Quaternions

Adrien Deloro
Şirince '21 Summer School

These are lecture notes with exercises for a two week course of $10 \times 2 = 20$ hours given in July 2021 at the Nesin Matematik Köyü in Şirince, Turkey. They should be accessible to a 3rd year student in mathematics; syllabus and prerequisites are described below.

If found with flaws or mistakes, this document compiled on 5th August 2021 should be returned to adrien.deloro@imj-prg.fr. Please mention the date.

Contents

Each section corresponds to a lecture of 2 hours, with the introduction fitting into § 1. The class can be taught in a different order as there are independences between lectures. Lectures § 1, § 2 (which can be made shorter) and § 4 form a first block; §§ 8–10 form another block. Lecture § 3 may be seen before or after § 4. Each of lectures § 5, § 6, § 7, and the block §§ 8–10, can be presented at any moment after § 4.

Introduction

Disclaimer. This is not really a course on quaternions, as the author does not care for quaternions. It is a review of elementary notions and methods in algebra, linear algebra, and geometry, together with an invitation to basic Lie theory. Quaternions are a pretext.

Hopefully you know what a field is. A *skew-field* is almost like a field, except that multiplication is *not* required to be commutative (addition remains commutative). Historically the first example of a skew-field was discovered by Hamilton; it is $\mathbb{H}$, the skew-field of *real quaternions*, to which this two-week course is devoted. Remarkably, $\mathbb{H}$ is a real vector space of dimension 4 with a certain magical multiplication. The algebraic structure on $\mathbb{R}^4$ actually encodes rotations of the Euclidean space $\mathbb{R}^3$, which makes quaternions useful when rotating computer-generated 3D-images.

Technically, quaternions are real linear combinations of the form $a\cdot 1+b\cdot i+c\cdot j+d\cdot k$, where the non-trivial relations are:

$$i^2 = j^2 = k^2 = -1; \quad ij = k; \quad ji = -k; \quad jk = i; \quad \text{etc.}$$

there is a notion of conjugate, namely $(a + bi + cj + dk)^* = a - bi - cj - dk$. We urge the reader to start manipulating quaternions now; in particular, they form a skew-field.

Hamilton had started with the following key observation.

Complex numbers can be used to code rotations in the vector space $\mathbb{R}^2$.

Namely, multiplying by complex number $z = e^{i\theta}$ amounts to rotating the real vector space $\mathbb{C} \simeq \mathbb{R}^2$ by θ. Complex numbers of this form are on the circle $\mathbb{S}^1 = \{z \in \mathbb{C} : |z| = 1\} \simeq \mathrm{SO}_2(\mathbb{R})$. Not all complex numbers have norm 1, but there is therefore a field 'enveloping' the group $\mathrm{SO}_2(\mathbb{R})$. It was then natural to try to generalise to $\mathbb{R}^3$.

Question. Is there a field structure which can be used for coding rotations of $\mathbb{R}^3$?

Since the group $\mathrm{SO}_3(\mathbb{R})$ (studied in § 9) of rotations of $\mathbb{R}^3$ is *not* commutative, this requires leaving commutative mathematics. To Hamilton's despair, he could *not* find a nice algebraic structure on $\mathbb{R}^3$ compatible with multiplication in $\mathrm{SO}_3(\mathbb{R})$. But to his surprise, he could in $\mathbb{R}^4$. Quaternions were born.

Prerequisites

The class is supposedly accessible to a third year student.

Analysis: Almost none. Real numbers are required but we never use their analytic properties. Instead we focus on *real closed fields*, which will be defined. The few topological arguments are all optional.

Algebraic structures: In theory there are little prerequisites; in practice, maturity is required. One needs to be very comfortable with general notions such as associativity, commutativity; groups, rings, domains, fields, and algebras over fields; morphisms and factoring kernels. Conjugation actions play an important role throughout. The first lecture could prove challenging to beginners in abstract algebra but things get better in § 2.

Geometry: Preliminar knowledge of the cross/wedge product $\times$ in $\mathbb{R}^3$ is recommended for § 8. Familiarity with rotations of the plane and of the space ($\mathrm{SO}_2(\mathbb{R})$, $\mathrm{SO}_3(\mathbb{R})$) and orientations of planes in $\mathbb{R}^3$ will also help, though all are recalled in § 9. Euler angles are *not* discussed in this class.

Linear algebra: One must know matrices, and eigenvalues. The orthogonal group is important. At some optional point, the definition of unitary groups is required.

Number theory: For one striking though optional application of quaternions in § 7, the notion of a Euclidean ring is needed.

1 Fields, skew-fields and algebras

This lecture provides background on fields, skew-fields, and algebras. Hopefully all definitions are already familiar. On skew-fields we say little, because skew-field theory is technically too challenging for an undergraduate course. On algebras there will be three main results to remember.

1. Only *commutative* fields can have non-trivial algebras (lemma 1.2.3).
2. Every associative, unital $\mathbb{K}$-algebra $\mathbb{A}$ embeds in some $\mathrm{End}_{\mathbb{K}}(V)$ (lemma 1.2.7, 'Wedderburn's representation theorem').
3. In this case, $\mathbb{K}$ embeds in the centre of $\mathbb{A}$ (lemma 1.2.10, 'central embedding').

1.1 Fields and skew-fields

1.1.1. Definition.

- A *field* $(\mathbb{F}; +, \cdot)$ is an associative, commutative, unital ring (= with 1) such that every element has a multiplicative inverse.
- A *skew-field* $(\mathbb{K}; +, \cdot)$ is an associative, unital ring such that every element has a 2-sided multiplicative inverse.

In order to stress the difference, we use the redundant phrase 'commutative field'.

1.1.2. Examples. The following are fields (under expected operations):

- the fields of rational numbers $\mathbb{Q}$, of real numbers $\mathbb{R}$, of complex numbers $\mathbb{C}$;
- for any field $\mathbb{K}$, the field of rational fractions $\mathbb{K}(X)$;
- for any prime p, the field with p-elements $\mathbb{F}_p \simeq \mathbb{Z}/p\mathbb{Z}$;
- it is a theorem by Galois that for any prime power $q = p^k$, there exists a unique (up to isomorphism) field of order q.

1.1.3. Examples. Quaternions form a skew-field denoted by $\mathbb{H}$.

The second list of examples is shorter. Of course there are other skew-fields; but they are fairly complicated objects. A whole book has been written on how to produce skew-fields[1]. Even the construction of a skew-field structure on $\mathbb{Q}^9$ takes a couple of pages.

Hence 'in nature' only quaternions appear without too many technicalities. There are two reasons.

1. In addition to the expected $\mathbb{R}$ and $\mathbb{C}$, there is only one skew-field which is finite-dimensional over the reals.

   Theorem (Frobenius). Let $\mathbb{A}$ be a skew-field which is finite-dimensional over the reals. Then $\mathbb{A} \simeq \mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$.

   We shall prove this theorem in § 6. It says that from the point of view of geometry, the only natural non-commutative skew-field is the one found by Hamilton.

2. Skew-fields do not arise at the finite level.

   Theorem (Wedderburn). Every finite skew-field is actually commutative.

   We shall *not* prove the latter. It says that from the point of view of discrete mathematics, the only natural skew-fields are commutative, and those found by Galois.

[1] Cohn, P., *Skew fields, theory of general division rings*, Encyclopedia of Mathematics and its Applications, vol. 57, Cambridge University Press, 1977

These explain why if you want another skew-field than the quaternions, you have to leave 'basic' mathematics. This is another indication that quaternions form a particularly robust structure, one which has to be investigated at least once in the course of your mathematical studies.

Vector spaces over skew-fields

If $\mathbb{K}$ is a skew-field, one can define left-vector spaces over $\mathbb{K}$ by the usual definition, which does not require commutativity.

Elementary linear algebra can be carried, and there still is dimension theory (existence of bases, well-definedness of the dimension). Of course other tools such as the determinant are lost.

One can *also* define right-vector spaces, letting scalars act from the right. In the commutative case, a left- and right-structure coincide; not so if $\mathbb{K}$ is a skew-field, because $\lambda \cdot (\mu \cdot x) = (\lambda\mu) \cdot x$ need not equal $(x \cdot \mu) \cdot \lambda = x \cdot (\mu\lambda)$. So when V is a left- and right-vector space over $\mathbb{K}$, the two linear structures can disagree strongly.

1.1.4. Remark. Left- and right-dimension theories need not agree. § 5.9 of Cohn's book contains the following result solving an old question by E. Artin.

Theorem (Schoefield, 1985). Let λ, ρ be two cardinals, finite or infinite, but both > 1. Then there exist skew-fields $\mathbb{K}_1 < \mathbb{K}_2$ such that the left-dimension of $\mathbb{K}_2$ over $\mathbb{K}_1$ is λ while its right-dimension is ρ.

(Nothing such can happen if $\mathbb{K}_1$ is commutative.)

All this shows that general skew-field theory is complicated beyond reason; we shall focus on quaternions and not discuss other skew-fields.

1.2 Algebras over fields; three lemmas

Hopefully the following is familiar as well.

1.2.1. Definition. Let $\mathbb{K}$ be a (commutative) field. An *algebra* over $\mathbb{K}$ is a $\mathbb{K}$-vector space $\mathbb{A}$ equipped with an inner bilinear map $\mathbb{A} \times \mathbb{A} \to \mathbb{A}$. Its *dimension* is its dimension as $\mathbb{K}$-vector space, denoted by $\dim_{\mathbb{K}} \mathbb{A}$.

Be careful that a 'dot product', e.g. $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, does *not* make $\mathbb{R}^n$ into a $\mathbb{R}$-algebra: the bilinear map must take values in $\mathbb{A}$, not in $\mathbb{K}$.

1.2.2. Remark. In this general definition, $\mathbb{A}$ is not assumed to be commutative *nor even associative* (see remark 1.2.5 below).

Only commutative fields have non-trivial algebras

The definition of an algebra may look too narrow as we restricted ourselves to a *commutative* base field. Here is why.

1.2.3. Lemma (proper skew-fields have no algebras)**.** *If $\mathbb{K}$ is a skew-field admitting a $\mathbb{K}$-algebra with non-zero multiplication, then $\mathbb{K}$ is actually commutative.*

Proof. Suppose that $\mathbb{K}$ is any skew-field, and that the product map on $\mathbb{A}$ is non-trivial. Let $\lambda, \mu \in \mathbb{K}$ and $a_1, a_2 \in \mathbb{A}$. For enhanced clarity we denote by $\lambda\mu$ the field product,

by $\lambda \cdot a$ the scalar action, and by $a_1 * a_2$ the inner algebra product.

Since the product map on $\mathbb{A}$ is non-trivial, we may assume $a_1 * a_2 \neq 0$. Computing with this pair one has for any $\lambda, \mu \in \mathbb{K}$:

$$\begin{aligned}
(\lambda\mu) \cdot (a_1 * a_2) &= \lambda \cdot (\mu \cdot (a_1 * a_2)) && \text{(def. of the scalar action)} \\
&= \lambda \cdot ((\mu \cdot a_1) * a_2) && \text{(left-linearity of } *) \\
&= (\mu \cdot a_1) * (\lambda \cdot a_2) && \text{(right-linearity of } *) \\
&= \mu \cdot (a_1 * (\lambda \cdot a_2)) && \text{(left-linearity of } *) \\
&= \mu \cdot (\lambda \cdot (a_1 * a_2)) && \text{(right-linearity of } *) \\
&= (\mu\lambda) \cdot (a_1 * a_2) && \text{(def. of the scalar action).}
\end{aligned}$$

Therefore $\lambda\mu = \mu\lambda$ for *any* λ and μ, which is commutativity of $\mathbb{K}$. □

Representation of associative, unital algebras

During undergraduate years, one tends to focus on 'associative' structures.

1.2.4. Definition. Let $\mathbb{A}$ be a $\mathbb{K}$-algebra.

- $\mathbb{A}$ is *associative* if its inner product $\mathbb{A} \times \mathbb{A} \to \mathbb{A}$ is associative, viz. satisfies the axiom $(\forall a, b, c \in \mathbb{A})[a(bc) = (ab)c]$.
- $\mathbb{A}$ is *unital* if there is an identity element $1_{\mathbb{A}}$, viz. with $(\forall a \in \mathbb{A})(a \cdot 1_{\mathbb{A}} = 1_{\mathbb{A}} \cdot a = a)$.

1.2.5. Remark. Associativity is not always assumed in mathematics. A Lie $\mathbb{K}$-algebra is a certain form of $\mathbb{K}$-algebra which is *not* associative. One hundred fifty years after the introduction of *continuous transformation groups* (in modern terms: Lie groups), the central role of Lie algebras in mathematics cannot be denied.

For the moment you know only one Lie algebra, without knowing it is one: $(\mathbb{R}^3, \times)$, where $\times$ is the cross/wedge-product with definition:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$

The class focuses on *associative* algebras; we shall come near Lie algebras when exploring geometric aspects in § 8.

1.2.6. Example. The set $M_n(\mathbb{K})$ of $n \times n$-matrices over $\mathbb{K}$ (more intrinsically, the set $\mathrm{End}_{\mathbb{K}}(V)$ of $\mathbb{K}$-endomorphisms of a $\mathbb{K}$-vector space V) is an associative $\mathbb{K}$-algebra.

Indeed, matrix multiplication (resp. composition of functions) is associative and bilinear. There is an identity element, namely I_n (resp. Id).

The matrix example is typical by the following crucial representation result.

1.2.7. Theorem (Wedderburn's representation theorem)**.** *Every associative, unital* $\mathbb{K}$*-algebra* $\mathbb{A}$ *can be represented as a subalgebra of* $\mathrm{End}_{\mathbb{K}}(V)$ *for some vector space* V.

Proof. Let V be the underlying vector space of $\mathbb{A}$ (i.e. when talking about V we simply forget that there was a multiplication).

For $a \in \mathbb{A}$ consider the left-mutiplication map:

$$\begin{array}{rccc} \lambda_a\colon & V & \to & V. \\ & x & \mapsto & a \cdot x \end{array}$$

This is a $\mathbb{K}$-linear map, in symbols $\lambda_a \in \mathrm{End}_{\mathbb{K}}(V)$.

Now consider the *left-representation*:

$$\begin{array}{rccc} \Lambda\colon & \mathbb{A} & \to & \mathrm{End}_{\mathbb{K}}(V). \\ & a & \mapsto & \lambda_a \end{array}$$

Quite interestingly, one has $\Lambda(a+b) = \Lambda(a) + \Lambda(b)$. Also $\Lambda(1_{\mathbb{A}}) = \mathrm{Id}_V$. Moreover, for any $x \in V$, one finds by associativity:

$$\begin{aligned} (\Lambda(a) \circ \Lambda(b))(x) &= \lambda_a(\lambda_b(x)) \\ &= \lambda_a(b \cdot x) \\ &= a \cdot (b \cdot x) \\ &= (a \cdot b) \cdot x \\ &= \lambda_{a \cdot b}(x) \\ &= (\Lambda(a \cdot b))(x). \end{aligned}$$

So *as functions*, $\Lambda(a \cdot b) = \Lambda(a) \circ \Lambda(b)$ and therefore $\Lambda\colon \mathbb{A} \to \mathrm{End}_{\mathbb{K}}(V)$ is a morphism of $\mathbb{K}$-algebras.

Finally, Λ is injective: we simply prove $\ker \Lambda = \{0\}$. Indeed if $a \in \ker \Lambda$ then $\lambda_a = 0$, so $a = a \cdot 1_{\mathbb{A}} = \lambda_a(1_{\mathbb{A}}) = 0$. So the left-regular representation ('action of $\mathbb{A}$ on itself from the left') injects $\mathbb{A}$ into $\mathrm{End}_{\mathbb{K}}(V)$ as $\mathbb{K}$-algebras, where V is the underlying vector space. □

1.2.8. Remarks.

- This is another instance of an idea also present in 'Cayley's theorem' (abstract groups embed into symmetric groups): the regular representation, obtained by letting a structure act on itself, is faithful. (This also underlies the Yoneda embedding in category theory.)
- This theorem of great historical significance reduced the growing catalogue of so-called 'hypercomplex number systems' (viz. real associative algebras) to one clear theory: matrix algebra.

  Apparently not everyone is aware of Wedderburn's theorem: on wikipedia there are pages for: split complex numbers, split quaternions, bi-quaternions, coquaternions, tessarines, … All actually represent as real matrices and should be treated as such.

The central embedding

1.2.9. Definition. Let $\mathbb{K}$ be a field and $\mathbb{A}$ be an associative algebra. The *centre* of $\mathbb{A}$ is $Z(\mathbb{A}) = \{a \in \mathbb{A} : (\forall x \in \mathbb{A})(a \cdot x = x \cdot a)\}$.

The centre (of an associative algebra) is a subalgebra, and is commutative. Our last lemma explains the position of the base field in an associative, unital algebra: in the centre.

1.2.10. Lemma (central embedding of the base field)**.** *Let $\mathbb{K}$ be a commutative field. Let $\mathbb{A}$ be an associative, unital $\mathbb{K}$-algebra. Then up to isomorphism, $\mathbb{K} \leq Z(\mathbb{A})$.*

Proof. Consider the map:

$$\begin{array}{rccc} \Lambda\colon & \mathbb{K} & \to & \mathbb{A} \\ & \lambda & \mapsto & \lambda \cdot 1_{\mathbb{A}} \end{array}$$

and let $\mathbb{K}'$ be its image. It is easy to see that $\Lambda\colon \mathbb{K} \simeq \mathbb{K}'$ is a field isomorphism. So we may suppose $\mathbb{K}' = \mathbb{K}$, so that $\mathbb{K} \leq \mathbb{A}$. As a matter of fact since $1_{\mathbb{A}}$ is central and $Z(\mathbb{A})$ is a vector $\mathbb{K}$-subspace of $\mathbb{A}$, this gives $\mathbb{K} \leq Z(\mathbb{A})$. The proof used associativity freely. □

1.2.11. Remark. This need not be true if $\mathbb{A}$ has no identity element, or if $\mathbb{A}$ is not associative.

As a form of converse to lemma 1.2.10, we make a simple observation.

1.2.12. Observation. Suppose $\mathbb{A}$ is an associative, unital algebra and $\mathbb{L} \leq \mathbb{A}$ is a subalgebra which happens to be a skew-field.

- Then $\mathbb{A}$ is an $\mathbb{L}$-vector space.
- In general $\mathbb{A}$ need not be an $\mathbb{L}$-algebra. (Typically $\mathbb{C} \leq \mathbb{H}$ but $\mathbb{H}$ is no $\mathbb{C}$-algebra.) However, *if* $\mathbb{L} \leq Z(\mathbb{A})$, then $\mathbb{A}$ is an $\mathbb{L}$-algebra.

1.3 Exercises

1.3.1. Exercise. *Let $\mathbb{K}$ be a skew-field. Prove that its centre $Z(\mathbb{K})$ is a* commutative *subfield of $\mathbb{K}$, and that $\mathbb{K}$ is an associative $Z(\mathbb{K})$-algebra.*

1.3.2. Exercise. *Let $\mathbb{K}$ be any field and $\mathbb{A}$ be a finite-dimensional, associative $\mathbb{K}$-algebra. Prove that $\mathbb{A}$ is a skew-field iff it is a domain.*

2 The real and complex fields

Historically the terminology of 'real' and 'imaginary' numbers goes back to Descartes; only calling the latter 'complex numbers' is due to Gauß. (We believe that none of these is good terminology.)

The main contents of this lecture should be familiar. However we shall make a couple of additions to the basic high school material.

- One can work over other fields than $\mathbb{R}$, the so-called 'real closed fields' of definition 2.1.1. (The course can be followed with only $\mathbb{R}$ in mind.)
- The formalisation of $\mathbb{C}$ as a 2-dimensional algebra over $\mathbb{R}$ is better understood with the tools of linear algebra (proposition 2.2.2).
- In particular, and in accordance with Wedderburn's representation theorem, complex numbers can be viewed as 2×2 matrices over $\mathbb{R}$ (proposition 2.2.5).
- This in turn gives rise to an isomorphism between the 1-sphere $\mathbb{S}^1$ and the special orthogonal group $\mathrm{SO}_2(\mathbb{R})$, finally explaining the polar decomposition $c = re^{it}$ without refering to function 'e^{it}'.

2.1 Real numbers

In this class we *admit* basic properties of the real numbers. We aim neither at describing, nor at formalising them. Perhaps you already know some 'constructions' of real numbers (viz. formalisations from other mathematical objects that one would regard as prior, more elementary, or more reliable). There exist around 20 different 'constructions' in the literature.[2]

One may also regard $\mathbb{R}$ as a fundamental object, in which case 'constructing' it is a very immodest waste of time: there is no need to construct what enjoys sufficient uniqueness to be canonical, hence intuitively universal, hence philosophically safe.

The course never uses the property characterising $\mathbb{R}$ up to isomorphism, viz. order-completeness. Most of the class applies to other fields with the same algebraic properties as $\mathbb{R}$.

2.1.1. Definition. A *real closed field* is an ordered field $(\mathcal{R}; +, \cdot, <)$ such that:

- the non-negative numbers are exactly the squares:

$$(\forall x \in \mathcal{R})[x \geq 0 \leftrightarrow (\exists y \in \mathcal{R})(x = y^2)];$$

- every odd degree polynomial has a root: for each odd n,

$$(\forall a_0, \ldots, a_n \in \mathcal{R})[a_n \neq 0 \to (\exists x \in \mathcal{R})(a_n x^n + \cdots + a_0 = 0)].$$

2.1.2. Remark. There are other equivalent definitions in the literature; in any case:

- $\mathbb{R}$ is real closed;
- every real closed field has characteristic 0, since it is ordered;
- $\mathbb{R} \cap \overline{\mathbb{Q}}$ is real closed, but not isomorphic to $\mathbb{R}$ by a cardinality argument;
- there exist real closed fields which do not embed into $\mathbb{R}$, because they violate 'archimedianity' (in modern parlance, cofinality of the integers): they have infinitesimals, while $\mathbb{R}$ has none.

Here and there one could relax real closedness to weaker algebraic constraints; large parts of the whole course would still hold.

2.2 Complex numbers

It is hard for us to imagine that the complex numbers were controversial once; as a matter of fact among their first name was 'impossible numbers'. Consequently need was early felt for a proof that using them introduces no contradiction (it is a pity we may not dwell on logic in this class), a problem one may consider solved around 1800—if the reals are granted then complex numbers are *plane numbers*, viz. a certain algebra on $\mathbb{R}^2$.

In modern language, the well-known interpretation is as follows.

2.2.1. Definition. Equip the $\mathbb{R}$-vector space $\mathbb{R}^2$ with the operation:

$$(a_1, a_2) \cdot (b_1, b_2) := (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1).$$

Let $\mathbb{C}$ be the resulting structure.

[2] Weiss, I., 'The real numbers—a survey of constructions', *Rocky Mountain J. Math.* 45(3), 737–762, 2015

2.2.2. Proposition. *$\mathbb{C}$ is a 2-dimensional associative $\mathbb{R}$-algebra with multiplicative identity $(1,0)$, and actually a (commutative) field.*

Letting $i = (0,1)$ one sees that $\mathbb{C} = \mathrm{Vect}(1,i) = \mathbb{R}[i]$; remember that the latter denotes the $\mathbb{R}$-algebra generated by i. Since $i^2 = (-1,0) = -(1,0)$, one sometimes writes $\mathbb{C} = \mathbb{R}[\sqrt{-1}]$.

Proof. There are a couple of things to check.

- Multiplication is bilinear. Indeed it is bi-additive, and compatible with the scalar action in either variable.
- Left- and right-'distributivity' are special cases of bilinearity.
- Multiplication is associative, viz. $(\forall c_1, c_2, c_3 \in \mathbb{C})[(c_1 \cdot (c_2 \cdot c_3) = (c_1 \cdot c_2) \cdot c_3)]$.

  One should not rush to a computation. Since multiplication is bilinear, it suffices to take c_1, c_2, c_3 all in the basis $\{1, i\}$. But 1 never violates the associativity axiom so it suffices to deal with $c_1 = c_2 = c_3 = i$; it is now trivial.
- Multiplication is commutative, viz. $(\forall c_1, c_2 \in \mathbb{C})(c_1 \cdot c_2 = c_2 \cdot c_1)$.

  Same argument: it is enough to work with $c_1, c_2 \in \{1, i\}$. But since 1 is central this is obvious.
- $(1,0)$ is a multiplicative identity (obvious);
- Every $c \neq (0,0)$ has a (two-sided) inverse.

  This is interesting. Define the *conjugate* of $c = (a,b) \in \mathbb{R}^2$ as:

  $$c^* = (a, -b).$$

  One sees that * is an *involutive automorphism* of the unital $\mathbb{R}$-algebra $\mathbb{C}$, viz. it preserves addition, multiplication, fixes $\mathbb{R} \cdot 1$ pointwise, and satisfies $c^{**} = c$.

  In the above notation, $c \cdot c^* = a^2 + b^2 \in \mathbb{R}$, and if $a^2 + b^2 \neq 0$ then:

  $$c^{-1} := \frac{1}{cc^*} \cdot c^*$$

  is a two-sided inverse for c. □

2.2.3. Remark. To obtain a commutative field we only used that:

- $\mathbb{R}$ is a commutative field;
- In $\mathbb{R}$, ($a \neq 0$ or $b \neq 0$) implies ($a^2 + b^2 \neq 0$), viz. -1 is not a square, or alternatively $X^2 + 1$ is irreducible in $\mathbb{R}[X]$.

Hence the construction of $\mathcal{R}[i]$ goes well beyond the case of $\mathbb{R}$, or even of real closed fields: one needs only some algebraic information, and no analysis at all. We do not even need to take square roots.

We shall return to the idea of 'doubling' $\mathbb{R} \rightsquigarrow \mathbb{C}$ with the Cayley-Dickson construction in § 5.

By central embedding of the base field (obtained in lemma 1.2.10), the field $\mathbb{R}$ embeds into $Z(\mathbb{C}) = \mathbb{C}$. Reading the proof again, and since $\mathbb{C}$ is an associative $\mathbb{R}$-algebra with identity $(1,0)$, we get the following.

