How are the real numbers constructed from the rationals?

Extended essay
Mathematics HL
Pentti Sunila
May 2013
M13-D1292027
Tikkurilan lukio 1292
Word count: 2091

Abstract

The research question of this extended essay is how the real numbers are constructed from the rationals. The main method under discussion is viewing the reals as the metric completion of the rationals. We start by giving the axiomatic definition of the real numbers as the complete ordered field. This is then compared to the field of rational numbers. We then give the definitions required to construct the Cauchy completion of the rational numbers, and it is shown how this construction satisfies the axioms given in the introduction. The method of Dedekind cuts is also briefly discussed and compared with the Cauchy completion method.
Wordcount: 104

Contents

1 Introduction

We're all intuitively familiar with the real numbers, and used to using them throughout our education and everyday life, but what exactly are they and how are they constructed? In this essay we will detail the construction of the real numbers from the rational numbers using the method of Cauchy completion. First, let's look at the axiomatic definition of the real numbers, so we'll know what we're working towards in the rest of the essay.

Definition 1. *[1]* ***The real number system*** *is the complete ordered field, this means that it is a set* $\mathbb{R}$ *with two operations,* $+$ *and* $\cdot$*, that can be performed to any two of its elements such that for all* $a, b, c \in \mathbb{R}$*:*

1. *Addition is* ***associative****:* $a + (b + c) = (a + b) + c$
2. *Addition is* ***commutative****:* $a + b = b + a$
3. *There exists an* ***additive identity*** *0* $\in \mathbb{R}$*:* $a + 0 = a$
4. *Every element a has an* ***additive inverse*** *-a* $\in \mathbb{R}$ *such that:* $a + (-a) = 0$
5. *Multiplication is associative:* $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
6. *Multiplication is commutative:* $a \cdot b = b \cdot a$
7. *There exists a* ***multiplicative identity*** *1* $\in \mathbb{R}$*:* $1 \cdot a = a$
8. *Every nonzero element a has a* ***multiplicative inverse*** $\frac{1}{a} \in \mathbb{R}$*, such that:* $a \cdot \frac{1}{a} = 1$
9. *Multiplication* ***distributes*** *over addition:* $a \cdot (b + c) = a \cdot b + a \cdot c$
10. $\mathbb{R}$ *has an* ***ordering*** $<$*, such that:*
    (a) *for any* $x, y \in \mathbb{R}$*, one of the following holds:* $x < y$*,* $y < x$*, or* $x = y$
    (b) *if* $x < y$ *and* $y < z$*, then* $x < z$
    (c) *if* $y < z$*, then* $x + y < x + z$
    (d) *if* $0 < x$ *and* $0 < y$ *then* $0 < x \cdot y$
11. $\mathbb{R}$ *is* ***complete****. For every subset* $F \in \mathbb{R}$ *for which there exists an* ***upper bound*** $b \in \mathbb{R}$ *such that for all* $f \in F$*,* $f < b$ *or* $f = b$*, there exists a* ***least upper bound*** $s \in \mathbb{R}$ *that is smaller than every other upper bound.*

This definition of the real numbers might seem cryptic and daunting, but it can be proven that it completely characterizes the real numbers, and that the real numbers are the only system satisfying these properties, explaining why they're called *the* complete ordered field [4]. What this definition means will come clearer as we discuss similar properties on the rational numbers and eventually realize these axioms by constructing the real numbers from the rationals.

2 The rational numbers

A rational number is a number that can be expressed as a fraction of two integers, $q = \frac{a}{b},\ a, b \in \mathbb{Z},\ b \neq 0$. The rational numbers $\mathbb{Q}$ have a number of nice properties. They can be multiplied with and added to each other and the result is always another rational number, there exists an identity for both addition and multiplication, namely 0 and 1, as adding zero to any number or multiplying it by 1 results in the same number. Each number also has inverse elements with respect to addition and multiplication, the negative and the reciprocal, the addition and multiplication by which results in the identity. Both of these operations are also commutative and associative, meaning that order of evaluation doesn't matter: $a+(b+c) = (a+b)+c$ and $a+b = b+a$ and the same for multiplication.. And, as we know, multiplication distributes over addition: $a(b+c) = ab+ac$. We see that the rational numbers do indeed form a field. A field is then an abstracted number system that follows some of the familiar rules of arithmetic. The rationals are actually an ordered field as the familiar relation $<$ creates an ordering for the rational numbers. The rational numbers are lacking in one property though.

For example, consider the set of all rational number whose square is less than 3. This set can be ordered because the rational numbers are ordered by $<$, and it is bounded above by, for example, 2. Yet, the set does not have a least upper bound, because there is no rational number that squares to 3, or a smallest rational number whose square is greater than 3. Similarly for the set of rational numbers whose squares are less than two, say. These examples show that these sets, while bounded above, do not have a least upper bound. For every upper bound for these sets we can find an another rational number smaller than it that is also an upper bound for the set. So the rational numbers form an ordered field, but are not complete, as subsets that are bounded above do not necessarily have a least upper bound. We can in fact define infinite sequences as subsets of these sets by picking a starting point and following the ordering, in which the difference between the elements must keep getting smaller, because the set is bounded above, and see that we can have infinite sequences of rational numbers that seem to tend to some sort of limit, which doesn't exist.

