q^{-1}$. We observe that composing a function of the type (6.67) with one of these applications returns another function of the type (6.67), from which we deduce that, calling D the set of functions of the type (6.67) and C the set of conformal functions that preserve orientation: $CD \subset D$, from which $C \subset D$ and thus, united with the previous point $C = D$, i.e., a function from the compactification of the quaternions to itself is a conformal application that preserves orientation if and only if it is of the type (6.67). □

Once a form for quaternionic conformal functions has been obtained, we go on to prove the result already mentioned earlier:

Proposition 6.15. *Let $f : \mathbb{H} \to \mathbb{H}$ be a function of a quaternionic variable and let $\nu(q) = (aq + b)(cq + d)^{-1}$ be a conformal application. Define $M(\nu)f$ as:*

$$[M(\nu)f](q) = \frac{1}{|b - ac^{-1}d|^2} \frac{(cq + d)^{-1}}{|cq + d|^2} f(\nu(q))$$

Then, if f is regular at $\nu(q)$, $M(\nu)f(q)$ is regular at q.

Proof. The proof of this theorem follows as an immediate corollary of the other two results proved in this section once the following fact is shown: Given the functions:

$$\begin{cases} \nu_1(q) = cq(b - ac^{-1}d)^{-1} \\ \nu_2(q) = q + d(b - ac^{-1}d)^{-1} \\ \nu_3(q) = q^{-1} \\ \nu_4(q) = q + ac^{-1} \end{cases}$$

We have that $\nu(q) = \nu_4 \circ \nu_3 \circ \nu_2 \circ \nu_1$, i.e., the function ν is given by the composition of the above-mentioned ones. Now, we observe that the function $\nu_1(q)$ preserves regularity by proposition 6.13, as does the function ν_3 by proposition 6.12. Observing finally that even translations preserve regularity, we have that $M(\nu)f$ is regular at q if f is regular at $\nu(q)$. □

6.11 Homogeneous Quaternionic Functions

In this section, we will introduce some facts about regular and homogeneous harmonic functions of a quaternionic variable, preparing to construct a quaternionic analogue of Taylor and Laurent series expansions.
First, let's give the definition of a homogeneous quaternionic function of degree n (with respect to $\mathbb{R}$).

Definition 6.10 (Quaternionic $\mathbb{R}$-homogeneous function of degree n)**.** *We will say that a function of a quaternionic variable $f : \mathbb{H} \setminus \{0\} \to \mathbb{H}$ is regular in $\mathbb{R}$ of degree n, if $\forall \alpha \in \mathbb{R}$:*

$$f(\alpha q) = \alpha^n f(q)$$

(we have removed the origin from the domain to also consider negative n).

At this point, we introduce some naming conventions that we will use from now on: we call U_n the set of quaternionic functions $f : \mathbb{H} \setminus \{0\}$ regular homogeneous of degree n, while we call W_n the set of quaternionic functions $f : \mathbb{H} \setminus \{0\}$ harmonic and homogeneous of degree n.
For corollary 6.5, we know that every regular function is harmonic, and therefore $U_n \subset W_n$.
Furthermore, we can observe that the sets just constructed form, with the operations of pointwise addition and multiplication by a quaternionic scalar, pseudo-vector spaces on the right $\mathbb{H}$ (i.e., modules on a skew-field, which possess some extra properties compared to modules and therefore we will call them pseudo-vector spaces).
We can consider the restrictions of U_n and W_n to the 3-sphere $S^3 = \{q \in \mathbb{H} \,;\, |q| = 1\}$, which we will indicate with $\tilde{U}_n$ and $\tilde{W}_n$:

$$\tilde{U}_n := \{f|_{S^3} \,;\, f \in U_n\} \quad ; \quad \tilde{W}_n := \{f|_{S^3} \,;\, f \in W_n\}$$

We observe that:

$$\tilde{U}_n \cong U_n \quad ; \quad \tilde{W}_n \cong W_n$$

i.e U_n and W_n are isomorphic to their restriction on the 3-sphere S^3, $\tilde{U}_n$, $\tilde{W}_n$, under the following correspondence:

$$f \in U_n \Leftrightarrow \tilde{f} \in \tilde{U}_n \quad with \quad f(q) = r^n \tilde{f}(u)$$

$$f \in W_n \Leftrightarrow \tilde{f} \in \tilde{W}_n \quad with \quad f(q) = r^n \tilde{f}(u)$$

where here $r = |q| \in \mathbb{R}$ and $u = \operatorname{sgn}(q) = \dfrac{q}{|q|} \in S^3$.
Now we define the following quaternionic vector fields in such a way as to be able to express the Cauchy-Riemann-Fueter equation in "polar" form:

$$X_0 f = \frac{d}{d\theta}[f(qe^{\theta})]_{\theta=0} \tag{6.68}$$

$$X_i f = \frac{d}{d\theta} f[qe^{e_i\theta}]_{\theta=0} = \frac{d}{d\theta} f[q(\cos\theta + e_i \sin\theta)]_{\theta=0} \quad (i = 1, 2, 3) \tag{6.69}$$

Their relation to the Cartesian vector fields $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial x_i}$ is given by:

$$\begin{cases} X_0 = t\dfrac{\partial}{\partial t} + x_i\dfrac{\partial}{\partial x_i} \\ X_i = -x_i\dfrac{\partial}{\partial t} + t\dfrac{\partial}{\partial x_i} - \epsilon_{ijk}x_j\dfrac{\partial}{\partial x_k} \\ \dfrac{\partial}{\partial t} = \dfrac{1}{r^2}(tX_0 - x_iX_i) \\ \dfrac{\partial}{\partial x_i} = \dfrac{1}{r^2}(\epsilon_{ijk}x_jX_k + tX_i + x_iX_0) \end{cases} \tag{6.70}$$

where here as usual ϵ_{ijk} indicates the Levi-Civita symbol, and where moreover we are omitting the summations on mute indices using the Einstein convention.
Using the last 2 relations of equation (6.70), we can write the Laplacian operator $\triangle$ and the operator $\bar{\partial}_l$ in terms of the vector fields just introduced in the following way:

$$\begin{cases} \bar{\partial}_l = \dfrac{1}{2}\bar{q}^{-1}(X_0 + e_iX_i) \\ \triangle = \dfrac{1}{r^2}(X_iX_i + X_0(X_0 + 2)) \end{cases} \tag{6.71}$$

With that said, we are now ready to introduce some important results on the spaces U_n and W_n just introduced.

Theorem 6.13 (Properties of the $\mathbb{H}$-pseudo vector space of homogeneous harmonic functions of degree n)**.** *Let W_n be the $\mathbb{H}$-pseudo vector space of homogeneous harmonic functions of degree n introduced earlier and let $\tilde{W}_n$ be its reduction to the 3-sphere S^3. Then it has the following properties:*

- $\tilde{W}_n \cong \tilde{W}_{-n-2}$.
- $\dim(\tilde{W}_n) = (n+1)^2$.

- *The elements of W_n are polynomials in q.*

Proof. We prove the facts point by point, starting with the first:

- We start the proof from the following observation: the functions $f \in W_n$, being homogeneous, are eigenfunctions of the vector field X_0 with eigenvalue n; to convince oneself of this, it is enough to note that, being homogeneous of degree n, the following identity will be valid:

$$f(qe^{\theta}) = e^{n\theta} f(q)$$

from which then, computing $X_0 f$ with the formulas presented before, we will get $X_0 f = nf$, i.e f is an eigenfunction of X_0 with eigenvalue n.
However, we note that our function f, by hypothesis (being in W_n) is also harmonic, and therefore satisfies the Laplace equation:

$$\triangle f = \frac{1}{r^2}(X_i X_i + X_0(X_0 + 2))f = 0$$

From which, knowing that $X_0 f = nf$:

$$\triangle f = 0 = X_i X_i f + n^2 f + 2nf \implies X_i X_i f = -n(n+2)f$$

i.e f is also an eigenfunction of $X_i X_i$ with eigenvalue $-n(n+2)$.
We now observe that the vector fields X_i, $i = 1, 2, 3$ are tangent to the 3-sphere S^3, and therefore their restrictions on it, $\tilde{X}_i = X_i|_{S^3}$, are vector fields on S^3.
Therefore, if $\tilde{f} \in \tilde{W}_n$, then $\tilde{f} = f|_{S^3}$ for an $f \in W_n$, and therefore being f an eigenfunction of $X_i X_i$ with eigenvalue $-n(n+2)$, $\tilde{f}$ will be an eigenfunction of $\tilde{X}_i \tilde{X}_i$ with eigenvalue $-n(n+2)$. On the contrary, assuming that a function $\tilde{f} \in \tilde{W}_n$ is an eigenfunction of $\tilde{X}_i \tilde{X}_i$ with eigenvalue $-n(n+2)$, then:

$$\triangle[r^n \tilde{f}(u)] = \frac{1}{r^2}[r^n \tilde{X}_i \tilde{X}_i \tilde{f} + (X_0(X_0 + 2)r^n)\tilde{f}] = 0$$

where $r = |q|$ and $u = \mathrm{sgn}(q)$ as usual.
And therefore, $\tilde{W}_n$ is alternatively characterizable as the space of eigenfunctions of $\tilde{X}_i \tilde{X}_i$ with eigenvalue $-n(n+2)$. From this follows the assertion:

$$\tilde{W}_n \cong \tilde{W}_{-n-2}$$

- The proof of this fact, requiring the use of results that to be adequately developed would require a significant amount of time, will be only hinted at and some texts in the bibliography will be mentioned for a more accurate proof.
  First, we note that the space $\tilde{W}_n$ is the quaternionification of the complex vector space $W_n^{\mathbb{C}}$ of complex eigenfunctions on S of $\tilde{X}_i\tilde{X}_i$ with eigenvalue $-n(n+2)$. But such space is $(n+1)^2$ dimensional, and therefore:
  $$\dim_{\mathbb{H}} \tilde{W}_n = \dim_{\mathbb{C}} W_n^{\mathbb{C}} = (n+1)^2$$
  For an accurate proof of what has been said before, see [44] (p.71) in the bibliography.

- From the definition of W_n and homogeneous function given above, we will have that W_0 will consist only of constant functions.
  It is now enough to note that, given a function $f \in W_n$, its partial derivatives $\frac{\partial f}{\partial t}$ and $\frac{\partial f}{\partial x_i}$ $(i = 1, 2, 3)$ are members of W_{n-1}; to convince oneself of this, it is enough to note that:
  $$\alpha^n \frac{\partial f(q)}{\partial t} = \frac{\partial}{\partial t}(f \circ \rho_q) = \alpha \frac{\partial f(\alpha q)}{\partial t}$$
  where $\alpha \in \mathbb{R}$ and $\rho_q := \alpha q$.
  Repeating this reasoning, we come to the conclusion that the n-th derivative of $f \in W_n$ will be a member of W_0 (and therefore a constant function) and its (n+1)-th derivative will be equal to 0, from which follows the result we wanted to demonstrate.

□

Exercise 6.12. *Show that:*

$$[X_0, X_i] = 0 \quad ; \quad [X_i, X_j] = 2\epsilon_{ijk}X_k$$

where here X_0, X_1, X_2, X_3 are the vector fields introduced in equations (6.68) *and* (6.69)*, and where* [] *indicates the Lie product.*

We conclude the chapter with other results, this time on the space of regular homogeneous functions U_n.

Theorem 6.14 (Properties of the $\mathbb{H}$-pseudo vector space of regular homogeneous functions of degree n)**.** *Let U_n be the $\mathbb{H}$-pseudo vector space of regular homogeneous functions of degree n introduced earlier and W_n, as before, the space of homogeneous harmonic functions of degree n; then:*

- $\tilde{W}_n \cong \tilde{U}_n \oplus \tilde{U}_{-n-2}$.
- $\tilde{U}_n \cong \tilde{U}_{-n-3}$.
- $\dim U_n = \frac{1}{2}(n+1)(n+2)$.

where here, as before, $\tilde{U}_n$ and $\tilde{W}_n$ represent the restrictions of such spaces to the 3-sphere S^3.

Proof.
- Let f be a regular and homogeneous function of degree n, i.e., $f \in U_n$. As mentioned earlier, f will be an eigenfunction of X_0 with eigenvalue n, and furthermore, f will satisfy the Cauchy-Riemann-Fueter equation, $\bar{\partial}_l f = 0$. From this it follows that:

$$\bar{\partial}_l f = 0 = X_0 f + e_i X_i f \Longrightarrow e_i X_i f = -nf$$

that is, f is an eigenfunction of $e_i X_i$ with eigenvalue $-n$. Similarly, if f is an eigenfunction of $e_i X_i$ with eigenvalue $-n$, $f \in U_n$. Therefore, U_n can be alternatively characterized as the space of eigenfunctions of $e_i X_i$ with eigenvalue $-n$.
Through direct calculations, and using the results of exercise 6.12, we see that:

$$(e_i X_i)^2 - 2e_i X_i + X_i X_i = 0 \Longrightarrow (e_i \tilde{X}_i)^2 - 2e_i \tilde{X}_i + \tilde{X}_i \tilde{X}_i = 0$$

Let $\tilde{f} \in \tilde{W}_n$: then we know, in light of what has been said before, that $\tilde{X}_i \tilde{X}_i f = -n(n+2)f$, i.e., f is an eigenfunction of $\tilde{X}_i \tilde{X}_i$ with eigenvalue $-n(n+2)$. From this it follows that $\tilde{W}_n$ is the direct sum of the eigenspaces of $e_i X_i$ with eigenvalues n and $-n-2$ respectively, i.e.:

$$\tilde{W}_n = \tilde{U}_n \oplus \tilde{U}_{-n-2}$$

- This point can be proved by explicitly constructing an isomorphism between the two spaces: in particular, we note that the isomorphism in question is given by the application If introduced in proposition 6.12.

- From the first point of this theorem, for a well-known fact about the dimensions of direct sum spaces, we will have:

$$\dim W_n = \dim U_n + \dim U_{-n-2}$$

but, for proposition 6.13, $\dim W_n = (n+1)^2$, and therefore:

$$(n+1)^2 = \dim U_n + \dim U_{-n-2} \tag{6.72}$$

now for the second point of this theorem, we will have that the dimension of U_{-n-2} will be equal to the dimension of U_{n-1}, and substituting in equation (6.72), we will have the following recursive relationship:

$$\dim U_n + \dim U_{n-1} = (n+1)^2$$

whose solution, with initial value $\dim U_0 = 1$, is equal to

$$\dim U_n = \frac{1}{2}(n+1)(n+2)$$

we have thus proved what was stated in the theorem.

□

Having said that, we are now ready to extend the concepts of Taylor and Laurent series expansions to regular functions of a quaternionic variable, with which the reader is probably already familiar if they have taken an introductory course in complex analysis.

6.12 Quaternionic Taylor and Laurent Series

At this point, the reader, given the large number of analogies between the theory of $\mathbb{C}$-holomorphic functions and that of $\mathbb{H}$-regular functions, may wonder whether regular quaternionic functions in balls can be expanded into a quaternionic analogue of Taylor series, and similarly whether regular analytic functions in annular regions can be expanded into Laurent series. We now have all the machinery necessary to extend Taylor and Laurent series expansions to $\mathbb{H}$-regular functions (in certain domains); we only need to start by introducing some preliminary notions.
Let $\{i_1, i_2, \ldots, i_n\}$ be a set of non-ordered integers with repetitions such that $1 \le i_j \le 3\ \forall j \in [1, n] \cap \mathbb{N}$. We will call such a set ν: note that this set can also be specified by the number of instances of 1, 2, and 3, which we will call $n_1, n_2, n_3 \in \mathbb{N}$; thus we write $\nu = [n_1; n_2; n_3]$.
The total number of possibilities for a set ν of this type with a total length of $n = n_1 + n_2 + n_3$ is $\frac{1}{2}(n+1)(n+2)$; moreover, we will call σ_n the set of all sets of type ν of length n.

Exercise 6.13. *The reader writes all possible sets $\nu = \{i_1, i_2, \ldots, i_n\}$ for the following values of n:*

- $n = 2$
- $n = 3$

We now introduce some notation and definitions that will be useful in the paragraph:

Definition 6.11 (Mixed nth-order differential operator). *We call the mixed nth-order differential operator, indicated by ∂_ν, the following differential operator:*

$$\partial_\nu = \frac{\partial^n}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_n}} = \frac{\partial^n}{\partial x^{n_1} \partial y^{n_2} \partial z^{n_3}} \tag{6.73}$$

where, in accordance with the nomenclature introduced recently, $\nu = [n_1; n_2; n_3]$, $n = n_1 + n_2 + n_3$.