2.2.4. Remark. We may, and will, view $\mathbb{R}$ as the proper subring $\{(a,0) : a \in \mathbb{R}\} < \mathbb{C}$.

Matrix representation and polar decomposition

We now derive a matrix representation of complex numbers as elements of $M_2(\mathbb{R})$, and apply it to the polar representation. Be careful that we do *not* measure angles.

2.2.5. Proposition (matrix representation of the complex numbers)**.** *The set*

$$\mathbb{M} = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : (a,b) \in \mathbb{R}^2 \right\}$$

is a 2*-dimensional* $\mathbb{R}$*-algebra, and* $\mathbb{M} \simeq \mathbb{C}$ *as such.*

Proof. This is an instance of Wedderburn's representation theorem, lemma 1.2.7. Understanding its proof, we see that an isomorphism will be given by $c \mapsto \operatorname{Mat}_{\mathcal{B}} \lambda_c$, where $\mathcal{B}$ is a real basis of $\mathbb{C}$ and λ_c is left-multiplication by $c \in \mathbb{C}$.

Remember that through central embedding we identified any real number a with the complex number $(a,0)$. As a real vector space, $\mathbb{C}$ has basis $\mathcal{B} = (1, i)$. Say $c = a + ib$. Then:

- the image of 1 is $\lambda_c(1) = c = a + ib$, with coordinates $\begin{pmatrix} a \\ b \end{pmatrix}$ in $\mathcal{B}$;
- the image of i is $\lambda_c(i) = c \cdot i = ai - b$, with coordinates $\begin{pmatrix} -b \\ a \end{pmatrix}$ in $\mathcal{B}$.

Therefore,

$$\operatorname{Mat}_{\mathcal{B}} \lambda_c = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}.$$

Wedderburn's theorem predicts that taking $c = (a,b)$ to this matrix is an isomorphism of $\mathbb{R}$-algebras. (The reader with a doubt should check that it is the case indeed.) □

We build on this isomorphism to give a *second* presentation of complex numbers.

- Recall that $\mathrm{SO}_2(\mathbb{R})$, given by:

  $$\mathrm{SO}_2(\mathbb{R}) = \{M \in \mathrm{GL}_2(\mathbb{R}) : M \cdot M^t = M^t \cdot M = I_2\},$$

  is a subgroup of $\mathrm{GL}_2(\mathbb{R})$. (We shall return to orthogonal groups in § 9, when we deal with geometric aspects of quaternions.)
- On the other hand, the unit circle is $\mathbb{S}^1 = \{c \in \mathbb{C}^\times : cc^* = 1\}$, clearly a subgroup.

2.2.6. Lemma. *As groups,* $\mathbb{S}^1 \simeq \mathrm{SO}_2(\mathbb{R})$.

Proof. For $c \in \mathbb{C}$ let $|c| = \sqrt{c^* c}$. If $c = a + bi$, then $|c| = \sqrt{a^2 + b^2}$: therefore $|\cdot|$ is the standard Euclidean norm on the real vector space $\mathbb{C} \simeq \mathbb{R}^2$.

Now since $|\cdot|$ is multiplicative, one has for $c, x \in \mathbb{C}$:

$$|\lambda_c(x)| = |c \cdot x| = |c| \cdot |x|.$$

In particular, λ_c is a linear isometry of $\mathbb{R}^2$ iff $|c| = 1$. This rephrases into: $\lambda_c \in \mathrm{SO}_2(\mathbb{R})$ iff $c \in \mathbb{S}^1$. The isomorphism of proposition 2.2.5 thus restricts to the desired isomorphism. □

2.2.7. Corollary (polar decomposition in $\mathbb{C}^\times$)**.** *The multiplicative group* $\mathbb{C}^\times = (\mathbb{C}\smallsetminus\{0\}, \cdot)$ *is isomorphic to* $\mathbb{R}_{>0} \times \mathrm{SO}_2(\mathbb{R})$.

Proof. The map:

$$\begin{array}{ccc} \mathbb{C}^\times & \to & \mathbb{R}_{>0} \times \mathrm{SO}_2(\mathbb{R}) \\ c & \mapsto & \left(|c|, \frac{c}{|c|}\right) \end{array}$$

is now an isomorphism. □

2.2.8. Remarks.

- For $\mathcal{R}$ an arbitrary real closed field, let $\mathcal{C} = \mathcal{R}[i]$ and $\mathcal{S}^1 = \{c \in \mathcal{C} : |c| = 1\}$. Then still: $\mathcal{S}^1 \simeq \mathrm{SO}_2(\mathcal{R})$ (with the matrix definition) and $\mathcal{C}^* \simeq \mathcal{R}_{>0} \times \mathcal{S}^1$.

  (True in any field where every sum of nonzero squares is a nonzero square. This is needed to take $\sqrt{cc^*}$ in corollary 2.2.7; such a field is called a *Pythagorean* field.)

- Actually lemma 2.2.6 and corollary 2.2.7 even give isomorphisms *of topological groups*. Since every ordering induces a topology, this makes sense and remains true over an arbitrary real closed field $\mathcal{R}$.

- Using more than the algebraic structure on $\mathbb{R}$, namely trigonometry, one may also represent any element of $\mathbb{S}^1$ as e^{it} with $t \in \mathbb{R}/2\pi\mathbb{Z}$. This is known as *measuring angles*. But to do this one essentially needs the complex exponential map:

  $$\begin{array}{ccc} \mathbb{R} & \to & \mathbb{S}^1 \\ t & \mapsto & e^{it} \end{array}.$$

  This function is specific to $\mathbb{R}$ and no such miraculous morphism is present in an arbitrary real closed field. For instance if $\mathcal{R} = \mathbb{R} \cap \overline{\mathbb{Q}}$, the complex exponential does *not* restrict as a map $\mathcal{R} \to \mathcal{S}^1$.

- I do not know what is the class of those real closed fields $\mathcal{R}$ admitting a continuous, surjective homomorphism $\mathcal{R} \to \mathrm{SO}_2(\mathcal{R})$.

- In short, 'rotation terms' in $\mathrm{SO}_2(\mathcal{R}) \simeq \mathcal{S}^1$ need not be measurable by numbers of the base field $\mathcal{R}$. Confusion between 'linear numbers' $r \in \mathcal{R}$ and 'circular numbers' $c \in \mathcal{S}^1$ is sadly common. But measuring angles is possible in $\mathbb{R}$ by miracle.

We shall comment further on angle measurements (and how to avoid them) in § 9. Matrix representations of quaternions, with a polar decomposition, will be given in § 4.3.

2.3 Exercises

2.3.1. Exercise. *Let $\mathbb{K}$ be a field in which for any x, $1+x^2$ is a square. Prove that any sum of squares is a square.*

Solution. A mere induction. It suffices to treat the case $n = 2$. Now $a^2 + b^2 = a^2(1 + (a^{-1}b)^2$ is a product of squares, hence a square.

2.3.2. Exercise. *Construct different orderings on the field $\mathbb{Q}(X)$.*

Solution. For any transcendental real number α, we can order $\mathbb{K}$ by letting $X = \alpha$. We can also order $\mathbb{K}$ by $\mathbb{Q} < X$.

2.3.3. Exercise (the formally real fields are the orderable fields)**.**

Definition. A field is *formally real* if no sum of non-zero squares equals 0.
The purpose of this exercise is to prove the following.

Theorem (Artin). A field is formally real iff it is orderable.

1. *Prove that a field is formally real iff -1 is not a sum of squares. Deduce that every formally real field has characteristic* 0 *(of course the converse fails: $\mathbb{C}$).*

2. *A* positive cone *is a subset $P \subseteq \mathbb{K}$ such that:*

   - $-1 \notin P$;
   - $(\forall x, y \in P)(x + y \in P \wedge x \cdot y \in P)$;
   - $(\forall x \in \mathbb{K})(x^2 \in P)$.

   Prove that a positive cone is closed under inversion of non-zero elements, and contains the natural integers.

3. *Prove that an ordering on $\mathbb{K}$ is the same as a maximal positive cone, or as a positive cone P with $(\forall x \in \mathbb{K})(x \in P \vee -x \in P)$.*

4. *Deduce that every formally real field can be ordered.*

Solution.

1. Trivial: $\sum a_i^2 = -1$ iff $1^2 + \sum a_i^2 = 0$. In particular, if $\mathbb{K}$ is formally real, then $1 + \cdots + 1 \neq 0$, so $\mathbb{K}$ does not have positive characteristic.

2. Let $a \in P$ be nonzero. Then $a^{-1} = a \cdot (a^{-1})^2 \in P$. Moreover $0^2 = 0$ and $1^2 = 1$ are in P; from there one uses addition to get $\mathbb{N} \subseteq P$.

3. Suppose $\leq$ is an ordering on $\mathbb{K}$ and let $P = \{x \in \mathbb{K} : x \geq 0\}$. Then P is a positive cone and we contend it is maximal as such. Indeed, if $P \subset P'$ is another positive cone and $x \in P' \setminus P$, then $x < 0$. Then $-x > 0$ so $-x \in P \subseteq P'$. But then $-1 = -x \cdot \frac{1}{x} \in P'$, a contradiction.

   Now suppose P is a maximal positive cone. Let $x \in \mathbb{K}$ such that $(x \notin P) \wedge (-x \notin P)$; we prove a contradiction. Let $P' = \{a + bx : (a, b) \in P^2\}$, which contains P. Actually since $0, 1 \in P$ one has in $x \in P'$, so P' strictly contains P. The set P' is clearly stable under $+$. It also is under $\cdot$, since in obvious notation:

$$(a_1 + b_1 x)(a_2 + b_2 x) = \underbrace{(a_1 a_2 + b_1 b_2 x^2)}_{\in P} + (a_1 b_2 + a_2 b_1)x.$$

If $-1 \notin P'$ then P' is a positive cone strictly extending P: a contradiction. Therefore there are $a, b \in P$ with $a+bx = -1$, and clearly $b \neq 0$. But then, $-x = (1+a)\cdot b^{-1} \in P$, a contradiction.

Finally suppose P is a positive cone satisfying $(\forall x \in \mathbb{K})(x \in P \vee -x \in P)$. Write $a \leq b$ if $b-a \in P$. By assumption on P, the ordering is total. It clearly is compatible with $+$ and $\cdot$: therefore, a field ordering.

4. Let $\mathbb{K}$ be formally real and $P = \{\sum_{i=1}^{n} a_i^2 : n \in \mathbb{N}, a_1, \ldots, a_n \in \mathbb{K}^n\}$. This is stable under addition, and also under product. Moreover it contains all squares, but not -1 as otherwise $\mathbb{K}$ is not formally real.

   Hence there exists a positive cone; using Zorn' lemma, take a maximal one. It is an ordering on $\mathbb{K}$.

2.3.4. Exercise. *Prove that up to isomorphism there are exactly three associative, unital, 2-dimensional $\mathbb{R}$-algebras:*

$$\mathbb{R}[i] \text{ with } i^2 = -1, \quad \mathbb{R}[j] \text{ with } j^2 = 1, \quad \mathbb{R}[\varepsilon] \text{ with } \varepsilon^2 = 0.$$

Which properties of $\mathbb{R}$ are needed?

Solution. Let $\mathbb{A}$ be such an algebra; use central embedding to assume $\mathbb{R} < \mathbb{A}$. Let $\alpha \in \mathbb{A} \smallsetminus \mathbb{R}$. Then $1, \alpha, \alpha^2$ cannot be linearly independent, so there is a non-trivial relation $a_0 + a_1\alpha + a_2\alpha^2 = 0$. Of course $a_2 \neq 0$, since otherwise $\alpha \in \mathbb{R}$. Since $\mathbb{R}$ is a field, we may assume $a_2 = 1$. Therefore α satisfies an equation: $\alpha^2 + a_1\alpha + a_0 = 0$. Up to considering $\alpha - \frac{a_1}{2}$, we may assume $a_1 = 0$.

So there is $\alpha \in \mathbb{A} \smallsetminus \mathbb{R}$ such that $\alpha^2 + a_0 = 0$. There are three cases. If $a_0 > 0$, up to considering $\frac{1}{\sqrt{a_0}}\alpha$ we may suppose $a_0 = 1$ and reach case i. If $a_0 < 0$, same reasoning and reach case j. If $a_0 = 0$ this is case ε.

2.3.5. Exercise. *Represent $\mathbb{R}[j]$ and $\mathbb{R}[\varepsilon]$ of exercise 2.3.4 as matrix algebras. Also determine their multiplicative groups.*

(The multiplicative group of an associative, unital algebra $\mathbb{A}$ is $\mathbb{A}^\times = \{x \in \mathbb{A} : (\exists y \in \mathbb{A})(xy = yx = 1)\}$, in general much smaller than $\mathbb{A} \smallsetminus \{0\}$).

3 Two limitation theorems

This lecture could also be read *after* § 4; we prefer to give it before, as it motivates Hamilton's construction.

Can one find fields extending $\mathbb{C}$? Yes, e.g. $\mathbb{C}(X)$. But can one find fields extending $\mathbb{C}$ *and finite-dimensional over* $\mathbb{R}$? After understanding irreducible polynomials of $\mathbb{R}[X]$ (corollary 3.1.3), we run into two obstructions:

1. a finite-dimensional $\mathbb{R}$-algebra which is a field must be isomorphic to $\mathbb{R}$ or $\mathbb{C}$ (theorem 3.2.2);

2. no $\mathbb{R}$-algebra structure on $\mathbb{R}^3$ is a domain (theorem 3.3.1).

These imply that in order to find a finite-dimensional skew-field over $\mathbb{R}$, one must drop commutativity and use dimension at least 4.

3.1 Real and complex polynomials

Complex numbers emerged naturally in the search for polynomial equations: indeed they form the algebraic closure of the reals. More intrinscally, $\mathbb{C}$ is an *algebraically closed field*, one where every non-constant polynomial has a root (equivalently, one where every polynomial of degree d has d roots, counting multiplicities).

3.1.1. Theorem ('D'Alembert-Gauß theorem'; first full proof by Argand, 1806)**.** $\mathbb{C} = \mathbb{R}[i]$ *is algebraically closed.*

3.1.2. Remarks.

- The proof uses a little Galois theory, for which we may refer to Serge Lang's *Algebra* (Chap. VI, §2). All we need is that $\mathbb{R}$ is a real closed field. Topological proofs (for instance using real, or even complex analysis), however elegant, tend to hide this.
- A remarkable form of converse is the *Artin-Schreier theorem*: if $\mathbb{K}$ is any (commutative) field such that $[\overline{\mathbb{K}} : \mathbb{K}] < \infty$, then $\mathbb{K}$ is either algebraically closed (in which case $\overline{\mathbb{K}} = \mathbb{K}$) or real closed (in which case $\overline{\mathbb{K}} = \mathbb{K}[i]$). This is remarkable since there are no assumptions on the characteristic!

  The proof of the Artin-Schreier theorem involves even more Galois theory and is definitely not in the scope of this class.

As a consequence of theorem 3.1.1, irreducible polynomials of $\mathbb{C}[X]$ have degree 1. One can also study $\mathbb{R}[X]$.

3.1.3. Corollary. *The irreducible polynomials of* $\mathbb{R}[X]$ *are exactly:*

- *all polynomials of degree* 1*;*
- *those polynomials of degree* 2 *with a negative discriminant.*

Proof. One direction is obvious. For the converse, recall that whenever $\mathbb{K}$ is a field, the polynomial ring $\mathbb{K}[X]$ allows unique Euclidean division (the degree being the decreasing function). As a corollary to uniqueness, if $\mathbb{K} \subseteq \mathbb{L}$ is a field extension and $P, Q \in \mathbb{K}[X]$ are such that $Q|P$ in $\mathbb{L}[X]$, then $Q|P$ in $\mathbb{K}[X]$ already. Bear this in mind.

Let $P \in \mathbb{R}[X]$ be irreducible; we may suppose that its leading coefficient is 1. In $\mathbb{C} = \mathbb{R}[i]$, which is algebraically closed, write $P = \prod(X - \alpha_k)$.

If one of the roots, say α_1, lies in $\mathbb{R}$, then $X - \alpha_1$ divides P in $\mathbb{C}[X]$, hence also in $\mathbb{R}[X]$: by irreducibility there, $P = X - \alpha_1$. So we may assume that P has no real root.

The conjugation map * (in abstract terms, the non-trivial element in the Galois group $\mathrm{Gal}(\mathbb{C}/\mathbb{R}) \simeq \mathbb{Z}/2\mathbb{Z}$) fixes the elements of $\mathbb{R}$ and only them; hence for any root, one has $(P(\alpha_k))^* = P(\alpha_k^*) = 0$ so $\alpha_k^* \neq \alpha_k$ is another root of P. Observe how $Q = (X - \alpha_k)(X - \alpha_k^*) = X - (\alpha_k + \alpha_k^*)X + \alpha_k\alpha_k^*$ divides P in $\mathbb{C}[X]$, and has coefficients in the fixed field of *, namely $\mathbb{R}$. Here again, Q divides P in $\mathbb{R}[X]$ and by irreducibility, $P = Q$ has degree 2. If the discriminant were a square, we would find a real root: it is therefore negative. □

3.1.4. Remark. The corollary vacuously holds over any real closed field; it essentially states that $\mathbb{R}$ is a real closed field.

3.2 No more commutative fields are finite-dimensional over $\mathbb{R}$

For the next theorem one needs a simple notion, which builds on principality of $\mathbb{K}[X]$ whenever $\mathbb{K}$ is a field.

3.2.1. Definition. Let $\mathbb{K}$ be a field, $\mathbb{A}$ be an associative, unital $\mathbb{K}$-algebra, and $\alpha \in \mathbb{A}$ be such that the subalgebra $\mathbb{K}[\alpha]$ is finite-dimensional over $\mathbb{K}$. Then the (unique) generator of the ideal $\{P \in \mathbb{K}[X] : P(\alpha) = 0\}$ with leading coefficient 1 is called the *minimal polynomial* of α over $\mathbb{K}$, and denoted by $\mathrm{Min}_{\mathbb{K}}^{\alpha}$.

The condition on α obviously holds if $\mathbb{A}$ itself is finite-dimensional over $\mathbb{K}$. We prove that the only commutative fields which are finite-dimensional over $\mathbb{R}$ are $\mathbb{R}$ and $\mathbb{C}$.

3.2.2. Theorem. *Let $\mathbb{A}$ be a finite-dimensional $\mathbb{R}$-algebra. If $\mathbb{A}$ is a commutative field, then $\mathbb{A} = \mathbb{R}$ or $\mathbb{A} \simeq \mathbb{C}$.*

Proof. Notice that $\mathbb{A}$ is associative and unital, being a field. By central embedding (lemma 1.2.10), we may identify $\mathbb{R}$ with $\mathbb{R} \cdot 1_{\mathbb{A}}$, so that $\mathbb{R} \leq \mathbb{A}$ is now a field extension. Of course we may suppose $\dim_{\mathbb{R}} \mathbb{A} > 1$, or we are done.

Take any $\alpha \in \mathbb{A} \setminus \mathbb{R}$ and let $P(X) = \mathrm{Min}_{\mathbb{R}}^{\alpha}(X)$ be its minimal polynomial with coefficients in $\mathbb{R}$. Since $\mathbb{A}$ is a domain, P must be irreducible. Indeed, suppose $P = Q_1 Q_2$. Since $\mathbb{A}$ is a domain, $Q_1(\alpha)Q_2(\alpha) = 0$ implies that one of Q_1, Q_2 vanishes at α. So by minimality one of them is already divisible by P, hence equal to P. (Same proof, fancier language: $\mathbb{A} \simeq \mathbb{K}[X]/(P)$ is a domain so (P) is a prime ideal in the factorial domain $\mathbb{K}[X]$; hence P is irreducible in $\mathbb{K}[X]$.)

By the classification of irreducible polynomials in $\mathbb{R}[X]$ (corollary 3.1.3), one has $\deg P \leq 2$; of course $\deg P \neq 1$ since otherwise $\alpha \in \mathbb{R}$. So we may write $P(X) = X^2 + pX + q$. Up to replacing α by $\alpha + \frac{p}{2}$, we may assume $\alpha^2 = r \in \mathbb{R}$, actually clearly in $\mathbb{R}_{\leq 0}$; up to rescaling we may assume $\alpha^2 = -1$. Then $\mathbb{A} \geq \mathbb{R}[\alpha] \simeq \mathbb{R}[i]$.

So we found an isomorphic copy of $\mathbb{C} = \mathbb{R}[i]$ inside $\mathbb{A}$. This makes $\mathbb{A}$ a $\mathbb{C}$-vector space, but here there is more (you may wish to read observation 1.2.12 again): *since $\mathbb{A}$ is commutative, $\mathbb{A}$ is actually an associative $\mathbb{C}$-algebra*. Now if $\beta \in \mathbb{A}$ then $\mathrm{Min}_{\mathbb{C}}^{\beta} \in \mathbb{C}[X]$ is irreducible, so by algebraic closedness of $\mathbb{C}$ its degree is 1: hence $\beta \in \mathbb{C}$, proving $\mathbb{A} = \mathbb{C}$. □

3.2.3. Remarks.

- In particular, there is no field extending $\mathbb{C}$ in a finite-dimensional way (of course one can always form transcendental extensions $\mathbb{C}(X)$, but this has infinite linear dimension over $\mathbb{C}$): which amounts to saying that $\mathbb{C}$ is algebraically closed.

  Theorem 3.2.2 is stronger because it supposes only an $\mathbb{R}$-algebra structure, not a $\mathbb{C}$-algebra structure.

- Return to the moment we turned $\mathbb{A}$ into a $\mathbb{C}$-algebra (and to observation 1.2.12). For right-linearity of $\cdot_{\mathbb{A}}$ over $\mathbb{C}$, commutativity of $\mathbb{A}$ is required.

  The argument does not work if $\mathbb{A}$ is merely assumed to be a skew-field. And indeed, we shall construct the skew-field $\mathbb{H} \neq \mathbb{R}, \mathbb{C}$.

Therefore if one wants *another* finite-dimensional real algebra which is a skew-field, one has to drop commutativity.

3.3 Can't multiply triplets

The following explains why Hamilton could not find a field structure on $\mathbb{R}^3$.

3.3.1. Theorem. *No 3-dimensional, associative $\mathbb{R}$-algebra is a domain.*

Proof. Let $\mathbb{A}$ be such an algebra and $\alpha \in \mathbb{A} \setminus \mathbb{R}$. Let $P = \mathrm{Min}_{\mathbb{R}}^{\alpha}$, which has degree 2 or 3. Since $\mathbb{A}$ is a domain, P is irreducible in $\mathbb{R}[X]$; since $\mathbb{R}$ is real closed and always by corollary 3.1.3, P has degree 2. Now $\mathbb{R}[\alpha]$ is a commutative domain and a finite-dimensional $\mathbb{R}$-algebra: it is a commutative field, and since $\alpha \notin \mathbb{R}$ we find $\mathbb{R}[\alpha] \simeq \mathbb{C}$ by theorem 3.2.2.

Hence $\mathbb{A}$ can be seen as a vector space over $\mathbb{C}$, and has a dimension as such, $\dim_{\mathbb{C}} \mathbb{A}$, which is an integer. Then $\dim_{\mathbb{R}} \mathbb{A} = 2 \dim_{\mathbb{C}} \mathbb{A}$ must be even, a contradiction. □

3.3.2. Remarks.

- In the proof $\mathbb{A}$ is turned into a $\mathbb{C}$-vector space, but not into an associative $\mathbb{C}$-algebra: one would need commutativity of $\mathbb{A}$, or at least $\mathbb{C} \leq Z(\mathbb{A})$ for this to hold. We do not assume it. The same situation will happen with quaternions.
- One can classify associative, unital $\mathbb{R}$-algebras of dimension 3 : exercise 3.4.1

The conclusion of theorem 3.3.1 is that in order to retrieve a domain (equivalently, a skew-field: see exercise 1.3.2) one has to look in dimension at least 4.

3.4 Exercises

3.4.1. Exercise (tedious but instructive)**.** *Classify associative, unital $\mathbb{R}$-algebras of dimension 3. You may proceed as follows:*

1. *If there is an element of degree 3, then $\mathbb{A}$ is commutative. Classify these up to isomorphism.*
2. *Otherwise choose $a \in \mathbb{A} \setminus \mathbb{R}$ and $b \in \mathbb{A} \setminus \mathbb{R}[a]$ such that $a^2, b^2 \in \{-1, 0, 1\}$; prove that -1 is not possible, then finish classification by hand.*

4 Quaternions: a first algebraic study

We opt for non-commutativity in dimension 4 and construct a skew-field structure on $\mathbb{R}^4$ (§ 4.1); pay attention to quaternion conjugation and norm (§ 4.2). We then move to geometric aspects. The 3-dimensional sphere $\mathbb{S}^3$ can be seen as a Lie group. One can represent quaternions both as real or complex matrices (§ 4.3), giving rise to a polar decomposition.

4.1 Constructing the quaternions

4.1.1. Definition (Hamilton, 1843)**.** Let $\mathbb{H} = \mathbb{R}^4$ with basis $(1, i, j, k)$ and for multiplication the unique bilinear map extending:

$\overset{\curvearrowright}{\cdot}$	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

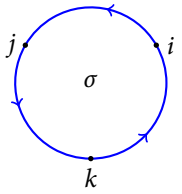
(Symbol $\curvearrowright$ in the table tells you in which order you must multiply. As the table is *not* symmetric, the order does matter: multiplication is not commutative.)

4.1.2. Proposition. *$\mathbb{H}$ is an associative $\mathbb{R}$-algebra with multiplicative identity 1 and centre $Z(\mathbb{H}) = \mathbb{R}$ (through central embedding).*

(It will remain to prove that it is a skew-field, in corollary 4.2.4.)

Proof. Bilinearity is by construction, and left- and right-distributivity immediately follow. Clearly 1 is the multiplicative identity. So we only have to prove associativity. But by bilinearity again, it suffices to check it for basic relations, i.e. for triples from the table. Since 1 cannot violate associativity, it is enough to check $a(bc) = (ab)c$ for elements $a, b, c \in \{i, j, k\}$; a priori there still are $3^3 = 27$ routine verifications.

