Question: What is a number?

Answer: $\mathbb{R}$, $\mathbb{C}$, $\mathbb{H}$, and sometimes $\mathbb{O}$.

(1.A) In order of logical dependence, we have "numbers that can be added and multiplied"

$$\mathbb{N} = \{0 < 1 < 2 < \cdots\},$$

followed by "numbers that can be subtracted"

$$\mathbb{Z} = \{\cdots < -2 < -1 < 0 < 1 < 2 < \cdots\},$$

followed by "numbers that can be divided"

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\},$$

where we declare that the abstract symbols $\frac{a}{b}$ and $\frac{c}{d}$ are equal if and only if $ad = bc$. In this book we also want to know when two numbers are "close together", so we need some kind of topology. The idea of topology arises in the transition from $\mathbb{N}$ to $\mathbb{Z}$. Recall that $\mathbb{Z}$ is defined as the set of ordered pairs

$$\mathbb{Z} := \mathbb{N}^2 = \{(a, b) : a, b \in \mathbb{N}\},$$

where we declare (a, b) and (c, d) equal if and only if $a + d = b + c$. We are supposed to think that $(a, b) =$ "$a - b$". Then the **absolute value** function $\mathbb{Z} \to \mathbb{Z}_{\geq 0}$ is defined by

$$|(a, b)| := \begin{cases} (a, b) & \text{if } a \leq b \\ (b, a) & \text{if } b \leq a \end{cases},$$

as you know. This function extends to $\mathbb{Q} \to \mathbb{Q}_{\geq 0}$ in the natural way by defining

$$\left|\frac{a}{b}\right| := \frac{|a|}{|b|} \in \mathbb{Q}_{\geq 0}.$$

Now things get a bit tricky. If $\mathbb{Q}$ is going to be our definition of number, then many interesting things like $\sqrt{2}$ are unfortunately not numbers. People puzzled over this for a long time until Cantor came up with the following (admittedly quite scary) solution.

We say that $(x_i)_i = x_1, x_2, x_3, \ldots \in \mathbb{Q}$ is a **Cauchy sequence** *if for each (presumably small) rational number $\varepsilon \in \mathbb{Q}_{>0}$ there exists a natural number $N_\varepsilon \in \mathbb{N}$ such that for all natural numbers $m, n > N_\varepsilon$ we have $|x_m - x_n| < \varepsilon$.*

The thing about a Cauchy sequence is that it seems to be going somewhere. Cantor's idea was to arbitrarily declare that it *is* going somewhere. We define the set

$$\mathbb{R} := \{(x_i)_i \text{ in } \mathbb{Q} : (x_i)_i \text{ is a Cauchy sequence}\},$$

and we declare that $(x_i)_i$ and $(y_i)_i$ are equal in $\mathbb{R}$ if their componentwise difference tends to zero—that is, if for each $\varepsilon \in \mathbb{Q}_{>0}$ there exists $N_\varepsilon \in \mathbb{N}$ such that for all $n > N_\varepsilon$ we have $|x_n - y_n| < \varepsilon$. We are supposed to think that $(x_i)_i =$ "x_∞".

Now this is a *real* number system.

(1.B) We say that $\mathbb{R}$ is the **topological completion** of $\mathbb{Q}$ with respect to the absolute value function $|\cdot| : \mathbb{Q} \to \mathbb{Q}_{\geq 0}$. More generally, a function $\|\cdot\| : \mathbb{Q} \to \mathbb{Q}$ is called a **norm** if it satisfies the following three properties:

- $\forall\, x \in \mathbb{Q}, \|x\| = 0 \Leftrightarrow x = 0$,
- $\forall\, x, y \in \mathbb{Q}, \|xy\| = \|x\|\|y\|$, and
- $\forall\, x, y \in \mathbb{Q}, \|x + y\| \leq \|x\| + \|y\|$.

Exercise: Show that these three properties imply that $\|x\| \geq 0$ for all $x \in \mathbb{Q}$. One can also fiddle around with the definitions to show that the absolute value is a norm.

