Math 131AH - Week 2

Textbook pages: 1-11, 24-25. (Optional additional reading: 17-21, 52-54).

Topics covered:

- Some operations on the rationals
- Gaps in the rational numbers
- Cauchy sequences (of rationals)
- The real numbers
- Ordering and the reals
- The least upper bound property
- Cardinality of sets

* * * * *

Review of last week's notes

- In the last week's notes, we began the systematic and rigorous construction of our familiar number systems: the natural number system **N**, the integers **Z**, the rationals **Q**, and the real numbers **R**. (Incidentally, **N**, **Q**, and **R** stand for "natural", "quotient", and "real" respectively. **Z** stands for "Zahlen", the German word for number.) We defined the natural numbers using the five Peano axioms, and postulated that such a number system existed; this is intuitively reasonable, since the natural numbers correspond to the very intuitive and fundamental notion of *sequential counting*. Using that number system one could then recursively define addition and multiplication, and verify that they obeyed the usual laws of algebra. We then constructed the integers by taking formal differences of the natural numbers, $a - -b$. (*Formal* means "having the form of"; at the beginning of our construction the expression $a - -b$ did not actually *mean* the difference $a - b$, since the symbol $--$ was meaningless. It only had the *form* of a difference. Later on we defined subtraction and verified that the formal difference was equal to the actual difference, so this eventually became a non-issue, and our symbol for formal differencing was discarded). We then constructed the rationals by taking formal quotients of the integers, $a//b$, although we need to exclude division by zero in order to keep the laws of algebra reasonable. (You are of course free to design your own number system, possibly including one where division by zero is permitted; but you will have to give up one or more of the field axioms from last week's notes, among other things, and you will probably get a less useful number system in which to do any real-world problems).

- The rational system is already sufficient to do a lot of mathematics - much of high school algebra, for instance, works just fine if one only knows about the rationals. However, there is one important, and basic, area of mathematics where the rational number system does not suffice - that of *geometry*. For instance, a right-angled triangle with both sides equal to 1 gives a hypotenuse of $\sqrt{2}$, which we will see is an irrational number. Things get even worse when one starts to deal with the subfield of geometry known as *trigonometry*, when one sees numbers such as π or $\cos(1)$, which turn out to be in some sense "even more" irrational than $\sqrt{2}$. (These numbers are known as *transcendental numbers*, but to discuss this further would be far beyond the scope of this course). Thus, in order to have a number system which can adequately describe geometry - or even something as simple as measuring lengths on a line - one needs to replace the rational number system with the real number system. Since differential and integral calculus is also intimately tied up with geometry - think of slopes of tangents, or areas under a curve - calculus also requires the real number system in order to function properly.

- However, a rigorous way to construct the reals from the rationals turns out to be somewhat difficult - requiring a bit more machinery than what was needed to pass from the naturals to the integers, or the integers to the rationals. In those two constructions, the task was to introduce one more *algebraic* operation to the number system - e.g. one can get integers from naturals by introducing subtraction, and get the rationals from the integers by introducing division. But to get the reals from the rationals is to pass from a "discrete" system to a "continuous" one, and requires the introduction of a somewhat different notion - that of a *limit*. The limit is a concept which on one level is quite intuitive, but to pin down rigorously turns out to be quite difficult. (Even such great mathematicians as Euler and Newton had difficulty with this concept. It was only in the nineteenth century when mathematicians such as Cauchy and Dedekind figured out how to deal with limits rigorously.)

- The purpose of this weeks notes is to explore the "gaps" in the rational numbers, and how to fill them in using limits to create the real numbers. The real number system will end up being a lot like the rational numbers, but will have some new operations - notably that of *supremum*, which can then be used to define limits and thence to everything else that calculus needs.

* * * * *

Some operations on the rationals

- We have already introduced the four basic arithmetic operations of addition, subtraction, multiplication, and division on the rationals. (Recall that subtraction and division came from the more primitive notions of negation and reciprocal by the formulae $x - y := x + (-y)$ and $x/y := x \times y^{-1}$). We also have a notion of order $<$, and have organized the rationals into the positive rationals, the negative rationals, and zero. In short, we have shown that the rationals $\mathbf{Q}$ form an *ordered field*.

- One can now use these basic operations to construct more operations. There are many such operations we can construct, but we shall just introduce two particularly useful ones: absolute value and exponentiation.

- **Definition** If x is a rational number, the *absolute value* $|x|$ of x is defined as follows. If x is positive, then $|x| := x$. If x is negative, then $|x| := -x$. If x is zero, then $|x| := 0$.

- Let x and y be real numbers. The quantity $|x-y|$ is called the *distance between x and y* and is sometimes denoted $d(x, y)$, thus $d(x, y) := |x - y|$. For instance, $d(3, 5) = 2$.

- The basic properties of absolute value are the following:

- **Proposition 1.** Let x, y be rational numbers. Then $|x|$ is either positive or zero, and $|x| = 0$ if and only if x is 0. Also, we have the *triangle inequality*
$$|x + y| \leq |x| + |y|,$$
the bounds
$$-|x| \leq x \leq |x|,$$
and the multiplicative identity
$$|xy| = |x|\ |y|.$$
In particular we have
$$|-x| = |x|.$$
The distance $d(x, y)$ obeys the following properties. If x, y, z are rational numbers, then $d(x, y) = 0$ if and only if $x = y$. Also, we have the symmetry property $d(x, y) = d(y, x)$ and the triangle inequality
$$d(x, z) \leq d(x, y) + d(y, z).$$

- **Proof.** See Week 2 Homework. $\square$

- Absolute value is useful for measuring how "close" two numbers are. Let us make a somewhat artificial definition:

- **Definition.** Let $\varepsilon > 0$, and x, y be rational numbers. We say that y is *ε-close* to x iff we have $d(y, x) \leq \varepsilon$.

- **Examples.** The numbers 0.99 and 1.01 are 0.1-close, but they are not 0.01 close, because $d(0.99, 1.01) = |0.99 - 1.01| = 0.02$ is larger than 0.01. The numbers 2 and 2 are ε-close for every positive ε.

- We do not bother defining a notion of ε-close when ε is zero or negative, because if ε is zero then x and y are only ε-close when they are equal, and when ε is negative then x and y are never ε-close. (In any event it is a long-standing tradition in analysis that the Greek letters ε, δ should only denote positive (and probably small) numbers).

- Note: This definition is not standard in mathematics textbooks; I will use it as "scaffolding" to construct the more important notions of limits (and of Cauchy sequences) later on, and once we have those more advanced notions we will discard the notion of ε-close.

- Some basic properties of ε-closeness are the following.

- **Proposition 2.** Let x, y, z, w be rational numbers.

- (a) If $x = y$, then x is ε-close to y for every $\varepsilon > 0$. Conversely, if x is ε-close to y for every $\varepsilon > 0$, then we have $x = y$.

- (b) Let $\varepsilon > 0$. If x is ε-close to y, then y is ε-close to x.

- (c) Let $\varepsilon, \delta > 0$. If x is ε-close to y, and y is δ-close to z, then x and z are $(\varepsilon + \delta)$-close.

- (d) Let $\varepsilon, \delta > 0$. If x and y are ε-close, and z and w are δ-close, then $x+z$ and $y+w$ are $(\varepsilon+\delta)$-close, and $x-z$ and $y-w$ are also $(\varepsilon+\delta)$-close.

- (e) Let $\varepsilon > 0$. If x and y are ε-close, they are also ε'-close for every $\varepsilon' > \varepsilon$.

- (f) Let $\varepsilon > 0$. If y and z are both ε-close to x, and w is between y and z (i.e. $y \leq w \leq z$ or $z \leq w \leq y$), then w is also ε-close to x.

- (g) Let $\varepsilon > 0$. If x and y are ε-close, and z is non-zero, then xz and yz are $\varepsilon|z|$-close.

- (h) Let $\varepsilon, \delta > 0$. If x and y are ε-close, and z and w are δ-close, then xz and yw are $(\varepsilon|z| + \delta|x| + \varepsilon\delta)$-close.

- **Proof.** We only prove the most difficult one, (g); we leave (a)-(f) to the exercises. Let $\varepsilon, \delta > 0$, and suppose that x and y are ε-close. If we write $a := y - x$, then we have $y = x + a$ and that $|a| \leq \varepsilon$. Similarly, if z and w are δ-close, and we define $b := w - z$, then $w = z + b$ and $|b| \leq \delta$.

  Since $y = x + a$ and $w = z + b$, we have

  $$yw = (x + a)(z + b) = xz + az + xb + ab.$$

Thus

$$|yw - xz| = |az + bx + ab| \leq |az| + |bx| + |ab| = |a||z| + |b||x| + |a||b|.$$

Since $|a| \leq \varepsilon$ and $|b| \leq \delta$, we thus have

$$|yw - xz| \leq \varepsilon|z| + \delta|x| + \varepsilon\delta$$

and thus that yw and xz are $(\varepsilon|z| + \delta|x| + \varepsilon\delta)$-close. □

- One should compare statements (a)-(c) of this Proposition with the first three axioms of equality (see supplemental handout on logic).

- Now we define exponentiation for natural number exponents. Just like addition was recursive incrementation and multiplication was recursive addition, exponentiation is recursive multiplication (at least when the exponent is a natural number)

- **Definition** Let x be a rational number. To raise x to the power 0, we define $x^0 := 1$. Now suppose that recursively that x^n has been defined for some natural number n, then we define $x^{n++} := x^n \times x$.

- Thus for instance $x^1 = x^0 \times x = 1 \times x = x$; $x^2 = x^1 \times x = x \times x$; $x^3 = x^2 \times x = x \times x \times x$; and so forth. This recursive definition defines x^n for all natural numbers n.

- We have the following properties of exponentiation with natural number exponents:

- **Proposition 3.** Let x, y be rational numbers, and let n, m be natural numbers.

- (a) We have $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.

- (b) We have $x^n = 0$ if and only if $x = 0$.

- (c) If $x \geq y \geq 0$, then $x^n \geq y^n \geq 0$.

- (d) We have $|x^n| = |x|^n$.

- **Proof.** These are all easy applications of induction, and are left to the reader. □

- Now we define exponentiation for negative integer exponents.

- **Definition** Let x be a non-zero rational number. Then for any negative integer $-n$, we define $x^{-n} := 1/x^n$.

- Thus for instance $x^{-3} = 1/x^3 = 1/(x \times x \times x)$. We now have x^n defined for any integer n, whether n is positive, negative, or zero.

- Exponentiation with integer exponents has the following properties:

- **Proposition 4.** Let x, y be non-zero rational numbers, and let n, m be integers.

- (a) We have $x^n x^m = x^{n+m}$, $(x^n)^m = x^{nm}$, and $(xy)^n = x^n y^n$.

- (b) If $x \geq y > 0$, then $x^n \geq y^n > 0$ if n is positive, and $0 < x^n \leq y^n$ if n is negative.

