Math 2300H. Days 1-3. What are real numbers?
Real numbers are the numbers used to measure lengths. (They were essentially invented for this purpose, hence this is the best way to understand them, although later it turns they can also be used to measure other quantities such as areas.) Imagine an ideal line, infinitely long in both directions, straight, and continuous without breaks or gaps. Fix a point to begin at, called 0 (zero), and fix another point to be called 1 (one), which defiens a choice of "unit length". Then there should be exactly one real number for every point on this line, such that the number measures how far that point is from the point 0, assuming the point 1 is one unit away. Positive numbers correspond to points on the same side of 0 as 1, and negative numbers correspond to those points on the opposite side of 0 from the point 1. Then how do we represent real numbers by symbols? And how do we add and multiply these numbers using those symbols? Possibly the best way is using decimals.

A <u>finite</u> <u>decimal</u> is a finite sequence of form $a_1a_2a_3....a_n.b_1b_2....b_m$, where each a_i and each b_j is one of the ten digits {0,1,2,3,....9}. A finite decimal corresponds to a point on the real line as follows. For example, 14.63 corresponds to the point constructed like this: first lay off 14 copies of the unit length, the first one being at 1, the second one (called 2) being one unit on the opposite side of 1 from 0, and the third one (called 3) on the opposite side of 2 from 1, and so on, until we come to the 14 th point (called 14). Then lay off another unit ending at 15. Then subdivide the interval between 14 and 15 into ten equal parts, with the end points of the 6th subinterval being called 14.6 and 14.7. Then subdivide that 6th subinterval again into ten equal parts and go out to the 3rd subinterval. The initial point of that subinterval is the point corresponding to 14.63. In this way one can assign to any finite decimal a point on the real line.

Not every point of the real line occurs as one of the points corresponding in this way to finite decimals however. For instance the point (called 1/3) lying one third of the way between 0 and 1 does not correspond to a finite decimal. It lies to the right of the all points corresponding to finite decimals of form { .3, .33, .333, .3333, .33333, ...}, but to the left of any point of form { .4, .34, .334, .3334, .33334, ...}. However since the points of form { .3, .33, .333, .3333, .33333, ...} get arbitrarily close to the point 1/3, any point to the left of 1/3 will lie to the left of one ofthe points { .3, .33, .333, .3333, .33333, ...}. For example if we take a point which is 1/1000 to the left of 1/3, then it will be to the left of the point .3333, which is within 1/10,000 of 1/3. Thus 1/3 is “the leftmost point which is not to the left of any finite decimal of form { .3, .33, .333, .3333, .33333, ...}”, i.e. 1/3 is the “smallest number not smaller than any of the numbers { .3, .33, .333, .3333, .33333, ...}”, technically we say 1/3 is the “least upper bound (l.u.b.) of the numbers { .3, .33, .333, .3333, .33333, ...}”. Although 1/3 does not equal any one of these finite decimals, this is a description of the point 1/3 in terms of the whole infinite sequence { .3, .33, .333, .3333, .33333, ...} of finite decimals. It is usual to replace the infinite sequence { .3, .33, .333, .3333, .33333, ...} of finite decimals simply by the one infinite decimal .3333333........ (3’s continuing forever), sometimes denoted by .3333ä3.... where the bar over the last 3 indicates infinite repetition of that symbol.

In this way every point of the real line can be described by either a finite decimal or an infinite decimal. I.e. given a point x on the line, to the right of 0 for example, to get the integer part of the decimal measure off copies of unit interval starting at 0, until the next unit interval will go past the point x. If x lies strictly between the 5th and the 6th point, for instance, then the integer part of the decimal for x is 5. Then subdivide that interval again into ten equal parts and see whether x lies exactly on one of the subdivision points. If it does lie on say the 2nd subdivision point, then x corresponds to the finite decimal 5.2. if x does not lie on one of the subdivision points but lies between say the 2nd and the third subdivisions points, then the second decimal approximation to x is 5.2. Continue in this way to subdivide and approximate x by decimals. If eventually x lies exactly on some subdivision point then x corresponds to a finite decimal. if x never lies on any subdivision point, as was the case with 1/3, then x corresponds to an infinite decimal. Thus each point of the line can be represented by a finite or infinite decimal. We often call the finite ones infinite decimals also, where we assume they are made to look infinite by writing an infinite number of zeroes after they stop. This makes the language easier and we can just say "every point of the real line corresponds to an infinite decimal". (Not all infinite decimals can be obtained in this way from points on the line. Try to convince yourself that this procedure will never lead to an infinite decimal ending in all 9's repeating forever.)

