Characterizing $\mathbb{R}$ and $\mathbb{C}$

To get a better feeling for what makes $\mathbb{C}$ (the complex numbers) so special, let's take a look at why in the end we are almost forced to study $\mathbb{R}$ (the real numbers) in the first place.

To begin, we very naturally want to have a way to describe cardinal (more or less: size) and ordinal (more or less: ordering) information. The positive integers, $\mathbb{Z}_{>0} = \{1, 2, \dots\}$ do both of these things for us. The number 3 both stands in for having 3 things (like "a, b, c" is a list of "3" letters) and being in 3rd place ("c" is letter number "3" in this list). People have been using these kinds of numbers for as long as there have been people.

Next, people needed to deal with parts of a whole, so they developed fractions. But this wasn't enough, we needed irrational quantities too. For example: In a 45°-45°-90° triangle with legs of length 1, the hypotenuse's length is $\sqrt{2}$. By the time of the Greek's, it was known that this number cannot be expressed as the ratio of two whole numbers. Fast forward and 0 is introduced in India. While negative numbers experienced some acceptance in China, they were mostly refused (in the Western world) until Fibonacci's time. Fibonacci used negative numbers himself and advocated that they were useful in finance (positive = credit, negative = debt).

So by the end of the Renaissance, the Western world saw the need for and usefulness of a number system which allowed addition, subtraction, multiplication, and division. The concept of a field (which came later) is the abstract characterization of such a numeric system.

Definition: Let $\mathbb{F}$ be a non-empty set equipped with two operations: $+ : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ denoted $a + b$ for $a, b \in \mathbb{F}$ and $\cdot : \mathbb{F} \times \mathbb{F} \to \mathbb{F}$ denoted $a \cdot b$ or ab for $a, b \in \mathbb{F}$. [We have already required *closure*: If $a, b \in \mathbb{F}$, then $a + b, ab \in \mathbb{F}$.] We also require that the following properties hold:

Associativity: For all $a, b, c \in \mathbb{F}$, $(a + b) + c = a + (b + c)$ and $(ab)c = a(bc)$.

Identity: There are special elements $0, 1 \in \mathbb{F}$ such that $0 \neq 1$ and for all $a \in \mathbb{F}$, $0 + a = a = a + 0$ and $1a = a = a1$.

Inverses: For each $a \in \mathbb{F}$ there is some $-a \in \mathbb{F}$ such that $(-a) + a = 0 = a + (-a)$. If in addition, $a \neq 0$, there exists $a^{-1} \in \mathbb{F}$ such that $a^{-1}a = 1 = aa^{-1}$.

Commutativity: For all $a, b \in \mathbb{F}$, $a + b = b + a$ and $ab = ba$.

Distribution: For all $a, b, c \in \mathbb{F}$, $(a + b)c = ac + bc$ and $a(b + c) = ab + ac$.

Such a system, $\mathbb{F}$, is called a **field**.

We can see that the rational numbers ($\mathbb{Q}$), real numbers ($\mathbb{R}$), and complex numbers ($\mathbb{C}$) are all examples of fields. Of course, there are other examples of fields that look less familiar. For example, $\mathbb{Z}_p$ (integers modulo a prime p) and $\mathbb{R}(x)$ (rational functions with real coefficients) both form fields. This means that the real numbers must possess other structure which makes them stand out.

Definition: Let $\mathbb{F}$ be a field equipped with a relation $\leq$. Suppose that the following properties hold:

Reflexive: For all $a \in \mathbb{F}$, $a \leq a$.

Antisymmetric: For all $a, b \in \mathbb{F}$, if $a \leq b$ and $b \leq a$, then $a = b$.

Transitive: For all $a, b, c \in \mathbb{F}$, if $a \leq b$ and $b \leq c$, then $a \leq c$.