- (c) We have $|x^n| = |x|^n$.

- **Proof.** This follows easily from Proposition 3, and a division into cases depending on whether n and m are natural number or negative integers, and is left to the reader. □

* * * * *

Gaps in the rational numbers

- Imagine that we arrange the rationals on a line, arranging x to the right of y if $x > y$. (This is a non-rigorous arrangement, since we have not yet defined the concept of a line, but this discussion is only intended to motivate the more rigorous propositions below). Inside the rationals we have the integers, which are thus also arranged on the line. Now we work out how the rationals are arranged within the integers.

- **Proposition 5.** Let x be a rational number. Then there exists an integer n such that $n \leq x < n + 1$. In fact, this integer is unique (i.e. for each x there is only one n for which $n \leq x < n + 1$). In particular, there exists a natural number N such that $N > x$ (i.e. there is no such thing as a rational number which is larger than all the natural numbers).

- **Proof.** See Week 2 Homework. □

- We also know from last week's notes that between every pair of rational numbers there is another. This seems to give a lot of rational numbers, but it turns out there are still some "gaps". For instance, we will now show that the rational numbers do not contain any square root of two.

- **Proposition 6.** There does not exist any rational number x for which $x^2 = 2$.

- **Proof.** We only give a sketch of a proof; the gaps will be filled in the homework.

- Suppose for contradiction that we had a rational number x for which $x^2 = 2$. Clearly x is not zero. We may assume that x is positive, for if x were negative then we could just replace x by $-x$ (since $x^2 = (-x)^2$). Thus $x = p/q$ for some positive integers p, q, so $(p/q)^2 = 2$, which we can rearrange as $p^2 = 2q^2$.

- Define a natural number p to be *even* if $p = 2k$ for some natural number k, and *odd* if $p = 2k + 1$ for some natural number k. Every natural number is either even or odd, but not both (why? See Homework). If p is odd, then p^2 is also odd (why? See Homework), which contradicts $p^2 = 2q^2$. Thus p is even, i.e. $p = 2k$ for some natural number k. Since p is positive, k must also be positive. Inserting $p = 2k$ into $p^2 = 2q^2$ we obtain $4k^2 = 2q^2$, so that $q^2 = 2k^2$.

- To summarize, we started with a pair (p, q) of positive integers such that $p^2 = 2q^2$, and ended up with a pair (q, k) of positive integers such that $q^2 = 2k^2$. Since $p^2 = 2q^2$, we have $q < p$ (why?). If we rewrite $p' := q$ and $q' := k$, we thus can pass from one solution (p, q) to the equation $p^2 = 2q^2$ to a new solution (p', q') to the same equation which has a smaller value of p. But then we can repeat this procedure again and again, obtaining a sequence (p'', q''), (p''', q'''), etc. of solutions to $p^2 = 2q^2$, each one with a smaller value of p than the previous, and each one consisting of positive integers. But this contradicts the principle of infinite descent (see Homework). This contradiction shows that we could not have had a rational x for which $x^2 = 2$. □

- On the other hand, we can get rational numbers which are arbitrarily close to a square root of 2:

- **Proposition 7.** For every rational number $\varepsilon > 0$, there exists a non-negative rational number x such that $x^2 < 2 < (x+\varepsilon)^2$.

- **Proof.** Let $\varepsilon > 0$ be rational. Suppose for contradiction that there is no non-negative rational number x for which $x^2 < 2 < (x+\varepsilon)^2$. This means that whenever x is non-negative and $x^2 < 2$, we must also have $(x+\varepsilon)^2 < 2$ (note that $(x+\varepsilon)^2$ cannot equal 2, by Proposition 6). Since $0^2 < 2$, we thus have $\varepsilon^2 < 2$, which then implies $(2\varepsilon)^2 < 2$, and indeed a simple induction shows that $(n\varepsilon)^2 < 2$ for every natural number n. (Note that $n\varepsilon$ is non-negative for every natural number n - why?). But, by Proposition 5 we can find an integer n such that $n > 2/\varepsilon$, which implies that $n\varepsilon > 2$, which implies that $(n\varepsilon)^2 > 4 > 2$, contradicting the claim that $(n\varepsilon)^2 < 2$ for all natural numbers n. This contradiction gives the proof. $\square$

- For example, if $\varepsilon = 0.001$, we can take $x = 1.414$, since $x^2 = 1.999396$ and $(x+\varepsilon)^2 = 2.002225$. (Incidentally, we will use the decimal system for defining terminating decimals, for instance 1.414 is defined to equal the rational number 1414/1000. We defer more discussion on the decimal system to a supplemental handout).

- This Proposition indicates that, while the set **Q** of rationals do not actually have $\sqrt{2}$ as a member, we can get as close as we wish to $\sqrt{2}$. For instance, the sequence of rationals

$$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

  seem to get closer and closer to $\sqrt{2}$, as their squares indicate:

$$1.96, 1.9881, 1.99396, 1.99996164, 1.9999899241, \ldots$$

  Thus it seems that we can create a square root of 2 by taking a "limit" of a sequence of rationals. This is how we shall construct the reals. (There is another way to do so, using something called "Dedekind cuts"; this is the approach pursued in the textbook, but this is optional reading material for the course. One can also proceed using infinite decimal expansions, but there are some sticky issues when doing so, e.g. one has to make $0.999\ldots$ equal to $1.000\ldots$, and this approach, despite being the most familiar, is actually *more* complicated than other approaches).

* * * * *

Cauchy sequences

- Let us first define the concept of a sequence. Let m be an integer. A *sequence* $(a_n)_{n=m}^{\infty}$ *of rational numbers* is any function from the set $\{n \in \mathbf{Z} : n \geq m\}$ to $\mathbf{Q}$, i.e. a mapping which assigns to each integer n greater than or equal to m, a rational number a_n. More informally, a sequence $(a_n)_{n=m}^{\infty}$ of rational numbers is a collection of rationals a_n, a_{n+1}, $a_{n+2}, \ldots$.

- For instance, the sequence $(n^2)_{n=0}^{\infty}$ is the collection $0, 1, 4, 9, \ldots$ of natural numbers; the sequence $(3)_{n=0}^{\infty}$ is the collection $3, 3, 3, \ldots$ of natural numbers. These sequences are indexed starting from 0, but we can of course make sequences starting from 1 or any other number; for instance, the sequence $(a_n)_{n=3}^{\infty}$ denotes the sequence $a_3, a_4, a_5, \ldots$, so $(n^2)_{n=3}^{\infty}$ is the collection $9, 16, 25, \ldots$ of natural numbers.

- We want to define the real numbers as the limits of sequences of rational numbers. To do so, we have to distinguish which sequences of rationals are convergent and which ones are not. For instance, the sequence

  $$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

  looks like it is trying to converge to something, as does

  $$0.1, 0.01, 0.001, 0.0001, \ldots$$

  while other sequences such as

  $$1, 2, 4, 8, 16, \ldots$$

  or

  $$1, 0, 1, 0, 1, \ldots$$

  do not.

- To do this we use the definition of ε-closeness defined earlier. Recall that two rational numbers x, y are ε-close if $d(x, y) = |x - y| < \varepsilon$.

- **Definition.** Let $\varepsilon > 0$. A sequence $(a_n)_{n=0}^{\infty}$ is said to be *ε-steady* iff each pair a_j, a_k of sequence elements is ε-close for every natural number j, k. In other words, the sequence $a_0, a_1, a_2, \ldots$ is ε-steady iff $d(a_j, a_k) \leq \varepsilon$ for all j, k.

- (Again, this notion of ε-steadiness is is a temporary notion, which we will not use in later weeks notes).

- This definition is for sequences whose index starts at 0, but clearly we can make a similar notion for sequences whose indices start from any other number: a sequence $a_N, a_{N+1}, \ldots$ is ε-steady if one has $d(a_j, a_k) \leq \varepsilon$ for all $j, k \geq N$.

- **Example.** The sequence $1, 0, 1, 0, 1, \ldots$ is 1-steady, but is not 1/2-steady. The sequence $0.1, 0.01, 0.001, 0.0001, \ldots$ is 0.1-steady, but is not 0.01-steady (why?). The sequence $1, 2, 4, 8, 16, \ldots$ is not ε-steady for any ε (why?). The sequence $2, 2, 2, 2, \ldots$ is ε-steady for every $\varepsilon > 0$.

- The notion of ε-steadiness of a sequence is simple, but does not really capture the *limiting* behavior of a sequence, because it is too sensitive to the initial members of the sequence. For instance, the sequence

  $$10, 0, 0, 0, 0, 0, \ldots$$

  is 10-steady, but is not ε-steady for any smaller value of ε, despite the sequence converging almost immediately to zero. So we need a more robust notion of steadiness that does not care about the initial members of a sequence.

- **Definition.** Let $\varepsilon > 0$. A sequence $(a_n)_{n=0}^{\infty}$ is said to be *eventually ε-steady* iff the sequence $a_N, a_{N+1}, a_{N+2}, \ldots$ is ε-steady for some natural number $N \geq 0$. In other words, the sequence $a_0, a_1, a_2, \ldots$ is eventually ε-steady iff there exists an $N \geq 0$ such that $d(a_j, a_k) \leq \varepsilon$ for all $j, k \geq N$.

- **Example.** The sequence $a_1, a_2, \ldots$ defined by $a_n := 1/n$, (i.e. the sequence $1, 1/2, 1/3, 1/4, \ldots$) is not 0.1-steady, but is eventually 0.1-steady, because the sequence $a_{10}, a_{11}, a_{12}, \ldots$ (i.e. $1/10, 1/11, 1/12, \ldots$) is 0.1-steady. The sequence $10, 0, 0, 0, 0, \ldots$ is not ε-steady for any ε less than 10, but it is eventually ε-steady for every $\varepsilon > 0$ (why?).

- Now we can finally define the correct notion of what it means for a sequence of rationals to "want" to converge. (As I mentioned before, the notion of ε-steadiness is just scaffolding).

- **Definition.** A sequence $(a_n)_{n=0}^{\infty}$ of rational numbers is said to be a *Cauchy sequence* iff for every rational $\varepsilon > 0$, the sequence $(a_n)_{n=0}^{\infty}$ is eventually ε-steady. In other words, the sequence $a_0, a_1, a_2, \ldots$ is a Cauchy sequence iff for every $\varepsilon > 0$, there exists an $N \geq 0$ such that $d(a_j, a_k) \leq \varepsilon$ for all $j, k \geq N$.