The other direction is harder, i.e. if we start with an infinite decimal, does it always correspond to a point of the real line? We could try to find the point, starting from the decimal as follows. If we have a finite decimal like 3.7 there is no problem, it is easy to find the corresponding point. Just go out to the fourth unit interval after 0, between the points 3 and 4, subdivide into ten equal parts and take the 7th subdivision point to be 3.7. But if the decimal is infinite, it is not so obvious. Say we have the decimal D = .12122122212222......... Does this correspond to a point x? Well first we subdivide the interval between 0 and 1 into ten equal parts and we consider the first subdivision point called .1. Then we know x lies to the right of .1. then we subdivide again and take the 2nd subdivision point in the subinterval, the point 1.2, and we know x lies to the right of that point.

Continuing in this way we find an infinite number of points (if we live long enough, otherwise we must imagine it) and we know the point corresponding to x should lie to the right of all of them. But it should also be the closest point which is to the right of all of them., So we describe the point x corresponding to an infinite decimal D as "the leftmost point which is to the right of all points correspondijg to finite decimal approximations of D", i.e. x is the lub of all finite decimal approximations to D". But how do we know there is such a point? We do not. But it seems plausible at least if the real line is truly supposed not to have any holes in it, so we take this as an axiom, or unproved fact about the real line. This is called the "least upper bound axiom": <u>For every infinite decimal, the sequence of finite decimal approximations has a least upper bound on the real line.</u>

Stated as fact about real numbers, it is usual to assume it in the following more general form: **Least upper bound axiom**: "If a set of real numbers is non empty and has an upper bound, then it has a <u>least</u> upper bound".

This concept can be used to describe many familiar numbers and solutions to many problems:
Examples:
(i) (assuming we know how to find the length of line segments and hence the perimeter of a polygon), the number π can be described as the lub of the lengths of all polygons inscribed in the unit semi circle. I.e. if you inscribe any polygon in the unit semi circle, the perimeter of that polygon will not be greater than π, but if you take a polygon with small enough sides, its perimeter will be as close as you like to the number π, i.e. π is the smallest number not smaller than any of those perimeters.

But how can we calculate this number, i.e. how can we find some of its finite decimal

approximations?
(ii) If we know how to find the area of a triangle and hence of a polygon, we can define the area of a circle as the lub of the areas of all inscribed polygons. But how can we show that this area is actually equal to πr^2, where π is defined above and r is the radius of the circle?

(iii) If we want to know what is meant by the value of an infinite sum like $1 + 1/2 + 1/4 + 1/8 + 1/16 + ...$, we can say it is the lub of all the finite "partial" sums $\{ 1, 1 + 1/2, 1 + 1/2 + 1/4, 1 + 1/2 + 1/4 + 1/8, ... \}$. But how can we actually calculate this sum, i.e. can we find this least upper bound?

(iv) If we want to find the slope of the parabola $y = x^2$ at the point (1,1), we can say it is the lub of the slopes of all the secant lines drawn through points of the form (x,x^2) and (1,1) where $x < 1$. But can we actually calculate this slope?

(v) If we want to describe the "square root of 2" we can say it is the lub of all finite decimals whose square is less than 2. (Since the square of a finite decimal is never 2, as you can easily check, the square root of 2 is going to be an infinite decimal, and it is not so easy to even tell how to square an infinite decimal. In fact the only way we have to do that, is to say that the square of an infinite decimal is the lub of the squares of all its finite decimal approximations!) Can we compute, or at least approximate this infinite decimal?

(vi) The cosine function, in radians, is defined as follows: given a positive real number t, measure off an arc of length t along the unit circle, starting at (1,0) going counterclockwise. Then the x coordinate of the point reached is cos(t), and the y coordinate is sin(t). But can we actually calculate say cos(1)?

All these problems have answers provided by calculus. For example, cos(t) is given by the infinite formula $\cos(t) = 1 - x^2/2! + x^4/4! - x^6/6! \pm ...$, where n! = "n factorial" = (1)(2)(3)....(n) is the product of the numbers between 1 and n. Cos(t) can be computed to any desired degree of accuracy by taking enough terms of this formula. For example, cos(1) is the least upper bound of the sequence of approximations $\{1-1/2, 1 -1/2 + 1/24 - 1/720, ...\}$ formed as above by taking finite partial sums ending in a negative term.

