

Lecture 2: The real numbers

The purpose of this lecture is for us to develop the real number system. This might seem like a very strange thing for us to be doing. It must seem to you that you have been studying real numbers most of your life. However, some introspection is likely to reveal that not everything you have been told about the real numbers is entirely believable. (As an example, a recent 7th grade textbook explains that to add and multiply rational numbers, you should follow a set of rules you have been given. To add and multiply reals, you should plug them into your calculator.)

Because the real numbers will be the central focus of inquiry in this course, we will take this moment to specify exactly what they are. The central tenet of mathematics is that one must always tell the truth, and one can't be sure that one is doing this about real numbers, unless one is sure exactly what they are. There is more than one possible approach to doing this. Most mathematicians' treatment of this (see Apostol's book, or Dinakar Ramakrishnan's notes) focus on what one expects to do with real numbers. One is given a set of axioms to cover this. It should be possible to perform basic arithmetic on the reals (the field axioms), there should be a way of comparing the size of two real numbers (the total ordering axiom) and a lot of limits should exist (the least upper bound property). After one has written down these axioms, one is in a good position to start proving theorems about the real numbers. But one might be quite confused about what is going on? Are these really the same real numbers I've always heard about? Are the real numbers the only set of numbers satisfying these axioms? What are individual real numbers like? It is possible with some work to proceed from these axioms to answer that question, but the work is non-trivial.

We will take a slightly different approach. We will describe the real numbers in much the way they were described to you in grade school, as decimal expansions. (Mathematicians tend not to like this because there are arbitrary choices like the choice of the base ten.) Then because we'd like to use the real numbers, we will check that they satisfy the axioms allowing us to order them, take limits, and do arithmetic. It will turn out that doing arithmetic is the hardest part. (There's a reason you need a calculator!) While you have been trained that one can do arithmetic in real numbers since long before you had Calculus, in order to actually be able to perform any arithmetic operation on general real numbers, you have to take limits.

Before we start, perhaps a few words are required about the usefulness of this. You are right to be concerned. As scientists and engineers pursuing practical objectives, you will not encounter any typical real number. Sure, you might collect some data. But it will come to you as floating point numbers with implicit error intervals. Why then should we study something so abstract, so idealized, dare I say it so unreal as the real numbers? The answer is that quite happily, the processes which we use to draw conclusions about real numbers and especially to study limits of them (the main subject of this course) are exactly the same as those used to rigorously study floating point numbers with error intervals. It

might be wise to take this viewpoint about the whole course. But this requires thinking differently than one is used to about what are the main questions.

Now we begin formally. What is a real number? It will be an expression of the form

$$\pm a_1 a_2 a_3 \dots a_m . b_1 b_2 b_3 \dots$$

Here the \pm represents a choice between plus and minus. The digit a_1 is an integer between 0 and 9 inclusive (unless m is different from 1 in which case it is restricted to being between 1 and 9, since it is the leading digit.) All other digits are integers between 0 and 9 inclusive. One detail is that some real number have two such representations. Namely a terminating decimal

$$\pm a_1 a_2 \dots a_m . b_1 b_2 \dots b_n 000 \dots,$$

where here b_n is different from 0, is the same as

$$\pm a_1 a_2 \dots a_m . b_1 b_2 \dots (b_n - 1) 999 \dots,$$

a decimal with repeating 9's. (Note that the repeating 9's could start to the left of the decimal place just as well as to the right.) The set of real numbers, we will invariably refer to as \mathbf{R} . Hopefully, we have now described the real numbers as you have seen them since grade school. It is often pointed out that they can be visualized as populating a line. You can do this by first marking of the integers at equal distances on the line. Then the interval between any two consecutive integers can be cut into ten equal subintervals. The value of the first digit after the decimal describes which interval the real number lies in. One continues the process, subdividing each of those ten intervals into ten equal subintervals and so on.

