CHAPTER 7
When there is no change, we're in trouble.

LECTURE 7-1

As I turned the corner, I was brought up short by the presence of a high brick wall about 40 yards down the street and extending completely across it...

When motion is uniform, I have a function that relates time and distance traveled. This function is explicit in the sense that I can write down its rule explicitly using algebra.

When motion is non-uniform, I have a function that models position by relating time and distance traveled but the function is often implicit in the sense that I can't write the rule in any usable form. I feel fortunate that in the example of the falling rock, I have the function that models distance traveled explicitly.

When motion is uniform, I have a function that models motion by relating time and speed.

So what is the function that models motion in the non-uniform case? Where is the 'speed' function? And what does 'speed' even mean if the motion isn't uniform?

Basically, I want a number that describes motion at an instant of Ideal World time. I'm going to call this number 'speed' because it describes the same property of motion that speed defines if the motion is uniform. It is not clear to me, at the moment, how to define 'speed' for non-uniform motion but I suspect it will be conceptually different than speed in uniform motion.

I'm going to use '...' around the word, speed, when I am referring to the non-uniform motion, 'speed'. One might make an argument to use a different word altogether but the non-uniform 'speed' is enough like uniform speed that I have no problem using the same word.

If I have an interval of time where the motion is uniform, I can divide the distance traveled by the length of that interval to find the speed, which is the same at each point of the interval. If the motion is not uniform, the interval where the motion doesn't change reduces to an instant of time and, since its length is zero, I can't divide by it. This is the problem in <u>defining</u> 'speed' for non-uniform motion in the Ideal World.

In the case of uniform motion, the speed is the same at each point in an interval of time and I can multiply speed by the length of the interval to get the distance traveled in that time. In my opinion, the ability to give distance is the major characteristic of 'speed'. In the case of non-uniform motion, the 'speed' changes each instant and there is no single speed to multiply by the length of any interval of time and get distance traveled. This is the problem of <u>using</u> 'speed' for non-uniform motion in the Ideal world.

I recall driving through downtown St. Louis on a snowy winter evening, hitting all the lights red, and noticing that the speedometer started at zero, rose steadily to 20 miles/hour as I accelerated, stayed there for a few moments and then dropped back to zero as I stopped for the next light. It seemed to me that I must have gone 10 miles/hour in there somewhere; the 'principle of continuity' would seem to imply that. But what did '10 miles/hour' mean? Since I was going faster, my motion was greater at the end of any interval of time than it was at the beginning. There was no positive interval of time where my motion was uniform so that I could divide a distance by a positive length of time and get speed. I must have been traveling 10 mph for only an instant, for a zero length of time, and moved through a single point, that is, moved zero distance. I could only compute speed as 0/0, which is a meaningless symbol.

I realized then that understanding differential calculus was understanding what that 10 miles/hour meant. I was momentarily stunned that the basic idea of differential calculus was captured in this question, a question that I encountered every day if I thought about it. Now, of course, I do.

I was assuming that I was driving in the Ideal World St. Louis. In the Ideal World St. Louis, no matter how small the interval of time is, I can detect that I am moving faster at the end of the interval than at the beginning. Even if the length of the interval is $10^{-10000000000}$ seconds, it is still a positive length of time. In the Real World, this length would be zero, but in the Ideal World it is not particularly special, just an everyday positive interval of time.

I make a big deal of the positivity because if the length of the time interval is positive, in particular, if the length is not zero, I can divide by it. The only reason that being non-zero is important for a number is that then it can be a divisor. If I am told that a number is non-zero, I know that sometime, somewhere, I'm going to have to divide by it. Although, if I am going to a movie with a friend and she says she has zero money, I doubt that I will have to divide by that particular zero.

The problem with 'zero length' instants of time is an Ideal World problem. Real World instants of time have a positive length. The Real World approach to 'speed' is to use average speed, which gives a global description of non-uniform motion, and use it on small intervals. If the averaging time interval is small enough, the motion has no <u>measurable</u> change, the motion appears uniform on the interval, and the <u>average speed is the same as speed</u>. If [2 , 2.01] were such an interval, then it would make sense to say that the average speed over this interval gives the 'speed' at t = 2. This 'speed' is an average speed and so has units of distance/time. It has the 'feel' of speed.

This is a useable idea in the Real World. At a drag strip, two cars or motorcycles run a quarter mile from a standing start as fast as they can. Two numbers are given the winner of the race, the elapsed time and the miles per hour at the end of the quarter mile. This last number is supposed to be the 'speed' at the instant of time when the winning vehicle hits the finish line.
One way of doing this uses a device that shoots a ray of light across the track into a

photo-electric cell. When the ray of light is broken, the photo-electric cell can either stop or start a clock. One of these is put at the finish line and another, say, a foot beyond the finish line. When the vehicle breaks the first ray, a clock is started, and when the second ray is broken, the clock is stopped. One foot or 1/5280 miles is divided by the time in hours it took to travel this foot, and this average speed is given as the speed at the finish line.

The motion isn't going to change much during this foot so the motion is very close to uniform and this number seems a reasonable choice for speed. They give the speed to an accuracy of .01 miles per hour at a drag strip and for this degree of accuracy, they could put the second timing device 10 or 20 feet behind the finish line; or in front of the finish line for that matter.

It seems to me that no matter how erratic the motion is, if I pick any time during the motion, say 2 seconds, then there is some interval of time both before and after 2 seconds, where I can measure no change in motion. My perception is that during these intervals before and after 2 seconds, the motion is uniform, and I feel OK about using the average speed over these intervals to give the 'speed' at 2 seconds.

If the motion is changing very rapidly, then there may be problems finding the very small intervals of time where the motion appears uniform. If the object is very small, there may be problems measuring the distances traveled during these small intervals of time. The movement of glaciers has its own set of problems.

I have described a technique to find a number that gives the 'speed' of an object in the Real World at a Real World point in time. Since it is, at root, an average speed, the units are distance per unit time. I have a number for 'speed' so why go into the Ideal World at all? Why not define 'speed' in the Real World using one of these average speeds.

There are a couple of reasons why not.

The length of the Real World instant of time changes when I change my measuring instrument. As the accuracy of the measurement gets better, I can detect changes in motion over smaller time intervals so the interval where the motion looks uniform gets smaller. When the length of the interval gets smaller, the average speed over the interval changes, and the 'speed' changes. It may not change much, but it changes. This method does not lead to a number where I can say, "This number and only this number describes the motion at t = 2, or whenever." Also, since the Real World instant is not a fixed, well defined object, it is not clear exactly where or when I am describing motion if I use Real World instants. I don't think that an instantaneous description of motion using a single number is possible in the Real World. I don't think the description of any physical quantity using a single number is possible in the Real World.

Suppose I want to define the concept of length. In the Real World I measure a board with a meter stick that is only accurate to a few decimal places. The end of the board falls on a mark of the meter stick, which has width, or it falls between two marks. Either way, an exact, repeatable length as a single, well defined number is unattainable in the Real World.

In the Ideal World, I can lay my ideal board down on a number line and measure it <u>exactly.</u> This is my definition of length in the Ideal World.

If I have a Real World board, I assume there is an Ideal World board that models it and that my measurements approximate the length of the Ideal World board. I behave as if the length of the Real World board is the length of the Ideal World board. I never know exactly what that ideal length is because my only access to it is through measurements that approximate it..

I think that it is an act of faith for a person to believe that the board has an ideal length for measurements to approximate, although some would say that experiment supports this belief. My own personal belief is that there is not one length but a statistical distribution of lengths, much like the energy state of an atom. Since these variations in length are too small to be measured, the measurements of the real board seem to cluster around one number which I take to be the length of the ideal board that models it. If, someday, I can measure these variations, then perhaps I'll change my model, but until then I'll go with the 'one unique length' model. My model has to be consistent with the measurements I do have, and not based on measurements that might happen. Of course, the number my measurements cluster around isn't there, it's in the Ideal World.

The fact is, I do have confidence that what I find out about Ideal World objects will be more or less true about their counterparts in the Real World under similar circumstances. A rational reason for this confidence comes from over 300 years of fairly successful prediction. Universities have Colleges of Engineering and Departments of Physics based on the Ideal World model. My radio usually works and airplanes seem to fly successfully over my house.

Because of this confidence I spend more and more time in the Ideal World. I can make simplifying assumptions in the Ideal World that make problems easier. I have all the numbers of the Ideal World available to me and I can assume measurements of infinite accuracy. I can devise algorithms that give me numbers that would be both difficult and expensive in time and money to find using Real World experiments. I spend some time in the Real World modeling and interpreting, but most of my time is spent defining and solving in the Ideal World. I find it very comfortable there.

My 'point of view' about 'speed' in the Real World is similar to my 'point of view' about the length of a board in the Real World and I make no claim for its "truth" because I don't know what the "truth" really is. I do, however, need a 'point of view' and I have one that I like, true or not. I do not believe in the Darwinian theory of evolution but, having nothing to offer in its place, I think it's fun to assume its truth for the sake of argument.

It is my 'point of view' that the experimental average speeds do, indeed, get close to <u>one</u> number if I average over smaller and smaller intervals of time. This agrees with my experience in measuring ‘speed’ but it is an act of faith because my measurements may not be good enough to tell me what is really going on. The change from the physics of Newton to the physics of Einstein is an example where the model was changed because of better measurements.

I am going to define ‘speed’ in the Ideal World in such way that that Real world average speeds will approximate it if I make the averaging intervals small. This decision makes the Ideal World seem more real because I am saying that the definition of ‘speed’, and hence its reality, lies in the Ideal World. If I am not careful, I will start believing that the Ideal World is reality.

I grew up being aware of movement and I felt that things had speed. I was aware of speedometers and speed limit signs. If someone had asked me if I knew what speed was I would have said, “Sure.” It came as a shock to me that when I thought about it, I had no verbal or mathematical way to say what speed was. I soon will.

This number is hidden from me because it is in the Ideal World and I do not have the technology that will take me there. I will never get the exact value of ‘speed’ any more than I will ever write down all the digits of pi. Perfection always lies just beyond my reach. The good news is that I can approximate the number in the Real World, and that is all I ever do in the Real World anyway.