Actually computing answers to problems
It is one thing to describe the answer to a problem as a lub of some set of numbers, but it is usually more desirable to actually find the answer in a nice simple form, or at least approximate it as well as we want. This is often not so easy, and may depend on the problem at hand. Thus there are two parts to solving most problems:

1) Describe the solution in precise terms, even if abstract ones.
2) Actually calculate that answer, say as a decimal, or at least show how to find as good a finite decimal approximation as we want. Sometimes we calculate the answer in terms of some other "known" number, such as when we say the area of a circle is πr^2, even if we may not know exactly how to calculate π.

Even step 1) above has two parts:
1a) decide whether the problem has a solution, and if so,

1b) describe it.

For example, if the solution of a problem is defined as the lub of some set of real numbers, to show it exists all we have to do by the lub axiom is prove the set is non empty and has some upper bound.
For example, to prove the infinite sum 1 + 1/2 + 1/4 + 1/8 +....... has a finite value, described as the lub of all the finite sums {1, 1 + 1/2, 1 + 1/2 + 1/4, ...} we must show there is an upper bound to these finite sums. But it is not hard to see these finite sums are never greater than 2, so 2 is an upper bound. Then the axiom tells us there is a least upper bound, which in fact turns out also to be 2.

The finite partial sums of the sequence 1 - 1/3 +1/5 - 1/7 + 1/9 - 1/11 ±....... are bounded above by 1, hence have a least upper bound, WAIT!! OOOOPS! The sum of this sequence is not the lub of all those finite partial sums since the minus signs cause the finite sums to go back and forth on both sides of the actual infinite sum. (Now is when we need the more general notion of "limit" instead of lub.) Anyway we can finesse this and say (correctly) the value of the infinite sum 1 - 1/3 +1/5 - 1/7 + 1/9 - 1/11 ±....... is the lub of the finite partial sums
{1 - 1/3, 1 - 1/3 +1/5 - 1/7, 1 - 1/3 +1/5 - 1/7 + 1/9 - 1/11, ...}.

I.e. if we are careful to always take partial sums which end in a negative term then they are actually smaller than the infinite sum we are trying to define. Thus we can say that 1 is an upper bound for THESE finite sums so there is a lub. But what is the lub ? It turns out to be $\pi/4$, rather amazing. In the case of the infinite sum 1 + 1/4 + 1/9 + 1/16 + 1/25 + ..., where the nth denominator is the square of the integer n, it is not even so easy to find any upper bound at all (until you know about how to compute area formulas by integral calculus). The least upper bound of these finite sums turns out to be $\pi^2/6$, incredibly. Not only that, Leonhard Euler knew this before the invention of calculus!!

Euler also knew how to evaluate the sum 1 + 1/16 + 1/81 + 1/ 243 + ..., where the nth denominator is the 4th power of the integer n, namely $\pi 4/90$, and he knew many more such even power sums and included them as essential material in his famous "PRECALCULUS" book! However I do not believe even today that anyone knows the value of 1 + 1/8 + 1/27 + 1/64 +..... where the nth denominator is the cube (or any other odd power) of n. I.e. these finite sums have an upper bound, and thus also a <u>least</u> upper bound, but no one knows how to describe this least upper bound in terms of any other known numbers.

Differential calculus is about how to:
1) describe the answer to the slope problem for the graph of a function in terms of "limits", and
2) how to actually calculate these limits to calculate the slope of y = f(x) at least as well as we know how to calculate f(x) itself.

Thus for a nice easy function like a polynomial $f(x) = 3x^2-6x+9$, we should be able to calculate the slope also as a polynomial. but for a trigonometric function like f(x) = cos(x) we will only be able to calculate the slope function as another trigonometric function. (In a later math course, when we know the infinite formula given above for cosine, we will also get an infinite formula for the slope of the graph of cosine.) For a more difficult function like 2^x, or $\log_2(x)$ (the logarithm "base 2" of x), the derivative will be also a challenge.

You have probably heard of "natural logarithms", or logarithms to the base "e". We will define this magic number "e" as the unique base such that the slope of the graph of $y = e^x$ at the point (0,1) equals 1. But then what is the number e? calculus can beused to give a very simple formula for the function $e^x = 1 + x + x^2/2! + x^3/3! + x^4/4! + \ldots\ldots.$, and this can be used to approximate e very well, by plugging in x = 1 and adding up a few terms. It turns out e is between 2.71828 and 2.71829.