It turns out that the norm axioms are the only properties needed in Cantor's construction of $\mathbb{R}$ from $\mathbb{Q}$. Thus, if one is worried about the ontological status of $\mathbb{R}$, one might wonder if there are other (inequivalent) norm functions on $\mathbb{Q}$. Indeed there are.

We define the **trivial norm** *on* $\mathbb{Q}$ *by*

$$\|x\|_0 := \begin{cases} 0 & x = 0 \\ 1 & x \neq 0 \end{cases}.$$

I wouldn't worry about that, but there are worse things.

Let $p \in \mathbb{N}$ *be a prime number. Then for each nonzero* $x \in \mathbb{Q}$ *we can factor* p *from the numerator and denominator to write* $x = p^n \frac{a}{b}$ *for some* $n \in \mathbb{Z}$ *with* a, b, p *coprime. We define the* p**-adic norm** *on* $\mathbb{Q}$ *by*

$$\|x\|_p := \begin{cases} 0 & x = 0 \\ p^{-n} & x \neq 0 \end{cases}.$$

Exercise: Show that the p-adic norm is a norm. We use the notation $\|\cdot\|_\infty$ for the usual absolute value, because it acts like that.

Ostrowski's Theorem (1916). The inequivalent norms on $\mathbb{Q}$ are precisely

$$\|\cdot\|_0, \quad \|\cdot\|_\infty, \quad \text{and } \|\cdot\|_p \text{ for } p \text{ prime}.$$

Let $\mathbb{Q}_s$ denote the topological completion of $\mathbb{Q}$ with respect to the norm $\|\cdot\|_s$. Then $\mathbb{Q}_0 = \mathbb{Q}$, $\mathbb{Q}_\infty = \mathbb{R}$, and $\mathbb{Q}_p$ is called the field of p**-adic numbers**. I will have no more to say about p-adic numbers, because I am not qualified.

The absolute value $|\cdot| = \|\cdot\|_\infty$ is sometimes called the **Archimedean norm** because it satisfies the additional (Archimedean) property

- $\forall\, 0 \neq x \in \mathbb{Q}, \exists\, n \in \mathbb{N}, \|\underbrace{x + x + \cdots + x}_{n \text{ times}}\| > 1$,

and the other norms don't. (Roughly, this says that there are no infinitesimal numbers.) So if you *are* worried about the ontological status of $\mathbb{R}$, you can say something like "it is the unique complete Archimedean field". But don't worry. (To paraphrase Hadamard:) All of this is merely to sanction and legitimize our intuition about the continuity of a piece of string, and there never was any other reason for it.

(1.C) Using $\mathbb{R}$ as a foundation, we can build three more amazing number systems. Here I will primarily follow Chapter 20 of Stillwell (2001), a book which I strongly recommend. For more references, see the NOTES.

The complex numbers $\mathbb{C}$ are the first **exceptional** structure in mathematics. (My use of the term "exceptional" will hopefully become clear later.) Apparently, they were first observed by Diophantus, who knew that

$$(a^2+b^2)(\alpha^2+\beta^2) = (a\alpha \mp \beta b)^2 + (b\alpha \pm \beta a)^2.$$

We will call this the **two-square identity**. In geometric terms, let (a,b) denote the right angled triangle with side lengths a and b. If (a,b) and (α,β) are two such triangles, then we have a rule for producing a third triangle

$$(a\alpha - \beta b, b\alpha + \beta a),$$

whose hypotenuse is the product of the hypotenuses of (a,b) and (α,β). It is literally unbelievable how much of modern mathematics comes from this simple rule. Following Hamilton (1835), we will define the complex numbers as pairs of real numbers

$$\mathbb{C} := \mathbb{R}^2 = \{(a,b) : a,b, \in \mathbb{R}\}$$

with componentwise addition and with the strange product

$$(a,b) \times (\alpha,\beta) := (a\alpha - \beta b, b\alpha + \beta a).$$

Exercise: Explain why the distributive law holds. We can extend the absolute value from $\mathbb{R}$ to $\mathbb{C}$ by defining $|(a,b)|^2 := a^2+b^2$, and the two-square identity tells us that this is a norm. Furthermore, we can define the **conjugation** map $(a,b)^* := (a,-b)$ and observe that