LECTURE 7-2

The wall had not been here the last time I passed this way and I figured that it must have been built on Thursday...

My plan is to consider an object moving in the Ideal World that models an object moving in the Real World. The distance traveled by this object is given by a function. I am going to use this function and the techniques of mathematics to come up with a number that I am going to <u>define</u> as ‘speed’ in the Ideal World. This is the number I will use to describe motion and it will have the property that it can be approximated by Real World average speeds.

At each point in time, there will be a unique number, called, ‘speed’ and I can define a new function which associates time and ‘speed’. The rule of this new function, evaluated at a time, will give the ‘speed’ of the object at that time This will be the function that models motion.

Here I go, ready or not.

An object moves along a straight line and I observe it for 10 seconds. I believe that there is a function that relates the distance, s, traveled to the time of travel, t, and I’m going to call it f. I know that the domain of the function is [0,10] but I have no idea what the rule is and if I want to refer to the rule, I will have to call it $s = f(t)$. I want to describe the motion at some time in the domain, say $t = 2$, and I want to describe it as a number. Want, want, want. Will I never stop?

While the definition of ‘speed’ will be made in the Ideal World, I want to be able to approximate it with Real World average speeds, so I’ll start by modeling the Real World average speeds .

The Real World average speed between, say, 2 and 2.01 is modeled by

$$f(2.01) - f(2.00) / (2.01 - 2.00) = f(2.01) - f(2.00) / .01.$$

This doesn’t really tell me much as it stands. I don’t know what the rule is so I don’t know anything about the numbers involved. I don’t know if the motion is nearly uniform over this interval or wildly changing. A graph would help me determine this because if the motion were nearly uniform, the graph would be nearly a line over this interval. But since I don’t have the rule, I don’t have the graph.

For want of anything better to do, I could determine if the motion is nearly uniform by taking the average speed over a smaller interval, say 2 seconds to 2.005 seconds. If this average speed is about the same as the average speed over the larger interval, then I would think that the motion is about uniform. Unfortunately I can't do this because I don't know the rule.

There are a lot of different kinds of motion that f could be modeling and for some of them 0.01 seconds is a long time. At Indy car 'speed's, 220 miles/hour, 0.01 seconds is about 3 feet, which is more than enough to win a race by.

In the Real World I want to take average speeds over smaller and smaller intervals to approximate 'speed' so I should probably look at several Ideal World average speeds and see what they do.

$f(2.01) - f(2.00) / (2.01 - 2.00) = f(2.01) - f(2.00) / .01$

$f(2.005) - f(2.00) / (2.005 - 2.00) = f(2.005) - f(2.00) / .005$

$f(2.001) - f(2.00) / (2.001 - 2.00) = f(2.001) - f(2.00) / .001$

$f(2.0001) - f(2.00) / (2.0001 - 2.00) = f(2.0001) - f(2.00) / .0001$

Since I don't know the rule of f, I find it hard to see what they are doing.

I am not sure how to proceed at this level of generality, so I'm going to look at a specific example and see if I get some insight into the problem. Fortunately, I have the function that models a falling rock, so I'll compute some average speeds using it and see what they do.

The rule of the function that models the distance, in feet, that a rock falls in t seconds is given by $f(t) = 16t^2$. I'll find average speeds over some intervals that I think are small and see if something suggests itself.

Now,

$$\begin{aligned} f(2.01) - f(2.00) / .01 &= 16(2.01)^2 - 16(2.00)^2 / .01 = \\ &16[(2.01)^2 - (2.00)^2] / .01 = \\ &16(2.01 + 2.00) (2.01 - 2.00) / .01 = \\ &16 (4.01) (0.01) / .01 = 16 \times 4.01 = 64.16 \end{aligned}$$

$$f(2.005) - f(2.00) / .005 = 16(2.005)^2 - 16(2.00)^2 / .005 = 64.08$$

$$f(2.001) - f(2.00) / .001 = 64.016$$

$$f(2.0001) - f(2.00) / .0001 = 64.0016$$

It looks as though the motion is uniform to one decimal place of accuracy, which is interesting, but this is not what catches my eye. It seems inescapable that as the length of the interval gets closer to zero, the average speed gets closer to 64. The average speeds are approximating 64. I would never have dreamed of a providence so divine. I could do the computation for even smaller intervals, but I've done enough to make me a believer. I wonder why the average speeds are behaving this way.

I'm going to do the computation again but this time I am going to keep a close eye on the .01.

$$
\begin{aligned}
f(2.01) - f(2.00) / .01 = f(2+.01) - f(2) / .01 &= 16(2 + .01)^2 - 16(2.00)^2 \;\; / .01 \\
&= 16[\,(2+.01)^2 - (2.00^2)] / .01 \\
&= 16(2+.01 + 2.00)\,(2+.01 - 2.00) / .01 \\
&= 16\,(4+.01)\,(0.01) / .01 \\
&= 16\,(4+.01) = 64+16(.01) \\
&= 64.16
\end{aligned}
$$

I find it visually difficult to follow the .01 through this computation and I think it would be helpful to call the .01 some letter, like h, which would set it apart.

The symbol, h, represents a positive number that gives the length of the interval of time I'm averaging over. This has the fortuitous by-product of giving me a formula for the average speed in terms of the length of the interval of time.

$$
\begin{aligned}
f(2+h) - f(2) / h = 16(2 + h)^2 - 16(2)^2 \;\; / h \\
&= 16(\,(2+h)^2 - 2.00^2\,)/ h \\
&= 16(2+h + 2)\,(2+h - 2) / h \\
&= 16\,(4+h)\,(h) / h \\
&= 16(4+h) \quad \text{if } h \neq 0 \\
&= 64+16\,h
\end{aligned}
$$

Now I can see why the average speeds get close to 64 as the length of the time interval, h, gets close to zero. In the Real World the length of the averaging interval, h, can only be made so small and no smaller. The nature of the Real World limits how small or how big something can be. In the Ideal World there is no such limitation and I can let h get as close to zero as I want, which is very close indeed.

I cannot, however, let h equal zero. The expression $f(2+h) - f(2) / h$ is not defined at $h = 0$. The equation,

$$f(2+h) - f(2) / h = 64 + 16h$$

holds only if $h \neq 0$, because in the process of going from the left side of the equation to the right side, I divided by h. I can't do that if h is zero. The 64 + 16h isn't there if h = 0.

$$f(2+h) - f(2) / h = 64 + 16h \quad \text{if } h \neq 0.$$

In the drag strip example, I could put the two timing lights at the finish line and a foot after the finish line or, at the finish line and a foot in front of the finish line. I would suppose that the average speed would be a little slower if I averaged over the interval in front of the finish line, but I would also suppose that my Real World instruments couldn't tell the difference.

If I average over the intervals before 2, here in the Ideal World, I think I'll get the same result. I'll try it.

If I again let h be the <u>positive</u> length of the interval, the average speed over the interval [2-h,2] is

$$\begin{aligned} f(2) - f(2-h) / h &= 16(2^2 - (2-h)^2) / h \\ &= 16[4 - (4 - 4h + h^2)] / h \\ &= 16(4h - h^2) / h \\ &= 64 - 16h, \text{ if } h \neq 0. \end{aligned}$$

As the length of the interval, h, goes to zero, the average speeds get close to 64. I seem to have thought correctly, which is gratifying, but I don't like the way the computation looks.

The problem here is similar to a problem I had with distance when setting up a coordinate system. Distance is inherently non-negative and I wanted to talk about distance to the left as different from distance to the right. I took care of this by introducing 'signed distance'.

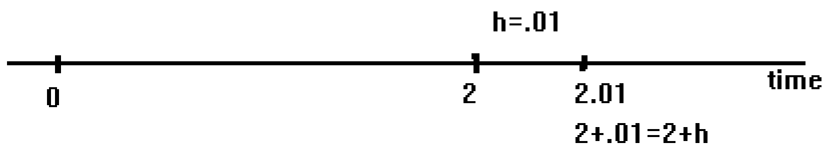
I have been considering the length of the time interval as non-negative but I'm going to back off from that. I can rewrite the expression, f(2) - f(2-h) / h, and <u>force</u> it to look like the expression, f(2+h) - f(2) / h. Here is what I mean.

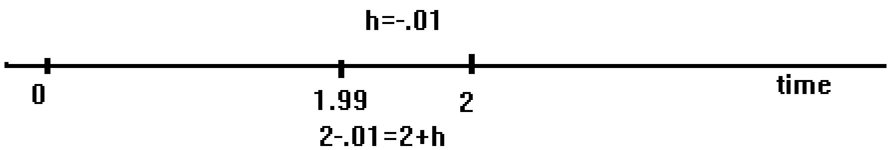
$$f(2) - f(2-h) / h = -(f(2-h) - f(2)) / h = f(2-h) - f(2) / -h.$$

The expression, f(2-h) - f(2) / -h, looks just like f(2+h) - f(2) / h with h replaced by -h. If I allow h to be both positive and negative, I have a pretty slick formula for the average speed over the time between 2 and 2+h,

$$\begin{array}{c} \text{average speed} \\ \text{over the interval} \\ \text{between 2 and } 2+h \end{array} = f(2+h) - f(2) / h = 64 + 16h.$$

If h is negative, it means that I am finding the average speed over an interval before t = 2.





As h goes to zero, through positive or negative values, the average speed gets close to 64. I vote for 64 as the 'speed' at t = 2 seconds.

As usual there is a little caveat here, 2 + h must lie in the domain of f, which in this example is [0,10]. There are two restrictions on h as it goes to 0; $h \neq 0$ and h + 2 must be in the domain of f.

Here are some reasons for my vote:

1. 64 is a single, well-defined number associated with the time t = 2. This is no 'wishy-washy' approximation. The big difference between the Real World and the Ideal World is that in the Ideal World the length of the averaging interval can go to zero. This gives me an exact number that the Ideal World average speeds get close to instead of a bunch of approximations.

2. The Ideal World average speeds get close to 64 as the averaging interval goes to zero and since the Ideal World average speeds model the Real World average speeds, I would bet that Real World average speeds get close to 64 too. Within the accuracy of modern measuring instruments, experiment shows that this is a good bet.