Let us make this a bit faster. Consider the cycle σ:



extended linearly: this is an automorphism of order 3 of the $\mathbb{R}$-vector space $\mathbb{H}$. Notice that σ preserves the multiplication table, viz. $\sigma(ab) = \sigma(a)\sigma(b)$ for $a, b, c \in \{i, j, k\}$. So σ is an automorphism of the (not yet associative) $\mathbb{R}$-algebra $\mathbb{H}$. Hence we may assume $a = i$; there are only $\frac{27}{3} = 9$ remaining verifications.

Now let τ do $\tau(i) = -i$, $\tau(j) = k$ and $\tau(k) = j$. Here again τ is an automorphism of $\mathbb{H}$ as an algebra; since -1 is central, the only remaining non-trivial verification is:

$$i(jk) = i^2 = -1 = k^2 = (ij)k,$$

and associativity holds, proving that $\mathbb{H}$ is an associative, unital $\mathbb{R}$-algebra.

In particular we may embed $\mathbb{R} \hookrightarrow Z(\mathbb{H})$ (see lemma 1.2.10) and now consider $\mathbb{R} \leq Z(\mathbb{H})$. The converse inclusion is seen in coordinates: let $q = a+bi+cj+dk \in Z(\mathbb{H})$; we prove $q \in \mathbb{R}$. By assumption $qi = iq$, with j-coordinate $d = -d$, viz. $d = 0$. But σ is an automorphism of $\mathbb{H}$, so it must stabilise $Z(\mathbb{H})$. Therefore $b = c = 0$ as well and $q \in \mathbb{R}$. (It is also possible to take the k-coordinate in $qi = iq$, then also use equality $qj = jq$.) □

4.1.3. Remarks.

- Be extremely careful that $\mathbb{H}$ is *not* a $\mathbb{C}$-algebra! Otherwise by central embedding (lemma 1.2.10) we would have $\mathbb{C} \hookrightarrow Z(\mathbb{H}) \simeq \mathbb{R}$, a contradiction. However, $\mathbb{H}$ is a $\mathbb{C}$-vector space.

  To feel the difference, just see that the product is *not* $\mathbb{C}$-bilinear. Indeed,

  $$j \cdot (i \cdot j) = j \cdot k = i \neq -i = i \cdot -1 = i \cdot (j \cdot j),$$

  so denoting by μ the multiplication we find $\mu(j, i \cdot j) \neq i \cdot \mu(j, j)$.

- For the construction of the associative algebra, all we need from $\mathbb{R}$ is to be a field. Notice that in characteristic 2 the resulting associative algebra *is* commutative.

- It is an interesting property that $\mathbb{R}$ can be recovered algebraically from $\mathbb{H}$. *This is not the case in* $\mathbb{C}$. Logicians say that $\mathbb{R}$ is *definable* in $(\mathbb{H}; +, \cdot)$ but not in $(\mathbb{C}; +, \cdot)$.

4.1.4. Example. Let $M = \mathbb{H}^2$, which is both a left- and right- $\mathbb{H}$-module. Let $u = (1, j)$ and $v = (i, k)$. Then in M as a left-module, they are linearly dependent through $v = i \cdot u$; in M as a right-module, they are linearly independent since $j \cdot i \neq k$.

4.2 The conjugate and norm of a quaternion

4.2.1. Definition. Let $q = a + bi + cj + dk$ with $a, b, c, d, \in \mathbb{R}$ be a quaternion.

- The *conjugate* of q is $q^* = a - bi - cj - dk$.
- The (number-theoretic) *norm* of q is $N(q) = qq^* = a^2 + b^2 + c^2 + d^2 \in \mathbb{R}$.
- The (geometric) *norm* of q is $\|q\| = \sqrt{N(q)}$.

4.2.2. Remarks.

- For conjugation, it is safe to use the same notation as for complex numbers: if one embeds $\mathbb{C}$ into $\mathbb{H}$ naturally, i.e. mapping $i_{\mathbb{C}}$ to $i_{\mathbb{H}}$, then quaternionic * extends the complex one.
- Depending on the purpose, one may prefer to work with $N(q)$ or $\|q\|$. The difference is essentially cosmetic.

  People in analysis or geometry will prefer $\|q\|$, which coincides with the usual Euclidean norm on $\mathbb{R}^4$ and satisfies the well-known triangle equality $\|q_1 + q_2\| \leq \|q_1\| + \|q_2\|$.

4.2.3. Properties. *The map $q \mapsto q^*$ is an additive, anti-multiplicative, involutive map, viz.:*

- $(\forall q \in \mathbb{H})(q^{**} = q)$;
- $(\forall q_1, q_2 \in \mathbb{H})((q_1 + q_2)^* = q_1^* + q_2^*)$;
- $(\forall q_1, q_2 \in \mathbb{H})((q_1 q_2)^* = q_2^* q_1^*)$ *(mind the reverse order).*

Moreover, $N(\cdot)$ is multiplicative, and $N(q) = 0$ iff $q = 0$.

One sometimes calls * an involutive *anti*-automorphism, or merely an *involution* of the $\mathbb{R}$-algebra $\mathbb{H}$. We shall return to $*$-algebras in § 5.

4.2.4. Corollary. $\mathbb{H}$ *is a skew-field.*

Proof. Let $q \in \mathbb{H} \smallsetminus \{0\}$. Then $N(q) \in \mathbb{R} \smallsetminus \{0\}$, so writing $q' = \frac{1}{N(q)} q^* \in \mathbb{H}$ makes sense. Then it is readily seen that $qq' = q'q = 1$, so q' is a two-sided inverse for q. □

4.2.5. Remarks.

- This will hold over any real field (the proof does *not* require $N(q)$ to be a square, only a non-zero element).
- Notice that $\pm i, \pm j, \pm k$ satisfy equation $x^2 + 1 = 0$. This does not contradict the fact that $\mathbb{H}$ is a skew-field: one can bound the number of roots of a polynomial only in a *commutative* domain.

We move to geometric aspects.

4.2.6. Corollary. *The 3-dimensional sphere* $\mathbb{S}^3 = \{v \in \mathbb{R}^4 : \|v\| = 1\}$ *(for the usual Euclidean norm on* $\mathbb{R}^4$*) can be equipped with an 'algebraic' group structure, viz. a group structure whose operation is given by polynomial functions in the coordinates.*

We give a trivial proof through quaternions, without which the result is hard.

Proof. Recall that for $q = a + bi + cj + dk$ one has $N(q) = a^2 + b^2 + c^2 + d^2$. So if one sees $\mathbb{H}$ as the $\mathbb{R}$-vector space $\mathbb{R}^4$ with the standard metric, then the quaternion sphere $\mathbb{S} = \{q \in \mathbb{H} : N(q) = 1\}$ becomes the usual 3-dimensional hypersphere sitting inside the 4-dimensional space, viz. $\mathbb{S} \simeq \mathbb{S}^3 \subset \mathbb{R}^4$ (isometrically).

Now since the norm is multiplicative, $\mathbb{S} \leq \mathbb{H}^\times$ is a subgroup. Multiplication on $\mathbb{R}^4 \simeq \mathbb{H}$ is clearly polynomial in the coordinates; we are done. □

4.2.7. Remark. Over $\mathbb{R}$ this is possible only for $\mathbb{S}^0 = \{\pm 1\}$, $\mathbb{S}^1$ (the circle), and $\mathbb{S}^3$. Dropping associativity one may still do something with $\mathbb{S}^7$ (using octonions, see § 5.3). This is all since deep results in differential geometry guarantee that over $\mathbb{R}$, the spheres $\mathbb{R}^0, \mathbb{R}^1, \mathbb{R}^3, \mathbb{R}^7$ are the only ones bearing a compatible algebraic structure.[3].

This does not seem to be known over arbitrary real closed fields, left alone real fields (the question may fail to make much sense over a fully arbitrary field).

4.3 Matrix representations and polar decomposition

We give *two* matrix representations of the quaternion algebra. The second is geometrically important as it allows for a polar decomposition.

4.3.1. Proposition. *As* $\mathbb{R}$*-algebras,*

$$\mathbb{H} \simeq \left\{ \begin{pmatrix} a & -b & -c & -d \\ b & a & -d & c \\ c & d & a & -b \\ d & -c & b & a \end{pmatrix} : (a,b,c,d) \in \mathbb{R}^4 \right\}.$$

[3] A beautiful and almost elementary proof *in even dimension*, using the notion of homology, can be found at http://www.raczar.es/webracz/ImageServlet?mod=publicaciones&subMod=revistas&car=revista62&archivo=p075.pdf.

Proof. This is an instance of Wedderburn's theorem, lemma 1.2.7. See $\mathbb{H}$ as an $\mathbb{R}$-vector space with basis $\mathcal{B} = (1, i, j, k)$. For any $q \in \mathbb{H}$ consider the left-multiplication map $\lambda_q(h) = qh$, and write its matrix in $\mathcal{B}$. It is an exercise. □

4.3.2. Proposition. *Denote by $c \mapsto c^*$ the conjugation in $\mathbb{C}$. Then as $\mathbb{R}$-algebras,*

$$\mathbb{H} \simeq \left\{ \begin{pmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{pmatrix} : (c_1, c_2) \in \mathbb{C}^2 \right\}.$$

Of course it is not an isomorphism of $\mathbb{C}$-algebras since $\mathbb{H}$ is *not* one as we noticed in Remarks 4.1.3.

Proof. This is more subtle, and important. Let $\mathbb{C} = \mathbb{R}[i] \subseteq \mathbb{H}$. One may certainly view $\mathbb{H}$ as a vector space over $\mathbb{C}$, more specifically as a *left*-vector space, viz. for the operation $(c, q) \mapsto c \cdot q$. Here left and right do matter since $\mathbb{H}$ is not commutative. Fix $q \in \mathbb{H}$. Then left-multiplication by q, viz. the map $\lambda_q(x) = qx$, is $\mathbb{R}$-linear, but not $\mathbb{C}$-linear: for instance, $\lambda_j(i \cdot 1) = ji = -k \neq k = i \cdot \lambda_j(1)$.

One may be tempted to then consider right-multiplication by q, viz. the map $\rho_q(x) = xq$. Then ρ_q is $\mathbb{C}$-linear by associativity: indeed,

$$\rho_q(\lambda \cdot x) = (\lambda \cdot x) \cdot q = \lambda(xq) = \lambda \cdot \rho_q(x).$$

This would lure us into considering the representation $q \mapsto \rho_q \in M_2(\mathbb{C})$. The catch is that qq' maps to $\rho_{qq'} = \rho_{q'}\rho_q$, so we shall *not* get a morphism of rings, merely an *anti-morphism*.

Since using left-multiplications λ_q seems to be non-negociable, we start over again by considering $\mathbb{H}$ as a *right*-vector space over $\mathbb{C}$: now $(c, q) = qc$ (no, there is no inverse on the right: it would be meaningless at 0 and no longer additive; besides, $\mathbb{C}$ is commutative). This is *not* the same structure as its left-vector space structure since $ij \neq ji$.

So treat $\mathbb{H}$ as a right-$\mathbb{C}$-vector space, with basis $\check{\mathcal{B}} = (1, j)$. With respect to this structure, left-multiplication by any $q \in \mathbb{H}$ is $\mathbb{C}$-linear: by associativity of $\mathbb{H}$. Now compute that if $q = c_1 + jc_2$ with $c_1 = a_1 + b_1 i$ and c_2 likewise in obvious notation:

$$\begin{aligned} \lambda_q(j) &= (c_1 + jc_2)j \\ &= (a_1 + b_1 i)j + j(a_2 + b_2 i)j \\ &= a_1 j + b_1 k - a_2 + b_2 i \\ &= -(a_2 - b_2 i) + j(a_1 - b_1 i) \\ &= -c_2^* + jc_1^*, \end{aligned}$$

so

$$\operatorname{Mat}_{\check{\mathcal{B}}} \lambda_q = \begin{pmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{pmatrix}.$$

The desired isomorphism is now obvious: let $\Lambda(q) = \operatorname{Mat}_{\check{\mathcal{B}}} \lambda_q$. □

4.3.3. Example. This must be understood before going any further. In the isomorphism

above, one has:

$$1 \mapsto \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}; \quad i \mapsto \begin{pmatrix} i & \\ & -i \end{pmatrix}; \quad j \mapsto \begin{pmatrix} & -1 \\ 1 & \end{pmatrix}; \quad k \mapsto \begin{pmatrix} & i \\ -i & \end{pmatrix}.$$

A unitary isomorphism; the polar decomposition

Building on the matrix representation, we prove a quaternionic version of the polar decomposition of non-zero complex numbers, corollary 2.2.7. This requires leaving orthogonal matrices and going to their complex analogues.

Recall that the *unitary group* of $\mathbb{C}$ over $\mathbb{R}$ (more precisely, with respect to complex conjugation $c \mapsto c^*$) is the group:

$$\mathrm{U}_n = \mathrm{U}_n(\mathbb{C}, *) = \{M \in \mathrm{GL}_n(\mathbb{C}) : M^{-t*} = M\}.$$

Being closed and bounded in $\mathrm{GL}_n(\mathbb{C})$, it is a compact group, the complex analogue of $\mathrm{O}_n(\mathbb{R})$. We then define its normal subgroup:

$$\mathrm{SU}_n(\mathbb{C}, *) = \mathrm{U}_n \cap \mathrm{SL}_n(\mathbb{C}).$$

4.3.4. Theorem (cf. lemma 2.2.6)**.** *The quaternion sphere* $\mathbb{S} = \{q \in \mathbb{H} : N(q) = 1\}$ *is isomorphic to the group* $\mathrm{SU}_2(\mathbb{C}, *)$.

Proof. This builds on the complex matrix representation of proposition 4.3.2. If $q = c_1 + jc_2$ with $c_1, c_2 \in \mathbb{C}$, then:

$$\Lambda(q) = \mathrm{Mat}_{\mathcal{B}} \lambda_q = \begin{pmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{pmatrix}.$$

This matrix is in $\mathrm{SU}_2(\mathbb{C}, *)$ iff $|c_1|^2 + |c_2|^2 = 1$. On the other hand, writing $c_1 = a_1 + ib_1$ and $c_2 = a_2 + b_2 i$ in real coordinates, one finds:

$$q = a_1 + b_1 i + a_2 j - b_2 k,$$

whence $N(q) = a_1^2 + b_1^2 + a_2^2 + b_2^2 = |c_1|^2 + |c_2|^2$. So q has norm 1 iff $|c_1|^2 + |c_2|^2 = 1$.

Hence:

$$\begin{array}{rccc} \Lambda\colon & \mathbb{S} & \to & \mathrm{SU}_2(\mathbb{C}, *) \\ & q & \mapsto & \mathrm{Mat}_{\mathcal{B}} \lambda_q \end{array}$$

is a group isomorphism. □

4.3.5. Remark. In notation $q = c_1 + jc_2$ and $\Lambda(q) = \begin{pmatrix} c_1 & -c_2^* \\ c_2 & c_1^* \end{pmatrix}$, one has:

$$\Lambda(q^*) = \Lambda(q)^*,$$

where the left star stands for quaternion conjugation and the right star for the Hermite-symmetric/complex adjoint of a matrix.

Like in the complex case, we deduce a polar representation.

4.3.6. Corollary (cf. corollary 2.2.7)**.** $\mathbb{H}^\times \simeq \mathbb{R}_{>0} \times \mathrm{SU}_2(\mathbb{C}, *)$ *as groups.*

4.3.7. Remarks.

- Here there is no temptation to 'measure elements of $\mathrm{SU}_2(\mathbb{C}, *)$ by angles'; this is meaningless (though precisely the theory of Euler angles).
- This shows that the group $\mathbb{H}^\times$ has a non-abelian, simple factor. More generally, Hua showed that in any non-commutative skew-field $\mathbb{K}$, the multiplicative group $\mathbb{K}^\times$ is non-soluble.

4.4 Exercises

4.4.1. Exercise. *Return to the proof of proposition 4.1.2. Describe σ and τ* geometrically, *viz. as transformations of* $\mathrm{Vect}(i, j, k) \simeq \mathbb{R}^3$.

4.4.2. Exercise. *For $q \in \mathbb{H}$ compute: $-\frac{1}{2}(q + iqi + jqj + kqk)$. Is there anything similar in the complex case?*

4.4.3. Exercise. *Prove that the quaternionic conjugation $q \mapsto q^*$ is definable in $(\mathbb{H}; +, \cdot)$. Hint: first define $\mathbb{R}$, then define* $\mathrm{Vect}(i, j, k)$.

4.4.4. Exercise. *Determine the set of solutions of $q^2 + 1 = 0$.*

Solution. Trivial but important. It is the set of those 'purely imaginary' quaternions (viz. in Vect i, j, k) which have norm 1.

4.4.5. Exercise. *For a ring R, its* opposite ring R^{op} *is the same underlying additive group with op-multiplication defined by $a * b = b \cdot a$. In general there is no reason for R and R^{op} to be isomorphic. Prove however that $\mathbb{H} \simeq \mathbb{H}^{\mathrm{op}}$.*

5 Cayley-Dickson construction of the octonion algebra

The purpose of this lecture is to go beyond $\mathbb{H}$ and discover the *non-associative* algebra of octonions $\mathbb{O}$. It still has some decent algebraic properties. The lecture can be read at any point, or not at all.

Prepare to lose associativity; we retain unitality, but odd objects will appear. § 5.1 is a double prologue: definitions in the non-associative case, and an alternative construction of $\mathbb{H}$, directly from the complex field (more honestly, from $\mathbb{C}$ with $*$). This suggests a general method which we investigate, the *Cayley-Dickson* construction, doing $\mathbb{R} \rightsquigarrow \mathbb{C}$ and $\mathbb{C} \rightsquigarrow \mathbb{H}$ (§ 5.2). Some properties are transferable under $\rightsquigarrow$, which will reveal the non-associative division algebra of *octonions* (§ 5.3).

This lecture is based on Baez' wonderful exposition[4]. (Our notation is different since we prefer to write $a \cdot i$ than $i \cdot a$. This has drammatic consequences when defining non-commutative multiplications.)

5.1 Double prologue: non-associative structures, and $\mathbb{H}$ from $\mathbb{C}$

The two introductory paragraphs are not related: there is one theoretical thread, and one practical thread. We shall combine them only at the end of § 5.2.

[4] Baez, J., 'The Octonions', *Bulletin of the American Mathematical Society*, 39(2), 145–205, 2002

Non-associative structures

5.1.1. Remark. Let $\mathbb{A}$ be a unital $\mathbb{R}$-algebra, not supposed to be associative. Since the centre $Z(\mathbb{A})$ need no longer be a subalgebra, the central embedding lemma 1.2.10 is not literally true.

However, $a \mapsto a \cdot 1_{\mathbb{A}}$ remains an embedding of $\mathbb{R}$ into $\mathbb{A}$. Then by bilinearity, elements of $\mathbb{R} \leq \mathbb{A}$ can be moved freely in products, even through parentheses.

What can it mean for a *non-associative* algebra to be 'field-like'? We give two possible formalisations.

5.1.2. Definition. Let $\mathbb{A}$ be a $\mathbb{K}$-algebra.

- $\mathbb{A}$ is an *algebra with inverses* if it is unital and every non-zero element has a two-sided inverse, viz.:

$$(\forall a \in \mathbb{A})(a \neq 0) \to [(\exists b \in \mathbb{A})(ab = ba = 1)].$$

- $\mathbb{A}$ is a *division algebra* if it is unital and has no zero divisors, viz.:

$$(\forall a, b \in \mathbb{A})((ab = 0) \to (a = 0 \vee b = 0)).$$

5.1.3. Remark. Both are considerably more general than being a skew-field; and they are in general not equivalent.

- If $\mathbb{A}$ is a finite-dimensional division algebra, then multiplication by a on the left λ_a has an inverse, and multiplication by a on the right ρ_a both have inverse maps. But these need not agree by lack of associativity.
- Conversely, if each $a \neq 0$ has a two-sided inverse, then a could still be a zero divisor by lack of associativity.

Hence in the absence of associativity, 'algebra with inverses' and 'division algebra' are unrelated properties. With associativity, 'with inverses' implies 'division'; the converse is true with associativity and finite-dimensionality.

So how much associativity is required to do mathematics? The next definition suggests that at some cost, one can do with less.

5.1.4. Definition. A $\mathbb{K}$-algebra $\mathbb{A}$ is *alternative* if the subalgebra generated by any two elements is associative.

This implies equations such as $a(ab) = (aa)b$, and so on. A characterisation is in exercise 5.4.1.

A complex approach to the quaternions

We give a construction of $\mathbb{H}$ from $\mathbb{C}$. View $\mathbb{H} = \mathbb{R}\cdot 1 + \mathbb{R}\cdot i + \mathbb{R}\cdot j + \mathbb{R}\cdot k$ as a left-vector space over $\mathbb{C} = \mathbb{R} \cdot 1 + \mathbb{R} \cdot i$. (Not studying representations, we are content with a left-vector space.)

Every quaternion q can be written as $q = c_1 + c_2 j$ with $c_1, c_2 \in \mathbb{C} = \mathbb{R} + \mathbb{R} \cdot i$.

5.1.5. Lemma. *Write $\mathbb{H} = \mathbb{C} \oplus \mathbb{C}j$. Then conjugation is given by $(c_1, c_2)^* = (c_1^*, -c_2)$ and multiplication by:*

$$(c_1, c_2) \cdot (d_1, d_2) = (c_1 d_1 - d_2^* c_2, c_2 d_1^* + d_2 c_1).$$

5.1.6. Remark. Before the proof: since $\mathbb{C}$ is commutative, there are other ways to write this identity. *However* this one generalises better as we shall see in definition 5.2.3.

Proof. Let $c_1 = r_1 + s_1 i$ with $r_1, s_1 \in \mathbb{R}$ and c_2 likewise. Then:

$$\begin{aligned}(c_1, c_2)^* &= (c_1 + c_2 j)^* \\ &= (r_1 + s_1 i + r_2 j + s_2 k)^* \\ &= r_1 - s_1 i - r_2 j - s_2 k \\ &= c_1^* - c_2 j \\ &= (c_1^*, -c_2).\end{aligned}$$

The key to the multiplication formula is to notice that if $c = r + si$ with $r, s \in \mathbb{R}$, then $cj = rj + sk = jr - jsi = j(r - si) = jc^*$. So clearly:

$$\begin{aligned}(c_1, c_2) \cdot (d_1, d_2) &= (c_1 + c_2 j)(d_1 + d_2 j) \\ &= c_1 d_1 + c_1 d_2 j + c_2 j d_1 + c_2 j d_2 j \\ &= c_1 d_1 + c_1 d_2 j + c_2 d_1^* j - c_2 d_2^* \\ &= (c_1 d_1 - c_2 d_2^*) + (c_1 d_2 + c_2 d_1^*) \\ &= (c_1 d_1 - d_2^* c_2, c_2 d_1^* + d_2 c_1).\end{aligned}$$

□

Actually, the *same* formula enable one to construct $\mathbb{C}$ from $\mathbb{R}$. In either case we start with a real algebra with some operation * and produce another one. This begs for a general definition. Notice that we demand unitality.

5.2 $*$-Algebras; the Cayley-Dickson construction

5.2.1. Definition. A $*$-*algebra* is a unital $\mathbb{R}$-algebra $\mathbb{A}$ with an involutive, $\mathbb{R}$-linear map * satisfying $(ab)^* = b^* a^*$.

5.2.2. Example. Obviously $\mathbb{R}$ (with trivial $^* = \mathrm{Id}$), $\mathbb{C}$, and $\mathbb{H}$ are $*$-algebras, even associative ones.

Also, $M_n(\mathbb{C})$ with Hermite-conjugation.

5.2.3. Definition. Let $\mathbb{A}$ be a $*$-algebra. Let $\hat{\mathbb{A}} = \mathbb{A}^2$, equiped with involution $(a_1, a_2)^* = (a_1^*, -a_2)$ and multiplication:

$$(a_1, a_2) \cdot (b_1, b_2) = (a_1 b_1 - b_2^* a_2, a_2 b_1^* + b_2 a_1).$$

5.2.4. Lemma. *Let $\mathbb{A}$ be a $*$-algebra. Then $\hat{\mathbb{A}}$ is a $*$-algebra, and $\mathbb{A}$ embeds into $\hat{\mathbb{A}}$ via $a \mapsto (a, 0)$.*

Proof. Clearly * is linear, and has order 2, since:

$$(a_1, a_2)^{**} = (a_1^*, -a_2)^* = (a_1^{**}, a_2).$$

Now the product is clearly bilinear, and $(1, 0)$ is its identity element; moreover, one

has:

$$(a_1, 0) \cdot (b_1, 0) = (a_1 b_1, 0),$$

proving that $a \mapsto (a, 0)$ embeds $\mathbb{A}$ into $\hat{\mathbb{A}}$ as $\mathbb{R}$-algebras. □

For the moment we have only given the algebraic description, not accounting for the behaviour of the norm function.

5.2.5. Definition. Let $\mathbb{A}$ be a $*$-algebra.

- Call $\mathbb{A}$ *real* if $a^* = a$ everywhere.
  (If $\mathbb{A}$ is real then it is commutative, since $ab = (ab)^* = b^* a^* = ba$.)
- $\mathbb{A}$ is *nicely normed* if for all $a \neq 0$ one has $a + a^* \in \mathbb{R}$ and $aa^* = a^* a \in \mathbb{R}_{>0}$.
- The *norm* of a is $N(a) = aa^*$ (especially useful when $\mathbb{A}$ is nicely normed).

Notice that we lose $M_n(\mathbb{C})$, which is not nicely normed.

5.2.6. Remark. Clearly, every nicely normed algebra has two-sided inverses (take $\frac{1}{aa^*} a^*$ in obvious notation).

But recall that in the absence of associativity, there is a difference between 'division algebra' and 'algebra with two-sided inverses'. So in order to get division algebras we need one more assumption. Recall that 'alternative' (definition 5.1.4) could be called 'locally associative': any two elements generate an associative subalgebra.

5.2.7. Lemma. *Every* alternative*, nicely normed $*$-algebra is a division algebra.*

Proof. Let $a, b \in \mathbb{A}$.

Step 1. a, b, a^*, b^* lie in an associative algebra.

Verification. Let $\mathrm{Re}(c) = \frac{1}{2}(c + c^*) \in \mathbb{R}$ and $\mathrm{Im}(c) = \frac{1}{2}(c - c^*)$. Notice that $\mathrm{Re}(c) \in \mathbb{R}$, and $c = \mathrm{Re}(c) + \mathrm{Im}(c)$.