$$(a,b) \times (a,b)^* = |(a,b)|^2.$$

It follows that we can divide by nonzero complex numbers

$$\frac{1}{(a,b)} = \frac{(a,b)^*}{|(a,b)|^2} = \left(\frac{a}{a^2+b^2}\,,\, \frac{-b}{a^2+b^2}\right),$$

hence $\mathbb{C}$ is a field. We can also think of $\mathbb{C}$ as a 2-dimensional vector space over $\mathbb{R}$ with basis $1 = (1,0)$ and $i = (0,1)$, so that $(a,b) = a1 + bi$. Then for each $(a,b) \in \mathbb{C}$, the multiplication map $(\alpha,\beta) \mapsto (a,b) \times (\alpha,\beta)$ is linear and we may identify (a,b) with the 2×2 matrix

$$(a,b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = a1 + bi.$$

Note that

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

We will call this a **Hermitian** structure on $\mathbb{R}^2$. If you already *believe* in linear algebra then this is the correct definition of $\mathbb{C}$ because it explains the properties of the conjugate $(a,b)^* = (a,b)^\top$ (transpose matrix) and norm $|(a,b)| = \det(a,b)$ (determinant). Taking all of this structure into account, we can say that $\mathbb{C}$ is a **real normed division algebra**.

(1.D) Hamilton was so fascinated by the complex numbers that he spent at least 13 years (from 1830 to 1843) trying to multiply *triples* of real numbers, in order to define *hypercomplex numbers*. Had he been a better student of number theory, he wouldn't have bothered. For example, we have

$$3 = 1^2+1^2+1^2 \quad \text{and} \quad 5 = 0^2+1^2+2^2,$$

but it is easy to show that $15 = 3 \cdot 5$ is *not* the sum of three integer squares. This implies that there is no such thing as a **three-square identity**, and hence it is impossible to give $\mathbb{R}^3$ the structure of a normed division algebra. But we are happy that Hamilton didn't know this. In frustration, he was eventually willing to throw away the commutative

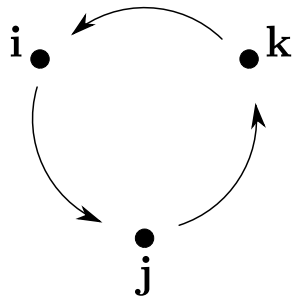
property of multiplication. On the afternoon of Monday, October 16, 1843, he was walking along the Royal Canal in Dublin with his wife, on the way to a meeting of the Royal Irish Academy, when he suddenly realized that there is a reasonable way to multiply *quadruples* of real numbers. After recording this insight in his notebook he got out his knife and carved the following formulas into the stone of the Brougham Bridge:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{ijk} = -\mathbf{1}. \qquad (\star)$$

Implicit in these formulas are the equations

$$\mathbf{ij} = -\mathbf{ji} = \mathbf{k}, \quad \mathbf{jk} = -\mathbf{kj} = \mathbf{i}, \quad \mathbf{ki} = -\mathbf{ik} = \mathbf{j},$$

which we can visualize in the following way:



Hamilton had just invented vector calculus—not to mention vectors.

To be specific, Hamilton defined the set of **quaternions**

$$\mathbb{H} := \mathbb{R}^4 = \{a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} : a, b, c, d \in \mathbb{R}\}.$$

The relations $(\star)$ induce an $\mathbb{R}$-bilinear associative product operation, which means that the quaternions can be represented as matrices. If you believe in the complex number $i \in \mathbb{C}$ then, following Cayley (1858), we can identify each quaternion with a complex 2×2 matrix:

$$\begin{aligned} a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} &= \begin{pmatrix} a1 + bi & c1 + di \\ -c1 + di & a1 - bi \end{pmatrix} \\ &= a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \end{aligned}$$