3. I was able to compute the number. This has less to do with whether 64 describes the physics correctly but rather is it a useable concept. It could be that I could define it but not be able to compute it. What a bummer that would be. Double bummer.

The number, 64, would get my total support if I could see how it was going to help me compute distance traveled and I don't see this right now. Not to worry. I'm sure it will come to me.

I just picked the number, 2, for value of time off the top of my head and I want to be able to compute the average speeds, and thus the 'speed', at any time in the domain of the function. I replace the 2 by the symbol 't' and do the computation again.

The average speed $= f(t + h) - f(t) / h = 16(t + h)^2 - 16(t)^2 / h$

$$\text{between } t \text{ and } t+h \quad = 16\{t^2 + 2th + h^2 - t^2\} / h$$
$$= 16(2th + h^2) / h$$
$$= 16(2t + h) \ \text{ for } h \neq 0.$$
$$= 32t + 16h.$$

The average speed between t and t+h gets close to 32t as h goes to zero. In my opinion, <u>32t is the 'speed' of the falling rock at every time, t, in the domain of f</u>.

Well, not quite all the values of t in the domain of f. There is trouble with $t = 0$ and $t = 10$. When $t = 0$ and h is negative, $t + h = h < 0$ and $t + h$ is not in the domain of f. When $t = 10$ and h is positive, $t + h = 10 + h > 10$ and $t + h$ is not in the domain of f. When I let h go to zero, I let it go to zero through <u>both</u> positive and negative values, so 0 and 10 are problems.

I only observed the falling object for 10 seconds and I don't know what happened before $t = 0$ or after $t = 10$. The object might have hit the floor at 10 seconds and then the motion would be uniform and the speed zero. It might have kept on falling and I could describe the motion as 320 ft. / sec. at $t = 10$. Maybe it fell into a swimming pool at $t = 10$ seconds and I haven't any idea what its motion was like. I can only describe the motion for $0 < t < 10$, and not at $t=0$ or $t=10$.

My plan was to use the function that models position and the techniques of mathematics to define a number I call 'speed' in the Ideal World. This number was to have the property that it could be approximated by Real world average speeds.

In particular, I took the function, f, that modeled the distance a rock falls as a function of time. The domain of this function is [0,10] and its rule is $s = f(t) = 16\,t^2$. I then defined 'speed' at each value of time, t, $0 < t < 10$ as the unique number the Ideal World average speeds approximate as the averaging interval becomes very small.

This is exactly the situation where functions are used; for each value of time there is a unique value of 'speed'. I even have an explicit rule that gives the 'speed'. I define a function, g, whose domain is $0 < t < 10$, and whose rule is

$$v = g(t) = 32t.$$

<u>This is the function that models the motion of the falling rock.</u> I have completed my plan for the example of the falling rock. How much harder can other examples be? I feel I am on my way home.

My first efforts with the falling rock involved increasingly elaborate measuring techniques to determine the relation between distance and time. The apparatus necessary to measure ‘speed’ would be even more so. Fortunately, there is no need. If I go to the edge of a high cliff and drop a rock, three seconds later its ‘speed’ would be $g(3) = 32 \times 3 = 96$ ft./sec.. The result of any experiment that I might do to compute ‘speed’ would only approximate 96. I think I’ll be lazy today, skip the experiment, and just use the formula.

LECTURE 7-3

Was this wall erected to impede my particular passage? I wondered...

A point that I find particularly interesting is that I have derived a new function from an old function working entirely in the Ideal World, and the new function can be interpreted in the Real World as modeling a significant property of an object in motion, its 'speed'. I have done things in the Ideal World that give me insight into the Real World. I like that.

My definition of 'speed' gives me a unique number for 'speed' but I am worried about the uniqueness of the definition itself. Could there be other definitions of 'speed' that would work as well or better? Beyond the fact that that I am able to give the definition, does it work at all?

How do I tell if I have the right definition of 'speed'? I have no *a priori* concept of speed to compare with my definition to see if they agree. The way it is now, my definition is the *a priori* concept of 'speed' that any other idea of 'speed' will be compared with it. What criterion should I use to judge if my definition is worthy?

Certainly, part of its worthiness would be its usefulness. Does it do anything beyond sit there and describe motion? I can say that the 'speed' of the rock is 64 ft./sec. at $t = 2$. So what? With 'speed' and a dollar I can buy a cup of coffee.

If the motion is uniform, knowledge of speed and time allows me to find the distance traveled and I would expect no less from 'speed' for non-uniform motion. If my instantaneous description of motion is going to be any good, it is going to have to allow me to compute the distance an object moves in a given time. This would be the proof of the pudding.

I will know I have the right definition of 'speed' when I can compute distance traveled from the knowledge of 'speed'.

Before I approach that problem, however, I'm going to try to understand 'speed' a little better, and before I do that, I'm going to work on language. I have some pretty awkward verbiage I have to clean up.

The phrase, '$f(2 + h) - f(2) / h$ gets close to 64 as h goes to zero' needs work. I am going to replace it by 'the limit of $f(2 + h) - f(2) / h$, as h goes to zero, is 64' and write

$$\lim_{h \to 0} f(2 + h) - f(2) / h = 64.$$

I must always remember that when I use an expression such as 'h goes to zero' or '$\lim_{h \to 0}$', that h never, ever equals zero.

The replacement phase may not seem much better than the original, but it introduces the word 'limit' and if there is one word that characterizes modern calculus, it is the word, 'limit'.

Here are some typical examples of how the notation is used:

1. $\lim_{h \to 0} 64 + 16h = 64$

2. $\lim_{t \to a} f(t) - f(a) / t - a = \lim_{t \to a} 16(t + a) = 32a.$

3. $\lim_{h \to 0} f(t + h) - f(t) / h = \lim_{h \to 0} 32t + 16h = 32t.$

The equality, $f(t + h) - f(t) / h = 32t + 16h$, only holds if $h \neq 0$, but I can apply

$$\lim_{h \to 0}$$

to both sides of the equation because h is never zero as h goes to zero. I can make a similar remark about the second example.

I mean something very specific when I talk about a limit, say,

$$\lim_{h \to 0} 64 + 16h = 64.$$

By making h close enough to zero, I can make the average speed, $64 + 16h$, as close to its limit, 64, as I want. For example, if I want the average speed to be within 0.0016 of 64, I use values of h in the interval $-.001 < h < .001$. The difference between the average speed and 64 is 16h and $64 + 16h$ is within 0.0016 of 64 if $-.001 < h < .001$.

The limit of the average speeds is not any particular average speed, but the <u>number that the average speeds get arbitrarily close to</u> as the length of the averaging interval goes to zero.

When I write

$$\lim_{h \to 0} \quad f(t + h) - f(t) / h$$

I have not started a continuing process that gets progressively closer to a number, I have written down a number; the number that $f(t + h) - f(t) / h$ gets close to as h goes to zero.

I can use the limit notation in situations other than considering average speeds. If F is a function and I write

$$\lim_{x \to a} \quad F(x) = L,$$

I mean that F(x) gets arbitrarily close to L as x goes to a. As always, ‘as x goes to a’ means that x gets close to ‘a’ through values both greater and smaller than ‘a’, but never equals ‘a’.

I walked over to the wall. The bricks were new and carefully mortared into place. Seeing that I was alone in the street, I began scraping at the mortar between two bricks with a mechanical pencil that I always carried...

I have defined 'speed' for the falling rock and now I am going to use the insight that I supposedly got from the specific example on the general problem.

To that end I consider a function, f, whose domain is the interval [a,b] and whose rule is s = f(t) where s represents the distance traveled in time t. I am not going to use a system of units. Who ever uses these ideas, should they turn out to be useful, will know the units they want to use for distance and time and I am not going to try to out guess them. I used units in the example because the rule of the function was derived on the assumption that I was using feet and seconds.

If t is some time between a and b, $a < t < b$, then the **average speed over the interval of time between t and t + h is**

$$f(t + h) - f(t) / h.$$

I allow h to be either positive or negative, but it can't be zero. h must also be small enough in magnitude so that t + h is in the interval [a,b], that is, t + h must be a time when the object was observed.

I <u>define</u> the Ideal World 'speed' of the object at a time t, $a < t < b$, as

limit as h goes to zero	of	average speed over the interval between t and t + h

or

$$\lim_{h \to 0} f(t + h) - f(t) / h.$$

I don't have 'speed' defined at t = a or t = b, but I do have it for every other time in the domain.

I have just gone from defining ‘speed’ for one specific example, the falling rock, to defining ‘speed’ for any moving object. I did it by defining ‘speed’ in general the same way I defined ‘speed’ in particular. No matter what rule I have, I use the same procedure on it that I did on the rule $f(t) = 16t^2$. The intuition is the same; the average speeds go to the ‘instantaneous speed’ as the averaging interval goes to zero.

The rule that relates distance and time for the falling leaf is forever lost to me, and hence also, the ‘instantaneous speed’. But, if a function models distance traveled and its rule is explicitly given, I can find a function whose rule gives the ‘instantaneous speed’. This may be a little over optimistic. After all, I do have to compute

$$\lim_{h \to 0} \quad f(t + h) - f(t) / h$$

and this may not be easy. But it may be not be too hard either. I did it for $f(t) = 16\ t^2$ and finding the ‘instantaneous speed’ for this function didn’t seem too hard. And even if it is hard, it at least gives me a direction to go and something to do while I am watching a ball game or listening to an opera.

It occurs to me that my definition of ‘speed’ is not at all like my intuitive idea of what speed is. If I consider accelerating my car from a stop to 20 mph., then I can say that I went 10 mph. at some time during the trip. But I went 10 mph. for a zero length of time and I traveled zero distance in that time. This is not the speed I used to know and love. That speed had a positive distance traveled in a positive time. This new ‘speed’ lasts for only an instant of time and I find it hard to mentally relate it to changes in distance and time. There aren’t any changes in distance and time. It feels to me like a new quantity all together, related to speed but different. It more or less acts like speed and it includes the case of uniform motion, but it goes beyond the concept I once had. It is an ‘ultra-speed’ and perhaps should be given a new name.