Now 1 (hence $\mathbb{R}$) never violates associativity relations, so it suffices to see that $\mathrm{Im}(a), \mathrm{Im}(b), \mathrm{Im}(a^*) = -\mathrm{Im}(a)$, and $\mathrm{Im}(b^*) = -\mathrm{Im}(b)$ lie in an associative algebra. We have reduced the question to only *two* elements and can use alternativity. ◇

Step 2. $N(c) = cc^* \colon \mathbb{A} \to \mathbb{R}_{\geq 0}$ is multiplicative.

Verification. For $a, b \in \mathbb{A}$ one can freely associate between a, b, a^*, b^*, and therefore:

$$N(ab) = (ab)(ab)^* = (ab)(b^* a^*) = abb^* a^* = aN(b)a^* = N(b)N(a). \diamond$$

We prove the lemma. If $ab = 0$, then $N(a)N(b) = 0$ an identity in $\mathbb{R}$, so one is 0. But $N(c)$ vanishes only at $c = 0$. □

We now transfer properties up from $\mathbb{A}$ to $\hat{\mathbb{A}}$.

5.2.8. Proposition (transfering properties)**.** *Let $\mathbb{A}$ be a $*$-algebra.*

(i) $\mathbb{A}$ is real iff $\hat{\mathbb{A}}$ is commutative.

(ii) $\mathbb{A}$ *is commutative and associative iff* $\hat{\mathbb{A}}$ *is associative.*

(iii) $\mathbb{A}$ *is nicely normed iff* $\hat{\mathbb{A}}$ *is nicely normed.*

(iv) $\mathbb{A}$ *is associative and nicely normed iff* $\hat{\mathbb{A}}$ *is alternative and nicely normed.*

Proof.

(i) If $\mathbb{A}$ is real then it is commutative. Now for $\alpha = (a_1, a_2)$ and $\beta = (b_1, b_2)$ one has:

$$\begin{aligned}\alpha\beta &= (a_1b_1 - b_2^* a_2, a_2 b_1^* + b_2 a_1)\\ &= (a_1b_1 - b_2 a_2, a_2 b_1 + b_2 a_1)\\ &= (b_1a_1 - a_2^* b_2, a_2 b_1 + b_2 a_1^*)\\ &= \beta\alpha\end{aligned}$$

so $\hat{\mathbb{A}}$ is commutative.

Conversely if $\hat{\mathbb{A}}$ is commutative, then:

$$(0, a) = (a, 0)(0, 1) = (0, 1)(a, 0) = (0, a^*),$$

proving that $\mathbb{A}$ is real.

(ii) What an ugly one! You have to compute $[(a_1, a_2)(b_1, b_2)](c_1, c_2)$, then compare with $(a_1, a_2)[(b_1, b_2)(c_1, c_2)]$. It is better to skip this one.

(iii) Suppose that $\mathbb{A}$ is nicely normed. Let $\alpha = (a_1, a_2) \in \hat{\mathbb{A}}$. Then $\alpha + \alpha^* = (a_1 + a_1^*, 0) \in \mathbb{R}$. Moreover,

$$\alpha\alpha^* = (a_1, a_2) \cdot (a_1^*, -a_2) = (a_1a_1^* + a_2^* a_2, a_2 a_1^{**} - a_2 a_1) = (a_1a_1^* + a_2^* a_2, 0).$$

Likewise, $\alpha^*\alpha = (a_1^* a_1 + a_2^* a_2, 0)$. Both are equal since $\mathbb{A}$ is nicely normed. And the result is in $\mathbb{R}_{>0}$ whenever a_1 or a_2 is non-zero.

The converse is obvious once you have noticed that * on $\hat{\mathbb{A}}$ extends * on $\mathbb{A}$.

(iv) One implication is clear: just compute in $\hat{\mathbb{A}}$ using associativity of $\mathbb{A}$. For the converse, we assume that $\hat{\mathbb{A}}$ is nicely normed and alternative. Notice that with the $(a, 0)$ embedding, $\mathbb{A} \leq \hat{\mathbb{A}}$; so $\mathbb{A}$ is alternative. Now let $a, b, c \in \mathbb{A}$, and compute:

$$\begin{aligned}(a, b) \cdot [(a, b) \cdot (0, c)] &= (a, b) \cdot (-c^* b, ca)\\ &= (-a(c^* b) - (a^* c^* b), -b(b^* c) + (ca)a)\\ [(a, b) \cdot (a, b)] \cdot (0, c) &= (aa - b^* b, ba^* + ba) \cdot (0, c)\\ &= (c^*(ba^*) - c^*(ba), c(aa) - c(b^* b).\end{aligned}$$

Taking the second coordinates, in view of alternativity of $\mathbb{A}$ we have $b(b^*)c = c(b^* b)$. Now taking first coordinates:

$$a(c^* b) + (a^* c^*)b = c^*(ba^*) + c^*(ba).$$

Since $\hat{\mathbb{A}}$ is nicely normed, this rewrites into:

$$a(c^*b) + (a^*c^*)b = c^*(b \cdot 2\,\mathrm{Re}(a)).$$

Now scalars can be moved freely by bilinearity, so the right-hand member equals $2\,\mathrm{Re}(a) \cdot (c^*b)$. Meanwhile the left-hand member equals:

$$a(c^*b) + a^*(c^*b) - a^*(c^*b) + a^*c^*)b = 2\,\mathrm{Re}(a)(c^*b) - a^*(c^*b) + (a^*c^*)b.$$

There remains only:

$$a^*(c^*b) = (a^*c^*)b,$$

which implies associativity since $*$ is a bijection. □

5.3 Octonions

5.3.1. Corollary. *$\mathbb{R}^8$ can be equipped with the structure of a unital, alternative division algebra.*

Proof. We see things more generally and iterate the Cayley-Dickson construction through proposition 5.2.8. First, $\mathbb{R}$ is a $*$-algebra with respect to the trivial $*$; it is real, associative, and commutative. Therefore $\mathbb{C} = \hat{\mathbb{R}}$ is associative and commutative. Therefore $\mathbb{H} = \hat{\mathbb{C}}$ is associative. At all stages, algebras were nicely normed and unital; now $\mathbb{O} = \hat{\mathbb{H}}$ is alternative and nicely normed. By lemma 5.2.7, it is a division algebra. □

$\mathbb{O}$ is called the algebra of *octonions*. It is *not* associative. Therefore the collection of invertible elements is not a group, more something like a 'non-associative group'. For completeness we include the definition.

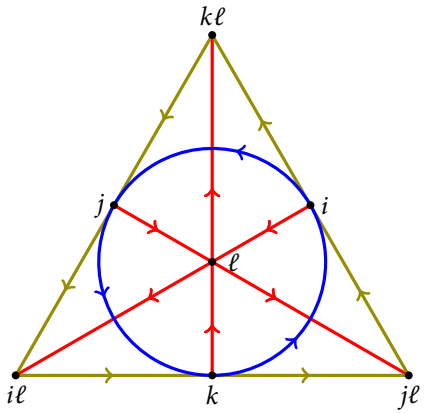
5.3.2. Definition. Definition not given in class.

- A *loop* is a structure $(L, *)$ such that:
  - $(\exists e \in L)(\forall x \in L)(x * e = x * e = x)$;
  - $(\forall a, b \in L)(\exists! x \in L)(a * x = b)$;
  - $(\forall a, b \in L)(\exists! x \in L)(x * a = b)$.
- A *Moufang loop* is a loop satisfying in addition, for all $x, y, z \in L$:
  - $z(x(zy)) = ((zx)z)y$;
  - $x(z(yz)) = ((xz)y)z$;
  - $(zx)(yz) = (z(xy))z$;
  - $(zx)(yz) = z((xy)z)$.

5.3.3. Corollary. *$\mathbb{S}^7 = \{v \in \mathbb{R}^8 : \|v\| = 1\}$ can be equipped with an algebraic Moufang loop structure (almost like a group, but associativity fails).*

Proof. $\mathbb{S}^7$ is the set of octonions with norm 1: which is stable under multiplication, since the norm is multiplicative as we just saw. □

Here is a useful way to picture octonions. Let $\mathbb{R}^8$ have basis $\{1, i, j, k, \ell, i\ell, j\ell, k\ell\}$. Multiplication is given by the following diagram.



Each edge is a triple (x, y, z); the arrow tells us $xy = z$. For instance, the left-most brown edge tells us $(k\ell) \cdot j = i\ell$.
Hence $(j\ell) \cdot i = k\ell$, while $j \cdot (\ell i) = -j \cdot (i\ell) = -k\ell$: associativity fails.

The algebra of 'sedenions' $\hat{\mathbb{O}}$ is a curiosity and not worth discussing; things get only worse and worse. Are octonions a curiosity or do they explain something?

5.3.4. Theorem (É. Cartan)**.** $\mathrm{Aut}(\mathbb{O}) \simeq G_2$, *the exceptional group discovered by Dickson.*

5.3.5. Remark. If you have followed the present lecture this far, you might as well start reading by yourself about Freudenthal's magic square. (The only reasonable conclusion to this class is: learn Lie theory.)

5.4 Exercises

5.4.1. Exercise (E. Artin)**.** *Let $\mathbb{A}$ be an algebra such that for all $a, b \in \mathbb{A}$:*

$$(\forall a, b \in \mathbb{A})\ [\ a(ab) = (aa)b \ \wedge\ (ba)a = b(aa)\ ].$$

Prove that $\mathbb{A}$ is alternative. Hint: it suffices to prove $a(ba) = (ab)a$. Use the associator $[\![x, y, z]\!] = (xy)z - x(yz)$.

Solution. Let $a, b \in \mathbb{A}$; we must show that the subalgebra $\langle a, b\rangle \leq \mathbb{A}$ generated by a and b is associative. Using bilinearity of the multiplication (and in the case of a unital algebra, since 1 never breaks associativity), it suffices to check only three identities: $a(ab) = (aa)b$, $a(ba) = (ab)a$, and $a(bb) = (ab)b$.

The first and last are given by assumption. So we introduce the associator $[\![x, y, z]\!] = (xy)z - x(yz)$. Notice that $[\![x, y, z]\!]$ iff $(xy)z = x(yz)$. Moreover, $[\![x, y, z]\!]$ is trilinear. Then we compute:

$$\begin{aligned}[\![a-b, a-b, a]\!] &= [\![a, a, a]\!] - [\![a, b, a]\!] - [\![b, a, a]\!] + [\![b, b, a]\!] \\ &= 0 - [\![a, b, a]\!] - 0 + 0.\end{aligned}$$

But the left-hand as well is 0, applying the axiom to $c = a - b$ and a. There remains $[\![a, b, a]\!] = 0$, as desired.

5.4.2. Exercise. *Prove the following extraordinarily ugly identity (Degen, 1818), where all numbers are real:*

$$\begin{aligned}(a_1^2+a_2^2+a_3^2+a_4^2+a_5^2+a_6^2+a_7^2+a_8^2)(b_1^2+b_2^2+b_3^2+b_4^2+b_5^2+b_6^2+b_7^2+b_8^2)\\ = \begin{cases} & (a_1b_1-a_2b_2-a_3b_3-a_4b_4-a_5b_5-a_6b_6-a_7b_7-a_8b_8)^2\\ + & (a_1b_2+a_2b_1+a_3b_4-a_4b_3+a_5b_6-a_6b_5-a_7b_8+a_8b_7)^2\\ + & (a_1b_3-a_2b_4+a_3b_1+a_4b_2+a_5b_7+a_6b_8-a_7b_5-a_8b_6)^2\\ + & (a_1b_4+a_2b_3-a_3b_2+a_4b_1+a_5b_8-a_6b_7+a_7b_6-a_8b_5)^2\\ + & (a_1b_5-a_2b_6-a_3b_7-a_4b_8+a_5b_1+a_6b_2+a_7b_3+a_8b_4)^2\\ + & (a_1b_6+a_2b_5-a_3b_8+a_4b_7-a_5b_2+a_6b_1-a_7b_4+a_8b_3)^2\\ + & (a_1b_7+a_2b_8+a_3b_5-a_4b_6-a_5b_3+a_6b_4+a_7b_1-a_8b_2)^2\\ + & (a_1b_8-a_2b_7+a_3b_6+a_4b_5-a_5b_4-a_6b_3+a_7b_2+a_8b_1)^2 \end{cases}\end{aligned}$$

Solution. No nead to read it entirely; just use the multiplicativity of the octonion norm.

6 Frobenius' classification theorem

This lecture can be followed at any point but requires knowledge of § 3.

Recall that the only finite-dimensional $\mathbb{R}$-algebras which are commutative fields are $\mathbb{R}$ and $\mathbb{C} = \mathbb{R}[i]$ (theorem 3.2.2). But dropping commutativity we also have $\mathbb{H}$, a finite-dimensional skew-field. One may wonder whether there are more such objects: and the answer is no. We prove this in § 6.1 and discuss some generalisations (without proving them) in § 6.2.

6.1 The original theorem

6.1.1. Theorem (Frobenius, 1877)**.** *Let $\mathbb{A}$ be a finite-dimensional, associative, unital $\mathbb{R}$-algebra which is a skew-field. Then as $\mathbb{R}$-algebras one has $\mathbb{A} \simeq \mathbb{R}$, $\mathbb{A} \simeq \mathbb{C}$, or $\mathbb{A} \simeq \mathbb{H}$.*

The associativity assumption is of course redundant, since $\mathbb{A}$ is supposed to be a skew-field. But it is good to keep track of it.

Proof. Theorem 3.2.2 will be used repeatedly, simply refering to 'the commutative case'.

Let $\mathbb{A}$ be as in the statement: being a skew-field, it is associative and has a multiplicative identity. By the commutative case, we may suppose that $\mathbb{A}$ is non-commutative.

Step 1. Finding i satisfying $i^2 = -1$.

Verification. Since $\mathbb{A}$ is non-commutative, $\dim_{\mathbb{R}} \mathbb{A} > 1$. Let $a \in \mathbb{A} \smallsetminus \mathbb{R}$; then $\mathbb{R}[a]$ is an associative and commutative $\mathbb{R}$-algebra of dimension > 1; it is a domain (hence a commutative field). By the commutative case, $\mathbb{R}[a] \simeq \mathbb{C}$; so there is $i \in \mathbb{R}[a] \leq \mathbb{A}$ with $i^2 = -1$. ◇

Let $\mathbb{C} = \mathbb{R}[i] \leq A$; it is not canonical and there are other copies of $\mathbb{C}$ inside $\mathbb{A}$ but we fix this one. We know that $\mathbb{A}$ is a left-vector space over $\mathbb{C}$. (It is *not* a $\mathbb{C}$-algebra; as

a matter of fact, if it were we would get $\mathbb{A} = \mathbb{C}$.)

Step 2. Finding j satisfying $j^2 = -1 \wedge ij = -ji$.

Verification. The first thing is to understand anti-commutation. Let:

$$\begin{array}{rccc} \rho_i\colon & \mathbb{A} & \to & \mathbb{A} \\ & a & \mapsto & ai \end{array}$$

be the right-multiplication by i. By associativity, it is $\mathbb{C}$-linear. Now $\rho_i^2 = -\mathrm{Id}$, so in $\mathrm{End}_{\mathbb{C}}(\mathbb{A})$, the linear map ρ_i is diagonalisable with eigenvalues $\pm i$. Let $\mathbb{A}_i$ and $\mathbb{A}_{-i}$ be the eigenspaces. Notice that $\mathbb{A}_i$ is the centraliser of i, viz. $\{x \in \mathbb{A} : xi = ix\}$, while $\mathbb{A}_{-i}$ is the subspace of elements anti-commuting with i. Our search for j will be inside $\mathbb{A}_{-i}$; before, we study the centraliser.

We contend that $\mathbb{A}_i = \mathbb{R}[i]$. On the one hand, $\mathbb{R}[i] \simeq \mathbb{C}$ is commutative, so certainly $\mathbb{R}[i] \leq \mathbb{A}_i$. Conversely if $a \in \mathbb{A}_i$, then a commutes with i: hence $\mathbb{R}[i, a]$ is a *commutative* subalgebra of $\mathbb{A}$, and still a domain; it is a finite-dimensional field extension of $\mathbb{R}[i]$ so by the commutative case, $\mathbb{R}[i, a] = \mathbb{R}[i]$ and $a \in \mathbb{R}[i]$. Therefore $\mathbb{A}_i = \mathbb{R}[i] \simeq \mathbb{C}$.

Since $\mathbb{A}$ is not commutative, $\mathbb{A}_i < \mathbb{A}$; since $\mathbb{A} = \mathbb{A}_i \oplus \mathbb{A}_{-i}$ by diagonalisability, there is $a \in \mathbb{A}_{-i} \smallsetminus \{0\}$. Notice that $a \notin \mathbb{A}_i$ since otherwise $ia = ai = -ia$. Always by the commutative case, $\mathbb{R}[a] \simeq \mathbb{C}$ (now a different copy). To find j we shall find inside the vector line $\mathbb{R} \cdot a$ an element squaring to -1. The argument is elementary but clever.

Let $L = \mathbb{R}[a] \cap \mathbb{R}[i]$, an $\mathbb{R}$-subspace of $\mathbb{R}[i]$. Since $1 \in L$ one has $\dim_{\mathbb{R}} L > 0$; since $a \notin \mathbb{R}[i]$ one has $\dim_{\mathbb{R}} L < 2$. So $\dim_{\mathbb{R}} L = 1$ and L exactly the vector line $\mathbb{R} = \mathbb{R} \cdot 1$. Return to $a \in \mathbb{A}_{-i}$; one has:

$$a^2 i = a \cdot ai = a \cdot -ia = -aia = ia^2,$$

so $a^2 \in \mathbb{R}[a] \cap \mathbb{A}_i = \mathbb{R}[a] \cap \mathbb{R}[i] = \mathbb{R}$.

Of course $a^2 \neq 0$. If $a^2 \in \mathbb{R}_{>0}$ then inside the commutative field $\mathbb{R}[a] \simeq \mathbb{C}$ we find $a \in \mathbb{R}$, which is a contradiction to $a \notin \mathbb{A}_i$. Hence $a^2 \in \mathbb{R}_{<0}$; rescaling we may assume $a^2 = -1$, and let $j = a \in \mathbb{A}_{-i}$. We have found j with $j^2 = -1$ and $ji = -ij$. ⋄

Step 3. Finding k and identifying.

Verification. Let $k = ij$. Observe how $k^2 = ijij = -i^2 j^2 = -1$; moreover $ik = -j$, $jk = jij = -ij^2 = i$, $ki = iji = j$, and $kj = -i$.

It is however not fully clear that $1, i, j, k$ are linearly independent over $\mathbb{R}$. But $1, i, j$ are, since $j \notin \mathrm{Vect}\, 1, i \simeq \mathbb{C}$. Now suppose that $k = a + bi + cj$ for some coefficients from $\mathbb{R}$. Then multiplying on the left by j one gets:

$$\begin{aligned} i &= aj - bk - c \\ &= aj - b(a + bi + cj) - c, \end{aligned}$$

whence $b^2 = -1$, a contradiction.

The above proves that $\mathrm{Vect}\, 1, i, j, k \leq \mathbb{A}$ is a subalgebra isomorphic to $\mathbb{H}$. To conclude it remains to prove equality: namely $\dim_{\mathbb{R}} \mathbb{A} = 4$. Consider the (left-)multiplication map by j, and restrict it to $\mathbb{A}_{-i}$. If $a \in \mathbb{A}_{-i}$ then $ai = -ia$; now

$jai = -jia = ija$ so $ja \in \mathbb{A}_i$. Thus λ_j injects $\mathbb{A}_{-i}$ into $\mathbb{A}_i$; a similar argument proves that it also injects $\mathbb{A}_i$ into $\mathbb{A}_{-i}$. As a consequence, $\mathbb{A}_i \simeq \mathbb{A}_{-i}$ as real vector spaces, and they have the same dimension. Therefore $\dim_{\mathbb{R}} \mathbb{A} = \dim_{\mathbb{R}} \mathbb{A}_i + \dim_{\mathbb{R}} \mathbb{A}_{-i} \leq 2 + 2 = 4$ and we are done. ◇

This completes the proof. □

6.1.2. Remarks.

- There are *commutative* counter-examples of infinite dimension, such as $\mathbb{C}(X)$.
- The theorem remains true over real closed fields, but not arbitrary fields. For instance there exist 9-dimensional $\mathbb{Q}$-algebras which are skew-fields. Constructing them is already fairly involved and better understood through Galois theory.

6.2 Generalisations of Frobenius' theorem

The associative world

Theorem 6.1.1 admits several extensions. One is about *normed* algebras; here, 'norm' is in the geometric sense (as in linear algebra). The theorem asserts that existence of a *submultiplicative* norm (one with $\|ab\| \leq \|a\| \cdot \|b\|$) is a strong geometric constraint, as strong as finite-dimensionality.

6.2.1. Theorem (Gelfand-Mazur Theorem; Mazur, 1938)**.** *Let $\mathbb{A}$ be an associative, unital $\mathbb{R}$-algebra which has a submultiplicative norm and is a skew-field. Then $\mathbb{A} \simeq \mathbb{R}$, $\mathbb{A} \simeq \mathbb{C}$, or $\mathbb{A} \simeq \mathbb{H}$.*

Proof in the case of a complex Banach algebra. We prove a very special case of the Gelfand-Mazur theorem:

> Let $\mathbb{A}$ be an associative, unital $\mathbb{C}$-algebra which has a submultiplicative norm $\|\cdot\|$ and is a skew-field. If $(\mathbb{A}, \|\cdot\|)$ is complete, then $\mathbb{A} \simeq \mathbb{C}$.

The completeness assumption is removed in exercise 6.3.2 (always for a $\mathbb{C}$-algebra).

Let $\mathbb{A}$ be given; by central embedding we suppose $\mathbb{C} \leq \mathbb{A}$; even suppose $\mathbb{C} < \mathbb{A}$ and fix some $a \notin \mathbb{C}$. Working in the closed subalgebras $C_{\mathbb{A}}(a)$, then $Z(C_{\mathbb{A}}(a))$, we may assume that $\mathbb{A}$ is commutative; actually this will play no role at all.

Consider the entire series:

$$(a - \lambda \cdot 1_{\mathbb{A}})^{-1} = a^{-1} \cdot \sum_{n \geq 0} \frac{\lambda^n}{a^n}.$$

The function $a - \lambda \cdot 1_{\mathbb{A}}$ is easily seen holomorphic. Since it does not vanish, the domain of its inverse, as a holomorphic function, is $\mathbb{C}$, and the right-hand is its expansion. Therefore the right-hand must have infinite radius.

However $\|1_{\mathbb{A}}\| = \|a^n \cdot a^{-n}\| \leq \|a\|^n \cdot \|a^{-n}\|$, so:

$$\sqrt[n]{\|a^{-n}\|} \geq \frac{1}{\|a\|}.$$

By the usual criteria on formal series (and completeness of $\mathbb{A}$), the series $a^{-1}\sum_{n\geq 0}\lambda^n\cdot a^{-n}$ has a *finite* convergence radius $\leq \|a\|$, a contradiction. □

The non-associative world

In § 5, relaxing 'associative skew-field' into 'alternative division algebra' has created the octonions $\mathbb{O}$; we may wish to characterise them uniquely. The relevant definitions are in § 5.1. As opposed to the Gelfand-Mazur theorem, Hurwitz' does require multiplicativity.

6.2.2. Theorem (Hurwitz, 1898)**.** *Let $\mathbb{A}$ be a (not necessarily associative) finite-dimensional, $\mathbb{R}$-division algebra with a multiplicative norm. Then $\mathbb{A}\simeq\mathbb{R},\mathbb{C},\mathbb{H},\mathbb{O}$.*

6.2.3. Theorem (Zorn, 1930)**.** *Let $\mathbb{A}$ be an alternative, unital, finite-dimensional, $\mathbb{R}$-division algebra. Then $\mathbb{A}\simeq\mathbb{R},\mathbb{C},\mathbb{H},\mathbb{O}$.*

And finally, the strongest to my knowledge.

6.2.4. Theorem (Hopf 1940, completed by Kervaire-Milnor 1958)**.** *Let $\mathbb{A}$ be a finite-dimensional $\mathbb{R}$-division algebra. Then $\dim_{\mathbb{R}}\mathbb{A}\in\{1,2,4,8\}$.*

Of course we have shifted from general algebraic structures to functional analysis and serious geometry, which are more interesting topics indeed.

6.3 Exercises

6.3.1. Exercise. *The purpose of this exercise is to prove Zorn's theorem using Hurwitz'. Let $\mathbb{A}$ be an alternative, unital, finite-dimensional $\mathbb{R}$-division algebra.*

1. *For $a\in\mathbb{A}\smallsetminus\mathbb{R}$, show that $\mathrm{Min}_{\mathbb{R}}^{a}$ makes sense and is of the form X^2+pX+q.*
2. *Let $N(a)=q$ as above (for real a, take $N(a)=a^2$). Prove that N is multiplicative.*
3. *Conclude using Hurwitz' theorem.*

Solution.

1. Let $a\in\mathbb{A}\smallsetminus\mathbb{R}$. The subalgebra $\langle a\rangle$ it generates is associative (by alternativity), commutative, and a finite-dimensional extension of $\mathbb{R}$. By theorem 3.2.2, one has $\langle a\rangle\simeq\mathbb{C}$. There, the minimal polynomial of a has degree 2, and the desired form.
2. Now let $a,b\in\mathbb{A}$. We may suppose that neither is in $\mathbb{R}$. By alternativity, $\langle a,b\rangle$ is associative; by finite-dimensionality, it is isomorphic at most to $\mathbb{H}$. But there, $N(a)$ and $N(b)$ coincide with the quaternion norm, so $N(ab)=N(a)N(b)$.
3. Still working in $\langle a,b\rangle$, we have $\sqrt{N(a+b)}\leq\sqrt{N(a)}+\sqrt{N(b)}$. So we have equipped $\mathbb{A}$ be a multiplicative norm. By Hurwitz' theorem, $\mathbb{A}\simeq\mathbb{R},\mathbb{C},\mathbb{H}$, or $\mathbb{O}$.