In addition, we will frequently use the following conventions:

$$G_\nu(q) = \partial_\nu G(q) \tag{6.74}$$

$$P_\nu(q) = \frac{1}{n!} \sum (te_{i_1} - x_{i_1}) \dots (te_{i_n} - x_{i_n}) \tag{6.75}$$

where in the second expression the sum is over all possible rearrangements of $1, 2$, and 3 equal to $\dfrac{n!}{n_1! n_2! n_3!}$.
We also observe that $G_\nu(q)$ will be a homogeneous function of degree $-n-3$.

Exercise 6.14. *Prove that given a $\nu = [n_1; n_2; n_3]$, $n_1 + n_2 + n_3 = n$, then the function $P_\nu(q)$ is a homogeneous function on $\mathbb{R}$ of degree n, i.e:*

$$P_\nu(\alpha q) = \alpha^n P_\nu(q) \quad \alpha \in \mathbb{R}$$

Theorem 6.15. *The polynomials P_ν $(\nu \in \sigma_n)$ are regular and form a basis for the $\mathbb{H}$ pseudo-vector space U_n.*

Proof: Let $f \in U_n$, i.e., let f be a regular homogeneous function of degree n. Since f is regular, it satisfies the Cauchy-Riemann-Fueter equation:

$$\bar{\partial}_l f = \frac{\partial f}{\partial t} + e_i \frac{\partial f}{\partial x_i} = 0 \tag{6.76}$$

and since it is homogeneous, it is an eigenfunction of X_0 with eigenvalue n, i.e:

$$X_0 f = t \frac{\partial f}{\partial t} + x_i \frac{\partial f}{\partial x_i} = n f(q) \tag{6.77}$$

Now, multiplying (6.76) by t and subtracting it from (6.77), we obtain:

$$n f(q) = \sum_i (x_i - te_i) \frac{\partial f}{\partial x_i}$$

Remember that, from considerations made in the previous section, we know that the function f in question is a quaternionic polynomial, and also that $\frac{\partial f}{\partial x_i} \in U_{n-1}$ (theorem 6.13). Thus, we can reapply the same reasoning to this latter n times obtaining:

$$f(q) = \frac{1}{n!}\sum_{i_1 \dots i_n}(x_{i_1} - te_{i_1})(x_{i_2} - te_{i_2})\dots(x_{i_n} - te_{i_n})\frac{\partial^n f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_n}}$$

$$\Longrightarrow f(q) = \sum_{\nu\in\sigma_n}(-1)^n P_\nu(q)\partial_\nu f(q)$$

Since f is a polynomial as already mentioned (third point of theorem 6.13), we have that the derivative $\partial_\nu f(q)$ will be a quaternionic constant, and we have concluded that the set $\{P_\nu \ ; \ \nu \in \sigma_n\}$ is a linear cover of the space U_n, i.e., every element of U_n can be written as a linear combination of elements of $\{P_\nu \ ; \ \nu \in \sigma_n\}$.
Define the right $\mathbb{H}$-pseudo-vector space as $Y_n := \operatorname{span}^r_{\mathbb{H}}\{P_\nu \ ; \ \nu \in \sigma_n\}$ (where here the r denotes the right span). Note that, since the elements of $U_n \subset W_n$ are polynomials (third point of theorem 6.13), then $U_n \subset Y_n$. But, at the same time, we know that $\dim Y_n \leq \frac{1}{2}(n+1)(n+2)$, which is the total number of polynomials of the type P_ν. But we see that $\frac{1}{2}(n+1)(n+2)$ is precisely $\dim U_n$, and therefore $\dim Y_n \leq \dim U_n$, from which $Y_n = U_n$. □

Theorem 6.16. *The expansions*

$$G(p-q) = \sum_{n=0}^{\infty}\sum_{\nu\in\sigma_n} P_\nu(q)G_\nu(p) = \sum_{n=0}^{\infty}\sum_{\nu\in\sigma_n} G_\nu(p)P_\nu(q) \tag{6.78}$$

are valid for $|q| < |p|$, and the series converges uniformly in all regions of the type $\{(p,q) \ ; \ |q| \leq r|p|\}$, with $r < 1$.

Proof: We proved in chapter 4 (example 4.11) that

$$(1-q)^{-1} = \sum_{n=0}^{\infty} q^n$$

for quaternions with a norm less than 1, and we also saw that for every ball $|q| \leq r$, with $r \in \mathbb{R}$, $r < 1$, the series converges uniformly. From this, it follows that we can expand $G(1-q) = (1-q)^{-2}(1-\bar{q})$ into a power series in q and $\bar{q}$, which converges uniformly in every ball with a radius less than 1. Since G has the multiplicative property:

$$G(q_1 q_2) = G(q_2)G(q_1)$$

it follows that the function $G(p-q)$ can be expanded into a power series of the variable $p^{-1}q$, multiplied by $G(p)$, and that the series will converge uniformly for $|p^{-1}q| \leq r$ with $r < 1$.
The multivariable Taylor expansion of G around p is given by:

$$G(p-q) = \sum_{r,s=0}^{\infty} \sum_{i_1 \dots i_s} \frac{(-1)^{r+s}}{(r+s)!} \frac{\partial^{r+s} G}{\partial t^r \partial x_{i_1} \dots \partial x_{i_s}}(p) t^r x_{i_1} \dots x_{i_s}$$

Being G regular, we can replace in the expansion the derivatives with respect to t of this latter with $-\sum_i e_i \frac{\partial G}{\partial x_i}$ obtaining:

$$G(p-q) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{r+s=n} \sum_{i_1 \dots i_s} \sum_{j_1 \dots j_r} (te_{j_1}) \dots (te_{j_r})(-x_{i_1}) \dots (-x_{i_s}) \frac{\partial^n G}{\partial x_{j_1} \dots \partial x_{j_r} \partial x_{i_1} \dots \partial x_{i_s}}(p)$$

But the expression just obtained is $\sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) G_\nu(q)$ and therefore:

$$G(p-q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) G_\nu(q) \tag{6.79}$$

Finally, to prove the validity of the identity $G(p-q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} G_\nu(q) P_\nu(q)$, the same procedure must be repeated, remembering that the function G is also right-regular: from the right Cauchy-Riemann-Fueter equation $\bar{\partial}_r G = 0$, the derivative of G with respect to t must be derived:

$$\frac{\partial G}{\partial t} = -\sum_i \frac{\partial G}{\partial x_i} e_i$$

and replace this latter in the previous formula.

□

We are now ready to demonstrate the quaternionic analogue of Taylor's theorem:

Theorem 6.17 (Quaternionic Taylor's theorem, Sudbery 1979)**.** *Let f be a regular function in a neighborhood of 0. Then, there exists a ball centered at 0 in which the function is represented by the following uniformly convergent series:*

$$f(q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) a_\nu \tag{6.80}$$

where the coefficients are given by:

$$a_\nu = \frac{1}{2\pi^2} \oiiint_{\partial B} G_\nu(q) Dq f(q) = (-1)^n \partial_\nu f(0)$$

Proof: Let S be the sphere centered at 0 contained in the regularity domain of f, and let B be a closed ball centered at 0 inside S. Then $\forall q \in B$, by the Cauchy-Fueter integral formula:

$$f(q) = \frac{1}{2\pi^2} \oiiint_S G(q' - q) Dq' f(q') \tag{6.81}$$

but by the previous theorem, for $|q| < |q'|$ we can expand the function $G(q' - q)$ into the uniformly convergent series

$$G(q' - q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) G_\nu(q')$$

But we see that if $q \in B$ and $q' \in S$, then necessarily $|q| < |q'|$, and therefore the expansion will be valid in the region considered in equation (6.81) substituting this series in (6.81), we obtain:

$$f(q) = \frac{1}{2\pi^2} \oiiint_S \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) G_\nu(q') Dq' f(q')$$

Now, knowing that the series converges uniformly over $B \times S$, we can invert the sign of summation and integration, obtaining:

$$f(q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) \Big(\frac{1}{2\pi^2} \oiiint_S G_\nu(q') Dq' f(q') \Big) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q) a_\nu$$

where $a_\nu = \frac{1}{2\pi^2} \oiiint_S G_\nu(q') Dq' f(q')$.

To conclude the proof it is sufficient to observe that, by an argument already used previously (proposition 1.17 of chapter 1), being $G_\nu Dqf$ a closed form, i.e $d(G_\nu Dqf) = 0$, we can substitute the contour S with ∂B.

$$a_\nu = \frac{1}{2\pi^2} \oiiint_{\partial B} G_\nu(q') Dq' f(q')$$

We also note that by applying the differential operator ∂_ν to the function f we obtain:

$$\partial_\nu f(q) = \partial_\nu \Big[\frac{1}{2\pi^2} \oiiint_{\partial B} G(q' - q) Dq' f(q') \Big] = \frac{(-1)^n}{2\pi^2} \oiiint_{\partial B} G_\nu(q' - q) Dq' f(q')$$

Putting $q = 0$ we have that $\partial_\nu f(0)$ is precisely $(-1)^n a_\nu$:

$$\partial_\nu f(0) = \frac{(-1)^n}{2\pi^2} \oiiint_{\partial B} G_\nu(q') Dq' f(q') = (-1)^n a_\nu$$

therefore we can also write a_ν as:

$$a_\nu = (-1)^n \partial_\nu f(0)$$

□

Corollary 6.7. *Let μ and ν be two sets of the type $\{i_1, \dots i_m\}$ and $\{j_1, j_2, \dots j_n\}$, whose number of $1, 2$ and 3 that generate them is respectively $[m_1; m_2; m_3]$ and $[n_1; n_2; n_3]$. Then:*

$$\frac{1}{2\pi^2} \oiiint_S G_\mu(q) Dq P_\nu(q) = \delta_{\mu\nu}$$

where $\delta_{\mu\nu}$ indicates, as usual, the Kronecker delta, S is a sphere that contains the origin, and the functions $G_\mu(q)$ and $P_\nu(q)$ are those introduced just now in the section.

We conclude this chapter with a proof of the quaternionic analogue of Laurent's theorem for regular functions of a quaternionic variable.

Theorem 6.18 (Quaternionic Laurent's theorem, Sudbery 1979)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a regular quaternionic function in an open set $U \subset \mathbb{H}$, except possibly at a point $q_0 \in U$. Then there is a neighborhood N of q_0 such that if $q \in N$ and $q \neq q_0$ then for $f(q)$ the following infinite series expansion applies:*

$$f(q) = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} [P_\nu(q - q_0) a_\nu + G_\nu(q - q_0) b_\nu] \tag{6.82}$$

which we call the Laurent series expansion of $f(q)$. This series converges uniformly in every hyper-annular region:

$$\{q \in \mathbb{H}\ ;\ r \leq |q - q_0| \leq R\} \quad \textit{with } r > 0, \textit{ inside } N$$

Furthermore, the coefficients a_ν and b_ν are given by:

$$a_\nu = \frac{1}{2\pi^2} \oiiint_C G_\nu(q - q_0) Dq f(q)$$

$$b_\nu = \frac{1}{2\pi^2} \oiiint_C P_\nu(q - q_0) Dq f(q)$$

where here C is any closed 3-chain in $U \backslash \{q_0\}$, homologous to the boundary ∂B of a ball B contained in U such that $q_0 \in B \subset U$ (such that C has a winding index around q_0 equal to 1).

Proof: Let $R_1 > 0$ be a real number chosen such that $B_1 = \{q\ ;\ |q - q_0| \leq R_1\}$ is contained in U; let us call $N = \operatorname{int}(B_1)$, and let S_1 be the boundary of the ball B_1, $S_1 = \partial B_1$. Given a quaternion $q \in N \setminus \{q_0\}$, we choose another positive real number $R_2 > 0$ such that $0 < R_2 < |q - q_0| < R_1$, and let S_2 be the sphere $\{q\ ;\ |q - q_0| = R_2\}$. Then, applying the Cauchy-Fueter integral formula:

$$f(q) = \frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_1} G(q' - q) Dq' f(q') - \frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_2} G(q' - q) Dq' f(q') \qquad (6.83)$$

Observe that for $q' \in S_1$, we have $|q' - q_0| < |q - q_0|$ and hence, by a reasoning analogous to that of the previous theorem, we expand $G(q' - q)$ into the series $\sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q - q_0) G_\nu(q' - q_0)$, and utilizing the uniform convergence of the series, by bringing the $\sum$ outside the integral, we obtain:

$$\frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_1} G(q' - q) Dq' f(q') = \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} P_\nu(q - q_0) a_\nu$$

where here a_ν is equal to:

$$a_\nu = \frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_1} G_\nu(q' - q_0) Dq' f(q')$$

The obtained series converges uniformly in every ball $|q - q_0| \leq R$ with $R < R_1$. Finally, being $G_\nu D q f$ a closed 3-form in $U \setminus \{q_0\}$ (i.e., a 3-form with external differential equal to 0), we can replace the integration contour S_1 with a closed 3-chain C homologous to S_1 (proposition 1.17 of chapter 1) and hence:

$$a_\nu = \frac{1}{2\pi^2} \oint\!\!\!\iiint_{C} G_\nu(q' - q_0) Dq' f(q')$$

For the second term of equation (6.83), we will follow a similar procedure with a slight detail (which we will see shortly). Again, as before, for $q' \in S_2$ we have $|q' - q_0| < |q - q_0|$ and hence we can expand the series $G(q' - q)$ as before. However, this time, we will expand it as

$$G(q' - q) = -\sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} G_\nu(q - q_0) P_\nu(q' - q_0)$$

exploiting the multiplicative property of the function $G(q)$ in the particular case $G(-q) = -G(q)\ \forall q \in \mathbb{H} \setminus \{0\}$. Substituting this expansion in the second term of equation (6.83) we obtain:

$$-\frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_2} G(q' - q) Dq' f(q') = \frac{1}{2\pi^2} \oint\!\!\!\iiint_{S_2} \sum_{n=0}^{\infty} \sum_{\nu \in \sigma_n} G_\nu(q - q_0) P_\nu(q' - q_0) Dq' f(q')$$

now as before, utilizing the uniform convergence of the series we can bring the summation outside the integral obtaining:

$$\frac{1}{2\pi^2}\oiiint_{S_2}\sum_{n=0}^{\infty}\sum_{\nu\in\sigma_n}G_\nu(q-q_0)P_\nu(q'-q_0)Dq'f(q')=\sum_{n=0}^{\infty}\sum_{\nu\in\sigma_n}G_\nu(q-q_0)b_\nu$$

where here b_ν is equal to:

$$b_\nu=\frac{1}{2\pi^2}\oiiint_{S_2}P_\nu(q'-q_0)Dq'f(q')$$

Finally, utilizing again the fact that $G_\nu Dqf$ is a closed 3-form in $U\setminus\{q_0\}$ and that $C\simeq S_2$ we can rewrite b_ν as:

$$b_\nu=\frac{1}{2\pi^2}\oiiint_{C}P_\nu(q'-q_0)Dq'f(q')$$

Substituting everything in (6.83), we obtain the following expansion for $f(q)$:

$$f(q)=\sum_{n=0}^{\infty}\sum_{\nu\in\sigma_n}[P_\nu(q-q_0)a_\nu+G_\nu(q-q_0)b_\nu]$$

Furthermore, this series, as seen in the proof, converges in the following annular domain contained in N:

$$\{q\in\mathbb{H}\ ;\ r\leq|q-q_0|\leq R\}$$

with $r,R>0$. $\square$

6.13 Final Considerations

In this chapter, we have seen that for quaternionic regular functions in the sense of Fueter, there are analogues of all the most important theorems for holomorphic functions on $\mathbb{C}$: the Cauchy-Riemann equations, Cauchy's theorem, Morera's theorem, Cauchy's integral formula, Taylor and Laurent expansions, Liouville's theorem, exc.
The theory thus has many strengths, but it also has weaknesses; in particular, as the reader may have noticed while reading the chapter, that the class of $\mathbb{H}$-regular functions defined in this way does not generally contain quaternionic polynomials, nor many other elementary functions.
A partial solution to this problem is that we have managed to define quaternionic analogues of classical polynomials, such as polynomials of the type $P_\nu(q)$ for unordered sets of integers 1,2,3 ν, with which we have managed to expand regular functions into infinite series. Similarly, for non-$\mathbb{H}$-regular functions like the quaternionic logarithm $\log(q)$, we have considered the most natural generalization of the latter that is a regular function, $\mathscr{L}(q)$, a multiple of the Laplacian of the function induced by the complex logarithm (which is a $\mathbb{C}$-holomorphic function) [section 9 of this chapter] (the function can also be seen alternatively as a primitive to the quaternionic cauchy kernel).
Although the theory presented in this chapter is generally considered the most "faithful" extension of analysis to quaternions to its complex counterpart, there are many other equally valid ones that compensate for some of the defects of Fueter's theory.
One of these, which we will see in the next chapter, manages to significantly broaden the class of $\mathbb{H}$-regular functions through a definition based on the concept of complex holomorphy [19].
Another, on the other hand, will define a differential analysis on quaternions by constructing a more general version of the classical Fréchet derivative (for the differentiation of functions between Banach spaces) to complete normed modules (which we will call Banach modules) [8] [7] [29] [28]. Before discussing this theory, however, I wanted to introduce a brief section presenting the fundamental results of analysis on Banach spaces, motivating this generalization gradually.