Like it or not, it is called ‘speed’.

I was waiting for a bus the other day and I noticed something written on a wall across the street. It said, “The function whose domain is [0,3] and whose rule is $s = f(t) = 25\ t^3$

models the number of feet a funny car travels in the first three seconds of its run as a function of time." I doubted that people who write graffiti have precise knowledge of funny cars, but you never know. Since I had some time to kill until the bus came, I thought I would see what the 'speed' of the car would be, in the unlikely event that the function was accurate.

$$f(t + h) - f(t) / h = (25(t + h)^3 - 25 t^3) / h = 25(t^3 + 3h t^2 + 3h^2 t + h^3 - t^3) / h$$

$$= 25(3h t^2 + 3h^2 t + h^3) / h$$

$$= 25(3 t^2 + 3h t + h^2) \text{ if } h \neq 0.$$

So,

$$\lim_{h \to 0} f(t + h) - f(t) / h = 75 t^2 ,$$

and the 'speed' in feet / second at any time, t, between 0 and 3 is evidently $75 t^2$. The speed of the funny car was modeled by the function, g, whose rule is $g(t) = 75 t^2$. After 2 seconds, the car would be going 300 ft. / sec. or about 204 miles / hr. As I recalled, a funny car goes close to 300 miles / hr. in 5 sec.. I wasn't sure if the speed was reasonable or not, but I was pleased at being able to do something constructive while waiting for the bus.

LECTURE 7-5

Day after day I stopped on my way home from work and scratched at the mortar with my pencil. Finally the day came when I removed the first brick...

It occurs to me that functions don't know what they model. The function whose rule is $f(t) = 16t^2$ and whose domain is [0,10], models the distance a rock that falls for times between 0 and 10 seconds. I know that, but the function doesn't. For all the function knows, it could be modeling the area inside a square whose side is 4t. So a function that models motion could also model other things as well, and I might wonder what

$$\lim_{h \to 0} f(t + h) - f(t) / h$$

<u>means</u> if f doesn't model motion but some other Real World physics. The function stays the same but the physics that the function models changes and the meaning of

$$\lim_{h \to 0} f(t + h) - f(t) / h$$

changes.

From another point of view, I can forget about physics altogether and think just about functions. If I have <u>any</u> function, k, I can form the quotients

$$k(x + h) - k(x) / h$$

where x and x + h are in the domain of k and see if, for a fixed x, these quotients approach a number as h goes to zero. If the quotients do approach a number as h goes to zero, I wonder what that number might mean.

The thought is father to the deed. I'm going to consider

$$k(x + h) - k(x) / h$$

as h goes to zero for any function k.

As always happens when I have a new investigation, I need some language so I can talk about it.

Suppose that for some fixed number, x, the quotients, $k(x + h) - k(x) / h$, approach some number, L, as h goes to zero.

They may not do this. It can happen that the quotients don't approach a single number, L, as h goes to zero. They could go to infinity; they could wander aimlessly around, never settling down to approach one fixed number, L; they could approach one number as h goes to zero from the right and another number as h goes to zero from the left.

But suppose they do get close to a number, L, which I think of as the limiting value of the quotients. I express the fact that the quotients get close to some number, for a fixed x, by saying that

the limit of $k(x + h) - k(x) / h$ exists as h goes to zero.

L is the number that $k(x + h) - k(x) / h$ gets close to ,as h goes to zero, and I express this fact by saying

' the limit of $k(x + h) - k(x) / h$, as h goes to 0, equals L'

or by

$$\lim_{h \to 0} k(x + h) - k(x) / h = L.$$

In the eyes of the function, k, the existence of ' $\lim_{h \to 0} k(x + h) - k(x) / h$ ' is

a very important property for a number, x, to have and the property should have a name.

If $\lim_{h \to 0} k(x + h) - k(x) / h$ exists, I will say that **k is differentiable at x.**

This is the traditional choice of word and I neither condemn nor defend it. I need a word; this one works.

I call $k(x + h) - k(x) / h$ a **difference quotient** of the function k at x. This is also traditional but makes a little more sense to me. If k models time and distance of a moving object, the difference quotient is an average speed.

If the limit of the difference quotients of k exists at the point x ,as h goes to zero, then the function k is **differentiable at x.**

I have to be a little fussy about the domain and which values of x in the domain of k I'm talking about. If the domain of k is the interval [a,b], then x can't be 'a' and it can't be 'b'.
The value of h can be positive or negative, so the points, x + h, lie on either side of x. This means that points that are close to x and on either side of x must be in the domain of k if I am going to let h go to zero in the expression, x + h.

How close is close? If (α,β), $\alpha < x < \beta$, is an interval <u>contained in the domain of k</u>, then all the points in this interval are close to x. I have called such an interval a neighborhood of x because I think of neighbors as being close. A function, k, can be differentiable at a point x in its domain only if x has a neighborhood contained in the domain. This assures that x will have points of the domain on both sides.

Edge points of domains, like the points 'a' and 'b' of a domain, [a , b], do not have neighborhoods that are contained in the domain. If a < x < b, then there is a point, c, between a and x, and there is a point, d, between x and b. The interval (c , d) is a neighborhood of x that lies in the domain of the function. It seems as though every 'x' that satisfies a < x < b is a possible point of differentiability for k.

I suppose that the domain of a function could be a very strange set, so strange that I can't even imagine what it would be like. But if a function models some Real World process, likely as not the domain will be an interval and I generally assume that is the case. It is certainly the case I run into most often.

The domain of a function k is an interval [a,b] and k is differentiable at all values of x in (a,b). I am going to follow the path I took with the function that modeled distance traveled for the falling rock.

The falling rock was modeled by the function, f, with domain [0,10] and rule $s = f(t) = 16\,t^2$. In the new language, f is differentiable for all t in (0,10). The Ideal World 'speed' was defined by, and the Real World 'speed' was modeled by, a function g that I derived from f. In particular, the domain of g was (0,10) and the rule of g was

$$g(t) = \lim_{h\to 0} \quad f(t + h) - f(t) / h = 32t.$$

I will define a function k` that is analogous to the function g.

The domain of k` is the set of all the numbers where k is differentiable and the rule of k` is given by

$$k`(x) = \lim_{h\to 0} \quad k(x + h) - k(x) / h.$$

Since I was working in a particular situation, g(t) was called the 'speed' of the object at time t. I am now working with a function that doesn't model anything at the moment, but I would still like a name for the function, k`.

I call k` the **derivative** of k.

I have never liked that name. I try not to let it bother me, but it does. I would have rather followed Newton, more or less, and called it the fluxion of k. Fluxion has fewer letters and I think it sounds better.

None the less, I follow the rest of the world and call k`(x) the derivative of k at x.

I have an arbitrary function, k, which could, possibly, model an object in motion. I have the derivative of k, k `, which could, possibly, model 'speed'. I can now write f ` in place of g and f `(t) in place of g(t). This new notation for 'speed' is convenient because it keeps the letter name of the distance function. It is commonly said that the derivative of distance is 'speed', and I guess that makes sense. Actually, the derivative of distance <u>defines</u> 'speed'.

I have begun with one function, k, and this function has given rise to a new function, k `. The domain of k` is the set of points where k is differentiable and the rule is

$$k`(x) = \lim_{h \to 0} \quad k(x + h) - k(x) / h.$$

I didn't say it was the easiest rule I ever saw.

LECTURE 7-6

I looked through the hole left by the brick I had removed. My view was restricted but I could tell that the street on the other side of the wall was very different than I remembered it. I began working on another brick...

I make a distinction between a function that models an object in motion and just any old function that is differentiable. I suppose, though, that it is possible that this distinction is not necessary.

"If k is a function whose domain is an interval, [a,b], and if k is differentiable in the interval (a,b), does k model some object in motion?"

While I can state the problem concisely, I can't answer it. The behavior of the derivative would have to be consistent with the behavior of 'speed' and I don't know how the 'speed' of an object in motion behaves. I can make some guesses but I need some further Real World investigation into the idea of 'speed'.

I can fall back on the 'principle of continuity' and assume that 'speed' should be continuous, but just because the derivative of k exists, does that also mean that the derivative is continuous? I don't know. It seems to be asking a lot and I doubt that it is true. And if k` is continuous, does that mean that k models some motion? I don't know. How frustrating.

It is my guess that, by and large, if k is differentiable, then k` is continuous and k does model some motion, but not all the time. It doesn't seem like anything happens all the time. Most of the functions that I run into are fairly well behaved but there are examples of functions that are differentiable at a point and their derivatives are not continuous there. I use these functions when I am showing someone that they exist, but they don't come up in applications because if a function models the Real World, its derivative is continuous.

Instead of idle speculation, I'm going to pick some of the fruits of my labor. I have another function that models objects in motion and I can look at it.

An object in uniform motion with speed v, in feet / second, is observed for 10 seconds. This is modeled by the function, f, whose domain is [0,10] and whose rule is s = f(t) = v t. I hope that when I compute 'speed' using the derivative, it will turn out to be v. Consistency here would be a wonderful thing.

Since f models an object in motion, I'll call the difference quotients, average speeds. The average speed between t and t + h is

$$\begin{aligned} f(t + h) - f(t) / h &= \\ v (t + h) - v t / h &= \\ v (t + h - t) / h &= \\ v h / h &= v \quad \text{if } h \neq 0. \end{aligned}$$

The average speeds are all the same and they are all equal to v.

Evidently the limit of the average speeds, as h goes to zero, exists and is equal to v. The function f is differentiable at every t in (0,10) and the rule of the derivative, f `, is

$$f`(t) = v.$$

This is a constant function and, as I had hoped, the derivative gives the speed of a uniformly moving object.