Rather than continuing to restrict ourselves to the concept of least upper bounds, it is more useful to use the concept of “limits”. These are harder to define precisely, and harder to prove the existence of, but easier to deal with intuitively. Thus in practice we will find it convenient to use this concept, since there are some good methods for actually computing these “limits”, using the notion of a “continuous function”. This is our next topic of study. For example, if we approximate the tangent line to $y = x^2$ at (1,1), by the secant line through the points (1,1) and (x,x^2), where $x < 1$, we can describe the slope of the tangent line as the lub of the slopes of all these secant lines, i.e. the lub of all numbers of form $(x^2-1)/(x-1)$ where $x < 1$. Simplifying the fraction gives x+1, and if x is any number < 1, the smallest number not smaller than any of the numbers x+1, is 2.

We might wonder though whether we get the same slope if we approximate from the right, looking at numbers of form $(x^2-1)/(x-1)$ where $x > 1$. These simply again to x+1, for $x > 1$, but this time the slope of our curve should be less than all these numbers. Thus we can describe our slope as the smallest number not smaller than any of the numbers x+1 for $x > 1$, i.e. as the greatest lower bound (glb) of these numbers. this is again 2.

However it is simpler to say that the slope of the tangent line is the number being approximated by the numbers $(x^2-1)/(x-1)$, when x is approximately 1, without worrying about whether the approximation is too small or too large. Thus again we are asking what number is approximated by x+1 when x is approximately 1. It seems clear that when x is approximately 1, then x+1 is approximately 2. Of course since we have not precisely defined what we mean by “approximately”, you may not feel this is so obvious. I will try to give you a feel for how to compute these limits, and will also give the rigorous precise definition of limit.

2300H Continuity and limits

We will take the idea of continuous functions as a basic notion. We will discuss it intuitively and state some properties of it which we hope are believable, and then derive some consequences. Afterwards we will go back and define continuity precisely and justify, i.e. prove, the properties we have assumed. To begin with, we say that a function f defined at a, is continuous at a if and only if the values of f(x) for $x \neq a$ are good approximations to the value f(a). For example we believe the area function of a circle πr^2 is continuous as function of boith π and r, since we believe that if we plug in 3.14 for π, which is only close to π, but not exactly equal to it, that we will get a close aproximation to the area. Similarly if the actual radius is $2^{1/2}$

and we plug in r = 1.414 which is only an approximation to $2^{1/2}$ we belive we will get a good approximation to the area.

So intuitvely a function f is continuous at a if plugging in values x which are close to a, gives us values f(x) which are close to f(a). This is not a precise definition of continuity since we have not defined what we mean by “close”, i.e. how close is “close”? Nonetheless we will assume we have some idea of what this means for now. For example, addition is a continuous opreation since if x and y are close to a and b respectively, then x+y is close to a+b. Multiplication is a continuous operation since if x and y are close to a and b respectively, then xy is close to ab. It follows that if f and g are continuous functions then so are f+g and fg. Now to begin with, it seems obvious that the function f(x) = x is continuous, since surely if x is close to a, then f(x) = x is also close to f(a) = a, and that this is true no matter we mean by “close”.

Thus if we assume that f(x) = x is continuous, and sums and products of continuous functions are continuous, then it follows that all functions made up using addition and multiplication starting from the function x are continuous. Thus $xx = x^2$ and $xx^2 = x^3$, ..., x^n, are all continuous. Moreover constant functions are surely continuous since if f(x) = c for all x, then for x near x certainly f(x) = c is close to f(a) = c, since surely c is close to c no matter what close means.

Thus all functions of form cx^n and also $cx^n + dx^m$ are continuous, and more generally all functions of form $f(x) = a_0x^n+a_1x^{n-1} +....a_{n-1}x + a_n$ are continuous, i.e. all polynomials are continuous. Then also note that division is continuous since if x is near a then 1/x is near 1/a, unless of course a is zero, since then 1/a is not defined. Moreover for x near 0 and x positive we have 1/x very large and positive, while if x is near zero but negative then 1/x is very large and negative. Thus if f and g are two polynomials then f/g is continuous at a unless g(a) = 0. I.e. all rational functions are continuous at points where the denominator is not zero.