Warning: $\mathbf{i} \in \mathbb{H}$ is not the same as $i \in \mathbb{C}$. If you don't believe in the complex number $i \in \mathbb{C}$, then everything can be said in terms of real 4×4 matrices using

$$\mathbf{i} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \text{ et cetera.}$$

As with the complex numbers, we can define quaternion "conjugation" and quaternion "absolute value". Given $q = a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \in \mathbb{H}$ we let

$$q^* := a\mathbf{1} - b\mathbf{i} - c\mathbf{j} - d\mathbf{k} \quad \text{and} \quad |q| := a^2 + b^2 + c^2 + d^2.$$

Exercise: Show that $qq^* = |q|^2$. From this it follows that we can divide by nonzero quaternions

$$\frac{1}{a\mathbf{1} + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}} = \frac{a}{|q|^2}\mathbf{1} - \frac{b}{|q|^2}\mathbf{i} - \frac{c}{|q|^2}\mathbf{j} - \frac{d}{|q|^2}\mathbf{k},$$

and we say that $\mathbb{H}$ is a **real normed division algbra**. In terms of 2×2 complex matrices we see that q^* is the complex conjugate transpose and $|q|$ is the determinant. The

multiplicative property of the determinant then implies that the absolute value $|\cdot| : \mathbb{H} \to \mathbb{R}$ is a norm. Writing this out explicitly, we obtain the **four-square identity**:

$$\begin{aligned}(a^2+b^2+c^2+d^2)(\alpha^2+\beta^2+\gamma^2+\delta^2) = \quad & (a\alpha - b\beta - c\gamma - d\delta)^2 \\ & + (\alpha\beta + b\alpha + c\delta - d\gamma)^2 \\ & + (a\gamma - b\delta + c\alpha + d\beta)^2 \\ & + (a\delta + b\gamma - c\beta + d\alpha)^2.\end{aligned}$$

Had Hamilton been a better student of number theory, he would already have known this. The four-square identity was discovered 100 years earlier by Euler (1743) and used by Euler and Lagrange to prove the famous theorem that every natural number is a sum of four squares.

Thus the quaternions enjoy all of the algebraic properties of the complex numbers except for commutative multiplication. In 1843 noncommutative multiplication was unusual but today it makes perfect sense: We can think of unit complex numbers as rotations of $\mathbb{R}^2$ (which commute) and we can think of unit quaternions as rotations of $\mathbb{R}^3$ (which don't commute). In particular, the unit quaternions **i**, **j**, and **k** represent rotations by 180° around the x-, y-, and z-axes in $\mathbb{R}^3$. I will have more to say later.

(1.E) Now here is a historical analogy. The algebraic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

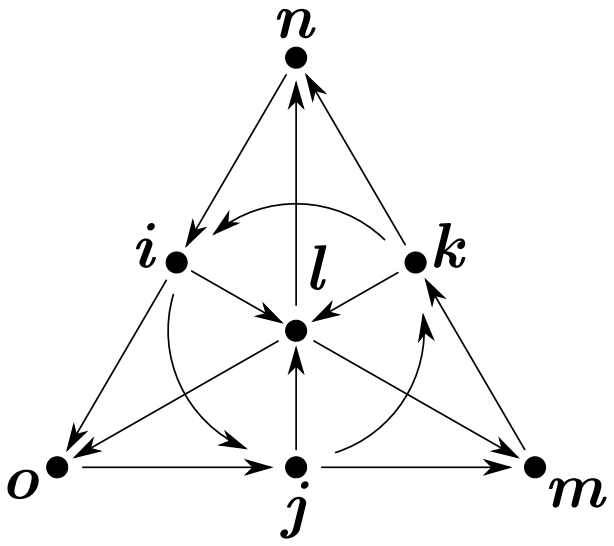
for solving the **quadratic equation** $ax^2 + bx + c = 0$ was known since antiquity. After Gerolamo Cardano learned the complete solution of the **cubic equation** (before 1545), he shared this information with his student Lodovico Ferrari. Almost immediately, the younger mathematician was able to extend Cardano's solution in order to solve the **quartic equation**. But then progress stalled (permanently, it turns out). Something similar happened in our story.