The functions that model uniform motion are members of the class of linear functions. The domain of a linear function is the set of all numbers, and its rule is of the form y = f(x) = ax + b. All of these functions have difference quotients that have a limit as h goes to zero, and so are differentiable. In particular, the difference quotients are

$$\begin{aligned} f(x + h) - f(x) / h &= \\ [a(x + h) + b - (ax + b)] / h &= \\ (ax + ah + b - ax - b) / h &= \\ ah / h &= a \quad \text{if } h \neq 0. \end{aligned}$$

The difference quotients are all the same and so their limit, as $h \rightarrow 0$, exists and equals 'a' for every value of x. The linear function, f, is differentiable at every number so the domain of the derivative is the set of all numbers. The rule of the derivative is

$$f`(x) = a.$$

A rather interesting coincidence happens here. If f is a linear function where f(x) = ax + b, then its graph is the line, y = ax + b, which has slope = a. It seems that the slope of the line is numerically equal to the derivative. <u>The graph of f is a line and the derivative of f is the slope of that line</u>. In the case of linear functions, I have a <u>geometric interpretation</u> of the very analytically defined derivative. It seems natural to see if there is some relation between the graph of a differentiable function and its derivative in general.

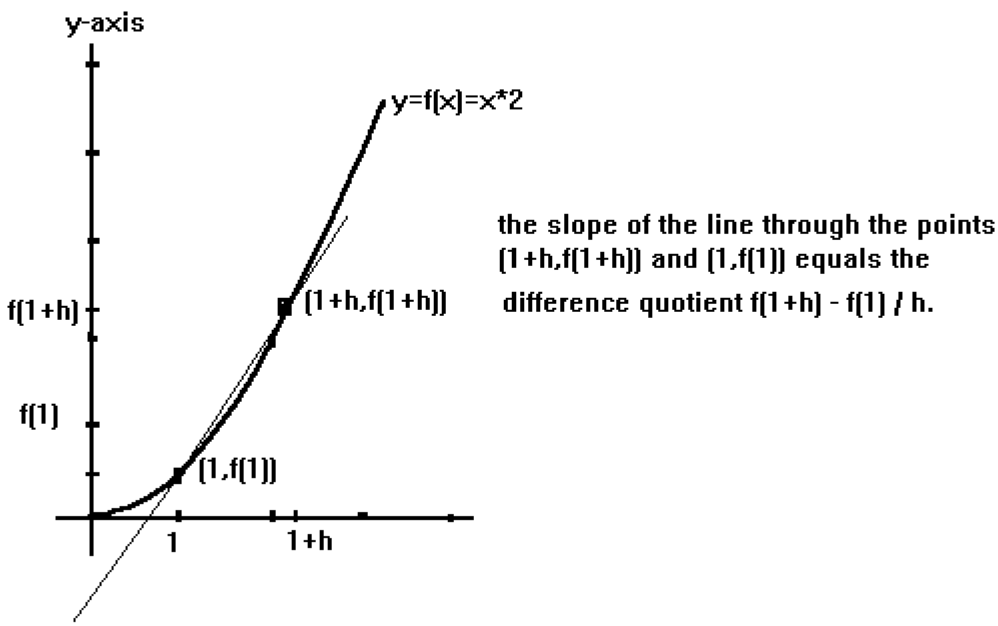
The idea of the graph of a function is geometric. If I can find some geometric significance of the derivative in the graph, I will be able to interpret an idea in one part of the Ideal World in terms of an idea in another part of the Ideal World.

LECTURE 7-7

After several weeks of scraping, my mechanical pencil had worn down to almost nothing and this made the work go slowly. Notwithstanding this difficulty, I eventually had a hole big enough to put my head through...

I am now after some geometrical significance of the derivative of a function. Since the derivative is the limit of difference quotients, it might not be a bad start to see what the geometric meaning of a difference quotient is, if any.

I could look at the difference quotients of the 'insight' function, $f(x) = 16x^2$, but it is hard to draw its graph anywhere near to scale and the numbers are too big. I am going to use $f(x) = x^2$ with domain [0,3]. I can't draw the graph of this function to scale either, but at least the numbers are smaller. I take what I can get, be grateful, and sketch the graph. The '16' of the original function was experimentally obtained and is determined by the gravitation of the planet. If I were on a smaller planet, the number would be smaller, and if the planet were exactly the right size, the number would be '1'. So for the moment, that's where I am; on a planet of exactly the right size.



The difference quotient

f(1 + h) - f(1) / h

is really

f(1 + h) - f(1) / (1 + h) - 1,

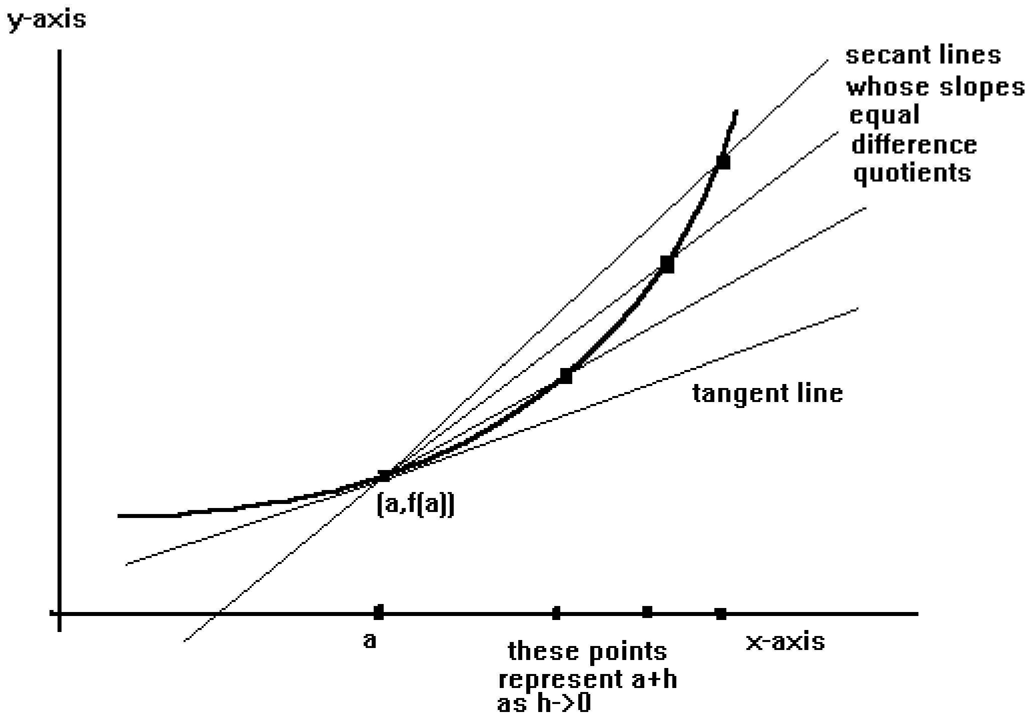
and this is the average speed over the interval [1 ,1 + h].

Since the 1's cancel in the denominator, I don't usually write them, but when I do write the difference quotient in this expanded form, I can <u>see</u> the two points (1,f(1)) and (1+h,f(1+h)) appear in the difference quotient as it changes into an average rate of change. The two points arise because average speed is computed from the distance, f(1+h), at
t = (1+h), and the distance, f(1), at t = 1. It seems natural to draw the line through them.

I suppose I could draw other curves through these two points, say circles, but there are a lot of circles that pass through them and only one line. I hate to make decisions. If I see two points I 'kick the can', and draw a line through them, not a circle. The <u>difference quotient is the slope of the line that passes through these two points.</u> I am pleased.

A line that passes through two points of a circle is called a secant. A line that passes through two points of a curve, as in the picture, is called a **secant line**. <u>The difference quotient equals the slope of a secant line.</u>

So that the picture can show me what is happening to the secant lines, as $h \to 0$, I need a lot of 'bow' in the curve, and the curve I am using doesn't have enough. This Gordian knot is cut by using an arc of the graph of some unknown function, which I will still call, f, and whose graph has lots of 'bow'. I don't know what the rule of f is nor do I care; <u>I'm interested in the picture, not the numbers</u>. I'll use the difference quotients f(a + h) - f(a) / h and the secant lines through the points (a + h, f(a + h)) and (a, f(a)).



As h goes to zero, the secant lines appear to pivot on the point (a, f(a)) and turn into a line that touches the graph of f at the single point, (a, f(a)).

This appearance is so real to me that I am willing to believe that it actually happens. At least I am willing to believe that the secant lines approach a line as h goes to zero, and that, for this particular function, they approach a line that passes through a single point on the graph of f.

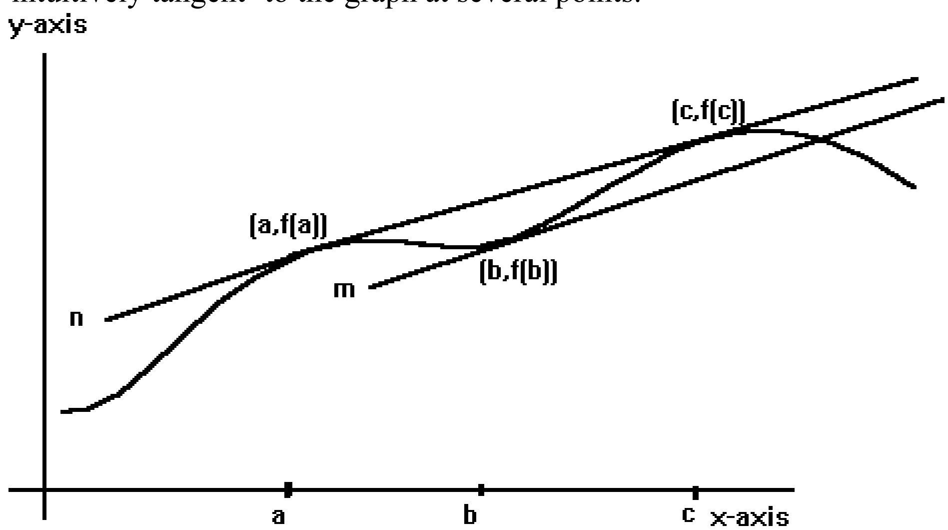
If a line passes through a single point on a circle, it is called a tangent line, and maintaining consistency, I call a line that passes through a single point of the graph of f , a **tangent line or a line tangent to the graph of f at that point.** This is my <u>naive definition</u> of what I mean by a line tangent to the graph of a function. This definition is geometric; there is no hint of numbers in it.

LECTURE 7-8

The street on the other side of the wall was lined with shops selling hot dogs, T-shirts, and surf boards. It ended at a small pier that reached out into an ocean. I was pretty sure that there was no ocean near Omaha...