Chapter 7

Alternative Analysis on $\mathbb{H}$

As mentioned in the final section of the last chapter, it is possible to construct theories that treat differential and integral calculus on quaternions alternatively (compared to the approach developed by Fueter, Moisil, Haefeli, and various other mathematicians between the beginning and middle of the twentieth century). Indeed, the theory discussed in the previous chapter presents some problems, which the theories we will briefly introduce aim to solve; unfortunately, as we will see, although some of these will succeed in solving some of the mentioned problems, they will lose some quite desirable properties.

7.1 Differential Calculus on Banach Spaces

Let's begin by seeing how we can build a theory of differential calculus on quaternions using their algebraic properties. Before proceeding, however, let's give a definition of a Banach space, as, to make the text as accessible as possible, the reader is not assumed to be familiar with functional analysis:

Definition 7.1 (Banach Space)**.** *Let V be a normed vector space, i.e., a vector space equipped with a norm $\|\ \| : V \to \mathbb{R}^+$, then we will call such a space a **Banach Space** if it is a complete metric space with respect to the metric topology induced by the norm $\|\ \|$.*

But, as seen in chapters 2 and 4, quaternions are a complete metric space with respect to the norm induced by their involution $*$, from which it follows that they are a Banach space with respect to this norm. A well-developed theory [20] of differential calculus on Banach spaces exists, which we will now introduce here; as always, before giving the main definitions, we introduce some preliminary definitions:

Definition 7.2 (Tangential Applications, Cartan)**.** *Let $\mathscr{B}$ and $\mathscr{B}'$ be two Banach spaces, and let $U \subset \mathscr{B}$ be an open set. We will say that two functions $\varphi_1 : U \to \mathscr{B}'$ and $\varphi_2 : U \to \mathscr{B}'$ are tangential at $a \in U$ if the function:*

$$m(r) := \sup_{||x-a|| \leq r} ||\varphi_1(x) - \varphi_2(x)||$$

satisfies the following equation:

$$\lim_{r \to 0} \frac{m(r)}{r} = 0 \tag{7.1}$$

i.e., $m(r)$ is an o-small of r, $m(r) = o(r)$.

It is easy to verify that this relationship forms an equivalence relation on the set of functions $\{\varphi : U \to \mathscr{B}'\}$. Thanks to the definition of tangential applications just introduced, we are ready to define what it means for an application between Banach spaces to be differentiable:

Definition 7.3 (Differentiable Application)**.** *A function between Banach spaces $\varphi : U \to \mathscr{B}'$ is differentiable at the point $a \in U$ if the following two properties are satisfied:*

- *φ is continuous at a*
- *there exists a linear application $g : \mathscr{B} \to \mathscr{B}'$ such that the applications $x \to \varphi(x) - \varphi(a)$ and $x \to g(x-a)$ are tangential at the point $a \in U$*

This can be rewritten, using the definition of tangential applications given earlier, as:

$$||\varphi(x) - \varphi(a) - g(x-a)|| = o(||x-a||)$$

Furthermore, we will say that a function $\varphi : U \to \mathscr{B}'$ defined in the same way as before is differentiable in U if it is differentiable at every point a in U. The definition of differentiability of an application between Banach spaces can equivalently be described by the following equation:

$$\lim_{x \to a} \frac{||f(x) - f(a) - g(x-a)||}{||x-a||} = 0$$

or, setting $||x-a|| = ||h||$:

$$\lim_{||h|| \to 0} \frac{||f(x+h) - f(x) - g(h)||}{||h||} = 0 \tag{7.2}$$

Functions belonging to the class of functions differentiable in this sense are commonly called "Fréchet-differentiable functions," and the linear application g is said to be the "Fréchet derivative of f at x", often written as $g = df(x)$. Let's now demonstrate some important properties of Fréchet derivatives, starting by noting that they are a direct generalization of derivatives of functions from $\mathbb{R}$ to $\mathbb{R}$:

Proposition 7.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a real function of a real variable; then f is differentiable in the classical sense if and only if it is differentiable in the Fréchet sense, and its Fréchet derivative $df_x(h)$ is equal to $f'(x)h$.*

Proof. Let's start by proving the direct direction, i.e., suppose that f is Fréchet differentiable. First, observe that since $df_x(h)$ is a linear function, we have $df_x(h) = hdf_x(1)$. Substituting this identity into expression (7.2) we obtain:

$$\lim_{||h|| \to 0} \frac{||f(x+h) - f(x) - df_x(h)||}{||h||} = 0 = \lim_{||h|| \to 0} \left| \frac{f(x+h) - f(x)}{h} - df_x(1) \right|$$

but the term $\frac{f(x+h)-f(x)}{h}$, for $h \to 0$ is precisely the derivative of f, $f'(x)$, from which $df_x(1) = f'(x)$. But again, using the linearity of df_x, and multiplying both sides by h we obtain:

$$df_x(h) = f'(x)h$$

For the reverse direction, instead, we will assume the existence of the derivative of f, $f'(x)$; to conclude it suffices to set $df_x(h) = f'(x)h$ and note that the limit (7.2) is equal to 0. □

Let's now work on some examples to solidify the intuition of the mathematical entities just defined, as well as to practice working computationally with them.

Example 7.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be defined as $f(x) = x^2$. We want to find a linear operator df_x such that the following limit is valid for $x \in \mathbb{R}$:*

$$\lim_{||h|| \to 0} \frac{||(x+h)^2 - x^2 - df_x(h)||}{||h||} = \lim_{||h|| \to 0} \frac{||2xh + h^2 - df_x(h)||}{||h||} = 0$$

from which it follows that the Fréchet derivative of x^2 is $df_x(h) = 2xh$.

Exercise 7.1. *Find the Fréchet derivative of the function $f(x) = x^3$.*

Let's now prove some general properties valid for the Fréchet derivative, which make it a good candidate for an extension of the derivative to more general spaces:

Proposition 7.2 (Linearity of the Fréchet Derivative)**.** *Let $f : U \to \mathscr{B}'$ and $g : U \to \mathscr{B}'$ be two functions between Banach spaces (real) differentiable at $x \in U$, and let $\lambda \in \mathbb{R}$ be a real scalar. Then:*

$$(f+g)'(x) = f'(x) + g'(x) \quad ; \quad (\lambda f)'(x) = \lambda f'(x)$$

Proof. Using the triangle inequality, we have:

$$\frac{|f(x+h) - f(x) - df_x(h) + g(x+h) - g(x) - dg_x(h)|}{|h|}$$

$$\leq \frac{|f(x+h) - f(x) - df_x(h)|}{|h|} + \frac{|g(x+h) - g(x) - dg_x(h)|}{|h|}$$

But the limit as $|h| \to 0$ on the right is 0, so:

$$\lim_{|h| \to 0} \frac{|f(x+h) + g(x+h) - (f(x) + g(x)) - (df_x(h) + dg_x(h))|}{|h|} = 0$$

For the second point, it suffices to observe that:

$$\lim_{h \to 0} \frac{|\lambda f(x+h) - \lambda f(x) - d(\lambda f)_x|}{|h|} = \frac{|\lambda||f(x+h) - f(x) - df_x|}{|h|} = 0$$

thus λf is differentiable: moreover, exploiting the linearity of the Fréchet derivative, we have $(\lambda f)'(x) = \lambda f'(x)$. □

Exercise 7.2. *Prove that for a constant function $f : U \to \mathscr{B}'$ defined as $f(x) := k \in \mathscr{B}'$ it will be:*

$$f'(x) = 0 \quad \forall x \in U$$

An extension of the chain rule for Fréchet derivatives of Fréchet-differentiable functions is generally valid.

Proposition 7.3 (Chain Rule for Fréchet Derivative)**.** *Let $\mathscr{B}$, $\mathscr{B}'$, and $\mathscr{B}''$ be three Banach spaces, and let $U \subset \mathscr{B}$ and $V \subset \mathscr{B}'$ be two open subsets of $\mathscr{B}$ and $\mathscr{B}'$ respectively. Let $f : U \to \mathscr{B}'$ and $g : V \to \mathscr{B}''$ be two continuous functions such that f is Fréchet-differentiable at a point $a \in U \subset \mathscr{B}$ and g is Fréchet-differentiable at $f(a)$, then the composite function $g \circ f$ is Fréchet-differentiable at a and its Fréchet derivative is given by:*

$$(g \circ f)'(a) = g'(f(a)) \circ f'(a) \tag{7.3}$$

Proof. Since f is differentiable (in the Fréchet sense) at a, we have:

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \tag{7.4}$$

and similarly, since g is Fréchet-differentiable at $f(a)$, we have:

$$g(x) = g(f(a)) + g'(f(a))(x - f(a)) + o(x - f(a))$$

Now let's calculate $(g \circ f)(x) - (g \circ f)(a)$:

$$g(f(x)) - g(f(a)) = g'(f(a))(f(x) - f(a)) + o(f(x) - f(a))$$

from which, substituting for $f(x)$ equation (7.4)

$$g(f(x)) - g(f(a)) = (g'(f(a)) \circ f'(a))(x-a) + g'(f(a))o(x-a) + o(f(x) - f(a))$$

From this form, it follows that, as the last two terms of the equation are tangential to 0, the Fréchet derivative of $g \circ f$ is precisely $g'(f(a)) \circ f'(a)$. □

Let's now see another important property of Fréchet-differentiable functions and their respective derivatives that will allow us to more easily prove some results. First of all, let's recall the definition of "equivalent metrics": two metrics on a topological space are said to be equivalent if they induce the same topology on it. Similarly, we will say that two norms on a vector space are equivalent if they induce equivalent metrics. Now let's state the following result:

Proposition 7.4. *If a function between Banach spaces, $f : \mathscr{B} \to \mathscr{B}'$ is differentiable at a point $a \in \mathscr{B}$ with respect to a norm $\| \, \|_1$, then it will be differentiable at the same point a for any other norm belonging to the equivalence class of $\| \, \|_1$, $[\| \, \|_1]$, i.e., with respect to any other norm equivalent to $\| \, \|_1$. Moreover, the derivative remains the same with respect to any norm in that equivalence class.*

Proof. By hypothesis, we have that the function f is differentiable at a certain $a \in \mathscr{B}$ with respect to $\| \, \|_1$. Let $\| \, \|_2$ be a norm equivalent to $\| \, \|_1$: then, for a known fact about metric spaces with equivalent norms (see [20] p.27), we will have that:

$$\frac{1}{\|x-a\|_2} \leq M \frac{1}{\|x-a\|_1}$$

where M is a constant greater than 0, $M > 0$. Similarly, we will have:

$$\|f(x) - f(a) - df_x(x-a)\|_2 \leq M' \|f(x) - f(a) - df_x(x-a)\|_1$$

for another positive real constant M'. Combining the two inequalities just obtained, we have:

$$\frac{\|f(x) - f(a) - df_x(x-a)\|_2}{\|x-a\|_2} \leq MM' \frac{\|f(x) - f(a) - df_x(x-a)\|_1}{\|x-a\|_1}$$

Since the limit as $x \to a$ of the right side of the inequality is 0, this will also be true for the left side:

$$\lim_{x \to a} \frac{\|f(x) - f(a) - df_x(x-a)\|_2}{\|x-a\|_2} = 0$$

This proves the assertion of the proposition. □

The fact just proved allows us to find an expression for the Fréchet derivative of various types of functions of the type $f : \mathbb{R}^n \to \mathbb{R}^m$. As in this case, we will see that the Fréchet derivative is closely related to the concept of the Jacobian matrix of an application of the type presented earlier. Let's start, for simplicity, from the simplest case $f : \mathbb{R} \to \mathbb{R}^m$.

Example 7.2. *Let $f : \mathbb{R} \to \mathbb{R}^m$ be a function in one real variable with an image in $\mathbb{R}^m$. The spaces $\mathbb{R}$ and $\mathbb{R}^m$ are Banach spaces with respect to the canonical Euclidean norm $|x| := \sqrt{\sum_{i=1}^n x_i^2}$, $x := (x_i)_{i=1}^n \in \mathbb{R}^n$, in the particular cases $n = 1$ and $n = m$.*
It is a well-known fact that the norm $|x|_1 := \sum_{i=1}^n |x_i|$ is equivalent to the canonical one (in general, for the vector space $\mathbb{R}^n$, it is true that every norm is equivalent).
Keeping this fact in mind, we will now prove that the Fréchet derivative of a function from $\mathbb{R}$ to $\mathbb{R}^m$, rewritable in terms of its component functions as $f(x) = \langle f_1(x), f_2(x), \ldots, f_m(x)\rangle$, is equal to:

$$df_x(h) = f'(x)h$$

where here $f'(x)$ is the vector function given by:

$$f'(x) = \langle f_1'(x), f_2'(x), \ldots, f_m'(x)\rangle$$

(in simpler terms the derivative of a function $f : \mathbb{R} \to \mathbb{R}^m$ in the "classical sense" of multivariable analysis). We will prove both directions, one at a time.
Let's start by assuming that f is differentiable "in the classical sense", i.e., there exists a vector function that we will call the derivative of f, defined as above:

$$f'(x) = \langle f_1'(x), f_2'(x), \ldots, f_m'(x)\rangle$$

Let's set $df_x(h) = f'(x)h$*; let's prove that with respect to this function, the limit*

$$\lim_{||h||\to 0} \frac{||f(x+h) - f(x) - df_x(h)||}{||h||} \tag{7.5}$$

is equal to 0:
Using proposition 7.4, let's consider the limit (7.5) *with respect to the norm* $|| \ ||_1$ *introduced at the beginning of the example (also called "taxi norm", "taxicab norm" or "Manhattan norm"). We have:*

$$\lim_{||h||\to 0} \frac{||f(x+h) - f(x) - df_x(h)||_1}{||h||}$$

$$= \lim_{||h||\to 0} \frac{||\langle f_1(x+h) - f_1(x) - f_1'(x), \ldots, f_m(x+h) - f_m(x) - f_m'(x)h\rangle||_1}{||h||} =$$

$$\lim_{||h||\to 0} \sum_{i=1}^{m} \left| \frac{f_i(x+h) - f_i(x) - f_i'(x)h}{h} \right| = \lim_{||h||\to 0} \sum_{i=1}^{m} \left| \frac{f_i(x+h) - f_i(x)}{h} - f_i'(x) \right| = 0$$

since $\forall i \in [1, m] \cap \mathbb{N}$, $\lim_{h\to 0} \frac{f_i(x+h) - f_i(x)}{h}$ *exists by hypothesis and is equal to* $f_i'(x)$.
Similarly, now assuming that f *is Fréchet differentiable, we have*

$$\lim_{||h||\to 0} \frac{||f(x+h) - f(x) - df_x(h)||_1}{||h||} = 0$$

from which, proceeding in a similar way as before:

$$\lim_{||h||\to 0} \sum_{i=1}^{m} \left| \frac{f_i(x+h) - f_i(x) - df_{xi}(h)}{h} \right| = \lim_{||h||\to 0} \sum_{i=1}^{m} \left| \frac{f_i(x+h) - f_i(x)}{h} - df_{xi}(1) \right| = 0$$

and thus $df_{xi}(1) = f_i'(x)$*, and exploiting the linearity of* df_x*, we will have the desired result, i.e* $df_{xi}(h) = f_i'(x)h$ *(Clarification about the notation: in the above limit* df_{xi} *indicates the i-th component of the linear application* df_x*).*