6.3.2. Exercise (complex Mazur theorem)**.** *The purpose of this longer exercise is an elementary proof of the following special case of the Gelfand-Mazur theorem.*[5]

Theorem. Let $\mathbb{A}$ be an associative, unital $\mathbb{C}$-algebra which has a submultiplicative norm $\|\cdot\|$ and is a skew-field. Then $\mathbb{A}\simeq\mathbb{C}$.

[5] Mazet, P., 'La preuve originale de S. Mazur pour son théorème sur les algèbres normées', *Gazette de la SMF*, 111, 5–11, 2007

By central embedding we assume $\mathbb{C} \leq \mathbb{A}$. Throughout, λ, μ denote complex numbers.

1. *Let $a \in \mathbb{A} \smallsetminus \mathbb{C}$ be fixed and $\varphi(\lambda) = \frac{1}{a-\lambda} : \mathbb{C} \to \mathbb{A}$.*

   (a) *Prove relation $\frac{\lambda}{X-\lambda} = \frac{X}{X-\lambda} - 1$, and deduce that $\|\varphi(\lambda)\| \xrightarrow[|\lambda|\to+\infty]{} 0$.*

   (b) *Using a similar rational expression in λ and μ, prove continuity of $\frac{1}{\|\varphi\|}$ and φ.*

2. *Let $j \in \mathbb{C}$ be a primitive third root of 1, viz. $j^3 = 1 \neq j$. For $f : \mathbb{C} \to \mathbb{A}$ we define the function of* two *complex variables:*

$$(\Delta f)(\lambda, \mu) = f(\lambda + \mu) + f(\lambda + j\mu) + f(\lambda + j^2\mu).$$

   (a) *Let $f = N : \mathbb{C} \to \mathbb{R} \leq \mathbb{A}$ which computes $N(\lambda) = \lambda\lambda^* = |\lambda|^2$. Show that $(\Delta N)(\lambda, \mu) = 3N(\lambda) + 3N(\mu)$.*

   (b) *Prove identity:*

$$\frac{1}{X-\mu} + \frac{1}{X-j\mu} + \frac{1}{X-j^2\mu} = \frac{3}{X} + \frac{3\mu^3}{X(X-\mu)(X-j\mu)(X-j^2\mu)},$$

   and use it to get that for fixed λ, $(\Delta\varphi)(\lambda, \mu) = 3\varphi(\lambda) + o(\mu^2)$.

3. *Let $\varepsilon \in \mathbb{R}_{>0}$ and $\chi_\varepsilon = \|\varphi\| + \varepsilon N$.*

   (a) *Fix λ and show that $(\Delta\chi_\varepsilon)(\lambda, \mu) > 3\chi_\varepsilon(\lambda)$ for small $\mu \neq 0$.*

   (b) *Let $R \in \mathbb{R}_{\geq 0}$ and $D_R = \{\lambda \in \mathbb{C} : |\lambda| \leq R\}$ be the closed disk of radius R. Prove that χ_ε attains its maximum in D_R on the boundary $C_R = \{\lambda \in \mathbb{C} : |\lambda| = R\}$.*

   (c) *Conclude that:*

$$\|\varphi(0)\| \leq \max_{C_R} \|\varphi\| + \varepsilon R^2,$$

   and a final contradiction.

Solution. Notice that we fix only one $a \in \mathbb{A}$, never two; so we do not require commutativity of $\mathbb{A}$. But the proof does require complex numbers, so $\mathbb{A}$ must be a $\mathbb{C}$-algebra for the present argument.

1. Since $a \notin \mathbb{C}$, for $\lambda \in \mathbb{C}$ the difference $a - \lambda$ is never 0; since $\mathbb{A}$ is a field, φ is well-defined.

   (a) The rational relation is trivial. We let $X = a$ and find $\frac{\lambda}{a-\lambda} = \frac{a}{a-\lambda} - 1$, equivalently $\lambda\varphi(\lambda) = a\varphi(\lambda) - 1$. This implies $|\lambda|\|\varphi(\lambda)\| \leq \|a\| \cdot \|\varphi(\lambda)\| + \|1\|$. So as $|\lambda|$ becomes large, $\|\varphi(\lambda)\| \leq \frac{\|1\|}{|\lambda| - \|a\|}$, which goes to 0.

   (b) We now guess $\frac{1}{X-\lambda} - \frac{1}{X-\mu} = \frac{\lambda-\mu}{(X-\lambda)(X-\mu)}$, with the effect that $\varphi(\lambda) - \varphi(\mu) = (\lambda - \mu)\varphi(\lambda)\varphi(\mu)$. Dividing and taking norms, $\left\|\frac{1}{\varphi(\mu)} - \frac{1}{\varphi(\lambda)}\right\| \leq |\lambda - \mu|$. Do not forget the 'other' triangle inequality $|\|a\| - \|b\|| \leq \|a - b\|$; here, $\left|\frac{1}{\|\varphi(\lambda)\|} - \frac{1}{\|\varphi(\mu)\|}\right| \leq |\lambda - \mu|$. This proves continuity of $\frac{1}{\|\varphi\|}$, and therefore of $\|\varphi\|$ and φ as well.

2. Recall that $1 + j + j^2 = 1$ and $j^* = j^2$.

(a) Simply compute:

$$\begin{aligned}(\Delta N)(\lambda,\mu) &= (\lambda+\mu)(\lambda+\mu)^* + (\lambda+j\mu)(\lambda+j\mu)^* + (\lambda+j^2\mu)(\lambda+j^2\mu)^* \\ &= \lambda\lambda^* + \lambda\mu^* + \mu\lambda^* + \mu\mu^* + \lambda\lambda^* + j^2\lambda\mu^* + j\mu\lambda^* + j^3\mu\mu^* \\ &\quad + \lambda\lambda^* + j\lambda\mu^* + j^2\mu\lambda^* + j^3\mu\mu^* \\ &= 3\lambda\lambda^* + (1+j^2+j)\lambda\mu^* + (1+j+j^2)\mu\lambda^* + 3\mu\mu^*,\end{aligned}$$

with desired vaninshing terms.

(b) Start with $\frac{3\mu^3}{X(X-\mu)(X-j\mu)(X-j^2\mu)}$, which has an expansion of the form:

$$\frac{3\mu^3}{X(X-\mu)(X-j\mu)(X-j^2\mu)} = \frac{c_0}{X} + \frac{c_1}{X-\mu} + \frac{c_2}{X-j\mu} + \frac{c_3}{X-j^2\mu}.$$

Taking the residual value at pole 0, we find $c_0 = \frac{3\mu^3}{-j^3\mu^3} = -3$. Then at pole μ, we obtain $c_1 = \frac{3\mu^3}{(1-j)(1-j^2)\mu^3} = 1$; find c_2 and c_3 likewise.
We apply the identity at $a - \lambda$ and get:

$$(\Delta\varphi)(\lambda,\mu) = 3\varphi(\lambda) + 3\mu^3\varphi(\lambda)\varphi(\lambda+\mu)\varphi(\lambda+j\mu)\varphi(\lambda+j^2\mu).$$

Since φ is continuous, for fixed λ the last term is $\mu^3\cdot(\varphi(\lambda)+o(1)) = o(\mu^2)$.

3. Notice that $f \mapsto \Delta f$ is linear.

(a) As a consequence, $\Delta\chi_\varepsilon = \Delta\|\varphi\| + \varepsilon\Delta N$. In particular, using the triangle inequality,

$$\begin{aligned}(\Delta\chi_\varepsilon)(\lambda,\mu) &= (\Delta\|\varphi\|)(\lambda,\mu) + \varepsilon(\Delta N)(\lambda,\mu) \\ &\geq \|\Delta\varphi\|(\lambda,\mu) + 3\varepsilon(N(\lambda)+N(\mu)) \\ &= 3\|\varphi(\lambda)\| + 3\varepsilon N(\lambda) + 3\varepsilon N(\mu) + o(\mu^2) \\ &= 3\chi_\varepsilon(\lambda) + 3\varepsilon N(\mu) + o(\mu^2).\end{aligned}$$

So if μ is a small non-zero complex number, we have $(\Delta\chi_\varepsilon)(\lambda,\mu) > 3\chi_\varepsilon(\lambda)$.

(b) By definition of $\Delta\chi_\varepsilon$, this implies:

$$\max\{\chi_\varepsilon(\lambda+\mu), \chi_\varepsilon(\lambda+j\mu), \chi_\varepsilon(\lambda+j^2\mu)\} > \chi_\varepsilon(\lambda).$$

As a consequence, χ_ε can never attain a maximum inside an open set $U \subseteq \mathbb{C}$. Indeed if $\lambda \in U$, then for μ small enough the three points $\lambda+\mu, \lambda+j\mu, \lambda+j^2\mu$ will be in U, contradicting maximality at λ.
Let D_R be as suggested; it is a compact set while χ_ε is continuous. By the above, χ_ε attains its maximum on the boundary C_R.

(c) Therefore:

$$\|\varphi(0)\| = \chi_\varepsilon(0) \leq \max_{C_R}\chi_\varepsilon = \max_{C_R}\|\varphi\| + \varepsilon R^2.$$

We first let $\varepsilon \longrightarrow 0$. This proves $\|\varphi(0)\| \leq \max_{C_R}\|\varphi\|$. Now we let $R \longrightarrow +\infty$; since $\|\varphi(\lambda)\| \xrightarrow[|\lambda|\to+\infty]{} 0$, we find $\|\varphi(0)\| = 0$. Hence $\varphi(0) = 0$, clearly a contradiction.

7 An application to Lagrange's four square theorem

This lecture, which can be followed at any point (or not at all), is a fascinating digression: quaternions found a striking application to an earlier result in number theory. Needless to say, in the present section we prefer to work with $N(q) = qq^* = q^*q = a^2+b^2+c^2+d^2$ (in obvious coordinates) instead of $|q| = \sqrt{N(q)}$.

Theorem (Lagrange, 1770). *Every integer is a sum of four squares.*

Remarks.

- No, 7 is *not* a sum of three squares.
- Euler (1749) had proved that n is a sum of two squares iff for each prime p congruent to 3 modulo 4 the p-adic valuation of n, viz. the highest power of p in n, is even; something announced by Fermat without a proof (as usual).

The proof starts here.

7.1 and multiplicative stability

We shall first prove that the problem reduces to prime numbers, and then deal with those. Both steps involve quaternions although the first is substantially easier.

7.1.1. Lemma. *The set of integers which are a sum of four squares is closed under product.*

Proof. Say $n_1 = a_1^2 + b_1^2 + c_1^2 + d_1^2$, and n_2 likewise. Let $q_1 = a_1 + b_1 i + c_1 j + d_1 k$, and q_2 likewise; both are quaternions with integer coefficients, viz. elements of the $\mathbb{Z}$-submodule of $\mathbb{H}$ generated by $1, i, j, k$, viz. elements of the ring $\mathbb{Z}[1, i, j, k] = \mathbb{Z}[i, j]$.

So is $q_3 = q_1q_2$. Hence $N(q_3) = N(q_1)N(q_2) = n_1n_2$ is a sum of four squares. □

7.1.2. Remark. In full form, the identity one could write is:

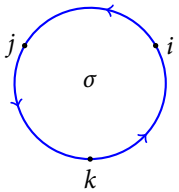
$$(a_1^2 + b_1^2 + c_1^2 + d_1^2) \cdot (a_2^2 + b_2^2 + c_2^2 + d_2^2) = \begin{cases} & (a_1a_2 - b_1b_2 - c_1c_2 - d_1d_2)^2 \\ + & (a_1b_2 + b_1a_2 + c_1d_2 - d_1c_2)^2 \\ + & (a_1c_2 - b_1d_2 + c_1a_2 + d_1b_2)^2 \\ + & (a_1d_2 + b_1c_2 - c_1b_2 + d_1a_2)^2 \end{cases}$$

This 'four-square identity' was discovered by Euler (1748)—without quaternions, such a discovery certainly required his virtuosity.

7.2 The ring of Hurwitz quaternions

The lemma 7.1.1 involves the ring $\mathbb{Z}[i, j] = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot i + \mathbb{Z} \cdot j + \mathbb{Z} \cdot k$; to continue the proof this would not suffice as will be clear in its end.

7.2.1. Definition (and notation). Let $\sigma = \frac{1+i+j+k}{2}$; this quaternion has norm 1 and acts like the 3-cycle we already met in the proof of proposition 4.1.2:



,

meaning that $\sigma i \sigma^{-1} = j$, $\sigma j \sigma^{-1} = k$, and $\sigma k \sigma^{-1} = i$. *(Take two minutes to check it.)*

Let A be the subring of $\mathbb{H}$ generated by i and σ, called the ring of *Hurwitz quaternions*. (We cannot use letter H because of Hamilton.)

7.2.2. Lemma. *$A = \mathbb{Z}\cdot 1 + \mathbb{Z}\cdot i + \mathbb{Z}\cdot j + \mathbb{Z}\cdot k + \mathbb{Z}\cdot\sigma = \mathbb{Z}[i,j] \cup (\sigma + \mathbb{Z}[i,j]) = \mathbb{Z}[i,\sigma]$ is an associative (but non-commutative) subring of $\mathbb{H}$. Moreover:*

(i) A is stable under *;

(ii) for every $a \in A$, one has $N(a) \in \mathbb{N}$ (recall that this is the number-theoretic *norm, in geometric terms $N(a) = |a|^2$);*

(iii) The group of invertible elements (also called units*) satisfies $A = \mathbb{Z}[i,j] \times A^*$.*

Proof.

Step 1. $A = \mathbb{Z}\cdot 1 + \mathbb{Z}\cdot i + \mathbb{Z}\cdot j + \mathbb{Z}\cdot k + \mathbb{Z}\cdot\sigma = \mathbb{Z}[i,j] \cup (\sigma + \mathbb{Z}[i,j]) = \mathbb{Z}[i,\sigma]$ is a ring stable under *.

Verification. Recall or check again that $\sigma i \sigma^{-1} = j$, $\sigma j \sigma^{-1} = k$, and $\sigma k \sigma^{-1} = i$. Then:

$$\mathbb{Z}\cdot 1 + \mathbb{Z}\cdot i + \mathbb{Z}\cdot j + \mathbb{Z}\cdot k + \mathbb{Z}\cdot\sigma \subseteq \mathbb{Z}[i,j] \cup (\sigma + \mathbb{Z}[i,j]) \subseteq \mathbb{Z}[i,\sigma].$$

We shall prove the missing inclusion; it is enough to see that the left-hand member is a subring. Let $A_0 = \mathbb{Z}\cdot 1 + \mathbb{Z}\cdot i + \mathbb{Z}\cdot j + \mathbb{Z}\cdot k + \mathbb{Z}\cdot\sigma$, a priori only a $\mathbb{Z}$-module.

Now observe that $N(\sigma) = \frac{1}{4}(1+1+1+1) = 1$, and $\sigma^* = \frac{1-i-j-k}{2} = 1-\sigma$; this already proves us that A_0 is closed under *. Moreover, $\sigma^2 = \sigma\sigma^{**} = \sigma(1-\sigma)^* = \sigma - N(\sigma) \in A_0$.

Then a quick computation gives:

$$i\sigma = \frac{i - 1 + k - j}{2} = \sigma - 1 - j.$$

We may now quickly check that $\mathbb{Z}\cdot 1 + \mathbb{Z}\cdot i + \mathbb{Z}\cdot j + \mathbb{Z}\cdot k + \mathbb{Z}\cdot\sigma$ is closed under $\cdot$. First,

$$\sigma \cdot i = (\sigma i)^{**} = (i^*\sigma^*)^* = (-i(1-\sigma))^* = (i\sigma - i)^* = (\sigma - 1 - i - j)^* \in A_0.$$

Since $\sigma i \sigma^{-1} = j$, we derive:

$$j\sigma = \sigma i \in A_0.$$

Similar arguments handle $\sigma j, k\sigma, \sigma k$. We already checked that $\sigma^2 \in A_0$. So A_0 is a subring of $\mathbb{H}$; since it contains i and σ, we find $A_0 = \mathbb{Z}[i,\sigma] = A$. ◇

An important consequence is that if $a \in A$, then either $a \in \mathbb{Z}[i,j]$, meaning that all coordinates of a are integers, or $a \in \sigma + \mathbb{Z}[i,j]$, meaning that all coordinates of a are in $\frac{1}{2} + \mathbb{Z}$.

Step 2. Norm properties.

Verification. Let $a \in A = \mathbb{Z}[i,j] \cup (\sigma + \mathbb{Z}[i,j])$. Then either a has integer coordinates in $(1,i,j,k)$, or it has *all* coordinates in $\frac{1}{2} + \mathbb{Z}$; in either case the norm is an integer.

We claim that $a \in A$ is invertible in A iff $N(a) = 1$. If a is invertible, then there is $b \in A$ with $ab = 1$, so $N(a)N(b) = N(ab) = N(1) = 1$; since all are integers, we

find $N(a) = 1$. Conversely if $N(a) = 1$ then the two-sided inverse of a is $a^* \in A$, by closure under $*$. ◇

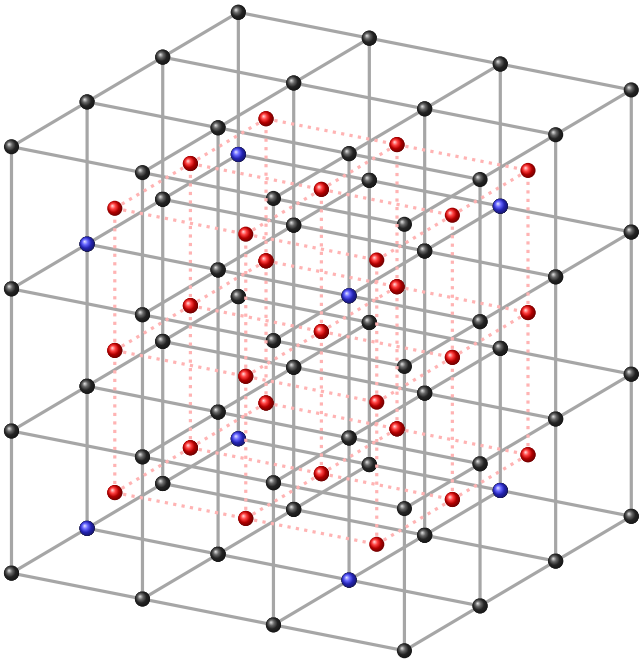
Step 3. The group of units.

Verification. Elements of A with norm 1 are exactly $\{\pm 1, \pm i, \pm j, \pm k\} \cup \left\{\frac{\pm 1 \pm i \pm j \pm k}{2}\right\}$, as easily seen. Now, if $a \in A$, then either it has all coordinates in $\mathbb{Z}$, or all coordinates in $\frac{1}{2} + \mathbb{Z}$.

In the former case we do nothing. In the latter, translation by some $\pm\frac{1}{2} \pm \frac{i}{2} \pm \frac{j}{2} \pm \frac{k}{2} \in A^\times$ will take it to a point with coordinates in $2\mathbb{Z}$, viz. $a = a_0 + u$ with $u \in A^\times$, $a_0 \in 2\mathbb{Z} \cdot 1 + 2\mathbb{Z} \cdot i + 2\mathbb{Z} \cdot j + 2\mathbb{Z} \cdot k$. Finally, $a = a_0 + u = (a_0 u^* + 1)u$, and $a_0 u^*$ has all coordinates in $\frac{1}{2}2\mathbb{Z} = \mathbb{Z}$: so $a \in \mathbb{Z}[i, j] \cdot A^\times$. ◇

This completes the proof. □

One may picture elements of $\mathbb{Z}[i, j]$ as those of the hypercubic lattice with integer coefficients. Adding σ has the effect of considering also the centres of those cubes. Below is a picture of the three-dimensional equivalent; bear in mind we are actually describing a four-dimensional object:



Red: points of $\sigma + \mathbb{Z}[i, j]$, viz. with coordinates in $\frac{1}{2} + \mathbb{Z}$.

Blue: points of $2\mathbb{Z} \cdot 1 + 2\mathbb{Z} \cdot i + 2\mathbb{Z} \cdot j + 2\mathbb{Z} \cdot k$, viz. with coordinates in $2\mathbb{Z}$.

The key lemma is however the following. Ordinary Euclidean division in commutative rings turns, in the non-commutative case, into left- and right- notions.

7.2.3. Definition. A ring R is *left-Euclidean* if there is a function $f\colon R \to \mathbb{N} \setminus \{0\}$ such that:

- $(\forall a, b \in R)[(ab \neq 0) \to (f(a) \leq f(ab))]$;
- $(\forall a, b \in R)[(b \neq 0) \to (\exists q, r \in R)(a = bq + r) \wedge (r = 0 \vee f(r) < f(b)))]$.

You may also remember from commutative algebra that a Euclidean ring is principal, viz. every ideal $I \lhd R$ is 1-generated, viz. there is $x \in I$ with $I = (x)$. Losing commutativity we have to introduce lateral versions.

7.2.4. Lemma. *A is left- and right-Euclidean (with respect to N), hence left- and right-principal.*

Proof. We prove left-Euclideanity. Recall from lemma 7.2.2 that elements of A are those either with all coordinates in $\mathbb{Z}$, or with all coordinates in $\frac{1}{2} + \mathbb{Z}$.

Step 1. $(\forall x \in \mathbb{H})(\exists y \in A)(N(x - y) < 1)$.

Verification. Recall that inside $\mathbb{H}$, A forms a lattice (not a cubic one though). Let $x \in \mathbb{H}$. Up to translating, we may assume that all coordinates of x (in the basis $(1, i, j, k)$) lie between 0 and 1. When one coordinate is $\leq \frac{1}{2}$, approximate it by 0; otherwise approximate by 1. The resulting approximation lies in A; the squared-distance is at most $4 \cdot \left(\frac{1}{2}\right)^2 = 1$, which is attained only when $x = \sigma$. But in that case the distance to A is 0. ◇

Step 2. A is left-Euclidean.

Verification. Recall that N is multiplicative, vanishes only at 0, and takes only integer values on A by lemma 7.2.2 (ii). Let $a, b \in A$ with $ab \neq 0$; in particular $b \neq 0$. Clearly $N(ab) = N(a)N(b) \geq N(a) \cdot 1 = N(a)$.

Now by the first step, let $q \in A$ be such that $N(ab^{-1} - q) < 1$; let $r = a - qb \in A$. Then one finds:

$$N(r) = N(a - qb) = N((ab^{-1} - q)b) = N(ab^{-1} - q) \cdot N(b) < N(b),$$

which proves the claim. ◇

The fact that left-Euclidean implies left-principal is like in commutative algebra. □

7.2.5. Remark. This explains why one had to use Hurwitz quaternions. Without σ we are just dealing with the hypercubic lattice $\mathbb{Z}[i, j]$. Then the centre $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ can be approximated by $(0, 0, 0, 0)$, with norm $4 \cdot \left(\frac{1}{2}\right)^2 = 1$.

Hence the first step becomes: $(\forall q \in \mathbb{H})(\exists a \in \mathbb{Z}[i, j])(N(q - a) \leq 1)$. But 1 is a possible value: an attempt at Euclidean division will result in $N(r) \leq N(b)$, possibly with equality. This is not enough for principality.

7.3 Proof of Lagrange's theorem

We want to prove that every positive integer is a sum of four squares. By lemma 7.1.1, it is enough to prove the following.

7.3.1. Proposition. *Every prime number p is a sum of four squares.*

Proof. Since $2 = 1^2 + 1^2 + 0^2 + 0^2$, we may assume that p is odd.

Step 1. There are $m, n \in \mathbb{N}$ with $p | m^2 + n^2 + 1$.

Verification. This is independent of our earlier work. Since p is odd, the set $\{v^2 : v \in \mathbb{F}_p\}$ has exactly $\frac{p-1}{2} + 1 = \frac{p+1}{2}$ elements. So does the set $\{-1 - v^2 : v \in \mathbb{F}_p\}$. Since $\mathbb{F}_p$ has order p, the two subsets must intersect: there are therefore $\mu, \nu \in \mathbb{F}_p$ such that $1 + \mu^2 + \nu^2 = 0$. Lifting to $\mathbb{N}$ we have the claim. ◇

We fix such m and n for the rest of the proof.

Step 2. Let $I = pA + (1 + mi + nj)A$. Then I is a left ideal with $pA < I$; moreover for any $x \in I$ one has $p | N(x)$.

Verification. It is a left ideal by construction. If $I = pA$, then there is $a \in A$ with $pa = (1 + mi + nj)$; the 1-coordinate of a must be $\frac{1}{p}$, but it is in $\frac{1}{2}\mathbb{Z}$ and $p \neq 2$: this is impossible.

Now let $x \in I$. Then x can be written $x = pa_1 + (1 + mi + nj)a_2$, so that:

$$\begin{aligned} N(x) &= (pa_1 + (1 + mi + nj)a_2)\,(pa_1 + (1 + mi + nj)a_2)^* \\ &= (pa_1 + \underbrace{(1 + mi + nj)a_2}_{b \in A})\,(pa_1^* + a_2^*(1 + mi + nj)^*) \\ &= p^2 N(a_1) + (\underbrace{1 + m^2 + n^2}_{\in p\mathbb{Z}})N(a_2) + p(\underbrace{a_3 a_1^* + a_1 b^*}_{r}) \end{aligned}$$

We contend that $r \in \mathbb{Z}$. Indeed, recall from lemma 7.2.2 that A is a ring stable under $*$, and that every element in A has either all coordinates in $\mathbb{Z}$, or all coordinates in $\frac{1}{2} + \mathbb{Z}$. So let $c = ba_1^* \in A$. Then $r = ba_1^* + a_1 b^* = c + c^* = 2\,\mathrm{Re}(c) \in \mathbb{Z}$, as claimed. Returning to the formula for $N(x)$, we find $p | N(x)$; this holds for any $x \in I$. ◇

(Since $1 \in A$ and $p \nmid 1$, it follows that I must be proper in A; we will not use this.)

Step 3. p is a sum of four squares.

Verification. Since A is left-principal as proved in lemma 7.2.4, there is $x \in I$ with $I = xA$. Since $A = \mathbb{Z}[i, j] \cdot A^\times$ by lemma 7.2.2 (iii), up to changing x by a unit, we may assume $x \in \mathbb{Z}[i, j]$. As we know from step 2, $p | N(x)$.