I am troubled by my definition of tangent line. Since one of the two points that a secant line passes through approaches the other as h goes to zero, it seemed to me that the limiting line would pass through the single stationary point. In geometry a line that passes through a single point of a curve is called a tangent line, as in tangents to circles, ellipses and so forth, so at first my definition seemed OK.

Unfortunately, I can think of instances where this definition might not work for tangents to graphs of functions. For example, the line could be ‘intuitively tangent’ to the graph at the point (a,f(a)) but pass through other points on the graph as well. Also, it could be ‘intuitively tangent’ to the graph at several points.



Neither of the lines **n** or **m** pass through a single point of the graph but

slope **n** = $f'(a) = f'(c)$

and

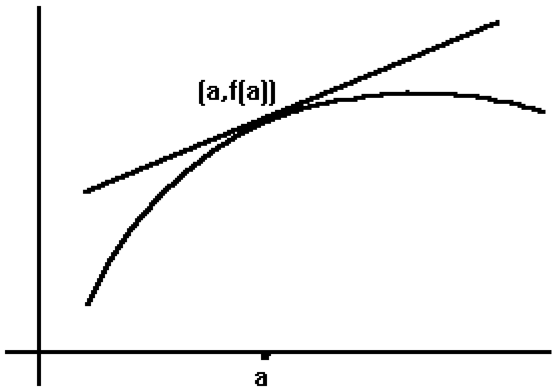
slope **m** = $f'(b)$.

There is also the obvious difficulty that a line can cut across a graph at a single point and not be tangent. Naive definitions often lead to these kinds of problems. There is an idea that I want to express but the most direct way to do it leaves lots of exceptions.

If the function f is linear, its graph is a line, say, **n**. It seems to me that if a line is tangent to **n** at some point, then it would be the line **n** itself. It is difficult to draw a picture because the picture can't distinguish between the graph, **n**, of the function and the tangent line.

In this case my definition of tangent line is totally demolished because instead of the tangent line passing through a single point of the graph, it passes through <u>all</u> the points of the graph.

While the picture of a tangent line in my mind is, and ever shall be, something like



,

the exceptions occur often enough to make me rethink my definition.

The tangent line was geometrically <u>defined</u> as a line passing through the single point, (a,f(a)) <u>before</u> the relationship to difference quotients and derivatives was considered. It was based on my previous experience in geometry with lines tangent to circles, ellipses, and so forth. The secant lines seemed to move to a line that looked like a tangent line to an ellipse looked.

This geometric definition gave me the 'little gem' I was looking for; that f `(a) = slope of the tangent line through the point (a,f(a)). This is what I wanted and I believe it to the depths of my very soul, regardless of the deficiencies of my definition of tangent line.

I have a new, improved definition of the line tangent to a graph in my pocket, and it will do away with the drudgery of exceptions. I am going to <u>define</u> the tangent line to the graph of f at the point (a,f(a)) in such a way that its slope is equal to the derivative of f at a. In this new, modern version, I cast geometry aside and define the tangent line analytically.

The line through the point (a,f(a)) with slope f `(a) is called the <u>line tangent</u> to the graph of f at the point (a,f(a)) or the <u>tangent line</u> to the graph of f at the point (a,f(a)) or something like that.

The equation of this line is

$$\mathbf{y - f(a) = f`(a) (x - a).}$$

This is my final definition of the line tangent to the graph of a function and this is the definition I'm going to have to live with.

Well, I am an 18th century mathematician at heart and I am going to hang onto my definition of tangent line. I am going to stick my head in the sand and pretend those pathological examples don't exist.

I have dumped the geometric definition of tangent line which fit so nicely with the limit of secant lines and which led the way to the geometric-analytic bridge: the slope of the tangent line equals f `(a). I have gone across that bridge and found an analytic hotshot to replace my tried and true friend. I can now write down the equation of the tangent line and not even think of a picture. I can, but I don't. I always think of the picture.

The linear function is now no problem at all. If f is a linear function, f(x) = mx + b, then f `(x) = m for all values of x. The line through the point (a,f(a)) with slope m, that is, the tangent line, is

$$\begin{aligned} y - f(a) &= m (x - a) \\ y &= m (x - a) + f(a) \\ y &= mx - ma + ma + b \\ y &= mx + b \end{aligned}$$

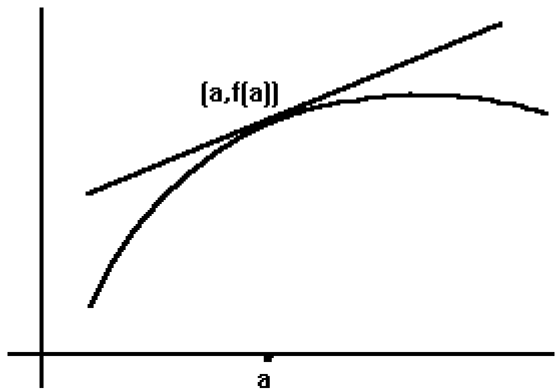
which is the line itself.

I am a little compulsive about definitions. It is not that I always make the tightest possible definition but when I don't, I feel a little guilty. Guilt requires a crime and I guess the crime is saying something that is not quite right. I find it very hard to be exactly correct, so there is always a little residual guilt which I try to rise above. I also find it much easier to understand ideas and concepts if I know exactly what the ideas and concepts are and this requires good definitions of the ideas and concepts.

I suppose that large parts of humanity live their lives productively and happily without being particularly fussy about the definitions in their lives. More power to them.

The derivative is captured geometrically as the steepness of a line. The steepness of a line is captured as the slope of a line. The slope of a line is captured in the equation of the line. Thus the geometry of the derivative is captured analytically in the equation of a line. I suppose that I should relate the geometry to this new definition of tangent line but that doesn't sound like much fun. My first definition gave me enough geometric insight so that I have no trouble believing that the line I get from the analytic definition looks like the line from geometry.

The geometric idea has the advantage in that it addresses what I actually mean by a line tangent to a graph. When I think about a line tangent to a graph, I think about a line passing through one point of the graph and I don't think about the exceptional cases. I think about the cases that best represent what tangent lines mean to me.



I can't write about the most important part of having a geometric interpretation of the derivative, or a geometric interpretation of anything else, because it is visual. It allows me to see something I can't write about. I can look at the tangent line, see its steepness and get a feeling about the size of the derivative, but I find it hard to verbalize steepness. My graphs often have tangent lines penciled in all over them so I can see the derivative better.

A definition should be formal enough to be used easily in a mathematical argument and used easily to construct the thing it defines, if that is appropriate. On the other hand, they should be embedded in enough intuition to give me some idea of what in the world the object of the definition is. I think the definition of tangent line is a good one.

LECTURE 7-9

I pulled my head back through the hole and thought about what I had seen. After putting the bricks back in the hole as best I could, I went back to my apartment...

I can describe uniform motion by an ordinary ratio of numbers which I called speed. The whole development of instantaneous speed was forced upon me because I couldn't describe non-uniform motion at an instant of time as an ordinary ratio of numbers. I defined the instantaneous speed as the <u>limit</u> of ordinary ratios of numbers, in particular, as the limit of the average speeds over time intervals that shrunk to a point. The first stage was ordinary ratios for uniform motion; the second stage was limits of ordinary ratios for non-uniform motion . I want to consider these ratios.

The average speed is a ratio of the change in distance to the change in time, so it is a ratio of changes. It is an average rate of change.

More generally, I am going to look at a Real World quantity, q, that changes with respect to a Real World quantity, x, where their relationship is modeled in the Ideal World by the function, f, with q = f(x). I call the difference quotient, f(x) - f(a) / x - a, an **average rate of change of q with respect to x.**

More generally yet, I can look at any old function at all, f, with y = f(x), and call the difference quotient, f(x) - f(a) / x - a, an **average rate of change of f with respect to x.**

If I call 'f(x) - f(a) / x - a' a difference quotient, I am emphasizing the algebraic nature of the expression. I am preparing to do some algebra on the difference quotient with the end in mind of putting it in a form where I can see what its limit is as $x \to a$.

If I call 'f(x) - f(a) / x - a' an average rate of change, I am emphasizing a certain quality of the ratio that allows it to model Real World processes.

The limit of the average speeds is called the 'instantaneous speed', and in line with this I call **f `(a) the instantaneous rate of change of q with respect to x, or the instantaneous rate of change of f with respect to x, depending on the context**.

If I call f `(x) the derivative, then I am thinking of f `(x) as the limit of difference quotients.

If I call f `(x) the instantaneous rate of change, I am thinking of a quantity that has the potential to model a Real World rate of change of one quantity with respect to another.

The derivative and the instantaneous rate of change are the same mathematical object, and I use the name that is appropriate to the context. When my neighbor, Pete, goes to work in the morning, I call him a pipe-fitter. When he works in his garage at night, I call him a mechanic. Same guy, different names.

Having made these distinctions, I usually use the word ‘derivative’ when I might better use ‘instantaneous rate of change’. I do it because it is shorter and because everybody else does. When I hear or use the word ‘derivative’, my head must supply the ‘instantaneous rate of change’. Personally, I always include ‘instantaneous rate of change’ when I am confronted by the derivative, even if I can conceive of no process that it could model.

Some instantaneous rates of change have names, others do not. Here are a couple that do.

If q is distance and x is time, then f `(a) is the instantaneous rate of change of distance with respect to time and I call this, **speed**.

If q is temperature and x is distance, f `(a) is the instantaneous rate of change of temperature with respect to distance and I call this, **temperature gradient**.

There are times when the actual numerical value of the derivative is of interest, but far and away its most important use is descriptive. My original goal was to describe motion, not to compute numbers. The laws of nature are there to describe the processes of nature and it is from these laws that I extract what I really want to know. The importance of the derivative lies in how it is used in these laws.

The voltage across an inductor is proportional to the instantaneous rate of change of current with respect to time.

The amount of heat that flows passed a point in a slender rod is proportional to the temperature gradient, that is, to the instantaneous rate of change temperature with respect to distance.

The amount of force needed to change the motion of a moving mass is proportional to the instantaneous rate of change of momentum.