It is possible, through a similar procedure, to find expressions for the Fréchet derivatives of more general functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. As in this case, we will see that the Fréchet derivative is closely related to the concept of the Jacobian matrix of an application of the type exposed before. We leave the verification of this fact to the reader, to let them exercise with the concept of Fréchet derivative:

Exercise 7.3. *Prove that, for a function $f : \mathbb{R}^n \to \mathbb{R}$, it will be Fréchet differentiable if and only if it is differentiable in the classical sense. Furthermore, its Fréchet derivative is:*

$$df_x(\vec{h}) = \vec{\nabla} f(x) \cdot \vec{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} h_i \tag{7.6}$$

where here $\vec{\nabla} f(x)$ indicates the gradient of $f(x)$,

$$\vec{\nabla} f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

Exercise 7.4. *Prove that, for a function $f : \mathbb{R}^n \to \mathbb{R}^m$, it will be Fréchet differentiable if and only if it is differentiable in the classical sense. Furthermore, its Fréchet derivative is:*

$$df_x(h) = J(f)h = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \dots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \dots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \dots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} \tag{7.7}$$

Thus, knowing that $\mathbb{H} \cong \mathbb{R}^4$ (forgetting its multiplicative superstructure, considering it only as a Banach space), we will have that the Fréchet derivative of a differentiable function of a quaternion variable $f : \mathbb{H} \to \mathbb{H}$ at a point $q \in \mathbb{H}$ is:

$$df_q(h) = \begin{bmatrix} \frac{\partial f_1}{\partial t}(q) & \frac{\partial f_1}{\partial x}(q) & \frac{\partial f_1}{\partial y}(q) & \frac{\partial f_1}{\partial z}(q) \\ \frac{\partial f_2}{\partial t}(q) & \frac{\partial f_2}{\partial x}(q) & \frac{\partial f_2}{\partial y}(q) & \frac{\partial f_2}{\partial z}(q) \\ \frac{\partial f_3}{\partial t}(q) & \frac{\partial f_3}{\partial x}(q) & \frac{\partial f_3}{\partial y}(q) & \frac{\partial f_3}{\partial z}(q) \\ \frac{\partial f_4}{\partial t}(q) & \frac{\partial f_4}{\partial x}(q) & \frac{\partial f_4}{\partial y}(q) & \frac{\partial f_4}{\partial z}(q) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix} \tag{7.8}$$

where here $q := t + xi + yj + zk$ and $f(q) = f_1(t,x,y,z) + f_2(t,x,y,z)i + f_3(t,x,y,z)j + f_4(t,x,y,z)k$ and $h = h_1 + h_2 i + h_3 j + h_4 k \in \mathbb{H}$ and where the function $df(q)$ that maps to each point of the quaternions where f is differentiable to its Fréchet derivative in q, df_q is an application of the type:

$$df : \mathbb{H} \to \mathscr{L}(\mathbb{H}; \mathbb{H})$$

where $\mathscr{L}(\mathbb{H}; \mathbb{H})$ indicates the set of linear applications from $\mathbb{H}$ to itself.

7.2 Fréchet Derivatives of Quaternionic Functions

Let's now see how to operationally calculate the Fréchet derivative of a quaternionic function $f : \mathbb{H} \to \mathbb{H}$ at a point q where it is differentiable (in the Fréchet sense). Let's re-express 7.8 in the following more suggestive form (by carrying out the matrix multiplication):

$$df_q(h) = \begin{bmatrix} \sum_{i=1}^4 \frac{\partial f_1}{\partial x_i} h_i \\ \sum_{i=1}^4 \frac{\partial f_2}{\partial x_i} h_i \\ \sum_{i=1}^4 \frac{\partial f_3}{\partial x_i} h_i \\ \sum_{i=1}^4 \frac{\partial f_4}{\partial x_i} h_i \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial t} h_1 + \frac{\partial f_1}{\partial x} h_2 + \frac{\partial f_1}{\partial y} h_3 + \frac{\partial f_1}{\partial z} h_4 \\ \frac{\partial f_2}{\partial t} h_1 + \frac{\partial f_2}{\partial x} h_2 + \frac{\partial f_2}{\partial y} h_3 + \frac{\partial f_2}{\partial z} h_4 \\ \frac{\partial f_3}{\partial t} h_1 + \frac{\partial f_3}{\partial x} h_2 + \frac{\partial f_3}{\partial y} h_3 + \frac{\partial f_3}{\partial z} h_4 \\ \frac{\partial f_4}{\partial t} h_1 + \frac{\partial f_4}{\partial x} h_2 + \frac{\partial f_4}{\partial y} h_3 + \frac{\partial f_4}{\partial z} h_4 \end{bmatrix}$$

$$= \frac{\partial f_1}{\partial t} h_1 + \frac{\partial f_1}{\partial x} h_2 + \frac{\partial f_1}{\partial y} h_3 + \frac{\partial f_1}{\partial z} h_4 + (\frac{\partial f_2}{\partial t} h_1 + \frac{\partial f_2}{\partial x} h_2 + \frac{\partial f_2}{\partial y} h_3 + \frac{\partial f_2}{\partial z} h_4) i$$

$$+ (\frac{\partial f_3}{\partial t} h_1 + \frac{\partial f_3}{\partial x} h_2 + \frac{\partial f_3}{\partial y} h_3 + \frac{\partial f_3}{\partial z} h_4) j + (\frac{\partial f_4}{\partial t} h_1 + \frac{\partial f_4}{\partial x} h_2 + \frac{\partial f_4}{\partial y} h_3 + \frac{\partial f_4}{\partial z} h_4) k \in \mathscr{L}(\mathbb{H};\mathbb{H})$$

Example 7.3. *Let's find the Fréchet derivative of the identity function on quaternions, $f(q) = q$: a straightforward calculation gives the following result:*

$$df_q(h) = h_1 + h_2 i + h_3 j + h_4 k = h \quad \forall q \in \mathbb{H}$$

where $h = h_1 + h_2 i + h_3 j + h_4 k \in \mathbb{H}$.

Example 7.4. *Let $f : \mathbb{H} \to \mathbb{H}$ be the function $f(q) = q^2$. By specifying the coefficients of the basis $\{1, i, j, k\}$ for f, we have:*

$$f(q) = q^2 = t^2 - x^2 - y^2 - z^2 + 2txi + 2tyj + 2tzk$$

$$\begin{cases} f_1(q) = t^2 - x^2 - y^2 - z^2 \\ f_2(q) = 2tx \\ f_3(q) = 2ty \\ f_4(q) = 2tz \end{cases}$$

All component functions are continuous and differentiable $\forall t, x, y, z \in \mathbb{R}$. Its Fréchet derivative at a point $q \in \mathbb{H}$ is:

$$df_q(h) = 2th_1 - 2xh_2 - 2yh_3 - 2zh_4 + (2xh_1 + 2th_2)i + (2yh_1 + 2th_3)j + (2zh_1 + 2th_4)k$$

where $h = h_1 + h_2 i + h_3 j + h_4 k \in \mathbb{H}$.

Example 7.5. *We want to calculate the Fréchet derivative at a point $q \in \mathbb{H}$, $q \neq 1$, of the function*

$$f(q) = (q-1)^{-1} = \frac{\overline{q}-1}{|q-1|^2} = \frac{(t-1) - xi - yj - zk}{(t-1)^2 + x^2 + y^2 + z^2}$$

Explicitly detailing the "component functions":

$$\begin{cases} f_1(q) = \dfrac{t-1}{(t-1)^2 + x^2 + y^2 + z^2} \\ f_2(q) = \dfrac{-x}{(t-1)^2 + x^2 + y^2 + z^2} \\ f_3(q) = \dfrac{-y}{(t-1)^2 + x^2 + y^2 + z^2} \\ f_4(q) = \dfrac{-z}{(t-1)^2 + x^2 + y^2 + z^2} \end{cases}$$

Let's calculate the Jacobian of f:

$$J(f) = \begin{bmatrix} \frac{-(t-1)^2+x^2+y^2+z^2}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2x(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2y(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2z(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} \\ \frac{2x(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & -\frac{-x^2+y^2+z^2+(t-1)^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2xy}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2xz}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \\ \frac{2y(t-1)}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2yx}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & -\frac{-y^2+x^2+z^2+(t-1)^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2yz}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \\ \frac{2z(t-1)}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2zx}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2zy}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & -\frac{(t-1)^2+x^2+y^2-z^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \end{bmatrix}$$

The fréchet derivative of our function is given by:

$$df_q(h) = \begin{bmatrix} \frac{-(t-1)^2+x^2+y^2+z^2}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2x(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2y(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & \frac{-2z(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} \\ \frac{2x(t-1)}{((t-1)^2+x^2+y^2+z^2)^2} & -\frac{-x^2+y^2+z^2+(t-1)^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2xy}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2xz}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \\ \frac{2y(t-1)}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2yx}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & -\frac{-y^2+x^2+z^2+(t-1)^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2yz}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \\ \frac{2z(t-1)}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2zx}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & \frac{2zy}{\left((t-1)^2+x^2+y^2+z^2\right)^2} & -\frac{(t-1)^2+x^2+y^2-z^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \end{bmatrix}$$

$$= \frac{-(t-1)^2 + x^2 + y^2 + z^2}{((t-1)^2 + x^2 + y^2 + z^2)^2} h_1 + \frac{-2x(t-1)}{((t-1)^2 + x^2 + y^2 + z^2)^2} h_2 + \frac{-2y(t-1)}{((t-1)^2 + x^2 + y^2 + z^2)^2} h_4 +$$

$$\frac{-2z(t-1)}{((t-1)^2 + x^2 + y^2 + z^2)^2} h_4 + \left(\frac{2x(t-1)}{((t-1)^2 + x^2 + y^2 + z^2)^2} h_1 - \frac{-x^2 + y^2 + z^2 + (t-1)^2}{\left((t-1)^2 + x^2 + y^2 + z^2\right)^2} h_2 + \right.$$

$$\left. \frac{2xy}{\left((t-1)^2 + x^2 + y^2 + z^2\right)^2} h_3 + \frac{2xz}{\left((t-1)^2 + x^2 + y^2 + z^2\right)^2} h_4 \right) i + \left(\frac{2y\,(t-1)}{\left((t-1)^2 + x^2 + y^2 + z^2\right)^2} h_1 \right.$$

$$+\frac{2yx}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_2-\frac{-y^2+x^2+z^2+(t-1)^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_3+\frac{2yz}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_4\Big)j+$$

$$\Big(\frac{2z\,(t-1)}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_1+\frac{2zx}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_2+\frac{2zy}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_3$$

$$-\frac{(t-1)^2+x^2+y^2-z^2}{\left((t-1)^2+x^2+y^2+z^2\right)^2}h_4\Big)k$$

Example 7.6. *Let's calculate the Fréchet derivative at a point $q \in \mathbb{H}$ of the exponential function e^q. Explicitly detailing the component functions, we have (for $|\operatorname{Vec}(q)| \neq 0$):*

$$e^q = e^t \cos\sqrt{x^2+y^2+z^2} + e^t \frac{xi+yj+zk}{\sqrt{x^2+y^2+z^2}} \sin\sqrt{x^2+y^2+z^2}$$

from which, using the convention whereby f_i, $i = 1,2,3,4$ are the components of 1,i,j, and k respectively of the function, we have:

$$\begin{cases} f_1(q) = e^t \cos\sqrt{x^2+y^2+z^2} \\ f_2(q) = \dfrac{xe^t \sin\sqrt{x^2+y^2+z^2}}{\sqrt{x^2+y^2+z^2}} \\ f_3(q) = \dfrac{ye^t \sin\sqrt{x^2+y^2+z^2}}{\sqrt{x^2+y^2+z^2}} \\ f_4(q) = \dfrac{ze^t \sin\sqrt{x^2+y^2+z^2}}{\sqrt{x^2+y^2+z^2}} \end{cases}$$

Calculating the partial derivatives of f_i with respect to x_j, $j = 1,2,3,4$, and using the derived formula for the Fréchet derivative for functions from $\mathbb{H}$ to $\mathbb{H}$ we get:

$$df_q(h) = J(f)h = e^t \cos\sqrt{x^2+y^2+z^2}h_1 - \frac{e^t x \sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_2-$$

$$\frac{e^t y \sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_3 - \frac{e^t z \sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_4+$$

$$\Big(\frac{e^t x \sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_1$$

$$+\frac{e^t\left(x^2\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}+y^2\sin\left(\sqrt{x^2+y^2+z^2}\right)h_1+z^2\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_2$$

$$+\frac{e^t x\left(y\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-y\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_3$$

$$+\frac{e^t x\left(z\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-z\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_4\Bigg)i+$$

$$\Bigg(\frac{ye^t\sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_1+\frac{e^t y\left(x\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-x\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_2$$

$$+\frac{e^t\left(y^2\cos\left(\sqrt{y^2+x^2+z^2}\right)\sqrt{y^2+x^2+z^2}+x^2\sin\left(\sqrt{y^2+x^2+z^2}\right)+z^2\sin\left(\sqrt{y^2+x^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_3$$

$$+\frac{e^t y\left(z\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-z\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_4\Bigg)j+$$

$$\Bigg(\frac{e^t z\sin\left(\sqrt{x^2+y^2+z^2}\right)}{\sqrt{x^2+y^2+z^2}}h_1+$$

$$\frac{e^t z\left(x\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-x\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_2$$

$$+\frac{e^t z\left(y\cos\left(\sqrt{x^2+y^2+z^2}\right)\sqrt{x^2+y^2+z^2}-y\sin\left(\sqrt{x^2+y^2+z^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_3+$$

$$\frac{e^t\left(z^2\cos\left(\sqrt{z^2+x^2+y^2}\right)\sqrt{z^2+x^2+y^2}+x^2\sin\left(\sqrt{z^2+x^2+y^2}\right)+y^2\sin\left(\sqrt{z^2+x^2+y^2}\right)\right)}{(x^2+y^2+z^2)\sqrt{x^2+y^2+z^2}}h_4\Bigg)k$$

This expression for the Fréchet derivative is valid for quaternions with a non-zero vector part: for quaternions with a zero vector part (i.e., real numbers), the Fréchet derivative will be the "standard" one for the real exponential function.

Exercise 7.5. *Determine the set of points in $\mathbb{H}$ where $\sin(q)$ and $\cos(q)$ are Fréchet differentiable, and find an expression for their Fréchet derivatives at a generic point q in that set.*

Exercise 7.6. *Calculate the Fréchet derivative of the function $f(q) = q^3$, $df_q(h)$, for $q = 1 - i - j$.*

Exercise 7.7. *Prove that, given two functions $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{H}$ Fréchet-differentiable at a point $q \in \mathbb{H}$ in their domain, the following holds:*

$$d(fg)_q(h) = \sum_{i=1}^{4}\left(\frac{\partial f_1}{\partial x_i}g_1+\frac{\partial g_1}{\partial x_i}f_1-\frac{\partial f_2}{\partial x_i}g_2-\frac{\partial g_2}{\partial x_i}f_2-\frac{\partial f_3}{\partial x_i}g_3-\frac{\partial g_3}{\partial x_i}f_3-\frac{\partial f_4}{\partial x_i}g_4-\frac{\partial g_4}{\partial x_i}f_4\right)h_i+$$

$$h_i\left(\frac{\partial f_1}{\partial x_i}g_2 + \frac{\partial g_2}{\partial x_i}f_1 + \frac{\partial f_2}{\partial x_i}g_1 + \frac{\partial g_1}{\partial x_i}f_2 + \frac{\partial f_3}{\partial x_i}g_4 + \frac{\partial g_4}{\partial x_i}f_3 - \frac{\partial f_4}{\partial x_i}g_3 - \frac{\partial g_3}{\partial x_i}f_4\right)i+$$

$$h_i\left(\frac{\partial f_1}{\partial x_i}g_3 + \frac{\partial g_3}{\partial x_i}f_1 - \frac{\partial f_2}{\partial x_i}g_4 - \frac{\partial g_4}{\partial x_i}f_2 + \frac{\partial f_3}{\partial x_i}g_1 + \frac{\partial g_1}{\partial x_i}f_3 + \frac{\partial f_4}{\partial x_i}g_2 + \frac{\partial g_2}{\partial x_i}f_4\right)j+$$

$$h_i\left(\frac{\partial f_1}{\partial x_i}g_4 + \frac{\partial g_4}{\partial x_i}f_1 + \frac{\partial f_2}{\partial x_i}g_3 + \frac{\partial g_3}{\partial x_i}f_2 - \frac{\partial f_3}{\partial x_i}g_2 - \frac{\partial g_2}{\partial x_i}f_3 + \frac{\partial f_4}{\partial x_i}g_1 + \frac{\partial g_1}{\partial x_i}f_4\right)k$$

Exercise 7.8. *Prove that, given two functions $f, g : \mathbb{H} \to \mathbf{P}$, i.e., two functions from the space of quaternions to pure quaternions, defined as:*

$$f(q) = f_1(t,x,y,z)i + f_2(t,x,y,z)j + f_3(t,x,y,z)k$$

$$g(q) = g_1(t,x,y,z)i + g_2(t,x,y,z)j + g_3(t,x,y,z)k$$

and differentiable (in the sense of Fréchet) both at a point q, then:

$$d_q(f \times g)(h) = d_q\varphi(h) - d_q\psi(h)$$

where $\times$ denotes the vector product of two pure quaternions, and the functions φ and ψ are defined as:

$$\varphi = f_2g_3i + f_3g_1j + f_1g_2k \quad ; \quad \psi = f_3g_2i + g_3f_1j + g_1f_2k$$

7.3 $\mathbb{H}_\Sigma$ Differentiable Functions

In the previous section, we presented the most "standard" method to develop differential calculus on a Banach space. However, as seen in the examples and problems, the Fréchet derivatives of almost all the most important elementary functions do not have a very "well-behaved" form, and moreover, although they are linear and a chain rule applies, there is no analogous product rule that has a form similar to that of derivatives of real and complex functions.