Since $p \in I$ there is $a \in A$ with $p = xa$; if $N(a) = 1$ then $a \in A^\times$, whence $I = xA = xaA = pA$, against step 2. Hence $N(a) > 1$. Now taking norms, $p^2 = N(x)N(a)$ is a factorisation of p^2 into two non-trivial integers: it follows $N(x) = N(a) = p$.

Thus $x \in \mathbb{Z}[i, j]$ has norm a sum of four squares, equal to p. ◇

This proves the proposition, and the four square theorem. □

7.4 Exercises

7.4.1. Exercise. *Prove that the (commutative) ring of Gauß integers $\mathbb{Z}[i]$ is Euclidean.*

7.4.2. Exercise. *Return to the proof of lemma 7.2.4 and prove this better approximation:*

$(\forall x \in \mathbb{H})(\exists y \in A)(N(x-y) \leq \frac{5}{8})$.

7.4.3. Exercise. *Prove right-Euclideanity of A using* *.

7.4.4. Exercise. *The number-theoretic observation '$p|n^2+m^2+1$' in step 1 of proposition 7.3.1 is actually a special case of a more general result.*

We use tuple notation: $\underline{X} = (X_1, \ldots, X_n)$, $\underline{a} = (a_1, \ldots, a_n)$.

Theorem (Chevalley-Warning). Let $\mathbb{F}$ be a finite field and $\{P_i(\underline{X}) : i = 1 \ldots r\}$ be polynomials with variables in $\underline{X}$ and coefficients in $\mathbb{F}$. Suppose that their (total) degrees satisfying:

$$\sum_j \deg P_i < n.$$

Then the cardinal of $\{\underline{a} \in \mathbb{F}_p^n : (\forall i = 1 \ldots r)(P_i(\underline{a}) = 0)\}$ is divisible by p.

1. *Let $q = \|\mathbb{F}|$. Prove that for any polynomial $Q(\underline{X})$ of degree $< n(q-1)$, one has $\sum_{\underline{a} \in \mathbb{F}^n} Q(\underline{a}) = 0$.*
2. *Deduce the theorem. Hint: consider $Q(\underline{X}) = \prod_i (1 - P_i^{q-1}(\underline{X}))$.*
3. *Retrieve step 1 of proposition 7.3.1 as an application.*

8 Quaternions and the cross-product algebra

Here begins a block of three lectures, §§ 8–10. It can be studied at any point after § 4. Moreover § 8 and § 9 are independent.

This section returns to basic notions. We show how the quaternion structure on $\mathbb{R}^4$ can explain the cross product on $\mathbb{R}^3$. The outline is simple: in § 8.1 we return to, and prove, the common properties of $u \times v$. In § 8.2 we provide other proofs, using the quaternion structure.

One must know the definition of $\mathrm{SO}_3(\mathbb{R})$: the group of linear isometries of $\mathbb{R}^3$ under the usual quadratic structure (linear maps preserving any of the following: orthogonality, the scalar product, the Euclidean norm). This group is investigated in § 9.

8.1 The cross product in $\mathbb{R}^3$

8.1.1. Definition. The *vector/cross/wedge product* of $u_1 = \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix}, u_2 = \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} \in \mathbb{R}^3$ is:

$$u_1 \times u_2 = \begin{pmatrix} b_1c_2 - c_1b_2 \\ c_1a_2 - a_1c_2 \\ a_1b_2 - a_2b_1 \end{pmatrix}.$$

This is obviously a very bad definition, as it is given extrinsically, in terms of coordinates. Is there something intrinsic?

8.1.2. Proposition (properties of $\times$)**.** $(\mathbb{R}^3, \times)$ *is a (non-associative, non-symmetric)* $\mathbb{R}$*-algebra, satisfying:*

• **algebraic orthogonality:** $\langle u_1 | u_1 \times u_2 \rangle = 0$;

• **determinant property:** $\langle u_1 | u_2 \times u_3 \rangle = \det(u_1, u_2, u_3)$;

- $\mathrm{SO}_3(\mathbb{R})$**-invariance:** *for* $f \in \mathrm{SO}_3(\mathbb{R})$*, one has* $f(u_1 \times u_2) = f(u_1) \times f(u_2)$*;*
- **Lagrange identity:** $u_1 \times (u_2 \times u_3) = \langle u_1 | u_3 \rangle\, u_2 - \langle u_1 | u_2 \rangle\, u_3$.

Proof. Notice that 'algebraic orthogonality' is a special case of the 'determinant property'.

Let M be the matrix with columns u_1, u_2, u_3. Let P be the matrix with columns $u_2 \times u_3, u_3 \times u_1, u_1 \times u_2$. By definition of the coefficients of a vector product, P is exactly the *comatrix* of M, viz. $P = \operatorname{com} M$. We know from determinant expansion that $M \cdot P^t = M \cdot \operatorname{com} M^t = \det M \cdot I_3$; in particular on the diagonal we find:

$$\langle u_1 | u_2 \times u_3 \rangle = \langle u_2 | u_3 \times u_1 \rangle = \langle u_3 | u_1 \times u_2 \rangle = \det(u_1, u_2, u_3),$$

this is the determinant property. (Notice that outside the diagonal we find $\langle u_1 | u_1 \times u_2 \rangle = 0$, viz. algebraic orthogonality.)

We move to $\mathrm{SO}_3(\mathbb{R})$-invariance. Let $f \in \mathrm{SO}_3(\mathbb{R})$. For any three vectors, one has:

$$\begin{aligned} \langle f(u_1) | f(u_2) \times f(u_3) \rangle &= \det(f(u_1), f(u_2), f(u_3)) \\ &= \det f \cdot \det(u_1, u_2, u_3) \\ &= \det(u_1, u_2, u_3) \\ &= \langle u_1 | u_2 \times u_3 \rangle \\ &= \langle f(u_1) | f(u_2 \times u_3) \rangle\,. \end{aligned}$$

In particular, the difference $f(u_2) \times f(u_3) - f(u_2 \times u_3)$ is in $(\operatorname{im} f)^\perp = (\mathbb{R}^3)^\perp = \{0\}$.

For the Lagrange identity, notice that everything here is trilinear, and the formula holds for all 3^3 choices $u_1, u_2, u_3 \in \{e_1, e_2, e_3\}$. □

A geometric proof of the Lagrange identity. The formula is trivial if (u_2, u_3) is not free. So we may suppose that $P = \operatorname{Vect} u_2, u_3$ is a plane. Let $v = u_2 \times u_3 \neq 0$. Then $v \in P^\perp$, which is a line, and $v^\perp = P^{\perp\perp} = P = \operatorname{Vect} u_2, u_3$. Hence $u_1 \times v \in v^\perp = \operatorname{Vect} u_2, u_3$. Therefore there are scalars a, b depending on u_1, u_2, u_3 such that $u_1 \times (u_2 \times u_3) = a u_2 + b u_3$.

Now take the scalar product with u_1; we find $0 = a \langle u_1 | u_2 \rangle + b \langle u_1 | u_3 \rangle$. In particular, there is a scalar λ such that:

$$u_1 \times (u_2 \times u_3) = \lambda(\langle u_1 | u_3 \rangle\, u_2 - \langle u_1 | u_2 \rangle\, u_3);$$

notice that this remains true if u_2 and u_3 are linearly dependent again. The problem is that a priori, λ depends on u_1, u_2, u_3.

8.1.3. Lemma. *Suppose* $f, g \colon V \to V$ *are two linear maps satisfying:* $(\forall v \in V)(\exists \lambda_v \in \mathbb{K})(f(v) = \lambda_v g(v))$*. Suppose in addition that* $\dim \operatorname{im} g \geq 2$*. Then* $(\exists \lambda \in \mathbb{K})(\forall v \in V)(f(v) = \lambda g(v))$.

Proof. Let $v, w \in V$ be such that $g(v)$ and $g(w)$ are independent. Then:

$$\begin{aligned} f(v+w) &= \lambda_{v+w} g(v+w) \\ &= \lambda_{v+w} g(v) + \lambda_{v+w} g(w) \\ = f(v) + f(w) &= \lambda_v g(v) + \lambda_w g(w). \end{aligned}$$

Since $g(v)$ and $g(w)$ are linearly independent, one finds $\lambda_v = \lambda_{v+w} = \lambda_w$, which we simply denote λ. Then every non-zero $g(u)$ will be independent from one of $g(v), g(w)$, and the same argument gives $\lambda_u = \lambda$ as well. There remains the case $g(u) = 0$. But then, $f(u) = 0$ so we can still take $\lambda_u = \lambda$. □

Let $g = \langle u_1 | u_3 \rangle u_2 - \langle u_1 | u_2 \rangle u_3$; see it as a function of u_1, then as a function of u_2, and then as a function of u_3. Since u_2 and u_3 are independent, in each case g is linear with image of dimension 2. Applying the lemma three times gives that λ does *not* depend on the vectors. We finish with $u_i = e_i$ to see $\lambda = 1$. □

8.1.4. Corollary. $(\mathbb{R}^3, \times)$ *is a* Lie algebra*, viz. it satisfies:*

- **bilinearity:** $\times$ *is left- and right-linear;*
- **antisymmetry:** $u_2 \times u_1 = -u_1 \times u_2$;
- **Jacobi identity:** $u_1 \times (u_2 \times u_3) + u_2 \times (u_3 \times u_1) + u_3 \times (u_1 \times u_2) = 0$.

Proof. Bilinearity, anti-symmetry are obvious. Using Lagrange's identity, so is Jacobi's. □

We have proved a number of identities without really explaining them. As always, lists of formulas tend to hide underlying structures.

8.2 Explaining geometry with quaternions

Quaternions can be used to encode the geometric structure on $\mathbb{R}^3$.

8.2.1. Notation. Let $\mathbb{P} = \operatorname{Vect} i, j, k$ be the space of *purely imaginary quaternions*.
Notice that $\mathbb{P} = \{q \in \mathbb{H} : q^* = -q\} = \{q \in \mathbb{H} : q^2 \in \mathbb{R}_{\leq 0}\}$.

The space $\mathbb{P} \simeq \mathbb{R}^3$ can be equipped with:

- the scalar product $\langle \cdot | \cdot \rangle : \mathbb{P} \times \mathbb{P} \to \mathbb{R}$;
- the cross product $\times : \mathbb{P} \times \mathbb{P} \to \mathbb{P}$;
- the quaternion product $\cdot : \mathbb{P} \times \mathbb{P} \to \mathbb{H}$.

(Only the cross product takes values in $\mathbb{P}$.) At first this begs for helpful notation. When we consider $y \in \mathbb{P}$ as a vector (subject to scalar and cross products), we write $\vec{y}$. Soon we shall drop the arrows.

8.2.2. Lemma. *Let* $y_1, y_2 \in \mathbb{P}$*. Then:*

$$y_1 \cdot y_2 = \vec{y}_1 \times \vec{y}_2 - \langle \vec{y}_1 | \vec{y}_2 \rangle .$$

Proof. The proof will not use Jacobi's or Lagrange's identity. All operations involved are bilinear, so it suffices to deal with elementary cases $y_1, y_2 \in \{i, j, k\}$. Then there are two cases:

- if $y_1 = y_2$, then $y_1 \cdot y_1^2 = -1$ while $\overrightarrow{y}_1 \times \overrightarrow{y}_2 = \overrightarrow{0}$ and $\left\langle \overrightarrow{y}_1 \middle| \overrightarrow{y}_2 \right\rangle = 1$, proving the formula;
- if $y_1 \neq y_2$, then $\left\langle \overrightarrow{y}_1 \middle| \overrightarrow{y}_2 \right\rangle = 0$ while both $y_1 \cdot y_2$ and $\overrightarrow{y}_1 \times \overrightarrow{y}_2$ compute the 'third' basic quaternion in direct order.

(One can also work in coordinates and expand, but this is tedious.) □

For instance, $\overrightarrow{i} \times \overrightarrow{j} = k + \left\langle \overrightarrow{i} \middle| \overrightarrow{j} \right\rangle = k$. Obviously, this is better understood without the arrows. We now decompose $\mathbb{H} = \mathbb{R} \oplus \mathbb{P}$; every quaternion can be written in a unique manner $a + u$ with $a \in \mathbb{R}$ and $u \in \mathbb{P}$.

8.2.3. Proposition. *Let (a, u) and (b, v) be quaternions. Then:*

$$(a, u) \cdot (b, v) = (ab - \langle u|v \rangle, av + bu + u \times v).$$

(The formula is perhaps worth learning; at least, one must remember that a short formula exists.)

Proof. Since quaternion multiplication is bilinear, hence biadditive, one may consider $(a, 0)$ and $(0, u)$ separately. Since real cases are obvious, the formula reduces to computing $(0, u) \cdot (0, v)$, which was done in the lemma. □

This multiplicative structure explains a number of otherwise mysterious properties. First, we retrieve corollary 8.1.4.

Quaternion proof that $(\mathbb{R}^3, \times)$ is a Lie algebra. Bilinearity of $\times$ is obvious since quaternion multiplication is bilinear. We prove antisymmetry of $\times$. Recall that $\mathbb{R} = \{q \in \mathbb{H} : q^* = q\}$ and $\mathbb{P} = \{q \in \mathbb{H} : q^* = -q\}$. In particular, if $q_1, q_2 \in \mathbb{P}$ then one has:

$$\begin{aligned}(q_1 q_2 + q_2 q_1)^* &= q_2^* q_1^* + q_1^* q_2^* \\ &= (-q_2) \cdot (-q_1) + (-q_1) \cdot (-q_2) \\ &= q_1 q_2 + q_2 q_1,\end{aligned}$$

implying $q_1 q_2 + q_2 q_1 \in \mathbb{R}$. Returning to $\mathbb{R} \oplus \mathbb{R}^3$, one has $(0, u) \cdot (0, v) = (-\langle u|v \rangle, u \times v)$. Therefore

$$(0, u) \cdot (0, v) + (0, v) \cdot (0, u) = (-2 \langle u|v \rangle, u \times v + v \times u) \in \mathbb{R},$$

so we find $u \times v + v \times u = 0$, as desired.

We turn to the Jacobi identity. It is a general fact that given an *associative* algebra $\mathbb{A}$, the operation $[\![a, b]\!] = ab - ba$ is a Lie bracket (viz. satisfies the definition of a Lie algebra). Indeed, bilinearity and antisymmetry are clear; for the Jacobi identity, one

computes using associativity:

$$\begin{aligned}
&[\![a,[\![b,c]\!]]\!]+[\![b,[\![c,a]\!]]\!]+[\![c,[\![a,b]\!]]\!]\\
&=a(bc-cb)-(bc-cb)a+b(ca-ac)-(ca-ac)b+c(ab-ba)-(ab-ba)c\\
&=abc-acb-bca+cba+bca-bac-cab+acb+cab-cba-abc+bac\\
&=0
\end{aligned}$$

Here, start with $\mathbb{H}$; then in earlier notation:

$$[\![(0,u),(0,v)]\!]=(-\langle u|v\rangle+\langle v|u\rangle\,,u\times v-v\times u)=(0,u\times v-v\times u).$$

By antisymmetry this also equals $(0,2u\times v)$, which is therefore a Lie bracket. This proves Jacobi's identity (without using it); we did not use Lagrange's either. □

We can also return to proposition 8.1.2 and give a better proof.

Quaternion proof of the properties of $\times$, except $\mathrm{SO}_3(\mathbb{R})$-invariance. We use associativity of quaternion multiplication. Compute:

$$\begin{aligned}
(0,u)\cdot[(0,v)\cdot(0,w)]&=(0,u)\cdot(-\langle v|w\rangle\,,v\times w)\\
&=(-\langle u|v\times w\rangle\,,-\langle v|w\rangle\,u+u\times(v\times w))\,,
\end{aligned}$$

and similarly:

$$\begin{aligned}
[(0,u)\cdot(0,v)]\cdot(0,w)&=(-\langle u|v\rangle\,,u\times v)\cdot(0,w)\\
&=(-\langle u\times v|w\rangle\,,-\langle u|v\rangle\,w+(u\times v)\times w)\,.
\end{aligned}$$

By associativity, they are equal, viz. we have the identity:

$$(-\langle u|v\times w\rangle\,,-\langle v|w\rangle\,u+u\times(v\times w))=(-\langle u\times v|w\rangle\,,-\langle u|v\rangle\,w+(u\times v)\times w)\,.$$

The real coordinate gives $\langle u|v\times w\rangle=\langle u\times v|w\rangle$. Doing $w=v$ and using antisymmetry, we deduce $\langle u\times v|v\rangle=0$, viz. algebraic orthogonality. Actually the map $\langle u|v\times w\rangle$ is trilinear, and alternative: hence a multiple of the determinant map. We can compute that $\langle i\times j|k\rangle=\langle k|k\rangle=1$, so the multiple is 1.

Leaving $\mathrm{SO}_3(\mathbb{R})$-invariance aside, there remains to prove the Lagrange identity. Return to the associative identity and take its pure quaternion component, getting:

$$-\langle v|w\rangle\,u+u\times(v\times w)=-\langle u|v\rangle\,w+(u\times v)\times w.$$

Introduce $f(u,v,w)=u\times(v\times w)$; notice that $(u\times v)\times w=-w\times(u\times v)=-f(w,u,v)$. Our equation becomes:

$$f(u,v,w)+f(w,u,v)=\langle v|w\rangle\,u-\langle u|v\rangle\,w.$$

The other two equations obtained by circular permutation are:

$$f(v,w,u)+f(u,v,w)=\langle w|u\rangle\,v-\langle v|w\rangle\,u$$

and

$$f(w,u,v)+f(v,w,u)=\langle u|v\rangle\,w-\langle w|u\rangle\,v.$$

Summing the first two and subtracting the last, there remains:

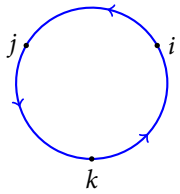
$$2f(u,v,w) = 2\langle u|w\rangle\, v - 2\langle u|v\rangle\, w,$$

as desired. (Notice that if we take circular permutations and then sum, we find another proof of the Jacobi identity.) □

However, we have *not* reproved $\mathrm{SO}_3(\mathbb{R})$-invariance through quaternions. This needs more tools, and will be completed in § 10.1.

8.3 Exercises

8.3.1. Exercise. *Prove that the rotation*



is obtained through conjugation by $\sigma = \frac{1+i+j+k}{2}$.

8.3.2. Exercise. *Give a geometric interpretation of the formula:*

$$\frac{1}{2}(\gamma_i + \gamma_j + \gamma_k - \mathrm{Id})(q) = q^*.$$

Solution. The equation is trivial on $\mathbb{R}$, so we focus on $\mathbb{P} \simeq \mathbb{R}^3$. There, γ_i is the half-turn of the space in $i^\perp$; likewise for the other two. Hence the sum $\gamma_i + \gamma_j + \gamma_k$ equals $-\mathrm{Id}$ (this can be seen matricially: our half-turns are diagonal in the canonical basis). But on $\mathbb{P}$, quaternion conjugation is $-\mathrm{Id}$.

8.3.3. Exercise. *Prove that for any two vectors* $u, v \in \mathbb{R}^3$ *one has* $\|u\|^2 \cdot \|v\|^2 = \|u \times v\|^2 + \langle u|v\rangle^2$.

Solution. Start with $u \cdot v = u \times v - \langle u|v\rangle$. Now the squared norm is $|u|^2|v|^2 = |u \times v|^2 + \langle u|v\rangle^2$, but for a pure quaternion one also has $|q|^2 = \|q\|^2$. (Direct proofs are possible.)

9 Orienting and rotating the real plane and space

We continue our investigation of low-dimensional geometry; this lecture is dedicated to Euclidean spaces of dimension 2 (§ 9.1) and 3 (§ 9.2). It does not build on § 8 and there are no quaternions. We shall:

- explain the notion of an orientation (definition 9.1.5);
- see generation by half-turns (an important phenomenon, corollary 9.2.6);
- describe rotations by their 'geometric elements' (corollary 9.2.9).

Recall a general definition.

Definition. Let $\mathbb{K}$ be any field and n be an integer.

- The *orthogonal group* is:

$$\mathrm{O}_n(\mathbb{K}) = \left\{ M \in \mathrm{GL}_n(\mathbb{K}) : M^{-t} = M \right\},$$

  the group of fixed points of the inverse-transpose automorphism (there are more obscure definitions).

- The *special orthogonal group* is $\mathrm{SO}_n(\mathbb{K}) = \mathrm{SL}_n(\mathbb{K}) \cap \mathrm{O}_n(\mathbb{K})$.

The definition is actually more intrinsic: if $(E, (\cdot,\cdot))$ is a Euclidean space, one can define the orthogonal group $\mathrm{O}(E)$ as the group of linear bijections preserving the bilinear form, viz. with $(f(x), f(y)) = (x, y)$ holding identically. In notation $O(E)$, the scalar product remains implicit; there is risk of confusion.

remark. It so happens that if $E \simeq \mathbb{R}^n$ is a Euclidean space, then fixing any orthonormal basis of E gives rise to a group isomorphism $\mathrm{O}(E) \simeq \mathrm{O}_n(\mathbb{R})$. But of course, there are many different isomorphisms $E \simeq \mathbb{R}^n$: this is the difference between something coordinatised (a coordinate system is given) and something coordinatisable (we still have the choice).

This remark will result in the subtle discussion of *orientations*.

A common misunderstanding

It is generally believed that an element of $\mathrm{SO}_3(\mathbb{R})$ is described by • an axis and • an angle. This is not correct. Consider the following problem.

> Facing the audience, the instructor holds a piece of paper vertically and rotates it by 90° *clockwise*. The students however see a *counterclockwise* rotation. *Why?*

9.1 In planes

We begin with the case of Euclidean spaces of dimension 2. All are isomorphic to $\mathbb{R}^2$, but when describing them we have to choose orientations. We explain this, starting from concrete $\mathbb{R}^2$.

Coordinatised plane

We first study the plane $\mathbb{R}^2$ equipped with the standard scalar product $(\cdot,\cdot)$.

9.1.1. Proposition. *Every non-trivial element of the orthogonal group* $\mathrm{O}_2(\mathbb{R})$ *is either:*

- *a* rotation*, with matrix of the form* $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ *with* $a^2 + b^2 = 1$*;*
- *an (orthogonal)* reflection*, with matrix of the form* $\begin{pmatrix} a & b \\ b & -a \end{pmatrix}$ *with* $a^2 + b^2 = 1$.

Proof. Let $f \in \mathrm{O}_2(\mathbb{R})$. Both columns of its matrix have Euclidean norm 1, and they are orthogonal. When the first column is fixed, there are exactly two choices for the second. □

Consequently :

- rotations of $\mathbb{R}^2$ form the group $\mathrm{SO}_2(\mathbb{R})$;
- an element in $\mathrm{SO}_2(\mathbb{R})$ is entirely described by two numbers a and b with $a^2+b^2=1$.

9.1.2. Corollary. *Every rotation of* $\mathrm{SO}_2(\mathbb{R})$ *is a product of two reflections; moreover for every pair of norm* 1 *vectors* $u_1, u_2 \in \mathbb{S}^1$ *(the unit circle inside* $\mathbb{R}^2$*), there is a unique rotation* f *with* $f(u_1)=u_2$*. Finally,* $\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{S}^1$.

In more algebraic terms, $\mathrm{O}_2(\mathbb{R}) = \mathbb{S}^1 \rtimes \mathbb{Z}/2\mathbb{Z}$ (semi-direct product with respect to inversion action).

Proof. Clearly,

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

so the latter is a product of two reflections. Now if u is a norm 1 vector, say $u = \begin{pmatrix} a \\ b \end{pmatrix}$ with $a^2+b^2=1$, we can always see it as the first column of a (unique) rotation matrix. Algebraically, there is a unique rotation doing $f(e_1)=u$. This carries to a pair of vectors: do $u_1 \mapsto e_1 \mapsto u_2$.

So fixing say e_1, the map $g \mapsto g \cdot e_1$ is a group isomorphism $\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{S}^1$. □

9.1.3. Remark (angle measurement is dubious methodology)**.** An orthogonal reflection is entirely determined by its axis, itself determined by a non-zero vector. In high school it is customary to describe a rotation by its *angle*, and write:

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

We shall not do this.

What is an angle? It is a bit less clear than you may think; in particular, the claim that one *measures* angles using the same real numbers as the ones on the line is quite ill-phrased. They are *not* the same numbers; a simple argument is that one must count modulo 2π.

One should keep in mind that *there are linear numbers, and circular numbers*. The real miracle is that linear numbers can be used to cover circular numbers, viz. that there is a surjective homomorphism $(\mathbb{R},+) \to (\mathbb{S}^1,\cdot)$. (This may look obvious when one writes it as $\mathbb{R} \to \mathbb{R}/2\pi\mathbb{Z}$, but try to explain π in purely algebraic terms to understand the subtlety of the question.) This involves some form of exponential function and will *fail* over other fields, such as the seemingly harmless $\mathbb{R} \cap \overline{\mathbb{Q}}$.

As a conclusion: in order to generalise to real-closed fields, one should avoid talking about angle measurement, and entirely decribe angles in terms of *elements of* $\mathrm{SO}_2(\mathbb{R})$. Not to mention the computational cost of trigonometric functions…

Abstract planes

We move to abstract planes (as opposed to the reference plane $\mathbb{R}^2$). This is surprisingly subtle. If P is a Euclidean plane, then $P \simeq \mathbb{R}^2$ though non-canonically. Therefore if (v_1, v_2) is an orthonormal basis, then so is (v_2, v_1); but there is no canonical way to prefer one over the other. This gives rise to the notion of an orientation of a plane.

9.1.4. Lemma. *Let P be a Euclidean plane, viz. $P \simeq \mathbb{R}^2$ with Euclidean structure.*

Then the action of $\mathrm{O}(P)$ on the set of orthonormal bases has exactly one orbit; the action of $\mathrm{SO}(P)$ has exactly two.

Proof. First understand the statement. Let β be the set of orthonormal bases of P. Suppose that $\mathcal{B} = (v_1, v_2) \in \beta$ and $f \in \mathrm{O}(P)$; then by definition of an isometry, $(f(v_1), f(v_2))$ is another orthonormal basis, so $f(\mathcal{B}) \in \beta$. Hence $\mathrm{O}(P)$ acts naturally on β; so does the subgroup $\mathrm{SO}(P) \leq O(P)$.