We see that there are "holes" in the set of rational numbers, sequences can be intuitively seen to tend to a number that doesn't exist. We will first have to formalize the notion of a limit of a sequence and a sequence tending to something, and then we will be able to "patch" these holes to create the real numbers.

3 Cauchy sequences

Definition 2. ***Convergent sequence, limit*** *Let (a_n) be a sequence. A number l is called the limit of the sequence (a_n) if for any $\epsilon > 0$, there exists a natural number N such that for every $n > N$, $|l - a_n| < \epsilon$. If the limit exists, the sequence (a_n) is called a convergent sequence.*

This makes our intuition of a sequence converging to a limit rigorous. The elements of the sequence get arbitrarily close to the limit. If $a = lim(a_n)$ and $b = lim(b_m)$, then $lim(a_n + b_m) = a + b$, because for any $\epsilon > 0$ there are $L, M \in \mathbb{N}$ so that for $m > L,\ n > N$, $|a - a_n| < \frac{\epsilon}{2}$ and $|b - b_m| < \frac{\epsilon}{2}$, so that $|a - a_n| + |b - b_m| = |(a+b) - (a_n + b_m)| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$. Similarly for multiplication: $lim(a_n b_m) = ab$ because for any $\epsilon > 0$, there is $M \in \mathbb{R}$ large enough so that for $m, n > M$, $|a - a_n| < \sqrt{\epsilon}$ and $|b - b_m| < \sqrt{\epsilon}$, which implies that:

$$
\begin{aligned}
|(a - a_n)(b - b_m)| &< \epsilon \\
|ab + a_n b_m - a_n b - a b_m| &< \epsilon \\
|ab + a_n b_m - a_n b - a b_m + (b_m(a - a_n) + a_n(b - b_m))| &< \epsilon + |b_m(a - a_n) + a_n(b - b_m)| \\
|ab - a_n b_m| &< \epsilon + |b_m(a - a_n) + a_n(b - b_m)|
\end{aligned}
$$

Where the coefficients of a_n and b_m are both less than the arbitrary epsilon, so the whole right hand side tends to zero with epsilon, showing that $lim(a_n b_m) = ab$. We've now established that convergent sequences work according to arithmetic rules.

Definition 3. ***Cauchy sequence*** *A sequence (a_n) is called a Cauchy sequence, if for any $\epsilon > 0$, there exists a natural number N such that for any $m, n > N$, $|a_n - a_m| < \epsilon$.*

So the elements of a Cauchy sequence get closer and closer to each other, but do not necessarily converge to a definite limit. Notice that every convergent sequence is a Cauchy sequence, because the distance between its elements keeps getting smaller, it can be set to be less than any arbitrary value, but not every Cauchy sequence is convergent, as we saw in the last section with the infinite sequences in a set that's bounded above, which were necessarily Cauchy, but did not converge to a limit. Notice also that the Cauchy sequences of the rational numbers form a field, this is easy to see since instead of adding and multiplying single rational numbers, we're adding and multiplying the corresponding elements of two sequences. With the help of the following definition, we're well on our way to constructing the real numbers.

Definition 4. ***Equivalence relation*** *A relation $\sim$ is called an equivalence relation on a set A if for all $a, b, c \in A$:*

1. *$a \sim b$ implies $b \sim a$*
2. *$a \sim a$*
3. *if $a \sim b$ and $b \sim c$, then $a \sim c$*

The set of all elements of A that are equal to a under the relation $\sim$ is called the equivalence class of a under $\sim$

Equivalence relations are useful because they can be used to partition sets by identifying subsets into elements, and therefore allow one to focus just on some particular properties. For example, we could define the equivalence relation $\sim$: $a \sim a + 2$ on the integers, with the result being the integers modulo 2, which separates the odd and even numbers.

4 Constructing the real numbers as a completion of the rationals

We now have all the definitions we need to construct the real numbers. We've seen that the rational numbers form an ordered field, but fail the completeness criterion, so to construct the real numbers, we need to make up for this shortcoming. This is exactly what we shall do in this section

Define two Cauchy sequences, $(a_n), (b_n)$ to be equal if for every $\epsilon > 0$, there exists a natural number N such that $|a_n - b_m| < \epsilon$ for all $n, m > N$. That is, if they eventually get arbitrarily close to each other. This is denoted with the usual equality sign: $(a_n) = (b_n)$. This forms an equivalence relation, because $|a_n - b_n| = |b_n - a_n|$, so $(a_n) = (b_n)$ implies $(b_n) = (a_n)$; obviously $a_n - a_n = 0$ for all n, so $(a_n) = (a_n)$. And if $|a_n - b_m| < \epsilon$ and $|b_m - c_k| < \epsilon$, then $|a_n - c_k| \leq |(a_n - b_m) + (b_m - c_k)| < 2\epsilon$ and because ϵ is arbitrary, the last term can be made arbitrarily small and it follows that $(a_n) = (c_n)$

Because this forms an equivalence relation for Cauchy sequences, we can construct its equivalence classes, and then form the quotient of the field of Cauchy sequences of rationals by these classes. What this means is that we identify all Cauchy sequences of rational numbers that eventually become arbitrarily close to each other and consider the resulting field of equivalence classes. Call the resulting field of equivalence classes $\mathbb{R}$. We will then have an element 5 corresponding to the equivalence class of the constant sequence 5, 5, 5, 5... and an element π corresponding to the set of sequences equal to 3, 3.1, 3.14...