Additionally, another problem with the theory just developed is that quaternions were algebraically contemplated as a complete normed real vector

space (real Banach space), while ideally, we would like to consider a structure in which we can multiply our vectors by quaternionic scalars, i.e., a module.
In this chapter, we will see how an alternative definition of differentiability on quaternions (exploiting their structure as a Banach D-module) will allow us to obtain a notion of derivative with many properties similar to those in the real and complex case. We will call the class of functions differentiable in this sense "$\mathbb{H}_\Sigma$-differentiable functions" to distinguish them from functions differentiable in the sense of Fréchet and regular functions in the sense of Fueter.

Definition 7.4 ($\mathbb{H}_\Sigma$-differentiable Functions, Dzagnidze)**.** *Let $f(q)$ be a function of a quaternionic variable $f : \mathbb{H} \to \mathbb{H}$, with $q = t + xi + yj + zk$, defined in a neighborhood G of a point q_0: we will say that f is $\mathbb{H}_\Sigma$-differentiable at $q_0 = t_0 + x_0 i + y_0 j + z_0 k$ if there exist two sequences, $A_k(q_0)$ and $B_k(q_0)$ such that the series $\sum_k A_k(q_0)B_k(q_0)$ is convergent and:*

$$f(q_0 + h) - f(q_0) = \sum_k A_k(q_0)hB_k(q_0) + \omega(q_0, h) \tag{7.9}$$

where

$$\lim_{h\to 0} \frac{|\omega(q_0, h)|}{|h|} = 0 \Longleftrightarrow |\omega(q_0, h)| = o(|h|)$$

and $q_0 + h \in G$. Finally, we will call the derivative of f calculated at q_0, indicated with $f'(q_0)$, the series:

$$f'(q_0) := \sum_k A_k(q_0)B_k(q_0)$$

Moreover, from now on, we will refer to the term $\omega(q_0, h)$ simply as $o(h)$, being that its norm is o-small compared to the norm of h.
This notion of derivative has many properties of differential calculus on real and complex numbers that the reader is familiar with. For example, a "power rule" for a monomial is generally valid for this derivative:

Proposition 7.5. *Let $f(q) = q^n$, where $q = t + xi + yj + zk \in \mathbb{H}$ and $n \in \mathbb{N}$. Then:*

$$f'(q) = (q^n)' = nq^{n-1}$$

Proof. We want to prove the validity for every $n \in \mathbb{N}$ of the following identity:

$$(q + h)^n - q^n = q^{n-1}h + q^{n-2}hq + \cdots + qhq^{n-2} + hq^{n-1} + o(h)$$

We proceed by induction; the identity is obviously true for $n = 1$ as:

$$q + h - q = h$$

For the inductive step, let's assume it's true for $n = k$. Then for $n = k + 1$, we have:

$$(q+h)^{k+1} - q^{k+1} = (q+h)(q+h)^k - q^{k+1} =$$
$$(q+h)(q^k + q^{k-1}h + \cdots + qhq^{k-2} + hq^{k-1} + o(h)) - q^{k+1} =$$
$$q^k h + q^{k-1}hq + \cdots + qhq^{k-1} + hq^k + o(h)$$

But we see that the right side of the equation is exactly of the form (7.9), where the sum $\sum_m A_m h B_m$ in this case is:

$$\sum_m A_m h B_m = \sum_m q^{k-m} h q^m$$

hence the derivative of f at q is:

$$(q^n)' = \sum_m A_m B_m = \sum_m q^{k-m} q^m = nq^{n-1}$$

□

Let's now prove a result that will be useful later for finding expressions of derivatives of important elementary quaternionic functions.

Lemma 7.1. *The following identities are valid for $|h| < 1$:*

$$\begin{cases} \dfrac{(q+h)^2 - q^2}{2!} = \dfrac{qh + hq}{2!} + A_2 \\ \dfrac{(q+h)^3 - q^3}{3!} = \dfrac{q^2h + qhq + hq^2}{3!} + A_3 \\ \qquad\vdots \\ \qquad\vdots \end{cases}$$

where $A_2 = \frac{h^2}{2}$, $A_3 = \frac{1}{3!}(qh^2 + hqh + h^2q + h^3), \ldots$ and so on, and where $A_n = o(h)$ for all $n \geq 2$.

Proof. To prove this, observe that:

$$|A_3| < \frac{2^3}{3!}(|q||h|^2 + |h|^3) < \begin{cases} \dfrac{2^3}{3!}|h|^2(1 + |h|) < \dfrac{2^3}{3!}|h|^2 \dfrac{1}{1 - |h|} & if \;\; |q| < 1 \\ \dfrac{2^3}{3!}|q||h|^2(1 + |h|) < \dfrac{2^3}{3!}|q||h|^2 \dfrac{1}{1 - |h|} & if \;\; |q| \geq 1 \end{cases}$$

$$|A_4| < \frac{2^4}{4!}(|q|^2|h|^2 + |q||h|^3 + |h|^4) < \begin{cases} \frac{2^4}{4!}|h|^2(1+|h|+|h|^2) < \frac{2^4}{4!}|h|^2\frac{1}{1-|h|} & if \ \ |q|<1 \\ \frac{2^4}{4!}|q|^2|h|^2(1+|h|+|h|^2) < \frac{2^4}{4!}|q|^2|h|^2\frac{1}{1-|h|} & if \ |q| \geq 1 \end{cases}$$

$$\vdots$$

$$\vdots$$

Proceeding in the same way, we obtain for a generic n:

$$|A_n| < \begin{cases} \frac{2^n}{n!}|h|^2\frac{1}{1-|h|} & if \ \ |q|<1 \\ \frac{2^n}{n!}|q|^{n-2}|h|^2\frac{1}{1-|h|} & if \ \ |q| \geq 1 \end{cases}$$

from which

$$\sum_{n=3}^{\infty}|A_n| < \begin{cases} |h|^2\frac{1}{1-|h|}\sum_{n=3}^{\infty}\frac{2^n}{n!} & if \ \ |q|<1 \\ |h|^2\frac{1}{1-|h|}\sum_{n=3}^{\infty}\frac{2^n}{n!}|q|^{n-2} & if \ \ |q| \geq 1 \end{cases}$$

but the series $\sum_{n=3}^{\infty}\frac{2^n}{n!}|q|^{n-2}$ and $\sum_{n=3}^{\infty}\frac{2^n}{n!}$ converge, thus:

$$\sum_{n=3}^{\infty}|A_n| = o(h)$$

for all $q \in \mathbb{H}, n \geq 2$. □

Let's demonstrate the first "familiar" corollary of lemma 7.1:

Proposition 7.6. *Let $f(q) = e^q$ be the quaternionic exponential function defined in chapter 5, then:*

$$f'(q) = (e^q)' = e^q$$

Proof. Using the expansion of the function e^q into its quaternionic infinite series (i.e., its definition given in chapter 5), we have:

$$e^{q+h} - e^q = h + \frac{(q+h)^2 - q^2}{2!} + \frac{(q+h)^3 - q^3}{3!} + \dots$$

Applying lemma 7.1 to the right side of the above equation, we get:

$$e^{q+h} - e^q = h + \frac{1}{2!}(qh + hq) + \frac{1}{3!}(q^2h + qhq + hq^2) + \cdots + o(h)$$

Therefore:

$$e^{q+h} - e^q = (1 + \frac{q}{2!} + \frac{q^2}{3!} + \ldots)h + (\frac{1}{2!} + \frac{q}{3!} + \frac{q^2}{4!} + \ldots)hq + (\frac{1}{3!} + \frac{q}{4!} + \frac{q^2}{5!} + \ldots)hq^2 + \cdots + o(h)$$

From which, by the definition of derivative given in definition 7.4, and noting that:

$$1 + \frac{q}{2!} + \frac{q^2}{3!} + \cdots + \frac{q}{2!} + \frac{q^2}{3!} + \frac{q^3}{4!} + \cdots + \frac{q^2}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{q^n}{n!} = e^q$$

we have that

$$(e^q)' = e^q$$

□

The following results are generally valid for the derivatives of quaternionic trigonometric functions:

Proposition 7.7. *Let $q \in \mathbb{H}$ be a quaternion and let $\sin(q)$ and $\cos(q)$ be the quaternionic sine and cosine functions. Then:*

$$(\sin(q))' = \cos(q)$$

$$(\cos(q))' = -\sin(q)$$

Proof. First, we find the derivative of the sine function:

$$\sin(q + h) - \sin(q) = (q + h) - \frac{(q + h)^3}{3!} + \frac{(q + h)^5}{5!} - \cdots - q + \frac{q^3}{3!} - \cdots =$$

$$h - \frac{(q + h)^3 - q^3}{3!} + \frac{(q + h)^5 - q^5}{5!} - \ldots$$

$$= h - \frac{1}{3!}(q^2h + qhq + hq^2) + \frac{1}{5!}(q^4h + q^3hq + q^2hq^2 + qhq^3 + hq^4) + \cdots + o(h) =$$

$$h + (-\frac{q^2}{3!} + \frac{q^4}{5!})h + qh(-\frac{q}{3!} + \frac{q^3}{5!}) + h(-\frac{q^2}{3!} + \frac{q^4}{5!}) + \cdots + o(h)$$

Thus, the function $\sin(q)$ is differentiable, and its derivative is equal to:

$$(\sin(q))' = 1 - \frac{q^2}{3!} + \frac{q^4}{5!} - \frac{q^2}{3!} + \frac{q^4}{5!} - \frac{q^2}{3!} + \frac{q^4}{5!} + \cdots = \cos(q)$$

Similarly, for the cosine, we observe that:

$$\cos(q+h)-\cos(q)=\sum_{n=0}^{\infty}\frac{(-1)^n(q+h)^{2n}}{(2n)!}-\frac{(-1)^n q^{2n}}{(2n)!}=1-\frac{(q+h)^2}{2!}+\frac{(q+h)^4}{4!}$$

$$-\cdots-1+\frac{q^2}{2!}-\frac{q^4}{4!}+\cdots=-\frac{(q+h)^2-q^2}{2!}+\frac{(q+h)^4-q^4}{4!}-\frac{(q+h)^6-q^6}{6!}+\cdots=$$

$$-\frac{1}{2!}(qh+hq)+\frac{1}{4!}(q^3h+q^2hq+qhq^2+hq^3)-\cdots+o(h)$$

We have reduced the difference $\cos(q+h)-\cos(q)$ to a form of type (7.9). From the obtained expression, by rearranging the terms in a similar way as done before, it follows that the derivative of our function is equal to:

$$(\cos(q))'=-\sin(q)$$

□

Exercise 7.9. *Prove that the "usual" relations on the derivatives of hyperbolic functions are valid for this new definition of derivative; i.e., show that, using the infinite quaternionic series expansions of* $\sinh(q)$ *and* $\cosh(q)$*, that:*

$$(\sinh(q))'=\cosh(q)\quad \textit{and}\quad (\cosh(q))'=\sinh(q)$$

As already mentioned before, for the $\mathbb{H}_\Sigma$ derivative of a function in a quaternionic variable, all the most important properties of derivatives of functions in a real and complex variable apply; let's prove it:

Theorem 7.1 (Properties of the $\mathbb{H}_\Sigma$ derivative)**.** *Let* f *and* g *be two functions of a quaternionic variable defined in a neighborhood of* $q\in\mathbb{H}$*. If both functions are* $\mathbb{H}_\Sigma$*-differentiable at* q*, then:*

- *For every* $p\in\mathbb{H}$, pf *and* fp *are differentiable at* q *and:*

$$(pf)'(q)=pf'(q)\quad (fp)'(q)=f'(q)p$$

- $f+g$ *is differentiable at* q *and:*

$$(f+g)'(q)=f'(q)+g'(q)$$

- fg *is differentiable at* q *and:*

$$(fg)'(q)=f'(q)g(q)+f(q)g'(q)$$

Proof. • For the first fact, we simply observe that if f is differentiable, we have:

$$f(q+h) - f(q) = \sum_k A_k(q)hB_k(q) + o(h)$$

Multiplying both sides from the left by the quaternionic constant p, we get

$$pf(q+h) - f(q) = \sum_k pA_k(q)hB_k(q) + o(h)$$

i.e., pf is differentiable and its derivative is equal to pf'. We obtain the same result on the right by multiplying p from the right.

- For the second result, we note that, under the assumption as before that f and g are differentiable at q, we have:

$$f(q+h) - f(q) = \sum_k A_k(q)hB_k(q) + o(h)$$

$$g(q+h) - g(q) = \sum_k C_k(q)hB_k(q) + o(h)$$

Hence:

$$f(q+h) + g(q+h) - (f(q) + g(q)) = \sum_k A_k(q)hB_k(q) + \sum_k C_k(q)hB_k(q) + o(h)$$

i.e., $f + g$ is differentiable and

$$(f+g)' = f' + g'$$

- Finally, for the product, we note that we can rewrite the difference between $f(q+h)g(q+h)$ and $f(q)g(q)$ as:

$$f(q+h)g(q+h) - f(q)g(q) = [f(q+h) - f(q)]g(q+h) + f(q)[g(q+h) - g(q)]$$

from which, using the given differentiability of f and g, we can equate to:

$$= \Big[\sum_k A_k(q)hB_k(q) + o(h)\Big]g(q+h) + f(q)\Big[\sum_k C_k(q)hD_k(q) + o(h)\Big] \quad (7.10)$$

but, still using the differentiability of g, remember that $g(q+h) - g(q) = \sum_k C_k(q)hB_k(q) + o(h)$, from which $g(q+h) = g(q) + \sum_k C_k(q)hB_k(q) + o(h)$. Substituting this relation into equation (7.10) we get

$$\Big[\sum_k A_k(q)hB_k(q) + o(h)\Big]\Big[g(q) + \sum_k C_k(q)hD_k(q) + o(h)\Big]$$

$$+f(q)\Big[\sum_k C_k(q)hD_k(q) + o(h)\Big]$$

distributing the products and noting that the terms $o(h)g(q)$, $o(h)\sum_k C_k(q)hD_k(q)$, $\sum_k A_k(q)hB_k(q)o(h)$ and $\sum_k A_k(q)hB_k(q)\sum_{k'} C_{k'}(q)hD_{k'}(q)$ are all $o(h)$, we have that the whole thing equals:

$$= \Big(\sum_k A_k(q)hB_k(q)\Big)g(q) + f(q)\Big(\sum_k C_k(q)hD_k(q)\Big) + o(h)$$

We have thus obtained an equation of the form (7.9) from which follows that:

$$(fg)'(q) = \Big(\sum_k A_k(q)B_k(q)\Big)g(q) + f(q)\Big(\sum_k C_k(q)D_k(q)\Big)$$

$$= f'(q)g(q) + f(q)g'(q)$$

□

There is also a formula for the $\mathbb{H}_\Sigma$ derivatives of quaternionic functions to compute the derivative of the reciprocal of a function. Let's demonstrate it below, but first, let's prove a preparatory lemma.