When dealing with the real numbers in practice, we very often approximate to a few decimal places. Strangely, there is no standard notation for this, so we introduce some. Given a real number

$$x = \pm a_1 a_2 \dots a_m . b_1 b_2 \dots b_n b_{n+1} b_{n+2} \dots,$$

we define $t_n(x)$, the truncation to n decimal places, as

$$t_n(x) = \pm a_1 a_2 \dots a_m . b_1 b_2 \dots b_n.$$

In order for t_n to be a well defined function on the reals, we must specify how it acts on reals with two decimal representations (the case of repeating zeroes and repeating nines). We specify that to apply t_n , we always take the representation with repeating zeroes. Thus given any real number, we uniquely map it with t_n to a terminating decimal, which we can also view as a rational number with denominator 10^n . We note that as n increases with x fixed, the truncation $t_n(x)$ increases.

We are now ready to define inequalities among real numbers. Given two real numbers x and y , we say that $x \geq y$ if $x = y$ or there is some n for which $t_n(x) > t_n(y)$. (Ask yourself

why we need the inequality between the truncations to be strict.) When presented with a new definition, it is often valuable to think about it in terms of algorithms. How do we check if the number x is greater than or equal to the number y . If x is actually greater, then we'll find out in a finite number of steps as we find a decimal place n where the truncation of x is actually bigger. If x and y are equal, we'll never find out, because we have to check all the truncations. While at first, this seems an unhappy state of affairs, it actually agrees with our intuition about approximations and error intervals. If two approximations are far apart so that their error intervals are disjoint, we can tell which one is bigger. Otherwise, we're not sure. Already, we see that in this way, that the real numbers which are an idealization, model reality well.

We now state as a proposition, that any two real numbers can be ordered.

Proposition 1 Given two real numbers x and y , then $x \geq y$ or $y \geq x$.

Proof: If for some n , we have $t_n(x) > t_n(y)$ or $t_n(x) < t_n(y)$, then we're done. The only case remaining is that $t_n(x) = t_n(y)$ for all n . In this case, x and y have the same decimal expansion and are therefore the same number. In this case, both $x \geq y$ and $y \geq x$ hold.

Thus we have completed one third of our project for defining the real numbers. They are ordered. Decimals are in fact quite helpful in the ordering which is basically alphabetical. (A more technical term for this kind of ordering is lexicographic.)

We are now prepared to establish the least upper bound property for the real numbers. Given a set A of real numbers, we say that a real number x is an upper bound for A if for every $y \in A$, we have that $x \geq y$. We say that x is the least upper bound for A if for every other upper bound z for A , we have that $x \leq z$. We are interested in least upper bounds as a kind of upper limit of the real numbers in the set A . An upper bound might miss being in A by a great deal. A least upper bound is just outside of A .

Proposition 2 Any nonempty set of real numbers A which has a real upper bound, has a least upper bound in the reals.

Proof We are given that A is nonempty. Let z be an element of it. We are given that it has an upper bound y . Now we are going to find the least upper bound x by constructing its decimal expansion. Since y is an upper bound, so is $a = t_n(y) + \frac{1}{10^n}$. The number a is an upper bound which also has a decimal expansion which terminates at the n th place. Moreover $a > t_n(z)$. In fact $10^n(a - t_n(z))$ is a positive natural number. We let B be the set of all natural numbers of the form $10^n(c - t_n(z))$ with c an upper bound for A with decimal expansion terminating at or before the n th place. This set B is a nonempty set of natural numbers which serves as a proxy for the set of upper bounds for A which terminate at n decimal places. To the set B , we may apply the Well Ordering Principle which we proved in the first lecture. The set B has a smallest element, b . Thus $10^{-n}b + t_n(z)$ is the smallest upper bound for A with an n -place decimal expansion. We

define $x_n = 10^{-n}b + t_n(z) - 10^{-n}$. Thus x_n just misses being an upper bound. If after some finite n , all x_m with $m > n$ are the same, we let x be this x_m . Otherwise, we let x be the real number so that $t_n(x) = x_n$. (We used the well ordering principle to construct the decimal expansion for x . (Question for the reader: Why did we treat the case of x with a terminating expansion separately. Hint: it was because of our definition for t_n .)

The above proof may be a little hard to digest. To understand it better, let us consider an example. Often, it is touted that one of the virtues of the real number system is that it contains $\sqrt{2}$. (You should be a little concerned that we haven't defined multiplication yet, but this example can be viewed as motivation for the definition.) How do we see that the real numbers contain $\sqrt{2}$? We find a least upper bound for all numbers whose square is less than 2. We do this in the spirit of the above proof. First, we find the small number with one decimal place whose square is more than 2. It is 1.5. We subtract .1 and record 1.4. Then we find the smallest number with two decimal places whose square is larger than 2. It is 1.42. We subtract .01 and record 1.41. Gradually, we build up the decimal expansion for $\sqrt{2}$, which begins 1.414213562. Our algorithm never terminates but we get an arbitrarily long decimal expansion with a finite number of steps.