- **Informal example.** Consider the sequence $1.4, 1.41, 1.414, \ldots$ mentioned earlier. This sequence is already 1-steady. If one discards the first element, i.e. $1.41, 1.414, 1.4142, \ldots$ then it is now 0.1-steady, which means that the original sequence was eventually 0.1-steady. Discarding the next element gives $1.414, 1.4142, \ldots$ which is 0.01-steady; thus the original sequence was eventually 0.01-steady. Continuing in this way it seems plausible that this sequence is in fact ε-steady for every $\varepsilon > 0$, which seems to suggest that this is a Cauchy sequence. However, this discussion is not rigorous for several reasons, for instance I have not precisely defined what this sequence $1.4, 1.41, 1.414, \ldots$ really is. An example of a rigorous treatment follows next.

- **Claim.** The sequence $a_1, a_2, a_3, \ldots$ defined by $a_n := 1/n$ (i.e. the sequence $1, 1/2, 1/3, \ldots$) is a Cauchy sequence.

- **Proof.** We have to show that for every $\varepsilon > 0$, the sequence $a_1, a_2, \ldots$ is eventually ε-steady. So let $\varepsilon > 0$ be arbitrary. We now have to find a number $N \geq 1$ such that the sequence $a_N, a_{N+1}, \ldots$ is ε-steady. Let us see what this means. This means that $d(a_j, a_k) \leq \varepsilon$ for every $j, k \geq N$, i.e.
$$|1/j - 1/k| \leq \varepsilon \text{ for every } j, k \geq N.$$

Now since $j, k \geq N$, we know that $0 < 1/j, 1/k \leq 1/N$, so that $|1/j - 1/k| \leq 1/N$. So in order to force $|1/j - 1/k|$ to be less than or equal to ε, it would be sufficient for $1/N$ to be less than ε. So all we need to do is choose an N such that $1/N$ is less than ε, or in other words that N is greater than $1/\varepsilon$. But this can be done thanks to Proposition 5. $\square$

- As you can see, verifying from first principles (i.e. without using any of the machinery of limits, etc.) that a sequence is a Cauchy sequence requires some effort, even for a sequence as simple as $1/n$. (And the part about selecting an N can be difficult for beginners - one has to think in reverse, working out what conditions on N would suffice to force the sequence $a_N, a_{N+1}, a_{N+2}, \ldots$ to be ε-steady, and then finding an N which obeys those conditions). Later we will develop the limit laws which allow us to determine when a sequence is Cauchy with more ease.

- We first prove a basic fact about Cauchy sequences: they are bounded.

- **Definition.** Let $M \geq 0$. A finite sequence $a_1, a_2, \ldots, a_n$ is *bounded by* M iff $|a_i| \leq M$ for all $1 \leq i \leq n$. An infinite sequence $(a_n)_{n=1}^\infty$ is *bounded by* M iff $|a_i| \leq M$ for all $i \geq 1$. A sequence is said to be *bounded* iff it is bounded by M for some $M \geq 0$.

- For instance, the finite sequence $1, -2, 3, -4$ is bounded (in this case, it is bounded by 4, or indeed by any M greater than or equal to 4). But the infinite sequence $1, -2, 3, -4, 5, -6, \ldots$ is unbounded. (Can you prove this? Use Proposition 5). The sequence $1, -1, 1, -1, \ldots$ is bounded (e.g. by 1), but not a Cauchy sequence. Note that a sequence is bounded by M if and only if it is M-close to the zero sequence $0, 0, 0, \ldots$ (why?).

- **Lemma 8**. Every finite sequence $a_1, a_2, \ldots, a_n$ is bounded.

- **Proof.** We prove this by induction on n. When $n = 1$ the sequence a_1 is clearly bounded, for if we choose $M := |a_1|$ then clearly we have $|a_i| \leq M$ for all $1 \leq i \leq n$. Now suppose that we have already proved the lemma for some $n \geq 1$; we now prove it for $n{+}{+}$, i.e. we prove every sequence $a_1, a_2, \ldots, a_{n++}$ is bounded. By the induction hypothesis we know that $a_1, a_2, \ldots, a_n$ is bounded by some $M \geq 0$; in particular, it must be bounded by $M + |a_{n++}|$. On the other hand, a_{n++} is also bounded by $M + |a_{n++}|$. Thus $a_1, a_2, \ldots, a_n, a_{n++}$ is bounded by $M + |a_{n++}|$, and is hence bounded. This closes the induction. □

- Note that while this argument shows that every finite sequence is bounded, no matter how long the finite sequence is, it does not say anything about whether an infinite sequence is bounded or not. (Infinity is not a natural number). However, we have

- **Lemma 9**. Every Cauchy sequence $(a_n)_{n=1}^\infty$ is bounded.

- **Proof.** See Week 2 homework. □

* * * * *

Equivalent Cauchy sequences

- Consider the two Cauchy sequences of rational numbers:

$$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

and

$$1.5, 1.42, 1.415, 1.4143, 1.41422, \ldots$$

Informally, both of these sequences seem to be converging to the same number, the square root $\sqrt{2} = 1.41421\ldots$ (though this statement is not yet rigorous because we have not defined real numbers yet). If we are to define the real numbers from the rationals as limits of Cauchy sequences, we have to know when two Cauchy sequences of rationals give the same limit, without first defining a real number (since that would be circular). To do this we use a similar set of definitions to those used to define a Cauchy sequence in the first place.

- **Definition.** Let $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ be two sequences, and let $\varepsilon > 0$. We say that the sequence $(a_n)_{n=0}^\infty$ is *ε-close* to $(b_n)_{n=0}^\infty$ iff a_n is ε-close to b_n for each $n \in \mathbf{N}$. In other words, the sequence $a_0, a_1, a_2, \ldots$ is ε-close to the sequence $b_0, b_1, b_2, \ldots$ iff $|a_n - b_n| \le \varepsilon$ for all $n = 0, 1, 2, \ldots$.

- **Example.** The two sequences

$$1, -1, 1, -1, 1, \ldots$$

and

$$1.1, -1.1, 1.1, -1.1, 1.1, \ldots$$

are 0.1-close to each other. (Note however that neither of them are 0.1-steady).

- **Definition** Let $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ be two sequences, and let $\varepsilon > 0$. We say that the sequence $(a_n)_{n=0}^\infty$ is *eventually ε-close* to $(b_n)_{n=0}^\infty$ iff there exists an $N \geq 0$ such that the sequences $(a_n)_{n=N}^\infty$ and $(b_n)_{n=N}^\infty$ are ε-close. In other words, $a_0, a_1, a_2, \ldots$ is eventually ε-close to $b_0, b_1, b_2, \ldots$ iff there exists an $N \geq 0$ such that $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$.

- **Example.** The two sequences

$$1.1, 1.01, 1.001, 1.0001, \ldots$$

and

$$0.9, 0.99, 0.999, 0.9999, \ldots$$

are not 0.1-close (because the first elements of both sequences are not 0.1-close to each other). However, the sequences are still eventually 0.1-close, because if we start from the second elements onwards in the sequence, these sequences are 0.1-close. A similar argument shows that the two sequences are eventually 0.01 close (by starting from the third element onwards), and so forth.

- **Definition** Two sequences $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ are *equivalent* iff for each rational $\varepsilon > 0$, the sequences $(a_n)_{n=0}^\infty$ and $(b_n)_{n=0}^\infty$ are eventually ε-close. In other words, $a_0, a_1, a_2, \ldots$ and $b_0, b_1, b_2, \ldots$ are equivalent iff for every rational $\varepsilon > 0$, there exists an $N \geq 0$ such that $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$.

- Thus the two sequences above appear to be equivalent. We now prove this rigorously.

- **Claim.** Let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be the sequences $a_n = 1 + 10^{-n}$ and $b_n = 1 - 10^{-n}$. Then the sequences a_n, b_n are equivalent.

- **Proof.** We need to prove that for every $\varepsilon > 0$, the two sequences $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ are eventually ε-close to each other. So we fix an

$\varepsilon > 0$. We need to find an $N > 0$ such that $(a_n)_{n=N}^\infty$ and $(b_n)_{n=N}^\infty$ are ε-close; in other words, we need to find an $N > 0$ such that

$$|a_n - b_n| \leq \varepsilon \text{ for all } n \geq N.$$

However, we have

$$|a_n - b_n| = |(1 + 10^{-n}) - (1 - 10^{-n})| = 2 \times 10^{-n}.$$

Since 10^{-n} is a decreasing function of n (i.e. $10^{-m} < 10^{-n}$ whenever $m > n$; this is easily proven by induction), and $n \geq N$, we have $2 \times 10^{-n} \leq 2 \times 10^{-N}$. Thus we have

$$|a_n - b_n| \leq 2 \times 10^{-N} \text{ for all } n \geq N.$$

Thus in order to obtain $|a_n - b_n| \leq \varepsilon$ for all $n \geq N$, it will be sufficient to choose N so that $2 \times 10^{-N} \leq \varepsilon$. This is easy to do using logarithms, but we have not yet developed logarithms yet, so we will use a cruder method. First, we observe using induction that 10^N is always greater than N for any $N \geq 1$. (Proof: This is true for $N = 1$: $10^1 > 1$. Now suppose inductively that $N \geq 1$, and we have already proved $10^N > N$. Multiplying this by 10, we obtain $10^{N++} > 10N$; since $10N > N++$ (why?), we thus have $10^{N++} > N++$). Thus $10^{-N} \leq 1/N$, and so $2 \times 10^{-N} \leq 2/N$. Thus to get $2 \times 10^{-N} \leq \varepsilon$, it will suffice to choose N so that $2/N \leq \varepsilon$, or equivalently that $N \geq 2/\varepsilon$. But by Proposition 5 we can always choose such an N, and the claim follows. $\square$

* * * * *

The construction of the real numbers

- We are now ready to construct the real numbers. We shall introduce a new (meaningless) symbol LIM, similar to the notations — and // defined earlier; as the notation suggests, this will eventually match the familiar operation of lim.

- **Definition.** A *real number* is defined to be an object of the form $\mathrm{LIM}_{n\to\infty} a_n$, where $(a_n)_{n=1}^\infty$ is a Cauchy sequence of rational numbers. Two real numbers $\mathrm{LIM}_{n\to\infty} a_n$ and $\mathrm{LIM}_{n\to\infty} b_n$ are said to be equal iff $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ are equivalent Cauchy sequences. The set of all real numbers is denoted $\mathbf{R}$.

- **Informal example.** Let $a_1, a_2, a_3, \ldots$ denote the sequence

$$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

and let $b_1, b_2, b_3, \ldots$ denote the sequence

$$1.5, 1.42, 1.415, 1.4143, 1.41422, \ldots$$

then $\mathrm{LIM}_{n\to\infty} a_n$ is a real number, and is the same real number as $\mathrm{LIM}_{n\to\infty} b_n$, because $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ are equivalent Cauchy sequences: $\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} b_n$.

- We will call $\mathrm{LIM}_{n\to\infty} a_n$ the *formal limit* of the sequence a_n. Later on we will define a genuine notion of limit, and show that the formal limit of a Cauchy sequence is the same as the limit of that sequence; after that, we will not need formal limits ever again. (The situation is much like what we did with formal subtraction — and formal division //).