Lemma 7.2. *Let $q_1, q_2 \neq 0$ be two quaternions. Then:*

$$q_1^{-1} - q_2^{-1} = q_1^{-1}(q_1 - q_2)q_2^{-1}(q_1 - q_2)q_2^{-1} - q_2^{-1}(q_1 - q_2)q_2^{-1}$$

Proof. Let's demonstrate it by algebraically manipulating the right side; doing so, we obtain

$$(1-q_1^{-1}q_2)q_2^{-1}(q_1q_2^{-1}-1)-(q_2^{-1}q_1-1)q_2^{-1} = (q_2^{-1}-q_1^{-1})(q_1q_2^{-1}-1)-(q_2^{-1}q_1q_2^{-1}-q_2^{-1})$$

$$= (q_2^{-1}q_1q_2^{-1} - q_2^{-1} - q_2^{-1} + q_1^{-1}) - (q_2^{-1}q_1q_2^{-1} - q_2^{-1}) = q_1^{-1} - q_2^{-1}$$

□

Proposition 7.8. *Let f be a function that is $\mathbb{H}_\Sigma$ differentiable at a point $q \in \mathbb{H}$ and such that $f(q) \neq 0$ in a neighborhood of q, then f^{-1} is also differentiable at q and:*

$$(f^{-1})' = \left(\frac{1}{f}\right)' = -\frac{1}{f(q)}f'(q)\frac{1}{f(q)}$$

Proof. Firstly let $f'(q) = \sum_k A_k(q)B_k(q)$ be the derivative of f at q: now let's analyze the difference $\frac{1}{f(q+h)} - \frac{1}{f(q)}$; using lemma 7.2 we get

$$\frac{1}{f(q+h)} - \frac{1}{f(q)} = \frac{1}{f(q+h)}(f(q+h) - f(q))\frac{1}{f(q)}(f(q+h) - f(q))\frac{1}{f(q)} -$$

$$\frac{1}{f(q)}(f(q+h) - f(q))\frac{1}{f(q)}$$

Remembering that $f(q+h) - f(q) = \sum_k A_k(q)hB_k(q) + o(h)$; substituting this identity in the above formula we have:

$$\frac{1}{f(q+h)}(\sum_k A_k(q)hB_k(q) + o(h))\frac{1}{f(q)}(\sum_k A_k(q)hB_k(q) + o(h))\frac{1}{f(q)} -$$

$$\frac{1}{f(q)}(\sum_k A_k(q)hB_k(q) + o(h))\frac{1}{f(q)}$$

distributing the various products, and being careful as before about which ones are $o(h)$, we finally get that the whole thing equals

$$= -\frac{1}{f(q)}\Big[\sum_k A_k(q)hB_k(q)\Big]\frac{1}{f(q)} + o(h)$$

from which:

$$\left(\frac{1}{f(q)}\right)' = -\frac{1}{f(q)}\Big[\sum_k A_k(q)B_k(q)\Big]\frac{1}{f(q)} = -\frac{1}{f(q)}f'(q)\frac{1}{f(q)}$$

□

Corollary 7.1. *Let $q \neq 0$, and let $f : \mathbb{H} \to \mathbb{H}$ be the function $f(q) = q^m$, where $m < 0$. Then:*

$$(q^m)' = mq^{m-1}$$

Proof. To prove this immediate corollary, it suffices to first apply proposition 7.5 ("power rule" for the $\mathbb{H}_\Sigma$ quaternionic derivative) and then proposition 7.8. □

Another important corollary of the proposition just proved will allow us to calculate, given 2 $\mathbb{H}_\Sigma$-differentiable quaternionic functions, the derivative of their ratio.

Corollary 7.2. *Let $f, g : \mathbb{H} \to \mathbb{H}$ be 2 quaternionic functions that are $\mathbb{H}_\Sigma$-differentiable at a point $q \in \mathbb{H}$, and let $g(q) \neq 0$ in a neighborhood of q. Then, the functions $f \cdot \frac{1}{g}$ and $\frac{1}{g} \cdot f$ are in turn $\mathbb{H}_\Sigma$-differentiable and their derivative is:*

$$\left(f \cdot \frac{1}{g}\right)'(q) = f'(q)\frac{1}{g(q)} - f(q)\frac{1}{g(q)}g'(q)\frac{1}{g(q)}$$

$$\left(\frac{1}{g} \cdot f\right)'(q) = -\frac{1}{g(q)}g'(q)\frac{1}{g(q)}f(q) + \frac{1}{g(q)}f'(q)$$

Proof. For the proof, it is sufficient to jointly apply the product rule for the $\mathbb{H}_\Sigma$ derivative (third point of theorem 7.1), and the proposition 7.8 just demonstrated. □

Concerning the properties of this alternative definition of derivative on the quaternionic skew-field, one last question remains; is there also for this one a ”chain rule” similar to that of the real and complex case? The answer is affirmative, albeit with some modifications.

Theorem 7.2 (Quaternionic Chain Rule, O. Dzagnize)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a function defined in an interval of a point $q_0 \in \mathbb{H}$ and let $g : \mathbb{H} \to \mathbb{H}$ be another function defined in an interval of the point $f(q_0) = w_0 \in \mathbb{H}$. Then, if f is $\mathbb{H}_\Sigma$-differentiable at q_0, and g is $\mathbb{H}_\Sigma$ differentiable at $w_0 = f(q_0)$, and their derivatives are given by:*

$$g'(w_0) = \sum_k A_k(w_0)B_k(w_0)$$

$$f'(q_0) = \sum_j C_j(q_0)D_j(q_0)$$

we will have that the composite function $g \circ f$ is differentiable at q_0, and its derivative is equal to:

$$(g \circ f)'(q_0) = \sum_k A_k(q_0)f'(q_0)B_k(q_0)$$

Proof. Let $q \in G \subset \mathbb{H}$, where G is an open set containing the point q_0. Let $w = f(q)$; then:

$$g(w) - g(w_0) = \sum_k A_k(w - w_0)B_k + \omega_1(w_0, w)$$

$$f(q) - f(q_0) = \sum_j C_j(q - q_0)D_j + \omega_2(q_0, q)$$

from which, explicitly calculating the same difference for the composite function:

$$g(f(q)) - g(f(q_0)) = \sum_k A_k(f(q) - f(q_0))B_k + \omega_1(f(q_0), f(q))$$

$$= \sum_k A_k\Big(\sum_j C_j(q - q_0)D_j + o(h)\Big)B_k + \omega_1(f(q_0), f(q))$$

$$= \sum_k A_k\Big(\sum_j C_j(q - q_0)D_j\Big)B_k + o(h) + \omega_1(f(q_0), f(q))$$

$$= \sum_k \sum_j A_kC_j(q - q_0)D_jB_k + o(h) + \omega_1(f(q_0), f(q))$$

But $\lim\limits_{q\to q_0} \dfrac{|\omega_1(f(q_0), f(q))|}{|q - q_0|} = 0$, i.e., $\omega_1(f(q_0), f(q)) = o(h)$, and thus:

$$g(f(q)) - g(f(q_0)) = \sum_k \sum_j A_kC_j(q - q_0)D_jB_k + o(h)$$

From this expression, we finally deduce the result of the assertion:

$$(g \circ f)'(q_0) = \sum_k A_k(q_0)\Big(\sum_j C_j(q_0)D_j(q_0)\Big)B_k(q_0) = \sum_k A_k(q_0)f'(q_0)B_k(q_0) \tag{7.11}$$

□

For the special case where the function $g(q) = q^n$, we have:

Corollary 7.3. *If f is a $\mathbb{H}_\Sigma$-differentiable function in a quaternion q, then:*

$$(f^n)' = f^{n-1}(q)f'(q) + f^{n-2}(q)f'(q)f(q) + f^{n-3}(q)f'(q)f^2(q) + \cdots + f'(q)f^{n-1}(q) \tag{7.12}$$

7.4 Differential Calculus on Banach D-modules

The theory of $\mathbb{H}_\Sigma$-differentiable functions of a quaternionic function developed in the previous section is a particular case of the theory of differential calculus on a more general class of algebraic structures (which include the quaternions $\mathbb{H}$) known as Banach D-modules (where D is a ring), i.e., normed D-modules whose metric topology induced by their norm is complete. The definition of a differentiable function and the derivative of a function is very similar, let's give it here below:

Definition 7.5 (Derivative of a function between Banach D-modules, A. Kleyn)**.** *Let A be a D-module of Banach with norm $\|\ \|_A$, and B a D-module of Banach with norm $\|\ \|_B$; then, an application*

$$f : A \to B$$

is said to be differentiable at $x \in A$ if its increment can be written as:

$$f(x+h) - f(x) = \frac{df(x)}{dx}(h) + o(h)$$

where

$$\frac{df(x)}{dx} : A \to B$$

is a linear application of D-modules, and $o : A \to B$ is a continuous function such that:

$$\lim_{h \to 0} \frac{\|o(h)\|_B}{\|h\|_A} = 0$$

The linear application $\frac{df(x)}{dx}$ is called the derivative of the function f.

The derivative of a function $f : A \to B$ can be expanded in components: recalling that the linear applications between two Banach D-modules A and B, indicated with $\mathcal{L}(D; A \to B)$, form a left $B \otimes B$-module, we can prove the following fact:

Proposition 7.9. *Given a base $\{F_k\}_k$ of the left $B \otimes B$-module $\mathcal{L}(D; A \to B)$, we can represent the derivative of an application $f : A \to B$ as:*

$$\frac{df(x)}{dx} = \sum_k \frac{d^k f(x)}{dx} \circ F_k$$

where $\frac{d^k f(x)}{dx} \in B \otimes B$.

The expression $\dfrac{d^k f(x)}{dx} = \dfrac{d^k_{s_k 0} f(x)}{dx} \otimes \dfrac{d^k_{s_k 1} f(x)}{dx} \in B \otimes B$ is a decomposable tensor and is called the k-th coordinate of the derivative $\dfrac{df(x)}{dx}$ with respect to the base $\{F_k\}_k$ of the module $\mathcal{L}(D; A \to B)$. In addition, the expressions

$\dfrac{d^k_{s_k 1}f(x)}{dx}$ and $\dfrac{d^k_{s_k 0}f(x)}{dx}$ are called components of the k-th coordinate of the derivative of f.
In terms of the above, the derivative can be written as:

$$\frac{df(q)}{dq} = \sum_k \frac{d^k_{s_k 0}f(q)}{dq} \otimes \frac{d^k_{s_k 1}f(q)}{dq} \tag{7.13}$$

We see that the concept of derivative given in the previous section is nothing but the derivative of functions between Banach D-modules introduced here calculated in $h = 1$. In fact, rewriting the $\mathbb{H}_\Sigma$ differentiability condition of a function $f : \mathbb{H} \to \mathbb{H}$ we notice that:

$$\sum_k A_k(q)hB_k(q) = \frac{df(q)}{dq}(h) = \sum_k \frac{d^k_{s_k 0}f(q)}{dq} h \frac{d^k_{s_k 1}f(q)}{dq}$$

where here the application under which it is calculated

$$\mathbb{H} \otimes \mathbb{H} \to \mathbb{H}$$

$$a \otimes b \to ahb$$

is often called "sandwich application" or "interposition of quaternions". Let's now see some examples:

Example 7.7. *Let $f : \mathbb{H} \to \mathbb{H}$ be the function $f(q) = q^2$. As already seen in section 7.3, we have:*

$$(q^2 + h) - q^2 = qh + hq + h^2 = qh + hq + o(h)$$

since $\lim_{h\to 0} \dfrac{|h|^2}{|h|} = 0$. *From this we obtain that:*

$$\frac{df(q)}{dq}(h) = qh + hq$$

i.e.

$$\frac{df(q)}{dq} = q \otimes 1 + 1 \otimes q$$

that is

$$\begin{cases} A_1 = \dfrac{d^1_{s_1 0} f(q)}{dq} = q \\ B_1 = \dfrac{d^1_{s_1 1} f(q)}{dq} = 1 \\ A_2 = \dfrac{d^2_{s_2 0} f(q)}{dq} = 1 \\ B_2 = \dfrac{d^2_{s_2 1} f(q)}{dq} = q \end{cases}$$

Example 7.8. *Let $f : \mathbb{H} \to \mathbb{H}$ be the function $f(q) = q^3$. The increment $f(q+h) - f(q)$ in this case is equal to:*

$$(q+h)^3 - q^3 = q^2h + hq^2 + qhq + hqh + h^2q + qh^2 + h^2q + h^3$$

but $hqh + h^2q + qh^2 + h^2q + h^3 = o(h)$, and thus:

$$(q+h)^3 - q^3 = q^2h + hq^2 + qhq + o(h)$$

The derivative of $f(q)$ is expressible as

$$\frac{df(q)}{dq} = q^2 \otimes 1 + 1 \otimes q^2 + q \otimes q$$

If calculated in $h \in \mathbb{H}$, (interposing h between the terms of the tensor products):

$$\frac{df(q)}{dq}(h) = q^2h + hq^2 + qhq$$

Our $\mathbb{H}_\Sigma$ derivative is this last function calculated in $h = 1$:

$$(q^3)' = 3q^2$$

Example 7.9. *Let $f(q) = e^q$ be the quaternionic exponential. As seen in the proof of proposition 7.6, we have:*

$$e^{q+h} - e^q = (1 + \frac{q}{2!} + \frac{q^2}{3!} + \ldots)h + (\frac{1}{2} + \frac{q}{3!} + \frac{q^2}{4!} + \ldots)hq + (\frac{1}{3!} + \frac{q}{4!} + \frac{q^2}{5!} + \ldots)hq^2 + \cdots + o(h)$$

from which we obtain the following expression for the derivative of e^q:

$$\frac{df(q)}{dq} = (1 + \frac{q}{2!} + \frac{q^2}{3!} + \ldots) \otimes 1 + (\frac{1}{2} + \frac{q}{3!} + \frac{q^2}{4!} + \ldots) \otimes q + (\frac{1}{3!} + \frac{q}{4!} + \frac{q^2}{5!} + \ldots) \otimes q^2 + \ldots$$

Or, rewriting it in a slightly more "clean" manner:

$$\frac{df(q)}{dq} = \sum_{n=0}^{\infty} \Big(\sum_{m=0}^{\infty} \frac{q^m}{(m+n+1)!} \Big) \otimes q^n \tag{7.14}$$

Exercise 7.10. *Find an expression for the derivative in the form of equation* (7.13) *for the function* $f(q) = q^4$.

Exercise 7.11. *Find an expression for the derivative in the form of equation* (7.13) *for the function* $f(q) = \sin(q)$ *and* $f(q) = \cos(q)$.

Exercise 7.12. *Find the component expansion of the derivative of the quaternionic hyperbolic functions* $f(q) = \sinh(q)$ *and* $f(q) = \cosh(q)$.

7.5 C-regular Functions

In this section, we will introduce an alternative approach to quaternionic analysis developed by Cullen and subsequently researched and expanded by other mathematicians [5] [19]. In particular, we will examine in this section the alternative formulation of the concept of ”quaternionic regularity” due to G.Gentili and D.C. Struppa [19].

Definition 7.6 (C-regular Function)**.** *Let $\Omega \subset \mathbb{H}$ be an open and connected set, and let $f : \Omega \to \mathbb{H}$ be a function ”differentiable in the classical sense of differential geometry/multivariable analysis”. Then, we say that f is C-regular if $\forall I \in S^2 := \{q = xi + yj + zk \in$ **P** *; $|q|^2 = 1\}$, the restriction f_I of f to the set $L_I := \mathbb{R} + \mathbb{R}I$, passing through the origin, 1 and I, is holomorphic (in the classical sense of complex analysis) on $\Omega \cap L_I$.*

From now on, we will simply denote the set $S^2 = \{q = xi + yj + zk \in$ **P** *;* $|q|^2 = 1\}$ as S.