Then $\leq$ is called a **partial order**. If in addition, for every $a, b \in \mathbb{F}$ we have either $a \leq b$ or $b \leq a$, then $\leq$ is called a **total order**. Suppose that $\leq$ is a total order and for all $a, b, c \in \mathbb{F}$ we have that $a \leq b$ implies $a + c \leq b + c$ as well as $0 \leq a$ and $0 \leq b$ implies that $0 \leq ab$. Then we call $\mathbb{F}$ an **ordered field**.

Define $a < b$ to mean $a \leq b$ but $a \neq 0$ and $a \geq b$ if $b \leq a$ and $a > b$ if $b < a$. We say a is positive if $a > 0$ and negative if $a < 0$. It isn't hard to prove that negative times negative is positive and positive times positive is positive. It isn't hard to show that $1 > 0$. This then implies that $0 = 1 - 1 > 0 - 1 = -1$. I won't go into more details, but essentially the relations $\leq$, $<$, etc. work just like we think they should. Notice that $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields with the usual inequality operators.

It is worth mentioning that once we have an ordering, we also have *topology*: we can discuss limits and continuity. For example: For an ordered field $\mathbb{F}$, we can define intervals such as $(a, b) = \{x \in \mathbb{F} \mid a < x < b\}$, $[a, b] = \{x \in \mathbb{F} \mid a \leq x \leq b\}$, $(a, \infty) = \{x \in \mathbb{F} \mid a < x\}$, etc. *Open* intervals then form a basis for a topology on $\mathbb{F}$. This means we can start to do some *analysis* and not just *algebra*.

Another consequence of being ordered is that $1, 1 + 1 = 2, 1 + 1 + 1 = 3, \dots$ must all be positive. In particular, $1 + 1 + \cdots + 1 \neq 0$. This means that an ordered field must have **characteristic** 0. For example: This means that the

fields $\mathbb{Z}_p$ cannot be ordered in a way compatible with its arithmetic. On the other hand, any field of characteristic 0 0 must contain an isomorphic copy of $\mathbb{Q}$ (the converse is also true).

Now is a good time to mention that $\mathbb{C}$ cannot be ordered either. While $\mathbb{C}$ is a field of characteristic 0, we have a problem with i. Suppose that $\mathbb{C}$ were ordered. Then either $i > 0$ or $i < 0$. But in either case, $-1 = i^2 > 0$ (which is impossible). The crux is the problem is that $\mathbb{C}$ has elements other than ± 1 with finite multiplicative order.

Now that we realize $\mathbb{C}$ does not have an ordered field structure, we can see why people balked at using the complex numbers for so long. If we think of *ordering* and *magnitude* as essential to being a *number*, then "things" such as $i = \sqrt{-1}$ cannot be "numbers". This leads to name calling: i is "not real" and/or "imaginary". How sad.

Let us return to characterizing $\mathbb{R}$. So far we've ruled out some fields by looking at orderings, but we still have that both $\mathbb{Q}$ and $\mathbb{R}$ are ordered fields. This is where analysis comes into play. The real numbers have an additional property:

Definition: Let $\mathbb{F}$ be an ordered field and $S \subseteq \mathbb{F}$. If $m \in \mathbb{F}$ and $m \leq x$ for all $x \in S$, we say that m is a **lower bound** for S. If S has a lower bound, then S is **bounded below**. Likewise, we could define **upper bound** and **bounded above**. If S is bounded both above and below, then S is **bounded**.

Next, suppose that g is a lower bound for S such that given any other lower bound m, we have $m \leq g$. In other words, if $g \leq x$ for all $x \in S$ and if $m \leq x$ for all $x \in S$, then $m \leq g$. Then g is called the **greatest lower bound** or **infimum** of S, denoted $\text{glb}(S)$ or $\inf(S)$, (one can show that if $\inf(S)$ exists, it is unique). Likewise, we could define **least upper bound** or **supremum** ($\text{lub}(S) = \sup(S)$) and show that if such a bound exists, it is unique.