- In order to ensure that this definition is valid, we need to check that the notion of equality in the definition obeys the first three laws of equality:

- **Proposition 10.** Let $x = \mathrm{LIM}_{n\to\infty} a_n$, $y = \mathrm{LIM}_{n\to\infty} b_n$, and $z = \mathrm{LIM}_{n\to\infty} c_n$ be real numbers. Then, with the above definition of equality for real numbers, we have $x = x$. Also, if $x = y$, then $y = x$. Finally, if $x = y$ and $y = z$, then $x = z$.

- **Proof.** See Week 2 homework. □

- Because of this Proposition, we know that our definition of equality between two real numbers is legitimate. (Of course, when we define other operations on the reals, we have to check that they obey the law of substitution: two real number inputs which are equal should give equal outputs when applying any operation on the real numbers).

- Now we want to give the real numbers all the usual arithmetic operations: addition, multiplication, etc. We begin with addition.

- **Definition** Let $x = \mathrm{LIM}_{n\to\infty} a_n$ and $y = \mathrm{LIM}_{n\to\infty} b_n$ be real numbers. Then we define the sum $x + y$ to be $x + y := \mathrm{LIM}_{n\to\infty} a_n + b_n$.

- Thus, for example, the sum of $\mathrm{LIM}_{n\to\infty} 1 + 1/n$ and $\mathrm{LIM}_{n\to\infty} 2 + 3/n$ should be $\mathrm{LIM}_{n\to\infty} 3 + 4/n$.

- We now check that this definition is valid. The first thing we need to do is confirm that the sum of two real numbers is in fact a real number:

- **Lemma 11.** Let $x = \mathrm{LIM}_{n\to\infty} a_n$ and $y = \mathrm{LIM}_{n\to\infty} b_n$ be real numbers. Then $x + y$ is also a real number (i.e. $(a_n + b_n)_{n=1}^{\infty}$ is a Cauchy sequence of rationals).

- **Proof.** We need to show that for every $\varepsilon > 0$, the sequence $(a_n + b_n)_{n=1}^{\infty}$ is eventually ε-steady. Now from hypothesis we know that $(a_n)_{n=1}^{\infty}$ is eventually ε-steady, and $(b_n)_{n=1}^{\infty}$ is eventually ε-steady, but it turns out that this is not quite enough (this can be used to imply that $(a_n + b_n)_{n=1}^{\infty}$ is eventually 2ε-steady, but that's not what we want). So we need to do a little trick, which is to play with the value of ε.

- We know that $(a_n)_{n=1}^{\infty}$ is eventually δ-steady for every value of δ. This implies not only that $(a_n)_{n=1}^{\infty}$ is eventually ε-steady, but it is also eventually $\varepsilon/2$-steady. Similarly, the sequence $(b_n)_{n=1}^{\infty}$ is also eventually $\varepsilon/2$-steady. This will turn out to be enough to conclude that $(a_n + b_n)_{n=1}^{\infty}$ is eventually ε-steady.

- Let's see how. Since $(a_n)_{n=1}^{\infty}$ is eventually $\varepsilon/2$-steady, we know that there exists an $N \geq 1$ such that $(a_n)_{n=N}^{\infty}$ is $\varepsilon/2$-steady, i.e. a_n and a_m are $\varepsilon/2$-close for every $n, m \geq N$. Similarly there exists an $M \geq 1$ such that $(b_n)_{n=M}^{\infty}$ is $\varepsilon/2$-steady, i.e. b_n and b_m are $\varepsilon/2$-close for every $n, m \geq M$.

- Let $\max(N, M)$ be the larger of N and M (we know from week 1 notes that one has to be greater than or equal to the other). If $n, m \geq \max(N, M)$, then we know that a_n and a_m are $\varepsilon/2$-close, and b_n and b_m are $\varepsilon/2$ close, and so by Proposition 2 we see that $a_n + b_n$ and $a_m + b_m$ are ε-close for every $n, m \geq \max(N, M)$. This implies that the sequence $(a_n + b_n)_{n=1}^{\infty}$ is eventually ε-close, as desired. □

- The other thing we need to check is the axiom of substitution: if we replace a real number x by another number equal to x, this should not change the sum $x + y$ (and similarly if we substitute y by another number equal to y).

- **Lemma 12.** Let $x = \mathrm{LIM}_{n\to\infty} a_n$, $y = \mathrm{LIM}_{n\to\infty} b_n$, and $x' = \mathrm{LIM}_{n\to\infty} a'_n$ be real numbers. Suppose that $x = x'$. Then we have $x + y = x' + y$.

- **Proof.** Since x and x' are equal, we know that the Cauchy sequences $(a_n)_{n=1}^\infty$ and $(a'_n)_{n=1}^\infty$ are equivalent, so in other words they are eventually ε-close for each $\varepsilon > 0$. We need to show that the sequences $(a_n + b_n)_{n=1}^\infty$ and $(a'_n + b_n)_{n=1}^\infty$ are eventually ε-close for each $\varepsilon > 0$. But we already know that there is an $N \geq 1$ such that $(a_n)_{n=N}^\infty$ and $(a'_n)_{n=N}^\infty$ are ε-close, i.e. that a_n and a'_n are ε-close for each $n \geq N$. Since b_n is of course 0-close to b_n, we thus see from Proposition 2 that $a_n + b_n$ and $a'_n + b_n$ are ε-close for each $n \geq N$. This implies that $(a_n + b_n)_{n=1}^\infty$ and $(a'_n + b_n)_{n=1}^\infty$ are eventually ε-close for each $\varepsilon > 0$, and we are done. □

- The above lemma verifies the axiom of substitution for the "x" variable in $x + y$, but one can similarly prove the axiom of substitution for the "y" variable. (A quick way is to observe from the definition of $x + y$ that we certainly have $x + y = y + x$, since $a_n + b_n = b_n + a_n$).

- We can similarly define multiplication:

- **Definition** Let $x = \mathrm{LIM}_{n\to\infty} a_n$ and $y = \mathrm{LIM}_{n\to\infty} b_n$ be real numbers. Then we define the product xy to be $xy := \mathrm{LIM}_{n\to\infty} a_n b_n$.

- The following Proposition ensures that this definition is valid, and that the product of two real numbers is in fact a real number:

- **Proposition 13.** Let $x = \mathrm{LIM}_{n\to\infty} a_n$, $y = \mathrm{LIM}_{n\to\infty} b_n$, and $x' = \mathrm{LIM}_{n\to\infty} a'_n$ be real numbers. Then xy is also a real number. Furthermore, if $x = x'$, then $xy = x'y$.

- **Proof** See Week 2 homework. □

- Of course we can prove a similar substitution rule when y is replaced by a real number y' which is equal to y.

- At this point we embed the rationals back into the reals, by equating every rational number q with the real number $\mathrm{LIM}_{n\to\infty} q$. For instance, if $a_1, a_2, a_3, \ldots$ is the sequence

$$0.5, 0.5, 0.5, 0.5, 0.5, \ldots$$

then we set $\lim_{n\to\infty} a_n$ equal to 0.5. This embedding is consistent with our definitions of addition and multiplication, since for any rational numbers a, b we have

$$\mathrm{LIM}_{n\to\infty} a + \mathrm{LIM}_{n\to\infty} b = \mathrm{LIM}_{n\to\infty} a+b; \quad \mathrm{LIM}_{n\to\infty} a \times \mathrm{LIM}_{n\to\infty} b = \mathrm{LIM}_{n\to\infty} ab;$$

this means that when one wants to add or multiply two rational numbers a, b it does not matter whether one thinks of these numbers as rationals or as the real numbers $\mathrm{LIM}_{n\to\infty} a$, $\mathrm{LIM}_{n\to\infty} b$. Also, this identification of rational numbers and real numbers is consistent with our definitions of equality: if a and b are equal, then $\mathrm{LIM}_{n\to\infty} a$ and $\mathrm{LIM}_{n\to\infty} b$ are equal, and if a and b are unequal, then $\mathrm{LIM}_{n\to\infty} a$ and $\mathrm{LIM}_{n\to\infty} b$ are unequal (because the Cauchy sequences $a, a, a, a, \ldots$ and $b, b, b, b \ldots$ are not equivalent (why?)).

- We can now easily define negation $-x$ for real numbers x by the formula

$$-x := (-1) \times x,$$

since -1 is a rational number and is hence real. Note that this is clearly consistent with our negation for rational numbers since we have $-q = (-1) \times q$ for all rational numbers q. Also, from our definitions it is clear that

$$-\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} - a_n$$

(why?). Once we have addition and negation, we can define subtraction as usual by

$$x - y := x + (-y),$$

note that this implies

$$\mathrm{LIM}_{n\to\infty} a_n - \mathrm{LIM}_{n\to\infty} b_n = \mathrm{LIM}_{n\to\infty} a_n - b_n.$$

- We can now easily show that the real numbers obey all the usual rules of algebra:

- **Proposition 14.** All the laws of Proposition 16 from last weeks notes hold not only for the integers, but for the reals as well.

- **Proof.** We illustrate this with one such rule: $x(y+z) = xy + xz$. Let $x = \mathrm{LIM}_{n\to\infty} a_n$, $y = \mathrm{LIM}_{n\to\infty} b_n$, and $z = \mathrm{LIM}_{n\to\infty} c_n$ be real numbers. Then by definition, $xy = \mathrm{LIM}_{n\to\infty} a_n b_n$ and $xz = \mathrm{LIM}_{n\to\infty} a_n c_n$, and so $xy + xz = \mathrm{LIM}_{n\to\infty} a_n b_n + a_n c_n$. A similar line of reasoning shows that $x(y+z) = \mathrm{LIM}_{n\to\infty} a_n(b_n + c_n)$. But we already know that $a_n(b_n + c_n)$ is equal to $a_n b_n + a_n c_n$ for the rational numbers a_n, b_n, c_n, and the claim follows. The other laws of algebra are proven similarly. $\square$

- The last arithmetic operation we need to define is reciprocal: $x \to x^{-1}$. This one is a little more subtle. The naive definition is to define

  $$(\mathrm{LIM}_{n\to\infty} a_n)^{-1} := \mathrm{LIM}_{n\to\infty} a_n^{-1},$$

  but there are a few problems with this. For instance, let $a_1, a_2, a_3, \ldots$ be the Cauchy sequence

  $$0.1, 0.01, 0.001, 0.0001, \ldots,$$

  and let $x := \mathrm{LIM}_{n\to\infty} a_n$. Then by this definition, x^{-1} would be $\mathrm{LIM}_{n\to\infty} b_n$, where $b_1, b_2, b_3, \ldots$ is the sequence

  $$10, 100, 1000, 10000, \ldots$$

  but this is not a Cauchy sequence (it isn't even bounded). Of course, the problem here is that our original Cauchy sequence $(a_n)_{n=1}^\infty$ was equivalent to the zero sequence $(0)_{n=1}^\infty$ (why?), and hence that our real number x was in fact equal to 0. So we should only allow the operation of reciprocal when x is non-zero.