Definition 7.7 (I-derivative of a Quaternionic Function)**.** *Let $\Omega \subset \mathbb{H}$, as before, be an open and connected subset of the quaternions, and let $f : \Omega \to \mathbb{H}$ be a function ”differentiable in the classical sense of differential geometry/multivariable analysis”. Then, $\forall I \in S$, and for every quaternion $q = x + yI \in \Omega \cap L_I$, $x, y \in \mathbb{R}$ we will call $\partial_I f(q)$ the I-derivative of f at q, defined as:*

$$\partial_I f(x + yI) = \frac{1}{2}\left(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y}\right) f_I(x + yI) \tag{7.15}$$

Definition 7.8. *Let $\Omega \subset \mathbb{H}$, as before, be an open and connected subset of the quaternions, and let $f : \Omega \to \mathbb{H}$ be a left C-regular function. Then, we will call the Cullen derivative of f, which we will write as $\partial_C f$ the following expression:*

$$\partial_C f(q) = \begin{cases} \partial_I f(q) & if \;\; q = x + yI \in \Omega \cap L_I \;, y \neq 0 \\ \dfrac{\partial f}{\partial x} & if \;\; q = x \in \mathbb{R} \end{cases} \tag{7.16}$$

Corollary 7.4. *Let $f : \Omega \to \mathbb{H}$ be a function of a quaternionic variable. Then f is C-regular in Ω if and only if $\overline{\partial}_I f(q) = 0 \; \forall q \in \Omega \cap L_I$.*

Proof. It suffices to remember the classical definition of a holomorphic function in complex analysis, and combine it with definitions 7.6 and 7.7. □

C-regular functions and Cullen derivatives have numerous desirable properties; for example, by summing C-regular functions we still obtain C-regular functions (we will not state this fact in a separate proposition as it is immediately deducible from corollary 7.4). Let's state some others below:

Proposition 7.10. *Let $f : \mathbb{H} \to \mathbb{H}$ and $g : \mathbb{H} \to \mathbb{H}$ be two C-regular functions at a point $q = x + yI \in \Omega \cap L_I$, then:*

- $\partial_C(f+g)(q) = \partial_C f(q) + \partial_C g(q)$.
- *The Cullen derivative of a C-regular function is itself C-regular, i.e.* $\overline{\partial}_I(\partial_C f) = 0$.

Proof. Both cases are a matter of direct computation:

- For the first point, we observe that:

$$\partial_C(f+g) = \frac{1}{2}(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y})(f_I(x+yI) + g_I(x+yI))$$

$$= \frac{1}{2}(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y})f_I(x+yI) + \frac{1}{2}(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y})g_I(x+yI)$$

- For the second:

$$\overline{\partial}_I(\partial_C f) = \frac{1}{4}\Big(\frac{\partial}{\partial x} + I\frac{\partial}{\partial y}\Big)\Big(\frac{\partial f_I}{\partial x} - I\frac{\partial f_I}{\partial y}\Big)$$

$$= \frac{1}{4}\Big(\frac{\partial^2 f_I}{\partial x^2} + \frac{\partial^2 f_I}{\partial y^2} + I\frac{\partial^2 f_I}{\partial y \partial x} - I\frac{\partial^2 f_I}{\partial x \partial y}\Big) = 0$$

□

Exercise 7.13. *Show that polynomials of the type $f(q) = \sum_{n=0}^{m} q^n a_n$ are C-regular functions.*

Lemma 7.3. *Let I, J be two orthogonal elements of S, and let $K = IJ$; then:*

- $K = IJ = -JI \in S$.
- *K is orthogonal to both quaternions I and J.*
- $JK = I = -KJ$; $KI = J = -IK$.

Proof. • Note that, by hypothesis, I and J are orthogonal, and thus $\langle I \mid J\rangle = 0$. From this, it follows that (as seen in chapter 2):

$$IJ = \langle I \mid J\rangle + I \times J = I \times J$$

and hence $I \times J = IJ = -J \times I = -JI$.

- Being $K = IJ = I \times J$, it will, by the properties of the vector product of pure quaternions, be orthogonal to both.

- $JK = J(-JI) = -J^2 I = I$; $-KJ = -IJ^2 = I$; $KI = K(-KJ) = J = -IK = -(JK)K = J$.

□

Thus, the set $\{I, J, K\}$ of elements meeting the requirements of lemma 7.3 forms a base on S.
Moreover, the set $\{1, I, J, K\}$ also forms a base for a $\mathbb{R}$-algebra quaternionica.
Let's now demonstrate another important preparatory result, known as the "separation lemma".

Lemma 7.4 (Separation Lemma)**.** *If f is a C-regular function on $B = B(0, R)$ $\Longrightarrow \forall I, J \in S$ (with J perpendicular to I), there exist two holomorphic functions $F, G : B \cap L_I \to L_I$ such that $\forall q = x + yI$:*

$$f_I(q) = F(q) + G(q)J$$

Proof. Let I, J be a pair of orthogonal vectors in S, and let $K = IJ = -JI$ as before. Write our function f as $f = f_0 + If_1 + Jf_2 + Kf_3$. Since f is C-regular, we will have $\left(\frac{\partial}{\partial x} + I\frac{\partial}{\partial y}\right)f_I(x + yI) = 0$, that is:

$$\frac{\partial f_0}{\partial x} + I\frac{\partial f_1}{\partial x} + J\frac{\partial f_2}{\partial x} + K\frac{\partial f_3}{\partial x} + I(\frac{\partial f_0}{\partial y} + I\frac{\partial f_1}{\partial y} + J\frac{\partial f_2}{\partial y} + K\frac{\partial f_3}{\partial y})$$

Using the properties of the used base proved in lemma 7.3 we will have:

$$\frac{\partial f_0}{\partial x} - \frac{\partial f_1}{\partial y} + I\left(\frac{\partial f_0}{\partial y} + \frac{\partial f_1}{\partial x}\right) + J\left(\frac{\partial f_2}{\partial x} - \frac{\partial f_3}{\partial y}\right) + K\left(\frac{\partial f_3}{\partial x} + \frac{\partial f_2}{\partial y}\right) = 0$$

From which

$$\begin{cases} \dfrac{\partial f_0}{\partial x} = \dfrac{\partial f_1}{\partial y} \\ \dfrac{\partial f_0}{\partial y} = -\dfrac{\partial f_1}{\partial x} \\ \dfrac{\partial f_2}{\partial x} = \dfrac{\partial f_3}{\partial y} \\ \dfrac{\partial f_3}{\partial x} = -\dfrac{\partial f_2}{\partial y} \end{cases}$$

i.e the functions f_0+If_1 and f_2+If_3 are holomorphic. Now, setting $F = f_0+If_1$ and $G = f_2 + If_3$ and noting with a direct calculation that $f_I(x + yI) = F(x + yI) + G(x + yI)J$ we obtain the desired result. □

Lemma 7.5. *Let $f : B \to \mathbb{H}$ be a C-regular function. Then, $\forall n \in \mathbb{N}$, its nth Cullen derivative is equal to:*

$$\partial_C^n f(x + yI) = \frac{\partial^n f}{\partial x^n}(x + yI)$$

Proof. We proceed by induction: for the base case $n = 1$ we see that:

$$\partial_C f(x + yI) = \frac{1}{2}\Big(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y}\Big)f(x + yI) = \frac{1}{2}\Big(\frac{\partial f}{\partial x} - I\frac{\partial f}{\partial y}\Big)(x + yI) = \frac{\partial f}{\partial x}(x + yI)$$

where here the last equality is a consequence of the Cauchy-Riemann equations $\frac{\partial f}{\partial x} = -I\frac{\partial f}{\partial y}$.
Now suppose it is true for a generic $n = k$, i.e.:

$$\partial_C^k f(x + yI) = \frac{\partial^k f}{\partial x^k}(x + yI)$$

to prove the case $n = k + 1$, observe that, since f is C-regular:

$$\Big(\frac{\partial}{\partial x} + I\frac{\partial}{\partial y}\Big)\Big(\frac{\partial^k f}{\partial x^k}\Big) = \frac{\partial^{k+1} f}{\partial x^{k+1}} + I\frac{\partial^{k+1} f}{\partial x^k \partial y} = \frac{\partial^k f}{\partial x^k}\Big(\frac{\partial f}{\partial x} + I\frac{\partial f}{\partial y}\Big) = 0$$

from which:

$$\frac{\partial^{k+1} f}{\partial x^{k+1}} = -I\frac{\partial^{k+1} f}{\partial x^k \partial y} \tag{7.17}$$

And so, using what was seen before and the inductive hypothesis for the case $n = k$:

$$\partial_C^{k+1} f = \partial_C\Big(\frac{\partial^k f}{\partial x^k}\Big) = \frac{1}{2}\Big(\frac{\partial}{\partial x} - I\frac{\partial}{\partial y}\Big)\Big(\frac{\partial^k f}{\partial x^k}\Big) = \frac{1}{2}\Big(\frac{\partial^{k+1} f}{\partial x^{k+1}} - I\frac{\partial^{k+1} f}{\partial x^k \partial y}\Big)$$

$$= \frac{1}{2}\left(2\frac{\partial^{k+1} f}{\partial x^{k+1}}\right) = \frac{\partial^{k+1} f}{\partial x^{k+1}}$$

(where here $\frac{\partial^{k+1} f}{\partial x^{k+1}} - I\frac{\partial^{k+1} f}{\partial x^{k}\partial y} = 2\frac{\partial^{k+1} f}{\partial x^{k+1}}$ due to equation (7.17)).

□

We can expand C-regular functions into an infinite power series, let's show the proof of this fact below:

Theorem 7.3 (G.Gentili D.C Struppa)**.** *A function $f : B(0, R) \to \mathbb{H}$ is C-regular if and only if it has a series expansion of the type:*

$$f(q) = \sum_{n=0}^{\infty} \frac{q^n}{n!}\frac{\partial^n f}{\partial x^n}(0) \tag{7.18}$$

Proof. Let Δ_I be the disk of the complex plane L_I with radius $a > 0$ smaller than R (the radius of the ball B). By lemma 7.4, we have:

$$f(q) = F(q) + G(q)J$$

where F, G are holomorphic on $B \cap L_I$. Moreover, for $\zeta, q \in B \cap L_I$, $\zeta \neq q$ we have

$$(\zeta - q)^{-1}F(q) = F(q)(\zeta - q)^{-1} = \frac{F(q)}{\zeta - q}$$

$$(\zeta - q)^{-1}G(q) = G(q)(\zeta - q)^{-1} = \frac{G(q)}{\zeta - q}$$

By the Cauchy integral formula, we have that $\forall q \in \Delta_I$:

$$f_I(q) = \frac{1}{2\pi I}\oint_{\partial\Delta_I} \frac{F(\zeta)}{\zeta - q}d\zeta + \left(\frac{1}{2\pi I}\oint_{\partial\Delta_I} \frac{G(\zeta)}{\zeta - q}d\zeta\right)J \tag{7.19}$$

Now, let's transform the 2 integrals into infinite series as follows:

$$\oint_{\partial\Delta_I} \frac{F(\zeta)}{\zeta - q}d\zeta = \oint_{\partial\Delta_I} \frac{1}{1 - \frac{q}{\zeta}}\frac{F(\zeta)}{\zeta}d\zeta = \oint_{\partial\Delta_I} \sum_{n=0}^{\infty}\left(\frac{q}{\zeta}\right)^n \frac{F(\zeta)}{\zeta}d\zeta = \sum_{n=0}^{\infty} q^n \oint_{\partial\Delta_I} \frac{F(\zeta)}{\zeta^{n+1}}d\zeta$$

and similarly for G:

$$\oint_{\partial\Delta_I} \frac{G(\zeta)}{\zeta - q}d\zeta = \oint_{\partial\Delta_I} \frac{1}{1 - \frac{q}{\zeta}}\frac{G(\zeta)}{\zeta}d\zeta = \oint_{\partial\Delta_I} \sum_{n=0}^{\infty}\left(\frac{q}{\zeta}\right)^n \frac{G(\zeta)}{\zeta}d\zeta = \sum_{n=0}^{\infty} q^n \oint_{\partial\Delta_I} \frac{G(\zeta)}{\zeta^{n+1}}d\zeta$$

But by the generalized Cauchy integral formula, and by the definition of the Cullen derivative we have:

$$\sum_{n=0}^{\infty} q^n \oint_{\partial\Delta_I} \frac{F(\zeta)}{\zeta^{n+1}} d\zeta = \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n F(0)$$

$$\sum_{n=0}^{\infty} q^n \oint_{\partial\Delta_I} \frac{G(\zeta)}{\zeta^{n+1}} d\zeta = \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n G(0)$$

Substituting everything into equation (7.19), and applying the linearity of the Cullen differential operator ∂_C we get:

$$f_I(q) = \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n F(0) + \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n G(0) J$$

$$= \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n (F + GJ)(0) = \sum_{n=0}^{\infty} q^n \frac{1}{n!} \partial_C^n f(0)$$

Now, using lemma 7.5 (i.e $\partial_C^n f(0) = \frac{\partial^n f}{\partial x^n}(0)$) we get:

$$f_I(q) = \sum_{n=0}^{\infty} \frac{q^n}{n!} \frac{\partial^n f}{\partial x^n}(0)$$

We have thus obtained the desired result. Moreover, since the coefficients $\frac{1}{n!}\frac{\partial^n f}{\partial x^n}(0)$ are independent of the choice of $I \in S$, this expansion will hold $\forall I \in S$. □

C-regular functions, moreover, satisfy a quaternionic analogue of the identity principle

Theorem 7.4 (Identity principle). *Let $f : B \to \mathbb{H}$ be a regular function. Let $Z_f := \{q \in B \mid f(q) = 0\}$ be the zero set of f (i.e the set of values for which the function yields 0). If $\exists I \in S$ for which $L_I \cap Z_f$ has an accumulation point, then $f = 0$ on all of B.*

Proof. By the splitting lemma, on $L_I \cap B$, we can write:

$$f(x + yI) = \phi_1(x + yI) + \phi_2(x + yI)J$$

where ϕ_1 and ϕ_2 are holomorphic functions on L_I. If $L_I \cap Z_f$ has an accumulation point, by the classical identity principle for holomorphic functions we will have that ϕ_1 and ϕ_2 will be identically 0 on $L_I \cap B$. This implies that the partial derivatives $\frac{\partial^n f(0)}{\partial x^n} = 0$ for all $n \in \mathbb{N}$, and thus f will be identically 0 on B, since the partial derivatives we considered are the coefficients of the power series expansion we derived in the previous theorem. □

From this result we get the following immediate corollary; in the literature the following is also often referred to as the **identity principle**.

Corollary 7.5. *Let f and g be quaternionic C-regular functions on a ball B. If $\exists I \in S$ such that $f = g$ on a subset of $L_I \cap B$ having an accumulation point in $L_I \cap B$ (i.e if f and g "agree" on a subset with an accumulation point in $L_I \cap B$), then $f = g$ everywhere on B.*

Proof. This follows immediately by taking h to be $h := f - g$. This is again a C-regular function, since it is the difference of 2 C-regular functions; we can then apply the previous theorem to h and get our desired result. □

C-regular functions also satisfy their analogue of Cauchy's integral formula, which we will prove and state now:

Theorem 7.5 (Cauchy integral formula for C-regular functions, G.Gentili D.C Struppa)**.** *Let $f : B(0,R) \to \mathbb{H}$ be a C-regular function and let $q \in B$. Then:*

$$f(q) = \frac{1}{2\pi I_q} \oint_{\partial\Delta_q(0,r)} \frac{d\zeta}{(\zeta - q)} f(\zeta) \tag{7.20}$$

where here I_q is defined as:

$$I_q := \begin{cases} \dfrac{\mathrm{Im}(q)}{|\,\mathrm{Im}(q)|} \in S \quad if \quad \mathrm{Im}(q) \neq 0 \\ I \in S \; any \quad if \quad \mathrm{Im}(q) = 0 \end{cases}$$

$\zeta \in L_{I_q} \cap B(0,R)$ and $r > 0$ is a positive real number selected in a way such that:

$$q \in \overline{\Delta_q(0,r)} := \{x + yI_q \; ; \; x^2 + y^2 \leq r^2\} \subset B(0,R)$$

i.e such that $\overline{\Delta_q(0,r)}$, the closure of $\Delta(0,r)$, is contained in the ball $B(0,R)$ and contains q.