We say that an ordered field has the **greatest lower bound property** if for every non-empty set S such that S is bounded below, $\text{glb}(S)$ exists. Likewise, we could define a **least upper bound property**. It turns out that these two properties are equivalent (if one holds, both hold). Finally, any ordered field having the greatest lower bound (or equivalently least upper bound) property is called a **complete ordered field**.

Notice that if $S = \{r \in \mathbb{Q} \mid \sqrt{2} \leq r\}$, then $\text{glb}(S) = \sqrt{2} \notin \mathbb{Q}$. This shows that $\mathbb{Q}$ does not have the greatest lower bound property – it is *not* a complete ordered field. On the other hand, $\mathbb{R}$ is complete. In fact, this characterizes $\mathbb{R}$: It is the only complete ordered field (up to isomorphism).

Theorem: Suppose that $\mathbb{F}$ and $\mathbb{K}$ are complete ordered fields. Then there exists a bijection $\varphi : \mathbb{F} \to \mathbb{K}$ such that $\varphi(a+b) = \varphi(a) + \varphi(b)$, $\varphi(ab) = \varphi(a)\varphi(b)$, and $a < b$ implies $\varphi(a) < \varphi(b)$ for all $a, b \in \mathbb{F}$. This means that any two complete ordered fields are isomorphic. We call any such field: $\mathbb{R}$, the real numbers.

Proof: As a sketch of a proof, take the $\mathbb{R}$ defined via Dedekind cuts or some such method and let $\mathbb{F}$ be a complete ordered field. Then we work on defining a map. To make sure φ is an isomorphism of fields we must let $\varphi(0) = 0$, $\varphi(1) = 1$, and so $\varphi(2) = \varphi(1+1) = \varphi(1) + \varphi(1) = 1 + 1 = 2$, etc. This eventually shows that φ maps $\mathbb{Q}$ in $\mathbb{R}$ to the corresponding copy of $\mathbb{Q}$ in $\mathbb{F}$. So far φ preserves orderings. Next, any real number can be defined by a "cut". Pick $r \in \mathbb{R}$ then $r = \inf(S)$ where $S = \{x \in \mathbb{Q} \mid r \leq x\}$. For things to work out, we must define $\varphi(r) = \inf(\varphi(S)) = \inf(\{\varphi(x) \mid x \in S\})$. It is not too difficult to show that since S is bounded below, so is $\varphi(S)$. Thus $\varphi(S)$ must have an infimum (because $\mathbb{F}$ is complete). We now have φ defined on all of $\mathbb{R}$. One now shows that it is an order preserving homomorphism and is one-to-one and onto. We now have that $\mathbb{R} \cong \mathbb{F}$ and so by transitivity any two complete ordered fields are isomorphic. ■

This now means that whether we construct the real numbers via Dedekind cuts or by equivalence classes of Cauchy sequences of rational numbers or by decimal sequences, these construction yield the same abstract system that we call $\mathbb{R}$.

Now what about $\mathbb{C}$? Well, notice that we cannot solve $x^2 + 1 = 0$ in $\mathbb{R}$. This means that while $\mathbb{R}$ is complete in some topological sense, it is not complete is an algebraic sense. We say that a field is **algebraically closed** if every non-constant polynomial has a root, or equivalently, every non-constant polynomial factors into linear factors. The **Fundamental Theorem of Algebra** (which we will prove later in this course) states that $\mathbb{C}$ is algebraically closed. It turns out that every field has an algebraic closure (every field is contained in an algebraically closed field and the smallest such field is its called its closure). Also, an algebraic closure of a field is unique up to isomorphism.

Theorem: The complex numbers, $\mathbb{C}$, are characterized as the algebraic closure of a complete order field: If $\bar{\mathbb{F}}$ is the algebraic closure of a complete ordered field $\mathbb{F}$, then $\bar{\mathbb{F}} \cong \mathbb{C}$ where this isomorphism extends an order preserving isomorphism between $\mathbb{F}$ and $\mathbb{R}$.

This explains why we are naturally led to the study of complex numbers whether we originally wanted to or not!