- However, even when we restrict ourselves to non-zero real numbers, we have a slight problem, because a non-zero real number might be the formal limit of a Cauchy sequence which contains zero elements. For instance, the number 1, which is rational and hence real, is the formal limit $1 = \mathrm{LIM}_{n\to\infty} a_n$ of the Cauchy sequence

  $$0, 0.9, 0.99, 0.999, 0.9999, \ldots$$

but using our naive definition of reciprocal, we cannot invert the real number 1, because we can't invert the first element (0) of this Cauchy sequence!

- To get around these problems we need to keep our Cauchy sequence away from zero. To do this we first need a definition.

- **Definition** A sequence $(a_n)_{n=1}^{\infty}$ is said to be *bounded away from zero* iff there exists a $c > 0$ such that $|a_n| \geq c$ for all $n \geq 1$.

- For instance, the sequence $1, -1, 1, -1, 1, -1, 1, \ldots$ is bounded away from zero (all the coefficients have absolute value at least 1). But the sequence $0.1, 0.01, 0.001, \ldots$ is not bounded away from zero, and neither is $0, 0.9, 0.99, 0.999, 0.9999, \ldots$. The sequence $10, 100, 1000, \ldots$ is bounded away from zero, but is not bounded.

- We now show that every non-zero real number is the formal limit of a Cauchy sequence bounded away from zero:

- **Lemma 15.** Let x be a non-zero real number. Then $x = \mathrm{LIM}_{n\to\infty} a_n$ for some Cauchy sequence $(a_n)_{n=1}^{\infty}$ which is bounded away from zero.

- **Proof**. Since x is real, we know that $x = \mathrm{LIM}_{n\to\infty} b_n$ for some Cauchy sequence $(b_n)_{n=1}^{\infty}$. But we are not yet done, because we do not know that b_n is bounded away from zero. On the other hand, we are given that $x \neq 0 = \mathrm{LIM}_{n\to\infty} 0$, which means that the sequence $(b_n)_{n=1}^{\infty}$ is NOT equivalent to $(0)_{n=1}^{\infty}$. Thus the sequence $(b_n)_{n=1}^{\infty}$ cannot be eventually ε-close to $(0)_{n=1}^{\infty}$ for *every* $\varepsilon > 0$. Therefore we can find an $\varepsilon > 0$ such that $(b_n)_{n=1}^{\infty}$ is NOT eventually ε-close to $(0)_{n=1}^{\infty}$.

  Let us fix this ε. We know that $(b_n)_{n=1}^{\infty}$ is a Cauchy sequence, so it is eventually ε-steady. Moreover, it is eventually $\varepsilon/2$-steady, since $\varepsilon/2 > 0$. Thus there is an $N \geq 1$ such that $|b_n - b_m| \leq \varepsilon/2$ for all $n, m \geq N$.

  On the other hand, we cannot have $|b_n| \leq \varepsilon$ for all $n \geq N$, since this would imply that $(b_n)_{n=1}^{\infty}$ is eventually ε-close to $(0)_{n=1}^{\infty}$. Thus there must be some $n_0 \geq N$ for which $|b_{n_0}| > \varepsilon$. Since we already know that $|b_{n_0} - b_n| \leq \varepsilon/2$ for all $n \geq N$, we thus conclude from the triangle inequality (how?) that $|b_n| \geq \varepsilon/2$ for all $n \geq N$.

This almost proves that $(b_n)_{n=1}^\infty$ is bounded away from zero. Actually, what it does is show that $(b_n)_{n=1}^\infty$ is *eventually* bounded away from zero. But this is easily fixed, by defining a new sequence a_n, by setting $a_n := \varepsilon/2$ if $n < N$ and $a_n := b_n$ if $n \geq N$. Since b_n is a Cauchy sequence, it is not hard to verify that a_n is also a Cauchy sequence which is equivalent to b_n (because the two sequences are eventually the same), and so $x = \mathrm{LIM}_{n\to\infty} a_n$. And since $|b_n| \geq \varepsilon/2$ for all $n \geq N$, we know that $|a_n| \geq \varepsilon/2$ for all $n \geq 1$ (splitting into the two cases $n \geq N$ and $n < N$ separately). Thus we have a Cauchy sequence which is bounded away from zero (by $\varepsilon/2$ instead of ε, but that's still OK since $\varepsilon/2 > 0$), and which has x as a formal limit, and so we are done. □

- Once a sequence is bounded away from zero, we can take its reciprocal without any difficulty:

- **Lemma 16.** Let $(a_n)_{n=1}^\infty$ be a Cauchy sequence which is bounded away from zero. Then the sequence $(a_n^{-1})_{n=1}^\infty$ is also a Cauchy sequence.

- **Proof.** Since $(a_n)_{n=1}^\infty$ is bounded away from zero, we know that there is a $c > 0$ such that $|a_n| \geq c$ for all $n \geq 1$. Now we need to show that $(a_n^{-1})_{n=1}^\infty$ is eventually ε-steady for each $\varepsilon > 0$. Thus let us fix an $\varepsilon > 0$; our task is now to find an $N \geq 1$ such that $|a_n^{-1} - a_m^{-1}| \leq \varepsilon$ for all $n, m \geq N$. But

$$|a_n^{-1} - a_m^{-1}| = |\frac{a_m - a_n}{a_m a_n}| \leq \frac{|a_m - a_n|}{c^2}$$

(since $|a_m|, |a_n| \geq c$), and so to make $|a_n^{-1} - a_m^{-1}|$ less than or equal to ε, it will suffice to make $|a_m - a_n|$ less than or equal to $c^2\varepsilon$. But since $(a_n)_{n=1}^\infty$ is a Cauchy sequence, and $c^2\varepsilon > 0$, we can certainly find an N such that the sequence $(a_n)_{n=N}^\infty$ is $c^2\varepsilon$-steady, i.e. $|a_m - a_n| \leq c^2\varepsilon$ for all $n \geq N$. By what we have said above, this shows that $|a_n - a_m| \leq \varepsilon$ for all $m, n \geq N$, and hence the sequence $(a_n^{-1})_{n=1}^\infty$ is eventually ε-steady. Since we have proven this for every ε, we have that $(a_n^{-1})_{n=1}^\infty$ is a Cauchy sequence, as desired. □

- We are now ready to make the definition of reciprocal:

- **Definition.** Let x be a non-zero real number. Let $(a_n)_{n=1}^\infty$ be a Cauchy sequence bounded away from zero such that $x = \mathrm{LIM}_{n\to\infty} a_n$ (such a sequence exists by Lemma 15). Then we define the reciprocal x^{-1} by the formula $x^{-1} := \mathrm{LIM}_{n\to\infty} a_n^{-1}$. (From Lemma 16 we know that x^{-1} is a real number.)

- We need to check one thing before we are sure this definition makes sense: what if there are two different Cauchy sequences $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ which have x as their formal limit, $x = \mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} b_n$. The above definition might conceivably give *two* different reciprocals x^{-1}, namely $\mathrm{LIM}_{n\to\infty} a_n^{-1}$ and $\mathrm{LIM}_{n\to\infty} b_n^{-1}$. Fortunately, this never happens:

- **Lemma 17.** Let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be two Cauchy sequences bounded away from zero such that $\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} b_n$ (i.e. the two sequences are equivalent). Then $\mathrm{LIM}_{n\to\infty} a_n^{-1} = \mathrm{LIM}_{n\to\infty} b_n^{-1}$.

- **Proof** Consider the following product P of three real numbers:

$$P := (\mathrm{LIM}_{n\to\infty} a_n^{-1}) \times (\mathrm{LIM}_{n\to\infty} a_n) \times (\mathrm{LIM}_{n\to\infty} b_n^{-1}).$$

If we multiply this out, we obtain

$$P = \mathrm{LIM}_{n\to\infty} a_n^{-1} a_n b_n^{-1} = \mathrm{LIM}_{n\to\infty} b_n^{-1}.$$

On the other hand, since $\mathrm{LIM}_{n\to\infty} a_n = \mathrm{LIM}_{n\to\infty} b_n$, we can rewrite P as

$$P = (\mathrm{LIM}_{n\to\infty} a_n^{-1}) \times (\mathrm{LIM}_{n\to\infty} b_n) \times (\mathrm{LIM}_{n\to\infty} b_n^{-1})$$

(cf. Proposition 13). Now if we multiply things out again, we get

$$P = \mathrm{LIM}_{n\to\infty} a_n^{-1} b_n b_n^{-1} = \mathrm{LIM}_{n\to\infty} a_n^{-1}.$$

Comparing our different formulae for P we see that $\mathrm{LIM}_{n\to\infty} a_n^{-1} = \mathrm{LIM}_{n\to\infty} b_n^{-1}$, as desired. □

- Thus reciprocal is well-defined (for each non-zero real number x, we have exactly one definition of the reciprocal x^{-1}). Note it is clear from the definition that $xx^{-1} = x^{-1}x = 1$ (why?); thus all the field axioms (Proposition 20 from last week's notes) apply to the reals as well as to the rationals. We of course cannot give 0 a reciprocal, since 0 multiplied by anything gives 0, not 1. Also note that if q is a non-zero rational, and hence equal to the real number $\mathrm{LIM}_{n\to\infty}q$, then the reciprocal of $\mathrm{LIM}_{n\to\infty}q$ is $\mathrm{LIM}_{n\to\infty}q^{-1} = q^{-1}$; thus the operation of reciprocal on real numbers is consistent with the operation of reciprocal on rational numbers.

- Once one has reciprocal, one can define division x/y of two real numbers x, y, provided y is non-zero, by the formula

$$x/y = x \times y^{-1},$$

just as we did with the rationals. In particular, we have the *cancellation law*: if x, y, z are real numbers such that $xz = yz$, and z is non-zero, then by dividing by z we conclude that $x = y$. Note that this cancellation law does not work when z is zero.

- We now have all four of the basic arithmetic operations on the reals: addition, subtraction, multiplication, and division, with all the usual rules of algebra. Next we turn to the notion of order on the reals.

* * * * *

Ordering the reals

- We know that every rational number is positive, negative, or zero. We now want to say the same thing for the reals: each real number should be positive, negative, or zero. Since a real number x is just a formal limit of rationals a_n, it is tempting to make the following definition: a real number $x = \mathrm{LIM}_{n\to\infty}a_n$ is positive if all of the a_n are positive, and negative if all of the a_n are negative (and zero if all of the a_n are zero). However, one soon realizes some problems with this definition. For instance, the sequence $(a_n)_{n=1}^{\infty}$ defined by $a_n := 10^{-n}$, thus

$$0.1, 0.01, 0.001, 0.0001, \ldots$$

consists entirely of positive numbers, but this sequence is equivalent to the zero sequence $0, 0, 0, 0, \ldots$ and thus $\mathrm{LIM}_{n\to\infty}a_n = 0$. Thus even though all the rationals were positive, the real formal limit of these rationals was zero rather than positive. Another example is

$$0.1, -0.01, 0.001, -0.0001, \ldots;$$

this sequence is a hybrid of positive and negative numbers, but again the formal limit is zero.