The day after Hamilton's discovery (on October 17, 1843) he sent a letter to his good friend John Graves (who had originally inspired Hamilton's interest in complex numbers). Graves was impressed by Hamilton's boldness, but he said: "There is still something in the system which gravels me. I have not yet any clear views as to the extent to which we are at liberty to arbitrarily create imaginaries, and to endow them with supernatural properties." He added: "If with your alchemy you can make three pounds of gold, why should you stop there?" Graves did not stop there. On December 26 he wrote to Hamilton detailing a new 8-dimensional number system which he called the *octaves*. Hamilton was slow to promote Graves' *octaves*; meanwhile, the young Arthur Cayley rediscovered the system and published his results in 1845. For this reason the system is sometimes known as the *Cayley numbers*.

Today we (nearly) follow Graves' notation, by defining the **octonions**

$$\mathbb{O} := \mathbb{R}^8 = \{a\boldsymbol{1} + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} + e\boldsymbol{l} + f\boldsymbol{m} + g\boldsymbol{n} + h\boldsymbol{o} : a, b, c, d, e, f, g, h \in \mathbb{R}\}.$$

The algebraic relations can be neatly summarized with the following mnemonic diagram of Freudenthal (1951):



Each oriented line in the diagram (and the one oriented circle $\boldsymbol{i} \to \boldsymbol{j} \to \boldsymbol{k}$) represents a Hamiltonian triple. For example, the oriented line $\boldsymbol{i} \to \boldsymbol{l} \to \boldsymbol{m}$ represents the relations

$$\boldsymbol{i}^2 = \boldsymbol{l}^2 = \boldsymbol{m}^2 = \boldsymbol{ilm} = -\boldsymbol{1}.$$

Thus the octonions contain many copies of the quaternions. The essential feature that makes this more than just an "arbitrary creation of imaginaries" is the existence of a multiplicative norm. Graves and Cayley independently observed that the absolute value

$$\left|a\boldsymbol{1} + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} + e\boldsymbol{l} + f\boldsymbol{m} + g\boldsymbol{n} + h\boldsymbol{o}\right| = a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2$$

satisfies $|u||v| = |uv|$ for all $u, v \in \mathbb{O}$. Writing this out in coordinates gives the **eight-square identity**:

$$\begin{aligned}
(a^2 + b^2 + c^2 + d^2 + e^2 + f^2 + g^2 + h^2)&(\alpha^2 + \beta^2 + \gamma^2 + \delta^2 + \varepsilon^2 + \zeta^2 + \eta^2 + \theta^2) \\
= \quad & (a\alpha - b\beta - c\gamma - d\delta - e\varepsilon - f\zeta - g\eta - h\theta)^2 \\
& + (a\beta + b\alpha + c\delta - d\gamma + e\zeta - f\varepsilon - g\theta + h\eta)^2 \\
& + (a\gamma - b\delta + c\alpha + d\beta + e\eta + f\theta - g\varepsilon - h\zeta)^2 \\
& + (a\delta + b\gamma - c\beta + d\alpha + e\theta - f\eta + g\zeta - h\varepsilon)^2 \\
& + (a\varepsilon - b\zeta - c\eta - d\theta + e\alpha + f\beta + g\gamma + h\delta)^2 \\
& + (a\zeta + b\varepsilon - c\theta + d\eta - e\beta + f\alpha - g\delta + h\gamma)^2 \\
& + (a\eta + b\theta + c\varepsilon - d\zeta - e\gamma + f\delta + g\alpha - h\beta)^2 \\
& + (a\theta - b\eta + c\zeta + d\varepsilon - e\delta - f\gamma + g\beta + h\alpha)^2.
\end{aligned}$$

At this point Graves searched the literature and discovered that Hamilton's four-square identity was known to Euler, and that his own eight-square identity was known to Ferdinand Degen in 1822.