Proof. Let's start the proof by manipulating the expression on the right side of the equality (7.20):

$$\frac{1}{2\pi I_q} \oint_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q} f(\zeta) = \frac{1}{2\pi I_q} \oint \frac{d\zeta}{\zeta - q} f_{I_q}(\zeta)$$

But by lemma 7.4 (separation lemma) we have:

$$\frac{1}{2\pi I_q} \oint_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q} f_{I_q}(\zeta) = \frac{1}{2\pi I_q} \oint_{\partial\Delta_q(0,r)} \frac{d\zeta}{\zeta - q} (F(\zeta) + G(\zeta)J)$$

$$= \frac{1}{2\pi I_q}\oint_{\partial\Delta_q(0,r)} \frac{F(\zeta)}{\zeta - q}d\zeta + \left(\frac{1}{2\pi I_q}\oint_{\partial\Delta_q(0,r)} \frac{G(\zeta)}{\zeta - q}d\zeta\right)J$$

and by the Cauchy integral formula:

$$\frac{1}{2\pi I_q}\oint_{\partial\Delta_q(0,r)} \frac{F(\zeta)}{\zeta - q}d\zeta + \left(\frac{1}{2\pi I_q}\oint_{\partial\Delta_q(0,r)} \frac{G(\zeta)}{\zeta - q}d\zeta\right)J = F(q) + G(q)J = f(q)$$

□

Theorem 7.6 (Cauchy Estimate)**.** *Let $f : B(0,R) \to \mathbb{R}$ be a C-regular function, and let $r < R$ be a positive real number smaller than R, $I \in S$ and $\partial\Delta_I(0,r) = \{x + yI\ ;\ x^2 + y^2 = r^2\}$. If $M_I = \max\{|f(q)|\ :\ q \in \partial\Delta_I(0,r)\}$ and if $M = \inf\{M_I\ ;\ I \in S\}$ then*

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \le \frac{M}{r^n} \qquad n \in \mathbb{N}$$

Proof. As seen in the proof of the previous theorem, we have (using jointly the generalized Cauchy integral formula and lemma 7.5)

$$\frac{1}{n!}\frac{\partial^n f}{\partial x^n} = \frac{1}{2\pi I}\oint_{\partial\Delta_I(0,r)} \frac{d\zeta}{\zeta^{n+1}} f(\zeta)$$

and then, through the following sequence of inequalities we obtain:

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| = \left|\frac{1}{2\pi I}\oint_{\partial\Delta_I(0,r)} \frac{d\zeta}{\zeta^{n+1}} f(\zeta)\right| \le \frac{1}{2\pi}\oint_{\partial\Delta_I(0,r)} \frac{M_I\, d\zeta}{r^{n+1}}$$

$$= \left(\frac{M_I}{2\pi r^{n+1}}\right)(2\pi r) = \frac{M_I}{r^n}$$

$\forall I \in S$, from which it follows that it will also be valid for M, obtaining the desired result:

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \le \frac{M}{r^n}$$

□

The fact just proved is important since we can prove an analogue of Liouville's theorem for C-regular functions as an immediate corollary of it:

Corollary 7.6 (Liouville's theorem for C-regular functions)**.** *Let $f : \mathbb{H} \to \mathbb{H}$ be a C-regular function at every point in $\mathbb{H}$. Then, if f is bounded in modulus (that is, there exists a positive real number M such that $|f(q)| \le M$ $\forall q \in \mathbb{H}$), f is constant.*

Proof. By the previous theorem (Cauchy's Estimate), we have $\forall r \in \mathbb{R}^+$:

$$\frac{1}{n!}\left|\frac{\partial^n f}{\partial x^n}(0)\right| \leq \frac{M}{r^n}$$

For $r \to \infty$, we will obtain that $\dfrac{\partial^n f}{\partial x^n}(0) = 0\ \forall n \in \mathbb{N}$, and from this, it follows that f will be a constant function, $f(q) = f(0)\ \forall q \in \mathbb{H}$. □

Moreover, an analogue of Morera's theorem is valid for C-regular functions, which can be proved as an immediate consequence of the very definition of C-regular function, let's state and prove it here below:

Theorem 7.7 (Morera's theorem for C-regular functions)**.** *Let $f : B(0,R) \to \mathbb{H}$ be a function "differentiable in the classical sense". If $\forall I \in S$ the differential form defined on $L_I \cap B(0,R)$, $f(z)dz$ (with $z = x + yI$ and $x, y \in \mathbb{R}$) is closed, then the function f is C-regular.*

Proof. By hypothesis, we have that $d(f(z)dz) = 0$, from which, by the classical Morera's theorem we have that $\forall I\ f_I(z)$ will be a holomorphic function. From this, by the definition of C-regular function, it follows that the function f is C-regular. □

Chapter 8

Notes

8.1 Introduction

- **Link 1**: *https://upload.wikimedia.org/wikipedia/commons/thumb/d/d5/Inscription_on_Broom_Bridge_%28Dublin%29_regarding_the_discovery_of_Quaternions_multiplication_by_Sir_William_Rowan_Hamilton.jpg/1920px-Inscription_on_Broom_Bridge_%28Dublin%29_regarding_the_discovery_of_Quaternions_multiplication_by_Sir_William_Rowan_Hamilton.jpg*

- **Note 1**: "B.0. Rodrigues, Des lois geometriques qui regissent les deplacements d'un systeme solide dans l'espace, et la variation des coordonnees provenant de ses deplacements consideres independamment des causes qui peuvent les produire, I. de Mathematiques Pures et Appliquees 5 (1840), 380-440"

- **Note 2**: R. P. Graves, "Life of Sir William Rowan Hamilton".

- **Note 3**: Kline, Morris. "A history of mathematical thought", volume 1. p. 253.

- **Note 4**: A. Cauchy "Mémoire sur les intégrales définies, prises entre des limites imaginaires" (1825).

- **Note 5**: Robert Argand, Essai sur une manière de représenter des quantités imaginaires dans les constructions géométriques, 2e édition, Gauthier Villars, Paris (1874).
  Whittaker, E. T.; Watson, G. N. (1927). A Course in Modern Analysis (Fourth ed.). Cambridge University Press.

- **Note 6**: ”Encyclopaedia Britannica” 11esima edizione.
  S.L Altmann ”Hamilton, Rodrigues, and the Quaternion Scandal” (1989).

- **Note 7**: P. Molenbroek, Shunkichi Kimura ”To Friends and Fellow Workers in Quaternions”(1895).

8.2 Chapter 1: Preliminary notions

- **Note 1.1**: Königsberg (today Kaliningrad, Russia) is crossed by the Pregel River and its tributaries and has two large islands that are connected to each other and with the two main areas of the city by seven bridges. The problem of the seven bridges of Königsberg is as follows: is it possible to take a walk following a path that crosses each bridge only once? The problem was solved in 1736 by Euler, who provided a negative answer to this question. The resolution of the problem, which was achieved by Euler by simplifying the problem in abstract terms, made people realize the importance of studying qualitative geometric properties of spaces, in contrast to quantitative properties.

8.3 Chapter 2: Quaternion algebra

- **Note 2.1**: ≪Tensors or signless numbers, such as $2, 3, 6, \frac{1}{2}, \sqrt{2}$, which operate only metrically on the lengths of the lines which they multiply, and which are to be combined among themselves, as factors, by arithmetical multiplication, or by the laws of the composition of ratios≫ Lectures on quaternions (1853), W.R Hamilton.

- **Note 2.2**: Leonhard Euler: Life, Work and Legacy, R.E. Bradley and C.E. Sandifer (eds), Elsevier, 2007, p. 193
  The theorem was proved by Lagrange and it is commonly known with the name of ”theorem of 4 squares” or ”Bachet's conjecture”.

- **Note 2.3**: The theorem of four squares by Lagrange was already present in the work ”Arithmetica” by Diophantus of Alexandria, which was translated in the early 1600s by Bachet (hence the common name for referring to the theorem in question, also called ”Bachet's conjecture”). The first proof of this theorem was provided by Lagrange in 1770 by exploiting Euler's four-square identity. Below is an

English translation of ”Arithmetica”:
https://archive.org/details/diophantusofalex00heatiala

- **Note 2.4**: ≪We propose to call the quotient, or the versor, thus obtained, the versor-element, or briefly, the *versor of the quaternion q*; and shall find it convenient to employ the same characteristic, U, to denote the operation of taking the versor of a quaternion, as that employed above to denote the operation of reducing a vector to the unit of length, without any change of its direction. On this plan, the symbol Uq will denote the versor of q≫
Hamilton, W. R. Elements of quaternions(London, Longmans Green, 1866), article 156 pg.135-136.

- **Note 2.5**: ≪The positive or negative quotient, $x = \frac{\beta}{\alpha}$, which is thus obtained by the division of one of two parallel vectors by another, including zero as a limit, may also be called a scalar; because it can always be found, and in a certain sense constructed, by the comparison of positions upon one common scale (or axis) [...] Such scalars are, therefore, simply the reals (or the real quantities) of Algebra≫
Hamilton, W. R. Elements of quaternions(London, Longmans Green, 1866), Book 1, chapter 2, article 17.

8.4 Chapter 4: Topology of quaternions

- **Nota 4.1**: P.S Alexandroff and H.Hopf, Topologie, Berlin, 1935 p. 477.

8.5 Chapter 5: Elementary functions

- **Note 5.1**:
  - Exponential function: https://www.youtube.com/watch?v=qWsIWGOvvsM
  - Hyperbolic cosine: https://www.youtube.com/watch?v=DYbFyoV0nJw
  - Hyperbolic sine: https://www.youtube.com/watch?v=OPM25hpbwmU
  - Sine: https://www.youtube.com/watch?v=sH_OmP9s_Dk

8.6 Chapter 6: Differential and integral calculus on $\mathbb{H}$

- **Note 6.1**: We are referring to the following theorem:

  Theorem 8.1 (Hartogs' Theorem)**.** *Let $U \subset \mathbb{C}^n$ be an open set and let $f : U \to \mathbb{C}$ be a function such that, for every $i \in [1, n] \cap \mathbb{N}$, for fixed $(n-1)$-tuples $z_1, z_2, \ldots, z_{i-1}, z_{i+1}, \ldots, z_n$, the function*

  $$w \to f(z_1, z_2, \ldots, z_{i-1}, w, z_{i+1}, \ldots, z_n)$$

  is analytic on:

  $$\{w \in \mathbb{C} \,; (z_1, \ldots, z_{i-1}, w, z_{i+1}, \ldots, z_n) \in U\}$$

  then f is continuous on U.

- **Note 6.2**: We are referring to the following theorem:

  Theorem 8.2. *Let U be an open set in $\mathbb{C}^n$, K a compact space, and μ a Radon measure on K. If the function $(x, y) \to f(x, y)$ is continuous on $U \times K$, and the function $x \to f(x, y)$ is holomorphic on U for every $y \in K$, then the function*

  $$\phi(x) = \int_K f(x, y) d\mu(y)$$

  is holomorphic in U.

Bibliography

[1] Simon L. Altmann. “Hamilton, Rodrigues, and the Quaternion Scandal”. In: *Mathematics Magazine* 62 (1989), pp. 291–308.

[2] Douglas B.Sweetser. “Doing Physics with Quaternions”. In: 2005.

[3] Harry Bateman. “The Solution of Partial Differential Equations by Means of Definite Integrals”. In: *Proceedings of The London Mathematical Society* (), pp. 451–458.

[4] Henri Paul Cartan. “Elementary Theory of Analytic Functions of One or Several Complex Variables”. In: 1963.

[5] Charles G. Cullen. “An integral theorem for analytic intrinsic functions on quaternions”. In: *Duke Mathematical Journal* 32 (1965), pp. 139–148.

[6] Cipher A. Deavours. “The Quaternion Calculus”. In: *American Mathematical Monthly* 80 (1973), pp. 995–1008.

[7] Omar P. Dzagnidze. “On Some New Properties of Quaternion Functions”. In: *Journal of Mathematical Sciences* 235 (2018), pp. 557–603.

[8] Omar P. Dzagnidze. “On the Differentiability of Quaternion Functions”. In: *arXiv: Complex Variables* (2012).

[9] A.D. Snyder E.B. Saff. “Fundamentals of Complex Analysis, with Applications to Engineering and Science”. In: 1987.

[10] Michael Eastwood. “Bateman’s formula”. In: (2002).

[11] Samuel Eilenberg and Ivan Morton Niven. “The “fundamental theorem of algebra” for quaternions”. In: *Bulletin of the American Mathematical Society* 50 (1944), pp. 246–248.

[12] V. C. A. Ferraro. “On Functions of Quaternions”. In: *Proceedings of the Royal Irish Academy. Section A: Mathematical and Physical Sciences* 44 (1937), pp. 101–108.

[13] Rudolf Fueter. “Über die analytische Darstellung der regulären Funktionen einer Quaternionenvariablen”. In: *Commentarii Mathematici Helvetici* 8 (1935), pp. 371–378.

[14] Rudolf Fueter. “Zur Theorie der Brandtschen Quaternionenalgebren”. In: *Mathematische Annalen* 110 (1935), pp. 650–661.

[15] Rudolf Fueter. “Zur Theorie der regulären Funktionen einer Quaternionenvariablen”. In: *Monatshefte für Mathematik und Physik* 43 (1936), pp. 69–74.

[16] Run Fueter. “Die Funktionentheorie der DifferentialgleichungenΔu=0 und$\Delta\Delta$u=0 mit vier reellen Variablen”. In: *Commentarii Mathematici Helvetici* 7 (1934), pp. 307–330.

[17] Stefania Gabelli. “Teoria delle Equazioni e Teoria di Galois”. In: Springer, 2008.

[18] Jean Gallier. “Geometric Methods and Applications”. In: Springer, 2001.

[19] Graziano Gentili and Daniele C. Struppa. “A New Theory of Regular Functions of a Quaternionic Variable”. In: *Advances in Mathematics* 216 (2007), pp. 279–301.

[20] H.Cartan. “Differential Calculus on Normed Spaces: A Course in Analysis”. In: Kershaw Publishing Company, 1971.

[21] Hans Georg Haefeli. “Hyperkomplexe Differentiale”. In: *Commentarii Mathematici Helvetici* 20 (1947), pp. 382–420.

[22] William Rowan Sir Hamilton. “Elements of Quaternions”. In: 1969.

[23] Allen Hatcher. “Algebraic Topology”. In: Cambridge University Press, 2002.

[24] A.S. Solodovnikov I.L Kantor. “Hypercomplex numbers”. In: Nauka, 1973.

[25] Andrew J.Hanson. “Visualizing Quaternions”. In: 2005.

[26] Yan-Bin Jia. “Quaternions and Rotations”. In: 2013.

[27] Wolfgang Sprößig João Pedro Morais Svetlin Georgiev. *Real Quaternionic Calculus Handbook*. Birkhäuser, 2014.

[28] Aleks Kleyn. “Derivative of Map of Banach algebra”. In: *arXiv: General Mathematics* (2015).

[29] Aleks Kleyn. “Introduction into Calculus over Banach algebra”. In: *arXiv: General Mathematics* (2016).

[30] John M. Lee. "Introduction to Smooth Manifolds". In: 2002.

[31] Eugene Lerman. "An introduction to Differential Geometry". In.

[32] Seymour Lipschutz. "Schaum's outline of theory and problems of general topology". In: 1965.

[33] M.Hervé. "Several Complex Variables: Local Theory". In: 1987.

[34] John M.Howie. "Real Analysis". In: Springer, 2001.

[35] Prerna Nadathur. "An Introduction to Homology". In: 2007.

[36] Ivan Morton Niven. "Equations in Quaternions". In: *American Mathematical Monthly* 48 (1941), pp. 654–661.

[37] L.S Pontryagin. "Foundations of Combinatorial Topology". In: Graylock Press, 1952.

[38] Carlo Presilla. "Elementi di Analisi Complessa". In: Springer, 2013.

[39] Walter Rudin. "Principles of mathematical analysis". In: 1964.

[40] James Ward Brown Ruel V.Churchill. "Complex variables and applications". In: McGraw-Hill, 1990.

[41] Takis Sakkalis, Kwang-mo Ko, and Galam Song. "Roots of quaternion polynomials: Theory and computation". In: *Theor. Comput. Sci.* 800 (2019), pp. 173–178.

[42] Christopher Stover. "A Survey of Quaternionic Analysis". In: 2014.

[43] A. Sudbery. "Quaternionic analysis". In: *Mathematical Proceedings of the Cambridge Philosophical Society* 85.2 (1979), pp. 199–225. DOI: 10.1017/S0305004100055638.

[44] M. Sugiura. "Unitary Representations and Harmonic Analysis". In: 1990.

[45] Edward Charles Titchmarsh. "The theory of functions". In: 1933.

[46] N. S. Topuridze. "On roots of quaternion polynomials". In: *Journal of Mathematical Sciences* 160 (2009), pp. 843–855.

[47] N. S. Topuridze. "On the Roots of Polynomials over Division Algebras". In: 2003.

[48] Alan Pollack Victor Guillemin. "Differential Topology". In: AMS Chelsea Publishing, 1974.

[49] Andrew H. Wallace. "An introduction to Algebraic Topology". In: Pergamon Press, 1957.

[50] Edmund Taylor Whittaker. “On the partial differential equations of mathematical physics”. In: *Mathematische Annalen* 57 (1903), pp. 333–355.