- The trick, as with the reciprocals in the previous section, is to limit one's attention to sequences which are bounded away from zero.

- **Definition** Let $(a_n)_{n=1}^\infty$ be a sequence of rationals. We say that this sequence is *positively bounded away from zero* iff we have a positive rational $c > 0$ such that $a_n \geq c$ for all $n \geq 1$ (in particular, the sequence is entirely positive). The sequence is *negatively bounded away from zero* iff we have a negative rational $-c < 0$ such that $a_n \leq -c$ for all $n \geq 1$ (in particular, the sequence is entirely negative).

- **Examples.** The sequence $1.1, 1.01, 1.001, 1.0001, \ldots$ is positively bounded away from zero (all terms are greater than or equal to 1). The sequence $-1.1, -1.01, -1.001, -1.0001, \ldots$ is negatively bounded away from zero. The sequence $1, -1, 1, -1, 1, -1, \ldots$ is bounded away from zero, but is neither positively bounded away from zero nor negatively bounded away from zero.

- It is clear that any sequence which is positively or negatively bounded away from zero, is bounded away from zero. Also, a sequence cannot be both positively bounded away from zero and negatively bounded away from zero at the same time.

- **Definition** A real number x is said to be *positive* iff it can be written as $x = \mathrm{LIM}_{n\to\infty} a_n$ for some Cauchy sequence $(a_n)_{n=1}^\infty$ which is positively bounded away from zero. x is said to be *negative* iff it can be written as $x = \mathrm{LIM}_{n\to\infty} a_n$ for some sequence $(a_n)_{n=1}^\infty$ which is negatively bounded away from zero.

- We now give the basic properties of positive and negative numbers.

- **Proposition 18.** For every real number x, exactly one of the following three statements is true: (a) x is zero; (b) x is positive; (c) x is negative. A real number x is negative if and only if $-x$ is positive. If x and y are positive, then so are $x + y$ and xy.

- **Proof.** See Week 2 Homework. □

- Note that if q is a positive rational number, then the Cauchy sequence $q, q, q, \ldots$ is positively bounded away from zero, and hence $\mathrm{LIM}_{n\to\infty} q =$

q is a positive real number. Thus the notion of positivity for rationals is consistent with that for reals. Similarly for the notion of negativity.

- Once we have defined positive and negative numbers, we can define order.

- **Definition** Let x and y be real numbers. We say that x is *greater than* y, and write $x > y$, if $x - y$ is a positive real number, and $x < y$ iff $x - y$ is a negative real number. We define $x \geq y$ iff $x > y$ or $x = y$, and similarly define $x \leq y$.

- Comparing this with the definition of order on the rationals from week 1 notes we see that order on the reals is consistent with order on the rationals, i.e. if two rational numbers q, q' are such that q is less than q' in the rational number system, then q is still less than q' in the real number system, and similarly for "greater than".

- **Proposition 19.** All the properties from Proposition 22 of last week's notes which held for rationals, continue to hold for real numbers.

- **Proof.** This is an easy consequence of Proposition 18, and is left to the reader. Example: suppose we have $x < y$ and z a positive real, and want to conclude that $xz < yz$. Since $x < y$, $y - x$ is positive, hence by Proposition 18 $(y-x)z = yz - xz$ is positive, hence $xz < yz$. The other analogues of Proposition 22 are proven in a similar manner (using Proposition 18 and the basic laws of algebra). □

- As an application of these propositions, we prove

- **Proposition 20.** Let x be a positive real number. Then x^{-1} is also positive. Also, if y is another positive number and $x > y$, then $x^{-1} < y^{-1}$.

- **Proof.** Let x be positive. Since $xx^{-1} = 1$, the real number x^{-1} cannot be zero (since $x0 = 0 \neq 1$). Also, from Proposition 18 it is easy to see that a positive number times a negative number is negative; this shows that x^{-1} cannot be negative, since this would imply that $xx^{-1} = 1$ is negative, a contradiction. Thus, by Proposition 18, the only possibility left is that x^{-1} is positive.

- Now let y be positive as well, so x^{-1} and y^{-1} are also positive. If $x^{-1} \geq y^{-1}$, then by Proposition 19 we have $xx^{-1} > yx^{-1} \geq yy^{-1}$, thus $1 > 1$, which is a contradiction. Thus we must have $x^{-1} < y^{-1}$. □

- Another application is that the laws of exponentiation (Proposition 4) that were previously proven for rationals, are also true for reals; see the Appendix at the end of these notes.

- We have already seen that the formal limit of positive rationals need not be positive; it could be zero, as the example $0.1, 0.01, 0.001, \ldots$ showed. However, the formal limit of *non-negative* rationals (i.e. rationals that are either positive or zero) is non-negative. (Eventually, we will see a better explanation of this fact: the set of non-negative reals is *closed*, whereas the set of positive reals is *open*).

- **Proposition 21.** Let $a_1, a_2, a_3, \ldots$ be a Cauchy sequence of non-negative rational numbers. Then $\mathrm{LIM}_{n\to\infty} a_n$ is a non-negative real number.

- **Proof.** We argue by contradiction, and suppose that the real number $x := \mathrm{LIM}_{n\to\infty} a_n$ is a negative number. Then by definition of negative real number, we have $x = \mathrm{LIM}_{n\to\infty} b_n$ for some sequence b_n which is negatively bounded away from zero, i.e. there is a negative rational $-c < 0$ such that $b_n \leq -c$ for all $n \geq 1$. On the other hand, we have $a_n \geq 0$ for all $n \geq 1$, by hypothesis. Thus the numbers a_n and b_n are never $c/2$-close, since $c/2 < c$. Thus the sequences $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ are not eventually $c/2$-close. Since $c/2 > 0$, this implies that $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ are not equivalent. But this contradicts the fact that both these sequences have x as their formal limit. □

- **Corollary 22.** Let $(a_n)_{n=1}^\infty$ and $(b_n)_{n=1}^\infty$ be Cauchy sequences of rationals such that $a_n \geq b_n$ for all $n \geq 1$. Then $\mathrm{LIM}_{n\to\infty} a_n \geq \mathrm{LIM}_{n\to\infty} b_n$.

- **Proof.** Apply Proposition 21 to the sequence $a_n - b_n$. □

- Note that the above Corollary does not work if the $\geq$ signs are replaced by $>$: for instance if $a_n := 1 + 1/n$ and $b_n := 1 - 1/n$, then a_n is always strictly greater than b_n, but the formal limit of a_n is not greater than the formal limit of b_n, instead they are equal.

- Once we have the notion of positive and negative real numbers, we can define absolute value $|x|$, and distance $d(x, y) := |x - y|$ just as we did for the rationals. In fact, Propositions 1 and 2 hold not only for the rationals, but for the reals; the proof is identical, since the real numbers obey all the laws of algebra and order that the rationals do.

- We now observe that the real numbers cannot get any "smaller" or "larger" than the rational numbers can:

- **Proposition 23.** Let x be a positive real number. Then there exists a positive rational number q such that $q \leq x$, and there exists a positive integer N such that $x \leq N$.

- **Proof.** Since x is a positive real, it is the formal limit of some Cauchy sequence $(a_n)_{n=1}^{\infty}$ which is positively bounded away from zero. Also, by Lemma 9, this sequence is bounded. Thus we have rationals $q > 0$ and r such that $q \leq a_n \leq r$ for all $n \geq 1$. But by Proposition 5 we know that there is some integer N such that $r \leq N$; since q is positive and $q \leq r \leq N$, we see that N is positive. Thus $q \leq a_n \leq N$ for all $n \geq 1$. Applying Corollary 22 we obtain that $q \leq x \leq N$, as desired. $\square$

- **Corollary 24 (Archimedean property).** Let x and ε be any positive real numbers. Then there exists a positive integer M such that $M\varepsilon > x$.

- **Proof.** The number x/ε is positive, and hence by Proposition 23 there exists a positive integer N such that $x/\varepsilon \leq N$. If we set $M := N{+}{+}$, then $x/\varepsilon < M$. Now multiply by ε. $\square$

- This property is quite important; it says that no matter how large x is and how small ε is, if one keeps adding ε to itself, one will eventually overtake x.

- **Proposition 25.** Given any two real numbers $x < y$, we can find a rational number q such that $x < q < y$.

- **Proof.** See Week 2 Homework.

- We have now completed our construction of the real numbers. This number system contains the rationals, and has almost everything that the rational number system has: the arithmetic operations, the laws of algebra, the laws of order. However, we have not yet demonstrated any *advantages* that the real numbers have over the rationals; so far, after much effort, all we have done is shown that they are *at least as good as* the rational number system. But in the next few sections we show that the real numbers can do more things than rationals: for example, we can take square roots in a real number system.

- One side remark: up until now, we have not addressed the fact that real numbers can be expressed using the decimal system. For instance, the formal limit of

$$1.4, 1.41, 1.414, 1.4142, 1.41421, \ldots$$

is more conventionally represented as the infinite decimal $1.41421\ldots$. We will address this in a supplemental handout, but remark that there are some subtleties in the decimal system, for instance $0.9999\ldots$ and $1.000\ldots$ are in fact the same real number.

* * * * *

The least upper bound property.

- We now give one of the most basic advantages of the real numbers over the rationals; one can take the *least upper bound* $\sup(E)$ of any subset E of the real numbers $\mathbf{R}$.

- For this we of course need some basic set theory, to make sense of such notions as "subset". This we leave to a supplemental handout of set theory. In the main sequence of notes we assume the reader is comfortable with sets, and related notions such as subset, union, and intersection, and the element relation $\in$.

- **Definition.** Let E be a subset of $\mathbf{R}$, and let M be a real number. We say that M is an *upper bound* for E, iff we have $x \leq M$ for every element x in E.

- **Example.** Let E be the interval $E := \{x \in \mathbf{R} : 0 \leq x \leq 1\}$. Then 1 is an upper bound for E, since every element of E is less than or equal to 1. It is also true that 2 is an upper bound for E, and indeed every number greater or equal to 1 is an upper bound for E. On the other hand, any other number, such as 0.5, is not an upper bound, because 0.5 is not larger than *every* element in E. (Merely being larger than *some* elements of E is not necessarily enough to make 0.5 an upper bound.)

- **Example.** Let $\mathbf{R}^+$ be the set of positive reals: $\mathbf{R}^+ := \{x \in \mathbf{R} : x > 0\}$. Then $\mathbf{R}^+$ does not have any upper bounds at all (why?).

- **Example.** Let $\emptyset$ be the empty set. Then every number M is an upper bound for $\emptyset$, because M is greater than every element of the empty set (this is a vacuously true statement, but still true).