This construction actually corresponds to the familiar real numbers, as we see below.

4.1 Verification of the axioms

Remember that the Cauchy sequences of rationals form a field themselves. This carries on to $\mathbb{R}$ by the arithmetic properties of limits we discussed in the previous section. What's left then are the axioms of order and completeness.

The ordering for the field $\mathbb{R}$ that we've constructed follows quite directly too. We say that $(a_n) < (b_n)$ if $(a_n - b_n) < 0$, i.e. their difference is equal to a sequence of negative numbers. For example, $3 < \pi$, because the equivalence class of the sequence $(3 - \pi)$ tends to the equivalence class of -0.1415... Obviously then any two elements of $\mathbb{R}$ are either equal or one of them is less than the other. From now on we will use single letters to signify elements of $\mathbb{R}$ even though they still represent equivalence classes of Cauchy sequences of rationals. If x, y, z are elements of $\mathbb{R}$ and $x < y$ and $y < z$ then $x - z = (x - y) + (y - z) < 0$. From the arithmetic that we've already established it also follows that if $y < z$ then $x + y < x + z$ and that if $0 < x$ and $0 < y$, then $0 < x \cdot y$

The completeness of $\mathbb{R}$ is a bit trickier to prove. Let $A \subset \mathbb{R}$ be a nonempty set of the form $A = \{x | x_0 \leq x \leq x_1\}$ for $x_0, x_1 \in \mathbb{R}$. Choose a rational number l_1 as an upper bound of A, and another rational number b_1 such that $b_1 < a$ for some $a \in A$. Set m_1 to be the middlepoint $m_1 = \frac{l_1 + b_1}{2}$. If m_1 in an upper bound for A, then set $l_2 = m_1$ and $b_2 = b_1$, if not, set $b_2 = m_1$ and $l_2 = l_1$. Then calculate $m_2 = \frac{l_2 + b_2}{2}$. By iterating this process, we arrive at a sequence (l_n) becoming ever closer to the boundary of A. (l_n) is always an upper bound for A while (b_n) never is, but the difference $l_n - b_n$ goes eventually to zero so the two sequences converge to the same limit l, which is a least upper bound for A.

5 Dedekind Cuts

There are also other methods of constructing the real numbers than the Cauchy completion method that we've used and the axiomatic one. One particularly widespread method is that of Dedekind cuts. The full construction of the reals by Dedekind cuts is quite long so we will only outline the main ideas of the construction, as knowing the main idea can be quite relevant to understanding the construction of the reals. There are also other, more arcane, methods of constructing the reals, such as A'Campo's construction given in [2], that uses functions between the integers to construct the reals, but we will not discuss those here.

The full construction of the real numbers by Dedekind cuts can be found for example in the appendix of chapter 1 in [3]. With this method, real numbers are defined as so-called cuts. A cut is a nonempty proper subset of the rationals that doesn't have a greatest element, and if a rational number p belongs to the cut, then all rational numbers $q < p$ also belong to it. A real number is therefore seen as the set of all rational numbers that are smaller than it. The sum of two cuts is the cut produced by the sums of all the elements in the two cuts. Similarly for the product of two cuts. While the principle of the Dedekind cut construction is easy to state, checking the axioms of the real number system can be quite cumbersome, so we will not go into that here. At it's heart, this approach is similar to the Cauchy completion we detailed earlier; it too considers sets of rationals without least upper bounds, and defines these sets to be that upper bound. Dedekind cuts use the sets of rationals less than a real number, while the Cauchy completion method uses the Cauchy sequences which tend towards it.

6 Conclusion

We have now seen how the real numbers can be constructed from the rational numbers, through the process of Cauchy completion. This method of construction is probably the easiest to motivate, though Dedekind cuts have their advantages too. It is easy to show how the rational numbers fail the least upper bound property and introduce Cauchy sequences and the Cauchy completion method through that. We basically compared the structureal differences of the fields of rational and real numbers and through that manipulated the rationals to create the real numbers. Dedekind cuts don't require the additional theory of Cauchy sequences and equivalence relations, a real number is determined by a single cut, rather than a whole equivalence class of different sequences. On the other hand, Dedekind cuts can be harder to motivate, as it is not immediately clear why one should consider the set of rational numbers less than a real number as its defining feature.

References

[1] Gelbaum & Olmsted. *Counterexamples in analysis*, Dover 2003

[2] Norbert A'Campo. "A natural construction for the real numbers" arXiv:math/0301015 [math.GN]

[3] W. Rudin. *Principles of mathematical analysis*, McGraw-Hill , 3rd edition 1976

[4] M. Spivak, *Calculus*, Cambridge University Press, 3rd edition 1994