I should confess right now that I cheated when I wrote the relation $\boldsymbol{ilm} = -\boldsymbol{1}$. It turns out to be okay because $(\boldsymbol{il})\boldsymbol{m} = \boldsymbol{m}^2 = -\boldsymbol{1} = \boldsymbol{i}^2 = \boldsymbol{i}(\boldsymbol{lm})$, but in general the octonions are *not associative*. To see this, observe that

$$\boldsymbol{i}(\boldsymbol{jm}) = \boldsymbol{io} = \boldsymbol{n}, \quad \text{but} \quad (\boldsymbol{ij})\boldsymbol{m} = \boldsymbol{km} = -\boldsymbol{n}.$$

(You'll need to prove for yourself that $\boldsymbol{n} \neq -\boldsymbol{n}$.) This means that the octions can *not* be represented by matrices, and so even *less* of the apparatus of linear algebra is available to

deal with them. Nevertheless, if we define octonion "conjugation" by

$$(a\boldsymbol{1} + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k} + e\boldsymbol{l} + f\boldsymbol{m} + g\boldsymbol{n} + h\boldsymbol{o})^* := a\boldsymbol{1} - b\boldsymbol{i} - c\boldsymbol{j} - d\boldsymbol{k} - e\boldsymbol{l} - f\boldsymbol{m} - g\boldsymbol{n} - h\boldsymbol{o},$$

it follows that $uu^* = |u|^2$. Hence $u^{-1} = u^*/|u|^2$, and we see that $\mathbb{O}$ is still a **normed division algebra**. This turns out to be enough to give $\mathbb{O}$ an important place in mathematics.

(1.F) In retrospect, each step in the chain of extensions

$$\mathbb{R} \to \mathbb{C} \to \mathbb{H} \to \mathbb{O}$$

is completely analogous. The following recursive construction is called **Dickson doubling** (1914). Let $\mathbb{A}_0 = \mathbb{R}$ and define a trivial conjugation $a^* = a$ for all $a \in \mathbb{R}$. Then for each $n \in \mathbb{N}$ consider the set $\mathbb{A}_{n+1} = \{(a, b) : a, b \in \mathbb{A}_n\}$ of ordered pairs from $\mathbb{A}_n$. We recursively define *addition*, *conjugation*, and *multiplication* operations on $\mathbb{A}_{n+1}$ by

- $(a, b) + (\alpha, \beta) := (a + \alpha, b + \beta)$,
- $(a, b)^* := (a^*, -b)$, and
- $(a, b) \times (\alpha, \beta) := (a \times \alpha - \beta^* \times b, b \times \alpha^* + \beta \times a)$.

Clearly we have $\mathbb{A}_1 = \mathbb{C}$. (This explains my previous notational choices.) With a little work we can also show that $\mathbb{A}_2 = \mathbb{H}$ and $\mathbb{A}_3 = \mathbb{O}$. Exercise: Verify that the doubling $\mathbb{H} \to \mathbb{O}$ is given by

$$\begin{array}{rclcrcl} \boldsymbol{1} & = & (\mathbf{1}, 0), & & \boldsymbol{l} & = & (0, \mathbf{1}), \\ \boldsymbol{i} & = & (\mathbf{i}, 0), & & \boldsymbol{m} & = & (0, \mathbf{i}), \\ \boldsymbol{j} & = & (\mathbf{j}, 0), & & \boldsymbol{n} & = & (0, \mathbf{j}), \\ \boldsymbol{k} & = & (\mathbf{k}, 0), & & \boldsymbol{o} & = & (0, \mathbf{k}). \end{array}$$