We now work in coordinates, in the orthonormal basis $\mathcal{B} = (v_1, v_2)$. Suppose (w_1, w_2) is *another* basis in β. Then in $\mathcal{B}$, w_1 has coordinates a norm 1 column; up to using an element of $\mathrm{SO}_2(\mathbb{R}) \simeq \mathrm{SO}(P)$, we may assume $w_1 = v_1$. Then in $\mathcal{B}$, w_2 has coordinates either $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, in which case $w_2 = v_2$, or coordinates $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Under the action of $\mathrm{O}(P)$ we may exchange v_2 and $-v_2$ fixing v_1 (this is a reflection); under the action of $\mathrm{SO}(P)$ we cannot, because the determinant is 1. □

9.1.5. Definition. An *orientation* of a Euclidean plane is the choice of an orbit of orthonormal bases under the action of $\mathrm{SO}(P)$ (there are two such choices by lemma 9.1.4).

A basis in the chosen orbit is *direct*; indirect if in the other orbit.

9.1.6. Corollary. *An element of $\mathrm{SO}(P)$ is entirely described by an orientation of P and an element of $\mathrm{SO}_2(\mathbb{R})$.*

Proof. Let $f \in \mathrm{SO}(P)$. A priori we only fix the orbit under $\mathrm{SO}(P)$ of one basis, not the basis itself. We must show that $\mathrm{Mat}_{\mathcal{B}} f \in \mathrm{SO}_2(\mathbb{R})$ does *not* depend on $\mathcal{B}$, provided $\mathcal{B}$ is a direct basis.

But by definition, if $\mathcal{B}, \mathcal{B}'$ are direct bases, there is $g \in \mathrm{SO}(P)$ such that $g(\mathcal{B}) = \mathcal{B}'$. Therefore, $\mathrm{Mat}_{\mathcal{B}} f$ and $\mathrm{Mat}_{\mathcal{B}'} f$ are conjugate (by g) inside $\mathrm{SO}(P)$. Since $\mathrm{SO}(P) \simeq \mathrm{SO}_2(\mathbb{R})$ is commutative, we find $\mathrm{Mat}_{\mathcal{B}} f = \mathrm{Mat}_{\mathcal{B}'} f$, as desired. □

9.1.7. Remarks (on orientations)**.**

- By lemma 9.1.4, a reflection reverts the orientation; you may have seen a mirror already.
- Another convenient way to understand the concept of change of orientation of a plane is to embed it 'as a piece of paper' in the space; looking at it from above or from below will change the orientation. (Which takes us to the space.)

9.2 In the space

We will work mostly with coordinatised $\mathbb{R}^3$, viz. study the concrete group $\mathrm{SO}_3(\mathbb{R})$.

9.2.1. Remarks (orientations, continued)**.**

- An abstract 3-dimensional Euclidean space has exactly two orientations (see exercise 9.3.1).
- The *physical* space we live in is oriented by convention following the *right-hand orientation*:

bending your right hand describes a direct basis.

Since you can rotate your hand freely but bend it only in one direction, your right hand can describe exactly the orbit of all *direct* bases. (You need the left hand for the indirect bases.)

The group of rotations

9.2.2. Remark. As opposed to $\mathrm{SO}_2(\mathbb{R})$, the group $\mathrm{SO}_3(\mathbb{R})$ is *not* commutative. This explains why when Hamilton was looking for a structure coding rotations of $\mathbb{R}^3$, he had to drop commutativity.

9.2.3. Example. Hold a stick of chalk in your hand, pointing it to the audience. Rotate your wrist to the left by a quarter of a turn; this is rotation ρ_1, which takes place in the vertical plane separating you from the audience, which we call V. After performing ρ_1, the chalk globally looks just the same. Then rotate your elbow to the left by a quarter of a turn; this is rotation ρ_2, which takes place in the horizontal plane H. The chalk arrives in horizontal position in $H \cap V$.

Now start again from the initial position. First perform ρ_2, i.e. rotate the elbow: the chalk is horizontal, in H. Of course your wrist has rotated with the rest of the arm, so you must return to the original definition of ρ_1: *rotating in* V. The stick of chalk now points downwards, illustrating $\rho_2\rho_1 \neq \rho_1\rho_2$.

It so happens that every element of $\mathrm{SO}_3(\mathbb{R})$ is a rotation (in the intuitive sense). This is made precise by the following. Recall that $L^\perp = \{x \in \mathbb{R}^3 : (\forall y \in L)((x, y) = 0)\}$.

9.2.4. Proposition. *Every rotation* $\rho \neq \mathrm{Id}$ *is uniquely described by a well-defined axis* $L = \ker(\rho - \mathrm{Id})$*, and an element of* $\mathrm{SO}(L^\perp)$.

Proof.

Step 1. 1 is an eigenvalue with multiplicity (both algebraic and geometric) equal to 1.

Verification. Let χ be the characteristic polynomial of ρ; it has degree 3 and coefficients in $\mathbb{R}$. Since $\mathbb{R}$ is real closed, there is a real root (eigenvalue) λ; moreover, the set of eigenvalues is invariant under $c \mapsto c^*$. Finally the product of eigenvalues is $\det \rho = 1$.

It is easy to see that $\lambda = \pm 1$: if x is an eigenvector, then $\|x\|^2 = (x, x) = (f(x), f(x)) = \lambda^2\|x\|^2 \neq 0$. Suppose that $\lambda = -1$. Then by the condition on the determinant, the other two roots (computed in $\mathbb{C}$) have -1 as a product, so they cannot be conjugate elements of $\mathbb{C} \setminus \mathbb{R}$; they are in $\mathbb{R}$ and equal to -1 and 1, so 1 was a root after all.

This proves that $\ker(\rho - \mathrm{Id}) \neq 0$. If $\dim\ker(\rho - \mathrm{Id}) = 3$ then $\rho = \mathrm{Id}$: a contradiction. If $\dim\ker(\rho - \mathrm{Id}) = 2$ then the orthogonal space $(\ker(\rho - \mathrm{Id}))^\perp$ has dimension 1. But is is ρ-invariant again, so it is an eigenline: the eigenvalue is ± 1, but by the determinant condition it must be 1, a contradiction. All this shows that $\dim\ker(\rho - \mathrm{Id}) = 1$. ◇

Let $L = \ker(\rho - \mathrm{Id})$, a line called the *axis*, and $P = L^\perp$, the *rotation plane*.

Step 2. Geometric conclusion.

Verification. Every isometry has the *orthogonal invariance property*: if $F \leq E$ is an f-invariant subspace (meaning $f(F) \leq F$), then $F^\perp = \{y \in E : (\forall x \in F)((x,y) = 0)\}$ is f-invariant as well.

Since L is ρ-invariant, so is P. Now the restriction $\rho_{|P}: P \to P$ is a linear isometry again, so $\rho_{|P} \in \mathrm{O}(P)$. Since ρ acts as 1 on L one has:

$$\det \rho_{|P} = 1 \cdot \det \rho_{|P} = \det \rho_{|L} \cdot \det \rho_{|P} = \det \rho = 1,$$

so $\rho_{|P} \in \mathrm{SO}(P)$.

Conversely, given an axis L and a linear isometry $\check{\rho}$ of $P = L^\perp$, there is a unique element in $\mathrm{SO}_3(\mathbb{R})$ acting as Id_L and extending $\check{\rho}$ on P. ◇

This completes the proof. □

Half-turns of the space

We derive an important corollary (which will also be used in the next lecture).

9.2.5. Definition. A *half-turn of the space* is a rotation of order exactly 2 (viz. with $\rho^2 = \mathrm{Id} \neq \rho$).

It is easily seen that every half-turn of the space is $\mathrm{O}_3(\mathbb{R})$-conjugate to $\begin{pmatrix} 1 & & \\ & -1 & \\ & & -1 \end{pmatrix}$.

9.2.6. Corollary. *Half-turns of the space generate* $\mathrm{SO}_3(\mathbb{R})$.

Proof. Let $f \in \mathrm{SO}_3(\mathbb{R})$; we may suppose $f \neq \mathrm{Id}$. By proposition 9.2.4, f has axis say L and plane $P = L^\perp$; using an orientation (hence isomorphism $P \simeq \mathbb{R}^2$), f corresponds to $\rho \in \mathrm{SO}_2(\mathbb{R})$. Now in the plane $P \simeq \mathbb{R}^2$, ρ is a product of two reflections $\rho = \sigma_2 \circ \sigma_1$ with axes say L_1 and L_2.

Let r_ℓ be the half-turn of the full space with axis L_ℓ. We claim that $f = r_2 \circ r_1$. Let $g = r_2 \circ r_1$. This reverses L twice, so g acts like Id on L. Moreover in P, g acts like $\sigma_2 \circ \sigma_1 = \rho$. Hence $g = f$. □

Further description

We return to the general problem of describing rotations, building on proposition 9.2.4. For a further description we also wish to use the isomorphism $\mathrm{SO}(P) \simeq \mathrm{SO}_2(\mathbb{R})$; but for this, we need an orientation of P as seen in corollary 9.1.6. Recall that $L \leftrightarrow L^\perp$ is a canonical correspondence between lines and planes in $\mathbb{R}^3$. Actually there is more.

9.2.7. Lemma (orthogonal orientation)**.** *Let S be any three-dimensional, Euclidean space with an orientation. Then there is a canonical correspondence between* oriented *lines and* oriented *planes.*

Proof. Fix an orientation, viz. a notion of a direct basis, in S. Given a line L and its orthogonal plane $P = L^\perp$, an orientation of L corresponds to an orientation of P: just see whether concatenation results in a direct global basis of S or not. □

9.2.8. Remark. Notice that neither the orientation of the space nor the orientation of the line gives one of the plane: *one really needs both*.

In terms of the right-hand rule, one needs both • the choice of the right hand (as opposed to the left hand), and • which vectors on the line are 'upwards' (to tell us which side of the plane, seen as a sheet of paper, is top and which is bottom).

9.2.9. Corollary (geometric elements)**.** *Every rotation* $\rho \in \mathrm{SO}_3(\mathbb{R}) \smallsetminus \mathrm{Id}$ *is uniquely described by a well-defined axis* $L = \ker(\rho - \mathrm{Id})$*, an orientation of the axis, and an element of* $\mathrm{SO}_2(\mathbb{R})$.

These are called the 'geometric elements' of ρ. For an abstract Euclidean space S, one also needs • an orientation of S.

Proof. The space $\mathbb{R}^3$ is oriented by the standard basis. Fix an orientation of L; by lemma 9.2.7 this naturally gives rise to an orientation of $P = L^\perp$. Then in the isomorphism $P \simeq \mathrm{SO}_2(\mathbb{R})$ of corollary 9.1.6, the restriction $\rho_{|P}$ is merely a circular element, some $\varphi \in \mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{S}^1$. □

Notice that when describing a half-turn of the space like in corollary 9.2.9, it is useless to specify the orientation of the axis: the direct and indirect half-turns are equal.

9.2.10. Remark (angle measurement, continued)**.** In $\mathbb{R}^3$ we are used to thinking of a rotation as described by its axis and 'angle θ'; here again the latter actually means 'angle measurement'. As we said in remark 9.1.3, one should avoid angle measurements as they are specific to $\mathbb{R}$ and do not carry to other relevant fields.

In other words, the description given in corollary 9.2.9 may be less usual than the one involving angle measurement $\theta \in \mathbb{R}$, but it has the advantage of carefully avoiding the 'linear-to-circular' exponentiel $e^{i\theta}$, and therefore of generalising to real closed fields.

As a conclusion, and returning to the introduction:

> It is generally believed that an element of $\mathrm{SO}_3(\mathbb{R})$ is described by • an axis and • an angle. This is not correct.

Here is the correct statement:

> An element of $\mathrm{SO}_3(\mathbb{R})$ is described by • an oriented axis and • a circular term $\check{\rho} \in \mathrm{SO}_2(\mathbb{R})$.

The circular term in $\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{S}^1$ need not be represented by an element of $\mathbb{R}/2\pi\mathbb{Z}$; actually, it is better not to if you want to avoid unnecessary trigonometry for a problem in linear algebra.

9.3 Exercises

9.3.1. Exercise. *Let E be a Euclidean space, viz. $E \simeq \mathbb{R}^n$ with Euclidean structure. Prove that the action of $\mathrm{O}(E)$ on the set of orthonormal bases has exactly one orbit while the action of $\mathrm{SO}(E)$ has exactly two.*

9.3.2. Exercise. *Let ρ_1, ρ_2 be two rotations of $\mathbb{R}^3$ with axes L_1, L_2. Show that $\rho_1\rho_2 = \rho_2\rho_1$ iff ($L_1 = L_2$ or ($L_1 \perp L_2$ and $\rho_1^2 = \rho_2^2 = \mathrm{Id}$)).*

9.3.3. Exercise. *Prove that $\mathrm{O}_3(\mathbb{R})$ is generated by its reflections, viz. the elements of order 2 in $\mathrm{O}_3(\mathbb{R}) \smallsetminus \mathrm{SO}_3(\mathbb{R})$.*

9.3.4. Exercise (classification of linear isometries of $\mathbb{R}^3$)**.** *Let $f \in \mathrm{O}(\mathbb{R}^3)$. Prove that f is one of the following:*

- Id *or* $-\mathrm{Id}$ *;*
- *a rotation, with* $\dim\ker(f-\mathrm{Id}) = 1$ *;*
- *a reflection, with* $\dim\ker(f-\mathrm{Id}) = 2$ *;*
- *an improper rotation (also known as a rotary reflection), that is the composition of a rotation and reflection in the same plane; then* $\dim\ker(f-\mathrm{Id}) = 0$.

Solution. The case of rotations is well-understood; we may suppose $\det f = -1$, so the product of eigenvalues is 1. Recall that f has at least one real eigenvalue.

If f has 3 real eigenvalues, then they must be $-1,-1,-1$ (so $f = -\mathrm{Id}$) or $-1,1,1$ (reflection). If f has exactly 2 real eigenvalues, the last must be real as well: a contradiction. If f has exactly 1 real eigenvalue, the other two are complex conjugate with product 1: so the real one is -1. Let $L = \ker(f+\mathrm{Id})$, which has dimension 1, and $P = L^\perp$. Let s be the reflection through P. Then fs is a rotation but fixes L: it is a rotation in P. We are done.

10 Quaternions and rotations of the space

This lecture builds on §§ 8–9 and explains the relations between quaternions and rotations. Quaternions code for rotations of $\mathbb{R}^3$; the statement is made precise in § 10.1; we explicitly compute the geometric elements attached to γ_q in § 10.2.

10.1 The theory: an orthogonal isomorphism

Recall two earlier notations:

- $\mathbb{P} = \mathrm{Vect}(i,j,k) = \{q \in \mathbb{H} : q^* = -q\} = \{q \in \mathbb{H} : q^2 \in \mathbb{R}_{\leq 0}\}$ is the space of pure quaternions;
- $\mathbb{S} = \{q \in \mathbb{H} : |q| = 1\}$ is the quaternion sphere.

Also recall that $\mathbb{S} \leq \mathbb{H}^\times$ is a subgroup, and $\mathbb{P} \leq \mathbb{H}$ is a hyperplane. Actually the linear isometry of normed spaces $(\mathbb{H}, |\cdot|) \simeq (\mathbb{R}^4, \|\cdot\|)$ restricts to $(\mathbb{P}, |\cdot|) \simeq (\mathbb{R}^3, \|\cdot\|)$. This gives $\mathbb{P}$ the structure of a Euclidean vector space.

10.1.1. Theorem. *As groups, one has* $\mathbb{H}^\times/Z(\mathbb{H}^\times) = \mathbb{H}^\times/\mathbb{R}^\times \simeq \mathbb{S}/\{\pm 1\} \simeq \mathrm{SO}_3(\mathbb{R})$.

(These are even isomorphisms of topological groups but we do not insist.)

10.1.2. Remark. $\mathbb{H}^\times/\mathbb{R}^\times$ may be seen as the 3-dimensional projective space over $\mathbb{R}$, with formal definition:

$$\Lambda^1(\mathbb{R}^4) = \{\text{vector lines in } \mathbb{R}^4\}.$$

As a corollary, the 3-dimensional projective space can be equipped with an algebraic group structure, which is non-trivial.

Proof.

Step 1. A continuous group homomorphism $\Gamma\colon \mathbb{H}^\times \to \mathrm{Aut}_{\mathbb{R}}(\mathbb{P})$.

Verification. The multiplicative group $\mathbb{H}^\times$ acts on the $\mathbb{R}$-vector space $\mathbb{H}$ by conjugation, i.e. for $q \in \mathbb{H}^\times$ one may introduce:

$$\begin{array}{rccc} \gamma_q\colon & \mathbb{H} & \to & \mathbb{H} \\ & x & \mapsto & qxq^{-1} \end{array},$$

which is $\mathbb{R}$-linear. Of course $\mathbb{R}\cdot 1 = Z(\mathbb{H})$ is fixed pointwise. Moreover the pure hyperplane $\mathbb{P} = \mathrm{Vect}(i,j,k) = \{x \in \mathbb{H} : x^2 \in \mathbb{R}_{\leq 0}\}$ is fixed setwise since $\gamma_q(x)^2 = \gamma_q(x)^2$. In particular we consider the legitimate restriction:

$$\begin{array}{rccc} \gamma_q\colon & \mathbb{P} & \to & \mathbb{P} \\ & x & \mapsto & qxq^{-1} \end{array},$$

and clearly $\gamma_q \in \mathrm{Aut}_{\mathbb{R}}(\mathbb{P}) \simeq \mathrm{GL}_3(\mathbb{R})$.

Hence we have a group homomorphism:

$$\begin{array}{rccc} \Gamma\colon & \mathbb{H}^\times & \to & \mathrm{Aut}_{\mathbb{R}}(\mathbb{P}). \\ & q & \mapsto & \gamma_q \end{array}$$

Notice that it is continuous (as one could work in coordinates). ◇

Let us determine its kernel and image.

Step 2. $\ker\Gamma = Z(\mathbb{H}^\times) = \mathbb{R}^\times$.

Verification. As we know, $\mathbb{R} = Z(\mathbb{H})$. This implies not only $Z(\mathbb{H}^\times) = \mathbb{R}^\times \leq \ker\Gamma$, but also that if $q \in \ker\Gamma$, then q centralises both $\mathbb{P}$ and $\mathbb{R}$, hence all of $\mathbb{H} = \mathbb{R}\oplus\mathbb{P}$: thus $q \in \mathbb{H}^\times \cap Z(\mathbb{H}) = \mathbb{R}^\times$, and we are done. ◇

Step 3. $\mathrm{im}\,\Gamma = \mathrm{SO}(\mathbb{P})$ for the quadratic structure induced by the quaternion norm.

Verification. First $\mathrm{im}\,\Gamma \leq \mathrm{O}(\mathbb{P})$. Indeed notice how $|\gamma_q(x)| = |qxq^{-1}| = |q|\cdot|x|\cdot|q|^{-1} = |x|$, so γ_q preserves the quaternion norm, which is equal to the Euclidean norm. So one already has $\mathrm{im}\,\Gamma \leq \mathrm{O}(\mathbb{P})$, the orthogonal group.

Next $\mathrm{SO}(\mathbb{P}) \leq \mathrm{im}\,\Gamma$. Indeed, recall that $\mathrm{SO}(\mathbb{P}) \simeq \mathrm{SO}_3(\mathbb{R})$ is generated by its half-turns by corollary 9.2.6. Let ρ be any half-turn of $\mathbb{P}$; we contend that $\rho \in \mathrm{im}\,\Gamma$. Let $L \leq \mathbb{P}$ be the axis of ρ; there is $q \in \mathbb{S}$ such that $L = \mathbb{R}q$. Recalling the behaviour of squares in $\mathbb{P}$, one has $q^2 \in \mathbb{S}\cap\mathbb{R}_{\leq 0} = \{-1\}$. Since Γ is a morphism, one finds $\Gamma(q)^2 = \Gamma(-1) = \mathrm{Id}_{\mathbb{P}} \neq \Gamma(q)$, so $\Gamma(q)$ is a half-turn, but each half-turn is determined by its axis. Since $\Gamma(q)(q) = qqq^{-1} = q$, the axis of $\Gamma(q)$ is $\mathbb{R}q$. This proves that $\rho = \Gamma(q) \in \mathrm{im}\,\Gamma$. We use corollary 9.2.6 to deduce $\mathrm{SO}(\mathbb{P}) \leq \mathrm{im}\,\Gamma$.

We finish by a simple connectedness argument: $\mathbb{S}$ is connected and Γ is continuous, so $\Gamma(\mathbb{S}) \leq \mathrm{O}(\mathbb{P})$ is connected. But $\mathrm{O}(\mathbb{P})$ is not connected (the determinant takes two values), so $\mathrm{im}\,\Gamma < \mathrm{O}(\mathbb{P})$. Since $[\mathrm{O}(\mathbb{P}) : \mathrm{SO}(\mathbb{P})] = 2$, we find $\mathrm{im}\,\Gamma = \mathrm{SO}(\mathbb{P})$. ◇

This proves the theorem. □

10.1.3. Remarks.

- We have seen in theorem 4.3.4 the isomorphism $\mathbb{S} \simeq \mathrm{SU}_2(\mathbb{C}, *)$. As a corollary, $\mathrm{SU}_2(\mathbb{C}, *)/\{\pm 1\} \simeq \mathrm{SO}_3(\mathbb{R})$; there are direct proofs without quaternions.
- The quaternion sphere $\mathbb{S}$ is therefore a *double cover* of $\mathrm{SO}_3(\mathbb{R})$; in order to determine which, one needs either more algebra (like in theorem 4.3.4) or more geometry/Lie theory, as follows.

  The group $\mathrm{SU}_2(\mathbb{C}, *)$ is simply connected; so is $\mathbb{S}$, like any hypersphere of dimension ≥ 2. These are therefore two connected, simply connected Lie groups with the same Lie algebra $\mathfrak{so}_3(\mathbb{R})$; by uniqueness of the simply connected form, they are Lie-isomorphic.
- The most interesting aspect here is that $\mathrm{SO}_3(\mathbb{R})$, though connected, is *not* simply connected. This is well illustrated by Dirac's 'cup trick'. Hold a cup, then circle your hand: once under your shoulder, then above your shoulder.

  The first turn brings the cup back to its initial position, but not your arm: the successive positions of the cup describe a closed path γ in $\mathrm{SO}_3(\mathbb{R})$ from Id to Id. But your arm is twisted: the path is not homotopic to the constant path 1. The second turn will free your arm: γ^2 is homotopic to 1; there is an element of order 2 in the fundamental group $\pi_1(\mathrm{SO}_3(\mathbb{R}))$, which actually generates it (something the cup trick does not give).
- How come you did *not* already know the 'spinor group' $\mathrm{SU}_2(\mathbb{C}, *)$? It is because it has no representation which is both faithful and irreducible. For instance, returning to the faithful real representation of proposition 4.3.2:

$$\left\{ \begin{pmatrix} c_1 & -c_2 \\ c_2^* & c_1^* \end{pmatrix} : (c_1, c_2) \in \mathbb{C}^2 \right\},$$

  one sees that $\left\{ \begin{pmatrix} z \\ -z^* \end{pmatrix} : z \in \mathbb{C} \right\}$ is a (real) invariant subspace.

The theorem also gives an explanation of $\mathrm{SO}_3(\mathbb{R})$-invariance of $\times$.

10.1.4. Corollary. *If $f \in \mathrm{SO}_3(\mathbb{R})$, then $f(u \times v) = f(u) \times f(v)$.*

Quaternion proof of $\mathrm{SO}_3(\mathbb{R})$-invariance of $\times$. We embed the problem in the algebra of quaternions (formula 8.2.3) and use theorem 10.1.1. Being a rotation, f is the conjugation γ_q by some $q \in \mathbb{S}$. So using that $\mathbb{R}$ is central in $\mathbb{H}$:

$$\begin{aligned} f(u) \times f(v) &= f(u) \cdot f(v) + \langle f(u) | f(v) \rangle \\ &= \gamma_q(u) \cdot \gamma_q(v) + \langle u | v \rangle \\ &= \gamma_q(u \cdot v) + \gamma_q(\langle u | v \rangle) \\ &= \gamma_q(u \cdot v + \langle u | v \rangle) \\ &= f(u \times v). \end{aligned}$$

□

10.2 Geometric elements

By theorem 10.1.1, quaternions code for elements of $\mathrm{SO}_3(\mathbb{R})$, because they act by conjugation on $\mathrm{Vect}(i,j,k) = \mathbb{P}$. Now according to corollary 9.2.9, every rotation of $\mathbb{R}^3$ is described by its 'geometric elements', viz. an *oriented* axis L and an element of $\mathrm{SO}_2(\mathbb{R}) \simeq \mathbb{S}^1$. Given a quaternion q, we shall give the 'geometric elements' of γ_q explicitly. The space $\mathbb{P}$ is oriented by taking (i,j,k) to be direct.

10.2.1. Proposition. *Let $q \in \mathbb{H} \smallsetminus \mathbb{R}$ and $\rho_q = (\gamma_q)_{|\mathbb{P}}$ be the associated rotation of $\mathbb{P} \simeq \mathbb{R}^3$ (mapping z to qzq^{-1}). Write $q = a + y$ with $a \in \mathbb{R}$, $y \in \mathbb{P} \smallsetminus \{0\}$.*

Then ρ_q is the rotation of $\mathbb{P} \simeq \mathbb{R}^3$ with:

- *axis $\mathbb{R}y$, oriented by y;*
- *circular term $\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$ with $c = \frac{a^2-|y|^2}{a^2+|y|^2}$, $s = \frac{2a|y|}{a^2+|y|^2}$, following the orientation of $(\mathbb{R}y)^\perp$ induced by that of $\mathbb{R}y$.*

Proof. First, ρ_q is a non-trivial rotation since $q \notin Z(\mathbb{H}^\times) = \mathbb{R}^\times$. So it has a well-defined axis. As we know, q commutes with y, so $\mathbb{R}y \leq \mathbb{P}$ is fixed by γ_q: it is the axis of ρ_q, and we naturally orient it by choosing y. The space $\mathbb{P}$ is oriented by deciding that (i,j,k) is a direct basis.