Furthermore if we look at the circle definition of the function cos, it seems clear that this function is continuous everywhere, since if two points q, w on the unit circle are close together, then their x coordinates, i.e. cos(q) and cos(w), are even closer together. Similarly their y coordinates, i.e. sin(q) and sin(w), are also closer together than are the points q and w, so sin is also continuous everywhere. Since tan = sin/cos, it follows that tan is continuous at any point where cos is not zero. It is not so obvious but the square root function, cube root function, and other root functions are continuous where defined, i.e. odd root functions (cube root, fifth root, seventh root,....) are continuous everywhere and even root functions (square root, fourth root,....) are continuous at all non - negative numbers.

Moreover, exponential functions such as a^x where a is any positive real number, are continuous everywhere, and log functions such as $\log_a(x)$ are continmuous at all positive numbers. These exponential and log functions take some work even to define carefully, and it is also a lot of work to prove they are continuous. So for now we will take our intuitive feel for continuity, state some true facts about continuity and use them to find some simple limits.

If f is a function defined near a but not necessarily at a, then it may or may not be possible to define f(a) so as to make f continuous at a. If it is possible, the unique value of f(a) that will make f continuous at a is called the limit of f as x approaches a. I.e. f is continuous at a if and only if the values f(x) for x near a are good approximations to f(a). Now we can look at the values of $f(x)$ for x near a and ask ourselves what value they are approximating as x approaches closer and closer to a. If there is no such value then f has no limit as $x \to a$. If there is such a value, say L, then seting $f(a) = L$ will make f continuous at a, and we say the limit of $f(x)$ as $x \to a$ is L.

Now f may or may not already be defined at a, and if f is defined at a, it may have the value L and it may have some other value. If it already has the value $f(a) = L$ then it is already continmuous, and if $f(a)$ has some other value then f is not continuous at a. None of this is relevant to the question of whether f has a limit at a. I.e. f has a limit at a if either of two things is true: either f is already continuous at a, or f can be made to become continmuous at a by defining or possibly redefining the value $f(a)$ appropriately. Thus certainly f has a limit at if f is continuous at a, and in that case the limit of $f(x)$ as x¨a is $f(a)$. But if $f(x) = 6$ say for all $x \neq a$, while $f(a) = 3$, then f is not continuous at a, since the values $f(x)$ for $x \neq a$ but x near a do not give good approximatiions to $f(a)$.

Our f would <u>become</u> continuous at a however if we were to redefine the value $f(a)$ to be 6. Thus we say f has a limit at a and that limit is 6. In all cases, f has at most one limit at a, i.e. if f is not continuous at a, then either f cannot be made continuous at a, or else there is exactly one value L which will make f continuous if we set $f(a) = L$. This is a very useful fact because it gives us a way to find limits of functions. I.e. if f is defined for $x \neq a$, but f is not defined at a, or if f is defined at a but still f is not continuous at a, and we wonder whether f has a limit at a, then if we can come up somehow with another function g such that $g(x) = f(x)$ for all $x \neq a$ and x near a, and g is continmuous at a, then g is the continuous function which we are trying to make f into, so the limit of $f(x)$ as x¨a must be $g(a)$. Thus we have two basic rules for finding limits:

1) Easy limits: If f is continuous at a then $\lim_{x \to a} f(x) = f(a)$.
Examples: If $f(x) = \cos(x)$, then $\lim_{x \to \pi} f(x) = \cos(\pi) = -1$. This is because we are taking it for granted for now that cos is continuous everywhere. If $f(x) = (x^4-7x+3)^{1/2}$, then $\lim_{x \to 2} f(x) = 5^{1/2}$, because polynomials are continuous everywhere and square roots are continuous at all positive numbers, and 5 is positive.

2) Limits which can be made into easy ones: If f is defined for all x near a (but not neceesarily at a), and if g is another function which is also defined for x near a and also at a, and if $g(x) = f(x)$ for all $x \neq a$ and x near a, and if also g is continuous at a, then f has a limit at a and $\lim_{x¨a} f(x) = g(a)$.

Examples: If $f(x) = \{x^2-1\}/(x-1)$, then $f(x)$ is equal to $g(x) = x+1$ everywhere except at $x = 1$. Hence these functions have the same limit as $x \to 1$. Since g is continuous we have $\lim_{x \to 1} \{x^2-1\}/(x-1) = \lim_{x \to 1} (x+1) = 2$. If $f(x) = \{(1/x) - 1\}/(x-1)$, and we multiply through by x/x and simplify we get $g(x) = -1/x$, where f and g are equal everywhere except at 1 and 0. Hence they have the same limit as x approaches 1. So $\lim_{x \to 1} \{(1/x) - 1\}/(x-1) = \lim_{x \to 1} (-1/x) = -1$.