In general, if we recursively define scalar multiplication $a \cdot (b, c) := (a \cdot b, a \cdot c)$ for $a \in \mathbb{R}$, then $\mathbb{A}_n$ becomes a **real algebra** of dimension 2^n. Is it always a division algebra? No: Given the octonion generators $\boldsymbol{i}, \boldsymbol{k}, \boldsymbol{l}, \boldsymbol{n} \in \mathbb{O}$, consider the pairs $(\boldsymbol{i}, \boldsymbol{n})$ and $(\boldsymbol{k}, \boldsymbol{l})$ in $\mathbb{A}_4$ and observe that

$$\begin{aligned} (\boldsymbol{i}, \boldsymbol{n}) \times (\boldsymbol{k}, \boldsymbol{l}) &= (\boldsymbol{ik} - \boldsymbol{l}^*\boldsymbol{n}, \boldsymbol{nk}^* + \boldsymbol{li}) \\ &= (\boldsymbol{ik} + \boldsymbol{ln}, -\boldsymbol{nk} + \boldsymbol{li}) \\ &= (-\boldsymbol{j} + \boldsymbol{j}, \boldsymbol{m} - \boldsymbol{m}) \\ &= (0, 0). \end{aligned}$$

Since $\mathbb{A}_4$ contains zero-divisors it is not a division algebra. It follows that the Euclidean absolute value $|\cdot| : \mathbb{A}_4 = \mathbb{R}^{16} \to \mathbb{R}$ is not multiplicative, since otherwise we would have zero-divisors in $\mathbb{R}$. Thus $\mathbb{A}_4$ is not even a normed algebra. It seems that the algebraic properties of $\mathbb{A}_n$ get worse as n gets larger. The situation is summarized by the following major theorems.

Frobenius' Theorem (1877). Every real associative division algebra is isomorphic to

$$\mathbb{R}, \mathbb{C}, \text{or } \mathbb{H}.$$

Hurwitz' Theorem (1898). Every real normed division algebra is isomorphic to

$$\mathbb{R}, \mathbb{C}, \mathbb{H}, \text{or } \mathbb{O}.$$

Like all classifications in mathematics, this one doesn't have sharp edges, but fades away in a fog of words like "power associative", "alternative", and "Moufang loop". At the present moment, the boundary between interesting and uninteresting in the sequence

$$\mathbb{A}_0, \mathbb{A}_1, \mathbb{A}_2, \mathbb{A}_3, \mathbb{A}_4, \ldots$$

seems to lie between $\mathbb{A}_3$ and $\mathbb{A}_4$. I don't know of a *really* good reason why the algebra $\mathbb{A}_4$ (called **sedenions**) is not interesting, but it seems not to be. Perhaps the key concept is not "normed division algebras", but "n-square identities". In fact, we could rephrase Hurwitz' Theorem in a more historically accurate way.

Hurwitz' Theorem (1898). There are no n-square identities except for $n = 1, 2, 4, 8$.

Finally, I will declare my philosophy. In this book I will discuss several examples of "classification" in mathematics, most of which are related to the mysterious labels "ADE". It is always difficult to know when a classification theorem is complete, i.e., when further extension is likely to lead to diminishing returns. Classification theorems usually have some "obvious" examples (like $\mathbb{R}$) and some "exceptional" examples (like $\mathbb{O}$) lying close to the interesting/uninteresting boundary. One of the great joys of mathematics is that "exceptional" objects in different areas frequently turn out to be related in deep ways. This is a huge source of motivation.

Philosophy: All exceptional structures in mathematics are related.

Exercises

(1.1) (Descartes, 1637) Let $\mathbb{F}$ be a field. For any two elements $a, b \in \mathbb{F}$ we have

$$a^m - b^m = (a-b)(a^{m-1} + a^{m-2}b + \cdots + ab^{m-2} + b^{m-1}).$$

Now let $f(x) \in \mathbb{F}[x]$ be any polynomial with coefficients in $\mathbb{F}$. For all $a \in \mathbb{F}$ show that $f(x) - f(a)$ is divisible by $(x-a)$ in the ring $\mathbb{F}[x]$. If $f(a) = 0$, conclude that $f(x)$ is divisible by $(x-a)$.

(1.2) Let $\mathbb{F}$ be a field and consider a polynomial $f(x) \in \mathbb{F}[x]$ of degree n. Use **(1.1)** and induction to show that $f(x)$ has *at most n roots* in $\mathbb{F}$.