- It is clear that if M is an upper bound of E, then any larger number $M' \geq M$ is also an upper bound of E. On the other hand, it is not so clear whether it is also possible for any number smaller than M to also be an upper bound of E. This motivates the following definition:

- **Definition.** Let E be a subset of $\mathbf{R}$, and M be a real number. We say that M is a *least upper bound* for E iff (a) M is an upper bound for E, and also (b) any other upper bound M' for E must be larger than or equal to M.

- **Example** Let E be the interval $E := \{x \in \mathbf{R} : 0 \leq x \leq 1\}$. Then, as noted before, E has many upper bounds, indeed every number greater than or equal to 1 is an upper bound. But only 1 is the *least* upper bound; all other upper bounds are larger than 1.

- **Example** The empty set does not have a least upper bound (why?).

- **Proposition 26.** Let E be a subset of $\mathbf{R}$. Then E can have at most one least upper bound.

- **Proof.** Suppose for contradiction that E has at least two least upper bounds, say M_1 and M_2, where $M_1 \neq M_2$. Since M_1 is a least upper bound and M_2 is an upper bound, then by definition of least upper bound we have $M_2 \geq M_1$. Since M_2 is a least upper bound and M_1 is an upper bound, we similarly have $M_1 \geq M_2$. Thus $M_1 = M_2$, a contradiction. □

- Now we come to an important property of the real numbers:

- **Theorem 27 (Least upper bound property).** Let E be a non-empty subset of $\mathbf{R}$. If E has an upper bound, (i.e. E has some upper bound M), then it must have exactly one *least* upper bound.

- **Proof. (Optional)** Let E be a non-empty subset of $\mathbf{R}$ with an upper bound M. By Proposition 26, we know that E has at most one least upper bound; we have to show that E has at least one least upper bound.

  Since E is non-empty, it contains at least one real number; let's say that E contains a real number x_0.

- Now let n be any positive integer. Since $x_0 - 1/n$ is less than x_0, we know that $x_0 - 1/n$ is NOT an upper bound for E. On the other hand, by the Archimedean property (Corollary 24), there is an integer K such that $x_0 + K/n > M$ (why?). Thus $x_0 + K/n$ is an upper bound for E. From these two statements, we can thus conclude that there exists some natural number i with $0 \leq i \leq K$ such that $x_0 + i/n$ is an upper bound for E, while $x_0 + (i-1)/n$ is not an upper bound for E. (Proof: if no such number i existed, then it is easy to use induction to show that $x_0 + i/n$ is not an upper bound for E for any $0 \leq i \leq K$. But this implies that $x_0 + K/n$ is not an upper bound for E, a contradiction). Since $x_0 + (i-1)/n < x_0 + i/n$, we can use Proposition 25 to find a rational number a_n such that

$$x_0 + (i-1)/n < a_n < x_0 + i/n.$$

  Thus $a_n + 1/n$ is an upper bound for E (being larger than the upper bound $x_0 + i/n$, but $a_n - 1/n$ is not an upper bound for E (since it is less than $x_0 + (i-1)/n$, which was not an upper bound for E).

- Now let n, m be two positive integers. Since $a_n + 1/n$ is an upper bound for E, but $a_m - 1/m$ is not, we must have $a_n + 1/n > a_m - 1/m$ (why? prove by contradiction). Similarly we have $a_m + 1/m > a_n - 1/n$. Thus we have

$$-1/n - 1/m < a_n - a_m < 1/n + 1/m.$$

  In particular, for any positive integer $N \geq 1$, we have

$$-2/N < a_n - a_m < 2/N \text{ for every } n, m \geq N. \qquad (1)$$

Thus a_n and a_m are $2/N$-close for all $n, m \geq N$, which implies that the sequence $(a_n)_{n=1}^{\infty}$ is eventually $2/N$-steady for every positive integer N. Thus the sequence $(a_n)_{n=1}^{\infty}$ is eventually ε-steady for every $\varepsilon > 0$ (since one can always find an integer N such that $N > 2/\varepsilon$, and hence $2/N < \varepsilon$). In particular, $(a_n)_{n=1}^{\infty}$ is a Cauchy sequence. We can thus construct the real number $x := \mathrm{LIM}_{n\to\infty} a_n$.

- From equation (1), we have in particular that

  $$-2/N < a_n - a_N < 2/N \text{ for every } n \geq N.$$

  From Corollary 22 (and a little modification; can you see what needs to be done?) we thus have

  $$-2/N < x - a_N < 2/N$$

  for any $N \geq 1$.

- We now show that x is the least upper bound for E. First, we show that x is an upper bound for E. Let y be any element of E; we have to show that $y \leq x$. But we already know that for any positive integer n, $a_n + 1/n$ is already an upper bound for E, thus $y \leq a_n + 1/n$. Since $-2/n < x - a_n$, we thus have $y \leq x + 3/n$. But this is true for all n, so this implies that $y \leq x$ as desired (why? Note that if $y > x$ then we can find an n such that $n > 3/(y - x)$, and thus $x + 3/n < y$). Thus x is an upper bound for E.

- Now we show that x is the least upper bound for E. Suppose we have some other upper bound z of E; we have to show that $x \leq z$. But we already know that for any positive integer n, $a_n - 1/n$ is not an upper bound for E; this implies that $a_n - 1/n < z$ (why?). Since $x - a_n < 2/n$, we thus have $x \leq z + 3/n$. But this is true for all n; arguing as before, this implies $x \leq z$ as desired. □

- Thus if E is non-empty and has some upper bound, we can talk about *the* least upper bound of E; we denote this by $\sup(E)$ or $\sup E$, and is also known as the *supremum* of E. Some additional notation: we introduce two new symbols, $+\infty$ and $-\infty$. If E has no upper bound, we set $\sup E = +\infty$; if E is empty, we set $\sup E = -\infty$. (At present,

$+\infty$ and $-\infty$ are meaningless symbols; we have no operations on them at present, and none of our results involving real numbers apply to $+\infty$ and $-\infty$, because these are not real numbers. Sometimes we add $+\infty$ and $-\infty$ to the reals to form the *extended real number system*, but this system is not as convenient to work with as the real number system, because many of the laws of algebra break down. (For instance, it is not a good idea to try to define $+\infty + -\infty$; setting this equal to 0 causes some problems). We will discuss the extended real number system more in the next set of notes.

- Now we give an example of how the least upper bound property is useful.

- **Proposition 28.** There exists a positive real number x such that $x^2 = 2$.

- This should be compared with Proposition 6; it shows that certain numbers are real but not rational. (Such numbers are called *irrational*). The proof of this proposition also shows that the rationals $\mathbf{Q}$ do not obey the least upper bound property, otherwise one could use that property to construct a square root of 2, which by Proposition 6 is not possible.

- **Proof.** Let E be the set $\{y \in \mathbf{R} : y \geq 0 \text{ and } y^2 < 2\}$; thus E is the set of all non-negative real numbers whose square is less than 2. Observe that E has an upper bound of 2 (because if $y > 2$, then $y^2 > 4 > 2$ and hence $y \notin E$). Also, E is non-empty (for instance, 1 is an element of E). Thus by the least upper bound property, we have a real number $x := \sup(E)$ which is the least upper bound of E. Then x is greater than or equal to 1 (since $1 \in E$) and less than or equal to 2 (since 2 is an upper bound for E). So x is positive. Now we show that $x^2 = 2$.

- We argue this by contradiction. We show that both $x^2 < 2$ and $x^2 > 2$ lead to contradictions. First suppose that $x^2 < 2$. Let $0 < \varepsilon < 1$ be a small number; then we have

$$(x + \varepsilon)^2 = x^2 + 2\varepsilon x + \varepsilon^2 \leq x^2 + 4\varepsilon + \varepsilon = x^2 + 5\varepsilon$$

since $x \leq 2$ and $\varepsilon^2 \leq \varepsilon$. Since $x^2 < 2$, we see that we can choose an $0 < \varepsilon < 1$ such that $x^2 + 5\varepsilon < 2$, thus $(x + \varepsilon)^2 < 2$. By construction of E, this means that $x + \varepsilon \in E$; but this contradicts the fact that x is an upper bound of E.

- Now suppose that $x^2 > 2$. Let $0 < \varepsilon < 1$ be a small number; then we have
$$(x - \varepsilon)^2 = x^2 - 2\varepsilon x + \varepsilon^2 \geq x^2 - 2\varepsilon x \geq x^2 - 4\varepsilon$$
since $x \leq 2$ and $\varepsilon^2 \geq 0$. Since $x^2 > 2$, we can choose $0 < \varepsilon < 1$ such that $x^2 - 4\varepsilon > 2$, and thus $(x - \varepsilon)^2 > 2$. But then this implies that $x - \varepsilon \geq y$ for all $y \in E$ (why? If $x - \varepsilon < y$ then $(x - \varepsilon)^2 < y^2 \leq 2$, a contradiction). Thus $x - \varepsilon$ is an upper bound for E, which contradicts the fact that x is the *least* upper bound of E. From these two contradictions we see that $x^2 = 2$, as desired. □

- Next week we will use the least upper bound property to develop the theory of limits, which allows us to do many more things than just take square roots.

- We can of course talk about lower bounds, and greatest lower bounds, of sets E; the greatest lower bound of a set E is also known as the *infimum* of E and is denoted $\inf(E)$ or $\inf E$. Everything we say about supremum has a counterpart for infimum; we will usually leave such statements to the reader.

* * * * *

Cardinality of sets

- We have spent a lot of effort patiently building up our number systems, starting with the natural number system and working our way all the way up to the real number system. (We could continue onward, to the complex number system and even beyond, but we will not do so here). We will of course need the real number system to do all sorts of things, starting with limits and continuing through infinite series and to derivatives and integrals. However, for now, we will backtrack and return to the natural numbers to clear up one point - the connection between natural numbers and cardinality of sets.

- We defined the natural numbers axiomatically, assuming that they were equipped with a 0 and an increment operation, and assuming five axioms on these numbers. Philosophically, this is quite different from one of our main conceptualizations of natural numbers - that of *cardinality*, or measuring *how many* elements there are in a set. Indeed, the Peano axiom approach treats natural numbers more like *ordinals* than *cardinals*. (The cardinals are One, Two, Three, ..., and are used to count how many things there are in a set. The *ordinals* are First, Second, Third, ..., and are used to order a sequence of objects. There is a subtle difference between the two, especially when comparing infinite cardinals with infinite ordinals, but this is beyond the scope of this course, and is dealt with in Math 112). We paid a lot of attention to what number came *next* after a given number n - which is an operation which is quite natural for ordinals, but less so for cardinals - but did not address the issue of whether these numbers could be used to count sets. The purpose of this section is to correct this issue by noting that the natural numbers *can* be used to count the cardinality of sets, as long as the set is finite.