Recall that a rotation matrix is conjugate to:

$$\begin{pmatrix} c & -s & \\ s & c & \\ & & 1 \end{pmatrix},$$

so its trace is $1 + 2c$. Let us compute $\operatorname{Tr} \rho_q$.

First notice that:

$$q^{-1} = \frac{q^*}{|q|^2} = \frac{a-y}{a^2+|y|^2},$$

so writing $y = bi + cj + dk$ in coordinates, and projecting back onto $\mathbb{R}i$ one finds:

$$\begin{aligned}
\pi_i\left[\rho_q(i)\right] &= \pi_i\left[qiq^{-1}\right] \\
&= \frac{1}{a^2+|y|^2}\pi_i\left[(a+bi+cj+dk)i(a-bi-cj-dk)\right] \\
&= \frac{1}{a^2+|y|^2}\pi_i\left[(ai-b-ck+dj)(a-bi-cj-dk)\right] \\
&= \frac{1}{a^2+|y|^2}(a^2+b^2-c^2-d^2)i.
\end{aligned}$$

From there it is easy to deduce that:

$$\begin{aligned}
\operatorname{Tr}\rho_q &= \frac{1}{a^2+|y|^2}(3a^2-b^2-c^2-d^2) \\
&= \frac{1}{a^2+|y|^2}(3a^2-|y|^2) \\
&= 1 + 2\frac{a^2-|y|^2}{a^2+|y|^2}.
\end{aligned}$$

Therefore $c = \frac{a^2-|y|^2}{a^2+|y|^2}$, but we still have to find s. A quick estimate gives:

$$s^2 = 1 - c^2 = \left(\frac{2a|y|}{a^2+|y|^2}\right)^2,$$

but determining s requires a sign information. In general this is achieved through estimating the sign of $\det(y, x, \rho(x))$, where y orients the rotation axis and $x \notin \mathbb{R}y$. If the sign is positive, the angle is $< \pi$ (in the chosen orientation).

Return to our earlier computation of $\rho_q(i)$; one finds:

$$\rho_q(i) = \frac{1}{a^2+|y|^2}\left[(a^2+b^2-c^2-d^2)i + 2(bc+ad)j + 2(bd-ac)k\right].$$

In particular,

$$\begin{aligned}\det(y, i, \rho_y(i)) &= \frac{1}{a^2+|y|^2}\begin{vmatrix} b & 1 & a^2+b^2-c^2-d^2 \\ c & 0 & 2(bc+ad) \\ d & 0 & 2(bd-ac)\end{vmatrix} \\ &= \frac{-2}{a^2+|y|^2}\begin{vmatrix} c & bc+ad \\ d & bd-ac\end{vmatrix} \\ &= \frac{2a(c^2+d^2)}{a^2+|y|^2},\end{aligned}$$

which has the sign of a and of $\frac{2a|y|}{a^2+|y|^2}$. But returning to the standard form of a rotation matrix, this should also be the sign of s: hence $s = \frac{2a|y|}{a^2+|y|^2}$. □

10.2.2. Remark. There are three traditional ways of describing rotations of $\mathbb{R}^3$: Euler angles, orthogonal matrices, and quaternions.

- Euler angles, though famous, have many disadvantages. They rely on trigonometry and lead to horrible computations. (The reason they were used and remain popular is that angle measurement is seemingly intuitive; in the real life, angles seem to be the most natural parameters to measure.) In short, Euler angles are decent for physical measurements, but hard to manipulate both for a human and a machine.
- Orthogonal matrices have the advantage of belonging to linear algebra; every human and machine knows about them. However, an element in $\mathrm{SO}_3(\mathbb{R})$ is then described by a 3×3-array: this is greedy in space, and leads to computations which are longer than really needed. Last, any approximation error in one of the entries of the matrix will result in progressive loss of orthogonality.
- Quaternions have mostly advantages: they use only 4 parameters (3 if one restricts to the sphere) and exhibit numerical stability. Of course, they are 'hidden': there is nothing immediately measurable in them.

10.3 Exercises

10.3.1. Exercise.

1. *Let $q \neq 0$; write the matrix of γ_q in the standard basis (i, j, k).*

2. *Conversely let* $M \in \mathrm{SO}_3(\mathbb{R})$. *Give explicitly* $\mathbb{R}^\times q_0$, *the family of quaternions such that* γ_q *has matrix* M *in* (i, j, k).

10.3.2. Exercise. *Using the classification of isometries in* $\mathrm{O}_3(\mathbb{R})$ *(exercise 9.3.4), give an alternative proof of in the end of theorem 10.1.1 that* $\operatorname{im}\Gamma \leq \mathrm{SO}_3(\mathbb{R})$.

Solution. Notice that $\mathbb{H}$ has no subalgebras of dimension 3 (see theorem 3.3.1), and that every subalgebra contains $\mathbb{R}$. As a result, for any subalgebra $\mathbb{A} \leq \mathbb{H}$, one has $\dim(\mathbb{A}\cap\mathbb{P}) = 1$ or 3.

Now take $q \in \mathbb{S}$. Then $\ker(\gamma_q - \mathrm{Id}) = C_{\mathbb{P}}(q) = C_{\mathbb{H}}(q) \cap \mathbb{P}$ is one such intersection, so it has dimension 1 or 3. In the latter case, $\gamma_q = \mathrm{Id}$; in the former, it is a rotation. But in either case, $\gamma_q \in \mathrm{SO}(\mathbb{P})$ and we are done.

10.3.3. Exercise. *Prove that* $(\mathbb{S}^3 \times \mathbb{S}^3)/\{(1,1),(-1,-1)\} \simeq \mathrm{SO}_4(\mathbb{R})$. *Hint: for* $(q_1, q_2) \in \mathbb{S}^3 \times \mathbb{S}^3$, *consider the map* $x \mapsto q_1 x q_2^{-1}$. *Prove that this defines a continuous morphism* $\mathbb{S}^3 \times \mathbb{S}^3 \to \mathrm{O}(\mathbb{H}, N) \simeq \mathrm{O}_4(\mathbb{R})$. *Compute its kernel, show that* $\mathrm{SO}_4(\mathbb{R})$ *is in the image, and finish using connectedness.*

Solution. For simplicity of notation, let $\mathbb{S} = \mathbb{S}^3$. We consider the suggested action, viz. $(q_1, q_2) \in \mathbb{S} \times \mathbb{S}$ acts on the vector space $\mathbb{R}^4 \simeq \mathbb{H}$ by:

$$\begin{array}{rccc} \beta_{q_1,q_2}\colon & \mathbb{R}^4 & \to & \mathbb{R}^4. \\ & x & \mapsto & q_1 x q_2^{-1} \end{array}$$

This map is linear, so is in $\mathrm{End}_{\mathbb{R}}(\mathbb{R}^4)$. It also preserves the quaternion norm since q_1, q_2 have norm 1; but the quaternion norm on $\mathbb{H}$ is the canonical Euclidean norm on $\mathbb{R}^4$. Therefore $\beta_{q_1,q_2} \in \mathrm{O}(\mathbb{R}^4)$ for the usual Euclidean structure. However, we are not quite sure about its determinant for the moment so it is unclear whether it is in $\mathrm{SO}(\mathbb{R}^4)$; this will require a final connectedness argument.

Now consider:

$$\begin{array}{rccc} \Phi\colon & \mathbb{S} \times \mathbb{S} & \to & \mathrm{O}(\mathbb{R}^4), \\ & (q_1, q_2) & \mapsto & \beta_{q_1,q_2} \end{array}$$

clearly a group homomorphism. Notice that its kernel is the set of pairs (q_1, q_2) such that identically $q_1 x = x q_2$; then $q_1 = q_2 \in Z(\mathbb{H}) = \mathbb{R}$, but also $q_1 \in \mathbb{S}$. There remains $\ker\Phi = \{\pm(1,1)\}$ (*not* $\{(\pm1,\pm1)\}$, be careful).

It remains to compute the image. We write $\mathcal{B} = (1, i, j, k)$ for the usual basis of $\mathbb{R}^4 \simeq \mathbb{H}$ as a real vector space. Let $\rho \in \mathrm{SO}(\mathbb{R}^4)$, so $\mathcal{B}' = (\rho(1), \rho(i), \rho(j), \rho(k))$ is a direct orthonormal basis of $\mathbb{R}^4$. First suppose $\rho(1) = 1$. Then $\rho(i), \rho(j), \rho(k)$ are in $1^\perp = \mathbb{P}$, and form a direct orthonormal basis of $\mathbb{P}$. But recall that $\mathbb{S}/\{\pm1\} \simeq \mathrm{SO}_3(\mathbb{R})$ when acting by conjugation on $\mathbb{P}$; in particular, it is transitive. Therefore, in case $\rho(1) = 1$, there is $q \in \mathbb{S}$ which conjugates $\mathcal{B}$ to $\mathcal{B}'$, viz. $qxq^{-1} = \rho(x)$ everywhere. For the general case, consider the basis $\mathcal{B}'' = (1, \rho(1)^{-1}\rho(i), \rho(1)^{-1}\rho(j), \rho(1)^{-1}\rho(k))$, which is orthonormal and direct again (one needs to compute or see something). There is $q \in \mathbb{S}$ such that conjugation by q performs $\mathcal{B} \to \mathcal{B}''$; now the map $x \mapsto \rho(1) q x q^{-1}$ takes $\mathcal{B}$ to $\mathcal{B}'$. This shows $\rho \in \operatorname{im}\Phi$, or $\mathrm{SO}_4(\mathbb{R}) \leq \operatorname{im}\Phi \leq \mathrm{O}_4(\mathbb{R})$.

We finish using a connectedness argument. The group morphism Φ is continuous and $\mathbb{S} \times \mathbb{S}$ is clearly path-connected. So $\operatorname{im}\Phi \leq \mathrm{O}(\mathbb{R}^4)$ is connected and contains $\mathrm{SO}(\mathbb{R}^4)$. Since $\mathrm{O}(\mathbb{R}^4)$ itself is not connected, we find $\operatorname{im}\Phi = \mathrm{SO}(\mathbb{R}^4)$.

10.3.4. Exercise. *The goal of this exercise is to show that every automorphism of the ring* $\mathbb{H}$ *is* inner, *viz. a conjugation automorphism* γ_q.

1. *Prove that* $\mathrm{Aut}_{ring}(\mathbb{R}) = \{\mathrm{Id}\}$. *Hint: a ring automorphism must preserve the ordering, hence the topology: it is continuous. Then use density of the rationals.*

   (Remark: this does *not* generalise to non-Archimedean real closed fields.)

2. *Digression: there are only* two *continuous automorphisms of the topological ring* $\mathbb{C}$.

   (Shocking remark: $\mathrm{Aut}_{\mathrm{ring}}(\mathbb{C})$ has cardinality is $2^{2^{\aleph_0}}$.)

3. *Now let* $\varphi \in \mathrm{Aut}_{ring}(\mathbb{H})$. *Prove that* φ *is inner as follows.*

   a. *Show that* φ *fixes* $\mathbb{R} \leq \mathbb{H}$ *setwise, hence pointwise.*

   b. *Show that* φ *fixes* $\mathbb{P}$ *setwise, and actually acts as an isometry* $\varphi_{|\mathbb{P}} \in O(\mathbb{P})$.

   c. *If* $\det\left(\varphi_{|\mathbb{P}}\right) = 1$, *show that* φ *is some conjugation automorphism.*

   d. *If* $\det\left(\varphi_{|\mathbb{P}}\right) = -1$, *contradict the fact that* $-\varphi_{|\mathbb{P}}$ *is multiplicative.*

Solution.

1. Let $\varphi \in \mathrm{Aut}_{\mathrm{ring}}(\mathbb{R})$, which is the same as $\mathrm{Aut}_{\mathrm{field}}(\mathbb{R})$. Then φ preserves the set of squares, which is exactly $\mathbb{R}_{\geq 0}$, hence also the order relation. Now φ being a field automorphism must fix the rationals pointwise; by Archimedeanity, these are dense in $\mathbb{R}$. If there is $x \in \mathbb{R}$ with $\varphi(x) \neq x$, then we may assume $\varphi(x) > x$ and take a rational q with $x < q < \varphi(x)$; then $\varphi(x) < \varphi(q) = q$, a contradiction.

   Remark. Generalising to other fields, one needs a couple of properties from $\mathcal{R}$: to be formally real, to be so-called Euclidean (every element is either a sum of squares, or the opposite of such a sum), to be Archimedan (the natural numbers are cofinal, equivalently the rational numbers are dense).

2. Same ideas. A *continuous* $\varphi \in \mathrm{Aut}_{\mathrm{ring}}(\mathbb{C})$ will fix the rationals pointwise, so by continuity it will also fix $\mathbb{R}$ pointwise. Then $\varphi(i) = \pm i$ gives rise to the two distinct possibilities: identity and conjugation.

   Without continuity, one can always permute a transcendence basis of $\mathbb{C}$, very much like in a vector space one can always permute a linear basis. Since any transcendence basis has cardinal $2^{\aleph_0}$, it admits $2^{2^{\aleph_0}}$ distinct permutations, which give rise to that many distinct field automorphisms. This is clearly the maximal possible cardinal for $\mathrm{Aut}_{\mathrm{ring}}(\mathbb{C}) \subseteq \mathbb{C}^{\mathbb{C}}$.

3. Let $\varphi \in \mathrm{Aut}_{\mathrm{ring}}(\mathbb{H})$ be a ring automorphism.

   (a) Being a ring automorphism, φ stabilises $Z(\mathbb{H}) = \mathbb{R}$, so $\varphi_{|\mathbb{R}} = \mathrm{Id}_{\mathbb{R}}$. In particular one now has $\varphi \in \mathrm{Aut}_{\mathbb{R}}(\mathbb{H})$, viz. φ is an automorphism of $\mathbb{H}$ as an $\mathbb{R}$-algebra.

   (b) Recall that the pure quaternionic hyperplane $\mathbb{P} = \mathrm{Vect}(i, j, k)$ is also $\mathbb{P} = \{q \in \mathbb{H} : q^2 \in \mathbb{R}_{\leq 0}\}$. So φ must stabilise $\mathbb{P}$, and we focus on the restriction $\varphi_{|\mathbb{P}}$; it is enough to find q with $\varphi_{|\mathbb{P}} = \gamma_{q|\mathbb{P}}$ since they have the same (trivial) action on $\mathbb{R}$. Slightly abusing notation we write φ for this restriction; this is harmless.

   On $\mathbb{P}$, the map $q \mapsto |q|^2$ is simply $q \mapsto -q^2$, which is preserved by φ; it follows that $\varphi \in \mathrm{O}(\mathbb{P})$.

(c) If $\det\varphi = 1$ then $\varphi \in \mathrm{SO}(\mathbb{P})$. But when proving the isomorphism $\mathbb{S}/\{\pm 1\} \simeq \mathrm{SO}_3(\mathbb{R})$ we showed that maps in $\mathrm{SO}(\mathbb{P})$ are obtained through quaternion conjugation: in this case, there is q with $\varphi = \gamma_q$ (a priori on $\mathbb{P}$, but then as we said, on $\mathbb{H} = \mathbb{R} \oplus \mathbb{P}$ as well).

(d) If $\det\varphi = -1$ then $-\varphi \in \mathrm{SO}(\mathbb{P})$; so there is q with $-\varphi = \gamma_q$. Then by multiplicativity:

$$\begin{aligned}\varphi(k) &= \varphi(ij)\\ &= -\gamma_q(ij)\\ &= -\gamma_q(i)\gamma_q(j)\\ &= -\varphi(i)\varphi(j)\\ &= -\varphi(k),\end{aligned}$$

which is a contradiction.

Remarks.

- In general, over an arbitrary real closed field, one can only factor φ into an inner automorphism of $\mathbb{H}$ (or the $\mathcal{R}$-analogue) and an automorphism induced by one of $\mathcal{R}$.
- Returning over $\mathbb{R}$, one therefore has:

$$\mathrm{Aut}_{\mathrm{ring}}(\mathbb{H}) \simeq \mathbb{H}^\times/Z(\mathbb{H}^\times) \simeq \mathrm{SO}_3(\mathbb{R}).$$

More surprisingly, $\mathrm{Aut}_{\mathrm{ring}}(\mathbb{O}) \simeq G_2$ (see § 5.3).

10.3.5. Exercise (Rodrigues' formula)**.** *Let ρ be the rotation of $\mathbb{R}^3$ with* oriented *axis $\mathbb{R}u$ for $\|u\| = 1$, and circular term $\rho \in \mathrm{SO}(u^\perp)$ (with respect to the orientation). In any direct orthonormal basis of $u^\perp$ let $\begin{pmatrix} c \\ s \end{pmatrix}$ be the first column of the matrix of ρ. Then for $x \in \mathbb{R}^3$:*

$$\rho(x) = (1-c)\,\langle u|x\rangle\, u + cx + su \times x.$$

1. *Give a direct proof. Hint: take $x \notin \mathbb{R}u$, let y be the orthogonal projection of x onto $\mathbb{R}u$, and let $z = x - y$. Work in the basis $\left(u, \frac{1}{\|z\|}z, \frac{1}{\|z\|}u \times z\right)$.*
2. *Give a proof using quaternions and the relationships between $\cdot$, $\times$, $\langle\cdot|\cdot\rangle$.*

Solution.

1. We may assume $x \notin \mathbb{R}u$, or there is not much to prove. Let $y \neq x$ be the orthogonal projection of x onto the rotation axis: since $\|u\| = 1$, we know that $y = \langle u|x\rangle\, u$. Of course $\rho(y) = y$.

   Now $z = x - y$ is the orthogonal projection of x onto the rotation plane P. Then $\mathcal{B} = \left(u, \frac{z}{\|z\|}, \frac{1}{\|z\|}u \times z\right)$ is a direct, orthonormal basis of $\mathbb{R}^3$ in which the matrix of ρ is obvious:

$$\mathrm{Mat}_{\mathcal{B}}\,\rho = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{pmatrix}.$$

In particular,

$$\rho(z) = \|z\|\rho\left(\frac{z}{\|z\|}\right) = \|z\|\left(c\frac{z}{\|z\|} + s\frac{1}{\|z\|}u \times z\right) = cz + su \times z.$$

Finally:

$$\begin{aligned}\rho(x) &= \rho(y) + \rho(z)\\ &= y + c(x-y) + su \times (x-y)\\ &= (1-c)y + cx + su \times x - su \times y\\ &= (1-c)\langle u|x\rangle u + cx + su \times x - s\langle u|x\rangle u \times u\\ &= (1-c)\langle u|x\rangle u + cx + su \times x\end{aligned}$$

2. Let $q = a + y$, with $a \in \mathbb{R}$ and $y \in \mathbb{P}$. We suppose $y \neq 0$. As in the proof of proposition 10.2.1, ρ_q denotes conjugation by q restricted to $\mathbb{P}$, that is: $\rho_q(x) = qxq^{-1} = \frac{1}{|q|^2}qxq^*$. It will be simpler to compute $|q|^2\rho_q(x)$.

   Recall the Lagrange identity: $u \times (v \times w) = \langle u|w\rangle v - \langle u|v\rangle w$. In particular, one has $y \times (y \times x) = \langle y|x\rangle y - |y|^2 x$. We are ready:

$$\begin{aligned}|q|^2\rho_q(x) &= [(a,y)\cdot(0,x)]\cdot(a,-y)\\ &= (-\langle y|x\rangle, ax + y \times x)\cdot(a,-y)\\ &= (-a\langle y|x\rangle + a\langle x|y\rangle + \langle y \times x|y\rangle,\\ &\qquad \langle y|x\rangle y + a^2x + ay \times x - ax \times y - (y \times x) \times y).\end{aligned}$$

   As expected, the real part is trivial. Now letting $u = \frac{y}{|y|}$, the vector term simplifies into:

$$\begin{aligned}|q|^2\rho_q(x) &= \langle y|x\rangle y + a^2x + 2ay \times x + \langle y|x\rangle y - |y|^2x\\ &= 2|y|^2\langle u|x\rangle u + (a^2 - |y|^2)x + 2a|y|u \times x.\end{aligned}$$

   Hence:

$$\rho_q(x) = \frac{2|y|^2}{a^2+|y|^2}\langle u|x\rangle u + \frac{a^2-|y|^2}{a^2+|y|^2}x + \frac{2a|y|}{a^2+|y|^2}u \times x,$$

   and knowing the geometric elements u, c, s of ρ_q from proposition 10.2.1, we recognise $\rho_q(x) = (1-c)\langle u|x\rangle x + cx + su \times x$.

10.3.6. Exercise (the quaternion exponential)**.** *In this exercise, as opposed to the rest of the notes, we* do *need the base field to be $\mathbb{R}$ (or at least to enjoy good topological properties which we do not wish to axiomatise), because we* do *use trigonometric functions.*

For $q \in \mathbb{H}$, the series $\sum \frac{1}{n!}q^n$ is absolutely convergent, hence convergent. We let $\exp(q)$*, also e^q, be its limit.*

1. *Let $\ell \in \mathbb{H}$ be such that $\ell^2 = -1$. Prove that for any $t \in \mathbb{R}$, one has $e^{t\ell} = \cos t + (\sin t)\ell$. (Hint: $\mathbb{R}[\ell] \simeq \mathbb{R}[i]$ naturally.)*

2. *Conversely prove that any quaternion of norm 1 can be written in the form $e^{t\ell}$ with $t \in \mathbb{R}$ and $\ell^2 = -1$.*

3. *Let $q = a + bi + cj + dk = a + y$ with $a \in \mathbb{R}$ and $y \in \mathbb{P}$; prove that:*

$$e^q = e^a \cdot \left(\cos|y| + \frac{\sin|y|}{|y|} y \right).$$

(As always one continuously extends the cardinal sine function by $\frac{\sin 0}{0} = 1$.)

4. *Let $\ell \in \mathbb{H}$ be such that $\ell^2 = -1$. Prove that conjugation by $e^{t\ell}$ is the rotation of $\mathbb{P}$ with axis $\mathbb{R}\ell$, and angle measure $2t \mod 2\pi$.*

Remark. One should not be too enthusiastic; I believe that there can be no extension to quaternions of the complex (holomorphic) calculus. For instance, due to lack of commutativity, Taylor expansions are simply unmanageable. In my opinion, attempts have not proved successful; an expository paper on the topic is not completely convincing.[6]

Solution.

1. Suppose $\ell^2 + 1 = 0$. Then the map:

$$\begin{array}{rccc} \varphi\colon & \mathbb{C} & \to & \mathbb{R}[\ell] \\ & a + bi & \mapsto & a + b\ell \end{array}$$

is an isomorphism of finite-dimensional $\mathbb{R}$-algebras, in particular is continuous. Let $E_n(q) = \sum_{k=0}^{n} \frac{q^k}{k!}$. By definition, $E_n(ti) \to e^{ti}$ in $\mathbb{C}$ while $E_n(t\ell) \to e^{t\ell}$ in $\mathbb{R}[\ell]$. By continuity, we get:

$$\varphi(e^{ti}) = \varphi\left(\lim_n E_n(ti) \right) = \lim_n \varphi\left(E_n(ti) \right) = \lim_n E_n(t\ell) = e^{t\ell}.$$

Computing in $\mathbb{C}$ on the other hand, $e^{ti} = \cos t + (\sin t)i$ by the usual Euler formulas. By linearity, we therefore have:

$$e^{t\ell} = \varphi(e^{ti}) = \varphi(\cos t + (\sin t)i) = \cos t + (\sin t)\ell.$$

2. Let q have norm 1. If q is real then either $q = 1 = e^{0i}$ or $q = -1 = e^{\pi i}$. If q is not real, then $\mathbb{R}[q] \simeq \mathbb{C}$; we work there. And there, q is a complex number with modulus 1, hence of the desired form.

3. In general $e^{q_1+q_2}$ and $e^{q_1} \cdot e^{q_2}$ are different; however equality holds if q_1 and q_2 commute. Moreover the quaternion exponential clearly extends the complex exponential; hence if $q = a + y$ with $a \in \mathbb{R}$ and $y \in \mathbb{P}$, then $e^q = e^a \cdot e^y$.

   So we are reduced to computing e^y for a *pure* quaternion y. We may assume $y \neq 0$. Let $y' = \frac{1}{|y|} y$, which squares to -1. Then:

$$e^y = e^{|y|y'} = \cos|y| + (\sin|y|)y' = \cos|y| + \frac{\sin|y|}{|y|} y.$$

4. We know that conjugation γ_q by $q \notin Z(\mathbb{H})$ is a rotation of $\mathbb{P}$. Here, $e^{t\ell}$ and ℓ commute, so $\gamma_{e^{t\ell}}$ centralises $\ell \neq 0$, therefore the axis is $\mathbb{R}\ell$. It remains to compute the angle.

[6] Deavours, C, 'The quaternion calculus', *American Mathematical Monthly*, 80 (9), 995-–1008, 1973

First suppose $\ell = i$. Then see that:

$$e^{ti} j e^{-ti} = (\cos t + \sin t i) j (\cos t - \sin t i) = (\cos^2 t - \sin^2 t) j + 2 \cos t \sin t k,$$

and $e^{ti} k e^{-ti} = (\cos^2 t - \sin^2 t) k - 2 \cos t \sin t j$ likewise. Therefore

$$\mathrm{Mat}_{(i,j,k)} \gamma_{e^{ti}} = \begin{pmatrix} 1 & & \\ & \cos^2 t - \sin^2 t & -2 \cos t \sin t \\ & 2 \cos t \sin t & \cos^2 t - \sin^2 t \end{pmatrix},$$

and we recognise the matrix of the rotation with oriented axis $\mathbb{R} i$ and angle $2t$.

For the general case, we let $\mathbb{S}$ act transitively by conjugation on the quaternion sphere: there is $x \in \mathbb{H}$ such that $x i x^{-1} = \ell$. But then, $\gamma_\ell = \gamma_x \circ \gamma_i \circ \gamma_x^{-1}$ is a conjugate of γ_i, so its geometric elements are: oriented axis $\gamma_x(\mathbb{R} i) = \mathbb{R}\ell$, angle measure $2t \mod 2\pi$, as desired.