The point here is not what voltage, inductor or momentum might mean, the point is that the derivative appears in the laws that describe whatever things these might be.

I have been told that the physical law that describes an electric circuit with a 5 henry inductor in series with a 10 ohm resistor and a 12 volt battery is modeled in the Ideal World by the equation

$$5\, I\,`(t) + 10\, I(t) = 12,$$

where I is the function that models the current in the circuit, and t is time. The goal is to determine the function, I. I am <u>not</u> given I(t) with the task to find I `(t).

I think of the above equation as describing the physical law of this circuit and I am looking for a function, I, that satisfies this description.

It is my opinion that the major role of the derivative is in description, such as in the previous equation. The really remarkable property of these descriptions is that they can be manipulated in the Ideal World in such a way as to actually obtain the function, I. In my opinion, that's a pretty neat property.

LECTURE 7-10

I spent a restless weekend thinking about the wall and the hole and the ocean. After work on Monday, I bought a new mechanical pencil and a swimming suit...

I began with an object in motion in the Real World and modeled it in the Ideal World with a function, f, that related distance traveled and time. This function gave rise to a new function, f `, where f `(x) is the 'instantaneous rate of change of f ' at x. This quantity was interpreted back in the Real World as the 'instantaneous speed' of the object that f models.

So far I have carried out this program for only two functions, the function that models an object in uniform motion and the function that models a freely falling body, namely a rock. While it is true that my technique seems to have the potential to find the 'speed' of objects in other kinds of motion, I can only use it if I already know the function that models distance and time. The two functions that model distance were not easily obtained and it is not clear to me how I am going to get them in any great number. Once I had the function that related distance and time, the computation that gave speed was not too bad. The method for computing the derivative looks a little formidable, but when I actually used it, I didn't have any trouble.

At first glance, it would seem that this is the way these ideas are used in the Real World; find the function that relates distance and time and then take the derivative of that function to find the 'speed'.

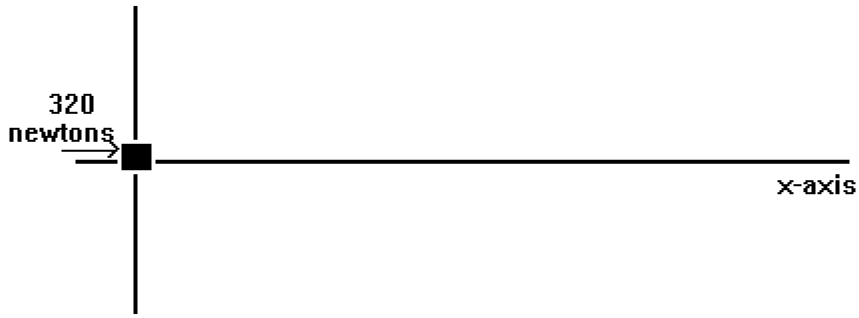
This appearance is misleading. In the beginning I had no concept of 'speed' or instantaneous rate of change, all I had was the concept of relating distance to time. It seemed natural to develop the concept I needed, 'speed', from the relation between distance and time that I had. But once I have the concept of instantaneous rate of change, I can step back and see how it actually arises in the Real World.

The concept of instantaneous rate of change is important because it is needed to state the laws of physics. The laws of the universe are relations between rates of change. A typical problem gives 'speed' as a function of time and the goal to find distance as a function of time.

It will eventually turn out that finding derivatives is rather routine, but it is often quite difficult to find the distance function from the knowledge of 'speed'.

Here is an example of an Ideal World model of a Real World problem.

An object whose mass is 10 kilograms is at rest at the origin of a number line which I am going to call the x-axis. A force of 320 newtons is applied constantly to the object causing it to move along the x-axis in the positive direction. I am going to measure distance in meters and time in seconds. I want to find a function, f, that relates the distance traveled on the x-axis to the time after the force is applied.



A form of one of Newton's laws describes this situation, namely

Force = mass $\times$ instantaneous rate of change of 'speed'.

Newton's contention was that the only way to change 'speed' is the application of force. In this particular case, it reads

320 = 10 $\times$ instantaneous rate of change of 'speed'.

If I denote the function that models the 'speed' of the object by g, then g(t) gives the 'speed' of the object at time t and g `(t) is the instantaneous rate of change of 'speed' at that time. The law now takes the form

$$32 = g`(t).$$

The first step on my way to finding distance as a function of time is to find the function g.

Now I know that the derivative of a linear function is a constant. In particular, the derivative of g(t) = 32t + b is 32, regardless of what b is. Looking at it geometrically, I don't see how the slope of the tangent to the graph of g could be constant without the graph being a line. Of course, just because I don't see how it could happen otherwise, is hardly a convincing proof that g is linear. Never-the-less, I think that it is reasonable to suppose that g(t) = 32t + b. Since 'speed' is zero at t = 0, g(0) = 0 = b, so g(t) = 32t. I have found the function that models motion, the function that gives 'speed'.

But ‘speed’ is an instantaneous rate of change, it is the derivative of distance. I am trying to find the function that relates distance and time and if I call this unknown function, f, I have g = f `. The function that I seek has made its way into an equation. I am blessed.

$$g(t) = f`(t) = 32t.$$

By a stroke of good fortune, I know a function whose derivative is 32t. It is the function that modeled the falling rock. While I am not constructing this function through a carefully reasoned argument, I see no advantage in forgetting that 32t is the derivative of the function, $f(t) = 16t^2$.

In my opinion, the function that relates distance and time for this object is f, where the domain of f is $0 \leq t$ and its rule is $f(t) = 16t^2$.

I’ll give some validity to this opinion by making sure that the function satisfies the conditions of the problem. If an object moves a distance of $f(t) = 16t^2$ meters in t seconds, then its ‘speed’ is surely f `(t) = 32t m./sec at the time, t.

Further, the instantaneous rate of change of ‘speed’ is 32 m./sec./sec. since the derivative of 32t is 32, and Newton’s law is satisfied. As a final check, f(0) = 0 agrees with the fact that the object is at the origin when t = 0 and, speed = 0 at t = 0, agrees with the fact that the object starts from rest. The function, f , is a solution to the problem.

There is a little question about uniqueness. Could there be two functions that satisfy the conditions of the problem? I certainly hope not and I am going to make another appeal to ‘justice in the universe’ that such an unfortunate occurrence could not happen. I also have a certain amount of worldly experience that tells me the solution of motion problems such as this are unique. Without this kind of uniqueness the sport of basketball could not exist. If pushed I could prove that the solution is unique but I don’t have a problem believing it so I think I’ll skip the proof.

If a 10 kg. object sitting at the origin is pushed to the right by a 320 newton force, the distance, s, from the origin at time, t, is given by the function, f, where $f(t) = 16t^2$. I think this is quite an accomplishment.

I did not take any derivatives is the process that led me from g`(t) = 32 to $f(t) = 16t^2$ . In fact, I moved the other way. Given a function g, I had to find a function, f, where f ` = g, not find g`. I was able to solve the problem because of previous knowledge gained from derivatives I had computed earlier. It is fortunate that I remembered them.

My goal was to describe Real World motion at an instant of time. In my opinion there are no ‘zero length’ instants of time in the Real World and no instantaneous speeds. There are very small intervals of time and average speeds over these intervals. But if I model in the Ideal World, I can let those intervals go to zero and there are instants of time and ‘instantaneous speed’s. I can find an Ideal World ‘instantaneous speed’ at a Real World time and interpret this as a Real World ‘instantaneous speed’. I solve the problem in the

Ideal World and interpret the result in the Real World. I have conceptualized 'instantaneous speed' in the Real World as Ideal World 'instantaneous speed'.

Not only have I solved the problem of describing motion, but I have also developed a concept of instantaneous rate change, the derivative, along the way. That seems like a pretty good day's work.

Instantaneous rate of change is the heart of the derivative.

I am no longer going to write 'speed' but will just use the usual speed. I am a professional now and know that speed means the Ideal World speed that describes non-uniform motion.

Further Considerations of Chapter 7

I have said that there are functions that have difference quotients that don't get close to any number and I want to back that up.

The first function I want to present is

$$f(x) = x \sin(1/x) \qquad \text{for } x \neq 0$$
$$f(0) = 0.$$

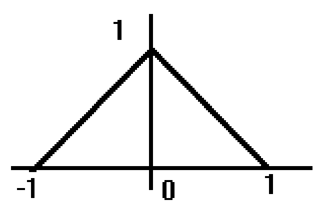
The difference quotients at $x = 0$ are

$$f(h+0) - f(0) / h = h \sin(1/h) / h = \sin(1/h) \text{ , for } h \neq 0,$$

which oscillate back and forth between -1 and 1 and don't approach any number at all as h goes to zero.

The next function I want to show is

$$f(x) = 1 - x \qquad \text{for } 0 \leq x \leq 1$$
$$f(x) = 1 + x \qquad \text{for } -1 \leq x < 0.$$



The difference quotients are

$$f(0+h) - f(0) / h = [(1 - h) - 1] / h = -1 \quad \text{if } h > 0$$

and

$$= [(1 + h) - 1] / h = 1 \quad \text{if } h < 0.$$

If h goes to 0 from the right, the difference quotients get close to -1 and as h goes to 0 from the left, the difference quotients get close to 1. They don't get close to a single number.

The last of the functions I want to exhibit is

$$f(x) = \sqrt{x} \qquad \text{for } 0 \le x$$
$$f(x) = \sqrt{-x} \qquad \text{for } x < 0.$$

The difference quotients at x = 0 are

$$f(h+0) - f(0) / h = \sqrt{h} / h = 1 / \sqrt{h} \qquad \text{for } h > 0$$
$$= \sqrt{-h} / h = 1 / \sqrt{-h} \qquad \text{for } h < 0,$$

and they go to infinity as h goes to 0. The problem here is that the 'h' in the denominator of the difference quotient is not absorbed in the numerator.

I have computed the limit of the difference quotients for two functions that give distance traveled and I would like to do this for the five basic functions.

The first of these is the power function, $f(x) = x^r$. I am not going to do a proof but just draw some conclusions from looking at some examples. I have found the derivative for r = 1 and r = 2. I'll find the derivative of $f(x) = x^3$.