(1.3) Now consider the polynomial $f(x) = x^2+1 \in \mathbb{H}[x]$ with coefficients in the quaterions. By definition we know that $f(\mathbf{i}) = f(\mathbf{j}) = f(\mathbf{k}) = 0$. By **(1.2)** this implies that $\mathbb{H}$ is not a field. (Where did the proof go wrong?)

(1.4) In fact, show that for *any* pure imaginary unit quaternion

$$u = a\mathbf{i} + b\mathbf{j} + c\mathbf{k} \quad \text{with} \quad a^2 + b^2 + c^2 = 1$$

we have $u^2 + 1 = 0$. Hence the polynomial $f(x) = x^2 + 1 \in \mathbb{H}[x]$ has *uncountably* many roots in $\mathbb{H}$ (a whole 3-sphere of them).

(1.5) Find all of the quaternion roots of the polynomial $f(x) = x^2 - 1 \in \mathbb{H}[x]$. [Hint: There are only two.]

(1.6)* What is the formal definition of an "n-square identity"? The fact that there is no three-square identity over $\mathbb{Q}$ is easy to see. The fact that there is no three-square identity over $\mathbb{R}$ is related to the fact that there is no nonvanishing smooth vector field on the 2-sphere. Is there a direct relationship between these two facts? That is, can one regard the fact that $15 = (1^2+1^2+1^2)(0^1+1^2+2^2)$ is not the sum of three integer squares as a *proof* that there is no nonvanishing smooth vector field on the 2-sphere?

Notes

I have mostly followed the notation of Stillwell (2001), Chapter 20. For more on the octonions and their history see the paper of Baez (2002). For much more on the octonions see the book of Conway and Smith (2003). For much, much more on the octonions and the concept of "number" in general, see the book of Ebbinghaus et al. (1990). In particular, see the chapter by Neukirch on the p-adic numbers. The story of the numbers $1, 2, 4, 8$ does not end with Hurwitz' Theorem. There is now a much stronger theorem of Kervaire and Milnor (1958) which says that the dimension of any real division algebra must be 1, 2, 4 or 8. The only known proofs use Bott periodicity and K-theory, which indicates that there is a deep relationship between "number" and "topology". For an exposition of this story see the chapter by Hirzebruch in *Numbers*.

Bibliography

Baez, John C. (2002). The Octonions. *Bull. Amer. Math. Soc.* **39**, 145–205.

Cayley, Arthur (1858). A memoir on the theory of matrices. *Phil. Trans. Roy. Soc. London*, **148**, 17–37. In his *Collected Mathematical Papers* 2: 475–496.

Conway, John H. and Smith, Derek A. (2003). *On Quaternions and Octonions: Their Geometry, Arithmetic, and Symmetry*. A K Peters, Natick, MA.

Dickson, Leonard Eugene (1914). *Linear Algebras*. Cambridge University Press, Cambridge.

Ebbinghaus et al. (1990). *Numbers, English Edition*. Springer-Verlag, New York.

Freudenthal, Hans (1951). *Oktaven, Ausnahmegruppen und Oktavengeometrie*. Mathematisch Instituut der Rijksuniversiteit te Utrecht, Utrecht.

Frobenius, Ferdinand Georg (1878). Über lineare Substitutionen und bilineare Formen. *J. reine und angew. Math.*, **84**, 1–63. In his *Gesammelte Abhandlungen* 1: 343–405.

Hamilton, William Rowan (1835). Theory of conjugate functions, or algebraic couples. Communicated to the Royal Irish Academy, 1 June 1835. In his *Mathematical Papers* 3: 76–96.

Hurwitz, Adolf (1898). Über die komposition der quadratischen Formen von bliebig vielen Variablen. *Göttinger Nachrichten*, pages 309–316. In his *Mathematische Werke* 2: 565–571.

Ostrowski, Alexander (1916). Über einige Lösungen der Funktionalgleichung $\varphi(x)\varphi(y) = \varphi(xy)$. *Acta Math.*, **41**, 271–284.

Stillwell, John (2001). *Mathematics and Its History, 2nd ed.* Springer-Verlag, New York.