The least upper bound property, while it is easy to prove using the decimal system, is a pretty sophisticated piece of mathematics. It is a rudimentary tool for taking limits, something we don't consider in school until we take Calculus. Adding and multiplying, though, is one of the first things we think of doing to real numbers. Perhaps we want a calculator to handle it but we imagine that nothing fancier is going on than our usual algorithms for adding and multiplying. Let's consider how this works. Let's say I want to add two typical real numbers. I write out their decimal expansions one above the other. Then I start at the right. Oops. The numbers have infinite decimal expansions, so I can never get to the right. This problem is not easily waved away. Through the process of carrying, quite insignificant digits of the summands can affect quite significant digits of the sum. In order to calculate, as a practical matter, an arithmetic operation performed on two numbers, we have to take a limit.

Luckily, we have established the least upper bound property. We can use it to define the arithmetic operations on the reals. Let x and y be two nonnegative real numbers. We let $A = \{t_n(x) + t_n(y)\}$ be the set of sums of truncations of x and y . We let $M = \{t_n(x)t_n(y)\}$ be the set of products of truncations of x and y . Note that both sets have upper bounds. (We can use $t_n(x) + t_n(y) + \frac{2}{10^n}$ as an upper bound for A (why?) and $(t_n(x) + \frac{1}{10^n})(t_n(y) + \frac{1}{10^n})$ as an upper bound for M . (Why?) Now we apply the least upper bound property to see that A and M have least upper bounds. We define $x + y$ to be the least upper bound for A and xy to be the least upper bound for M . We restricted to x and y positive, so that the expressions $t_n(x) + t_n(y)$ and $t_n(x)t_n(y)$ are increasing in n so that the least upper bounds are really what we want.

Since we have so far only defined addition and multiplication for positive numbers, defining subtraction of positive numbers seems a high priority. Again given x and y nonnegative real numbers. We define $S = \{t_n(x) - t_n(y) - \frac{1}{10^n}\}$. We subtracted $\frac{1}{10^n}$

from the n th element so that while we are replacing x by an underestimate $t_n(x)$, we are replacing y by an overestimate $t_n(y) + \frac{1}{10^n}$ and when we subtract, we have an underestimate for the difference. We define $x - y$ to be the least upper bound of S .

What about division? Let x and y be nonnegative real numbers. Let $D = \{z : x \geq yz\}$. Thus D consists of real numbers we can multiply by y to get less than x . These are the underestimates of the quotient. We define $\frac{x}{y}$ to be the least upper bound of D .

So how are we doing? We have defined the real numbers in a way that we recognize them from grade school. We have shown that this set of real numbers has an order, that it satisfies the least upper bound property and that we may perform arithmetic operations. Mathematicians might still not be entirely satisfied as these arithmetic operations still must be proven to satisfy the laws they should inherit from the rational numbers. This is not quite as easy as it looks. For instance, let's say we want to prove the distributive law. Thus if x, y , and z are nonnegative real numbers, we would like to show that $\{t_n(x + y)t_n(z)\}$ has the same least upper bound as $\{t_n(xz) + t_n(yz)\}$. It is true and it can be done. But to do it, it really helps to deal carefully with something we have completely set aside thus far. It helps to have estimates on how far away a truncated version of $(x + y)z$ actually is from the least upper bound.

This gets at an objection that a practical person could have for the way we've defined our operations thus far. Certainly, the least upper bounds exist. But they are the output of an algorithm that never terminates. To actually use real numbers as a stand in for approximations with error intervals, we need to be able at each step of a never terminating algorithm to have control on the error. Notice we did have that kind of control in the example with $\sqrt{2}$. When we had n decimal places, we knew we were within 10^{-n} of the answer. In the next lecture, we will get at both the practical and theoretical versions of this problem by introducing the definition of the limit. We will see that understanding that a limit exists is more than knowing what the limit is. It also involves estimating how fast the limit converges. In practical terms, this means calculating an error interval around the limitand.