I first write the difference quotients,

$$f(x+h) - f(x) / h = (x+h)^3 - x^3 / h$$
$$= (x^3 + 3x^2 h + 3xh^2 + h^3 - x^3) / h$$
$$= (3x^2 h + 3xh^2 + h^3)/ h$$
$$= h(3x^2 + 3xh + h^2) / h.$$

This is the critical spot in the process. The 'h' in the denominator must be absorbed by the numerator. One of the ways the 'h' is absorbed is by canceling an 'h' into the numerator, as is the case in this example. If the 'h' in the denominator isn't absorbed, the difference quotient goes to infinity as h goes to zero, and the derivative would not exist.

Since $h \neq 0$, I can cancel the h's and

$$f(x+h) - f(x) / h = 3x^2 + 3xh + h^2 .$$

It is now easy to see that

$$\lim_{h \to 0} f(x+h) - f(x) / h = 3x^2$$

and

$$f`(x) = 3x^2 .$$

I'll try one more, $f(x) = x^4$.

The difference quotients are

$$f(x + h) - f(x) / h = (x + h)^4 - x^4 / h$$
$$= (4x^3 h + 6x^2 h^2 + 4xh^3 + h^4)/h$$
$$= (4x^3 + 6x^2 h + 4xh^2 + h^3), \text{ if } h \neq 0.$$

$$f`(x) = \lim_{h \to 0} f(x + h) - f(x) / h = \lim_{h \to 0} (4x^3 + 6x^2 h + 4xh^2 + h^3) = 4x^3 .$$

If I write down the four derivatives I have, I can see a pattern.

$f(x) = x^1$ $\quad$ $f`(x) = 1\, x^0 = 1.$
$f(x) = x^2$ $\quad$ $f`(x) = 2\, x^1 = 2x$
$f(x) = x^3$ $\quad$ $f`(x) = 3\, x^2$
$f(x) = x^4$ $\quad$ $f`(x) = 4\, x^3$.

If n is a positive integer and $f(x) = x^n$, then $f`(x) = n\, x^{n-1}$.

If r is any number, I have to be a little careful about the values of x where I compute the derivative of $f(x) = x^r$, because the domain of f depends on what r is.

I can say that if

$$\mathbf{f\,(x) = x^r ,}$$

is differentiable at x, then

$$\mathbf{f\,`(x) = r\, x^{r-1} .}$$

I have tried to show by example why this rule works if r is a positive integer and have not even waved a hand at any other values for r. What kind of world would we live in if this rule for taking the derivative of $f(x) = x^r$ was only good if r was a positive integer. I would not entertain such a thought for even a moment. Of course it must be true for all values of r.

I'm going to try the exponential function next, $f_b\ (x) = \exp_b x = b^x$.

The derivative at x = 0 is

$$f_b\,`(0) = \lim_{h\to 0} f_b\,(h) - f_b\,(0) / h = \lim_{h\to 0} (b^h - b^0) / h = \lim_{h\to 0} (b^h - 1) / h .$$

I am going to assume that $\lim_{h\to 0} (b^h - 1) / h$ exits for any positive value of b. This

assumption is true but not obvious. The particular thing that keeps it from being obvious is how the 'h' in the denominator is absorbed into the numerator. Regardless of how 'h' is absorbed, it happens and $f_b\,`(0)$ exists for all positive values of 'b'.

The difference quotients for any value of x are

$$\begin{aligned} f_b\,(x + h) - f_b\,(x)\ / h &= b^{x+h} - b^x\ / h \\ &= b^x\ (b^h - 1) / h \\ &= b^x\ [\,(b^h - 1) / h\,]. \end{aligned}$$

I am at the critical step where the h in the denominator must be absorbed in the numerator. I can't help but notice that there isn't any h to cancel in the numerator. I don't have to cancel the h, however, I have to <u>absorb</u> it when I take the limit as h goes to zero. The h is absorbed when I take the limit of the term, $[\,(b^h - 1) / h\,]$, because

$$\lim_{h\to 0}\ [(b^h - 1) / h] = f_b\,`(0),$$

and I have assumed that this limit exists. There are those who might say that my assumption begs the question on the existence of these limits and they would be right.

Since all the limits are taken as h goes to zero, I can write 'lim' in place of $\lim_{h\to 0}$

and I will because I think it looks better on the page.

$$\begin{aligned} f`_b\,(x) &= \lim f_b\,(x + h) - f_b\,(x)\ / h \\ &= \lim b^x\ [\,(b^h - 1) / h\,] \\ &= b^x \times \lim [\,(b^h - 1) / h\,], \text{ because } b^x \text{ doesn't depend on h,} \\ &= b^x \times f_b\,`(0) \end{aligned}$$

I have the remarkable result that

$$\mathbf{f`_b\,(x) = f_b\,(x) \times f`_b\,(0) = exp_b\ x \times f`_b\,(0).}$$

This is nice, but it would be perfect if $f`_b(0) = 1$. The term $f`_b(0)$ changes if b changes and it is true that there is a value of b that makes $f`_b(0)$ equal to one. I love when something turns out to be perfect even if the reason for this perfection is not entirely evident. This value of b has the traditional name of 'e' and I would never dream of changing it. If b = e, I drop the subscript and write $\exp_b x = \exp x$.

If f(x) = exp x, then f `(x) = exp x.

This function is its own derivative. When I talk about the **exponential function**, this is the function I mean. Written in the exponential form, the rule looks like:

If $f(x) = e^x$, then $f`(x) = e^x$.

The number, e, was introduced as the value of 'b' that made

$$\lim_{h \to 0} [(b^h - 1) / h] = f_b`(0) = 1,$$

so e is the number such that

$$\lim_{h \to 0} [(e^h - 1) / h] = 1.$$

I introduced an 'e' earlier when I first presented the exponential function. Using notation I have since developed, I defined that 'e' as

$$e = \lim_{n \to \infty} (1 + 1/n)^n ,$$

where n is a positive integer. I could also have defined it as

$$e = \lim_{x \to \infty} (1 + 1/x)^x .$$

or

$$e = \lim_{x \to 0} (1 + x)^{1/x}.$$

It is certainly not self-evident to me that the two e's are the same, but they are. It is the job of mathematicians to take care of these details and they are very good at it. The exponential function, exp, I defined previously and the exponential function, exp, I defined here, are the same.

The natural logarithm, ln, is the sister function of exp. The function, ln, is defined by the four conditions:

$$\ln xy = \ln x + \ln y,$$
$$\ln a > 0 \text{ if } a > 1,$$
$$\ln a = 0 \text{ if and only if } a = 1.$$
$$\ln e = 1$$

The difference quotients are

$$\begin{aligned}[\ln(x + h) - \ln x] / h &= 1/h\, [\ln((x + h)/x)] \\ &= 1/x \times x/h \times [\ln(1 + h/x)].\end{aligned}$$

The quantity, $x/h = u$, goes to infinity as h goes to zero so it seems to me that

$$\begin{aligned}\lim_{h \to 0} x/h \times [\ln(1 + h/x)] &= \lim_{u \to \infty} [u \ln(1 + 1/u)] \\ &= \lim_{u \to \infty} \ln(1 + 1/u)^u \\ &= \ln \lim_{u \to \infty} (1 + 1/u)^u \\ &= \ln e \\ &= 1,\end{aligned}$$

which is gratifying.

If $f(x) = \ln x$, then

$$\begin{aligned}f`(x) &= \lim [\ln(x + h) - \ln x] / h \\ &= \lim 1/h\, [\ln(x + h)/x] \\ &= \lim 1/x \times x/h \times [\ln(1 + h/x)] \\ &= 1/x \times \lim \{x/h \times [\ln(1 + h/x)]\} \\ &= 1/x \times 1 \\ &= 1/x.\end{aligned}$$

If $f(x) = \ln x$, then $f`(x) = 1/x$.

The final two functions are sin x and cos x. They are defined by the relations

$$\sin(x + y) = (\sin x)(\cos y) + (\cos x)(\sin y)$$
$$\cos(x + y) = (\cos x)(\cos y) - (\sin x)(\sin y)$$

$\sin(\pi / 2 - x) = \cos x$.

The difference quotients for sin x are

$$\sin(x + h) - \sin x / h = [(\sin x) (\cos h) + (\sin h) (\cos x) - \sin x] / h$$
$$= \sin x ((\cos h) - 1) / h + [(\sin h) / h] (\cos x).$$

I have two terms which have an h in the denominator that has to be absorbed. This can be done if one happens to know that

$$\lim_{h \to 0} (\sin h) / h = 1,$$

which I do. I have to work with the other term a little bit.

$$((\cos h) - 1) / h = [(\cos h)^2 - 1] / [(\cos h) + 1] (\sin h)$$
$$= (\sin h)^2 / [(\cos h) + 1] (\sin h)$$
$$= (\sin h) / [(\cos h) + 1]$$

and

$$\lim_{h \to 0} ((\cos h) - 1) / h = \lim_{h \to 0} (\sin h) / [(\cos h) + 1]$$
$$= 0 / 2 = 0,$$

since sin 0 = 0 and cos 0 = 1 and sin and cos are both continuous.

If f(x) = sin x , then

$$f`(x) = \lim [\sin x ((\cos h) - 1) / h + [(\sin h) / h] (\cos x)]$$

$$= \lim (\sin h) / [(\cos h) + 1] + \lim [(\sin h) / h] (\cos x)$$

$$= 0 + 1 \cos x$$

$$= \cos x.$$

There is the derivative of the cosine left to do, but I think I have done enough. I wanted to give some idea of the computation of the derivative and what I have done is sufficient for that. The derivative of the cosine turns out to be - sine and I can take the derivative of the five basic functions.

$f(x) = x^r$; $f`(x) = r x^{r-1}$

$f(x) = e^x$; $f`(x) = e^x$

$f(x) = \ln x$; $f`(x) = 1/x$

$f(x) = \sin x$; $f`(x) = \cos x$

$$f(x) = \cos x \; ; \;\; f`(x) = -\sin x$$

The whole theory of taking derivatives would be complete if I showed how the derivative works with the five basic ways to combine functions, but that is another story for another day.