- The first thing is to work out when two sets have the same size: it seems clear that the sets $\{1, 2, 3\}$ and $\{4, 5, 6\}$ have the same size, but that both have a different size from $\{8, 9\}$. One way to define this is to say that two sets have the same size if they have the same number of elements, but we have not yet defined what the "number of elements" in a set is. Besides, this runs into problems when a set is infinite.

- The right way to define the concept of "two sets having the same size" is not immediately obvious, but can be worked out with some thought. One intuitive reason why the sets $\{1, 2, 3\}$ and $\{4, 5, 6\}$ have the same size is that one can match the elements of the first set with the elements in the second set in a perfect pairing: $1 \leftrightarrow 4$, $2 \leftrightarrow 5$, $3 \leftrightarrow 6$. (Indeed, this is how we first learn to count a set: we match the set we are trying to count with another set, such as a set of fingers on your hand). We will use this intuitive understanding as our rigorous basis for "having the same size".

- **Definition** Let $f : X \to Y$ be a function from one set X to another set Y. We say that f is *bijective* iff for every $y \in Y$ there is exactly one $x \in X$ such that $f(x) = y$.

- Being bijective is the same as being one-to-one and onto (see the supplemental handout on sets and functions). It is also the same as being *invertible*; if f is bijective, then there is an *inverse function* $f^{-1} : Y \to X$ such that $f(f^{-1}(y)) = y$ for every $y \in Y$, and $f^{-1}(f(x)) = x$ for every $x \in X$. (Indeed, we can define $f^{-1}(y)$ to be the unique element x for which $f(x) = y$).

- **Example.** Let $f : \{0, 1, 2\} \to \{3, 4\}$ be the function $f(0) := 3$, $f(1) := 3$, $f(2) := 4$. This function is not bijective because if we set $y = 3$, then there is more than one x in $\{0, 1, 2\}$ such that $f(x) = y$ (this is a failure of injectivity). Now let $g : \{0, 1\} \to \{2, 3, 4\}$ be the function $g(0) := 2$, $g(1) := 3$; then g is not bijective because if we set $y = 4$, then there is no x for which $g(x) = y$ (this is a failure of surjectivity). Now let $h : \{0, 1, 2\} \to \{3, 4, 5\}$ be the function $h(0) := 3$, $h(1) := 4$, $h(2) := 5$. Then h is bijective, because each of the elements 3, 4, 5 comes from exactly one element from 0, 1, 2.

- Remark. A common error is to say that a function $f : X \to Y$ is bijective iff "for every x in X, there is exactly one y in Y such that $y = f(x)$." This is not what it means for f to be bijective; it is what it means for f to be a *function*: each input gives exactly one output. A function cannot map one element to two different elements, for instance one cannot have a function f for which $f(0) = 1$ and also $f(0) = 2$. The functions f, g given in the previous example are not bijective, but they are still functions, since each input still gives exactly one output.

- **Definition** We say that two sets X and Y have *equal cardinality* iff there exists a bijection $f : X \to Y$ from X to Y.

- Thus, for instance, the sets $\{0, 1, 2\}$ and $\{3, 4, 5\}$ have equal cardinality, since we can find a bijection between the two sets. Note that we do not yet know whether $\{0, 1, 2\}$ and $\{3, 4\}$ have equal cardinality; we know that one of the functions f from $\{0, 1, 2\}$ to $\{3, 4\}$ is not a bijection, but we have not proven yet that there might still be some other bijection from one set to the other. (It turns out that they do not have equal cardinality, but we will prove this a little later). Note that this definition makes sense regardless of whether X is finite or infinite (in fact, we haven't even defined what finite means yet).

- Note that two sets having equal cardinality does not preclude one set containing the other. For instance, if X is the set of natural numbers and Y is the set of even natural numbers, then the map $f : X \to Y$ defined by $f(n) := 2n$ is a bijection from X to Y (why?), and so X and Y have equal cardinality, despite Y being a subset of X and seeming intuitively as if it should only have "half" of the elements of X.

- The notion of having equal cardinality is an equivalence relation:

- **Proposition 29** Let X, Y, Z be sets. Then X has equal cardinality with X. If X has equal cardinality with Y, then Y has equal cardinality with X. If X has equal cardinality with Y and Y has equal cardinality with Z, then X has equal cardinality with Z.

- **Proof.** See Week 2 Homework. □

- Let n be a natural number. Now we want to say when a set X has n elements. Certainly we want the set $\{i \in \mathbf{N} : 1 \leq i \leq n\} = \{1, 2, \ldots, n\}$ to have n elements. (This is true even when $n = 0$; the set $\{i \in N : 1 \leq i \leq 0\}$ is just the empty set). Using our notion of equal cardinality, we thus define:

- **Definition** Let n be a natural number. A set X is said to have *cardinality* n, iff it has equal cardinality with $\{i \in \mathbf{N} : 1 \leq i \leq n\}$. We also say that X *has* n *elements* iff it has cardinality n.

- One can use the set $\{i \in \mathbf{N} : i < n\}$ instead of $\{i \in \mathbf{N} : 1 \leq i \leq n\}$, since these two sets clearly have equal cardinality (why? What is the bijection?).

- **Example.** Let a, b, c, d be distinct objects. Then $\{a, b, c, d\}$ has the same cardinality as $\{i \in N : i < 4\} = \{0, 1, 2, 3\}$ or $\{i \in N : 1 \leq i \leq 4\} = \{1, 2, 3, 4\}$ and thus has cardinality 4. Similarly, the set $\{a\}$ has cardinality 1. The empty set $\{\}$ has cardinality 0 (why?).

- There might be one problem with this definition: a set might have two different cardinalities. But this is not possible:

- **Proposition 30**. Let X be a set with some cardinality n. Then X cannot have any other cardinality, i.e. X cannot have cardinality m for any $m \neq n$.

- Before we prove this Proposition, we need a lemma.

- **Lemma 31.** Suppose that $n \geq 1$, and X has cardinality n. Then X is non-empty, and if x is any element of X, then the set $X - \{x\}$ (i.e. X with the element x removed) has cardinality $n - 1$.

- **Proof.** If X is empty then it clearly cannot have the same cardinality as the non-empty set $\{i \in \mathbf{N} : 1 \leq i \leq N\}$, as there is no bijection from the empty set to a non-empty set (why?). Now let x be an element of X. Since X has the same cardinality as $\{i \in \mathbf{N} : 1 \leq i \leq N\}$, we thus have a bijection f from X to $\{i \in \mathbf{N} : 1 \leq i \leq N\}$. In particular, $f(x)$ is a natural number between 1 and N. Now define the function $g : X - \{x\}$ to $\{i \in \mathbf{N} : 1 \leq i \leq N - 1\}$ by the following rule: for any $y \in X - \{x\}$, we define $g(y) := f(y)$ if $f(y) < f(x)$, and define $g(y) := f(y) - 1$ if $f(y) > f(x)$. (Note that $f(y)$ cannot equal $f(x)$ since $y \neq x$ and f is a bijection). It is easy to check that this map is also a bijection (why?), and so $X - \{x\}$ has equal cardinality with $\{i \in \mathbf{N} : 1 \leq i \leq N - 1\}$. In particular $X - \{x\}$ has cardinality $n - 1$, as desired. □

- **Proof of Proposition 30.** We induct on n. First suppose that $n = 0$. Then X must be empty, and so X cannot have any non-zero cardinality. Now suppose that the Proposition is already proven for some n; we now prove it for $n++$. Let X have cardinality $n++$; and suppose that X also has some other cardinality $m \neq n++$. By Proposition 29, X is non-empty, and if x is any element of X, then $X - \{x\}$ has cardinality n and also has cardinality $m - 1$. By induction hypothesis, this means that $n = m - 1$, which implies that $m = n++$, contradiction. This closes the induction □

- Thus, for instance, we now know, thanks to Proposition 30 (and Proposition 29), that the sets $\{0, 1, 2\}$ and $\{3, 4\}$ do not have equal cardinality, since the first set has cardinality 3 and the second set has cardinality 2.

- **Definition** A set is *finite* iff it has cardinality n for some natural number n; otherwise, the set is called *infinite.*

- Thus, for instance, the sets $\{0, 1, 2\}$ and $\{3, 4\}$ are finite, as is the empty set (0 is a natural number). However, some sets are infinite:

- **Theorem 32.** The set of natural numbers $\mathbf{N}$ is infinite.

- **Proof.** Suppose for contradiction that the set of natural numbers $\mathbf{N}$ was finite, so it had some cardinality n. Then there is a bijection f from $\{i \in \mathbf{N} : 1 \leq i \leq n\}$ to $\mathbf{N}$. Now consider the sequence $f(1), f(2), \ldots, f(n)$. By Lemma 8, this sequence is bounded by some rational number, which then implies (e.g. by Proposition 23) that it is bounded by some natural number M. Thus $f(i) \leq M$ for all $1 \leq i \leq n$. But then the natural number $M + 1$ is not equal to any of the $f(i)$, contradicting the hypothesis that f is a bijection. □

- One can also use similar arguments to show that any unbounded set is infinite; for instance the rationals $\mathbf{Q}$ and the reals $\mathbf{R}$ are infinite. However, in the next week's notes we will show that the real numbers are in some sense "more" infinite than the rationals and natural numbers.

* * * * *

Appendix: Exponentiation of reals.

- In the beginning of this week's notes we defined exponentiation x^n when x is rational and n is a natural number, or when x is a non-zero rational and n is an integer. Now that we have all the arithmetic operations on the reals (and Proposition 19 assures us that the arithmetic properties of the rationals that we are used to, continue to hold for the reals) we can similarly define exponentiation of the reals.

- **Definition** Let x be a real number. To raise x to the power 0, we define $x^0 := 1$. Now suppose that recursively that x^n has been defined for some natural number n, then we define $x^{n++} := x^n \times x$.

- **Definition** Let x be a non-zero real number. Then for any negative integer $-n$, we define $x^{-n} := 1/x^n$.

- Clearly these definitions are consistent with the definition of rational exponentiation given earlier. We can then assert

- **Proposition 33.** All the properties in Proposition 3 and 4 remain valid if x and y are assumed to be real numbers instead of rational numbers.

- Instead of giving an actual proof of this proposition, we shall give a metaproof (an argument appealing to the nature of proofs, rather than the nature of real and rational numbers).

- **Meta-proof.** If one inspects the proof of Propositions 3 and 4 we see that they rely on the laws of algebra and the laws of order for the rationals (Propositions 20 and 22 of Week 1 notes). But by Propositions 14, 19, and the identity $xx^{-1} = x^{-1}x = 1$ we know that all these laws of algebra and order continue to hold for real numbers as well as rationals. Thus we can modify the proof of Proposition 3 and 4 to hold in the case when x and y are real. $\square$

- In the rest of the course we shall now just assume the real numbers to obey all the usual laws of algebra, order, and exponentiation.