CHAPTER 8

LECTURE 8-1

It was almost dark when I left and my car was a patch of gray across the parking lot. I really liked that car...

I have defined the concept of speed for non-uniform motion. I have taken the function that relates the distance an object travels to time and derived from it the function that relates time to the speed of the object. "I have drunk delight of battle with my peers on the ringing planes of windy Troy."

Having spent much time and effort developing a way to go from distance to speed, that is, developing the derivative, I must now admit that it is more usual to know the function that gives speed at each time and want to find the function that gives the distance traveled at each time. This is my next project.

If the motion is uniform, the speed is constant in time and there are two basic relationships:

1. speed = distance $\div$ time

2. distance = speed $\times$ time

Neither of these are true, as they stand, for non-uniform motion.

If the motion is non-uniform, the time intervals where the speed is constant reduce to instants and the space intervals where the speed is constant reduce to points.

A distance divided by time approach to speed for non-uniform motion led to 0/0, which is meaningless, and I had to define speed as a <u>limit</u> of quantities that are a distance divided by a time, a limit of the average speeds.

If the motion is non-uniform, a speed multiplied by time approach to distance fails because if the length of the time interval is positive, there is no fixed speed to multiply by the length of the time interval. The only interval of time where the motion is uniform and the speed constant is an instant. I could try to add up the distances traveled during all the instants of time, but at each instant of time, the distance traveled is zero. If I add the distances traveled in all the instants of time during the trip, I seem to be adding up a bunch of zeros, and this would appear to give zero for the distance traveled.

sum of distances traveled in each instant = sum of zeros = distance traveled = 0.

This can't be true, but the alternative is equally dubious, sum of distances traveled in each instant = distance traveled = sum of zeros $\neq 0$.

I made sense out of 0/0 when I was looking for speed and I am confident that I will make sense out of

$$\text{sum of zeros} \neq 0.$$

I am looking for a way to find the distance traveled during any interval of time using only speed. Once I have done this, I will be able to give the function that models distance as

$$f(t) = \text{distance traveled on } [0,t].$$

Since I go into this search knowing only the function that gives speed, I have to ask myself what speed tells me about distance. Since all I know about speed is its definition, I have to ask myself what the definition of speed tells me about distance. By definition, speed is a limit of average speeds, and I think I have come to the end of the trail. What does average speed tell me about distance?

I'll let g denote the function that gives the speed of an object during the interval of time, [a,b]. I am going to suppose, for the moment, that I have the function I'm looking for, the function that relates time and distance. If I call this function, f, the distance traveled between t = a and t = b is f(b) - f(a). Distance is no problem if I know the function, f. The problem is to relate f(b) - f(a) to speed.

To make it more concrete, I'm going to suppose that the object was observed at t = 2 and so 2 is in the domain of f. The average speed is the ratio of the change in distance to the change in time, so I see some possibility of sneaking distance into the discussion.

$$\text{As } t \to 2,\ f(t) - f(2) / (t - 2) \to g(2) = \text{'speed at } t = 2\text{'}.$$

If t is close to 2,

$$f(t) - f(2) / (t - 2) \text{ is close to } g(2)$$

and

$$f(t) - f(2) \text{ is close to } g(2) \times (t - 2).$$

It appears that the distance traveled between times t and 2 is close to $g(2) \times (t - 2)$. I'll see what I can do with $g(2) \times (t - 2)$ or at least something like this.

There is good news and bad news.

The good news is that I have an approximation of distance that involves only speed and the length of a time interval. Since speed and time are all that I know, this is indeed good news.

The bad news is that I don't get an exact distance, only an approximation, and I don't even know how close the approximation is to the actual distance. Further, the approximation arises from a limit as t goes to 2, so the approximation will be good only if t is close to 2, and the interval is pretty small. I had grander aspirations.

I seem to be stuck with approximations of distances traveled over very small intervals of time, and these approximations are of doubtful quality. This is a far cry from finding the exact distance traveled over any interval of time, but this is the hand I've been dealt, so I'll play it.

I am going to use approximations over small time intervals to find an approximation over any time interval. My plan is to break up a given interval of time into small intervals, approximate the distance traveled over the small intervals of time, and then add up my small interval approximations. My starting point for the small interval approximation will be $g(2) \times (t - 2)$. The error in the approximation over the given interval is the sum of the errors on each of the small intervals and I want to be able to make the error over the given interval as small as I want by making the 'small intervals' small.

I really don't care for the term 'small intervals' and can tolerate it no longer. I'm going to call them **'subintervals'**. I think of the small intervals as being contained in a big interval, so 'sub' is a reasonable prefix. The prefix, 'sub' is very popular in mathematics.

Because I have no equalities, only 'close to' relationships, I seem to be led into approximations and at present I see no way of getting an exact value for distance. As I think about it, though, the definition of speed just gave a way to approximate speed. The definition said that I could approximate speed as closely as I wished by average speeds. It seemed a happy accident that in the case of the falling rock, I could actually compute an exact Ideal World value for speed. Fortunately, this accident turns out to be the rule rather than the exception.

If I can do as well for distance as I did for speed and be able to approximate it as closely as I please, I will feel fortunate. And who knows? Maybe I'll get lucky once in a while be able to compute distance exactly. It could happen.

The motion has less chance to change if the time interval is smaller and it seems to me that the motion approaches uniform motion with speed, g(2), as the interval shrinks to zero. This intuition makes me think that $g(2) \times (t - 2)$ is going to be a good approximation.

In the Real World, I can find an interval of time so small that I can measure no change in motion, and the Real World computation of distance over such a small interval would be exact. In the Ideal World, no matter how small the interval is, I can always detect a change in motion and any speed that I might multiply by the time is only an approximation of distance.

Unfortunately, being able to approximate distance over the subintervals is not enough to ensure a good approximation over any interval.

If the error on each subinterval is ‘d’ and if there are ‘n’ subintervals, then the error on the large interval is $n \times d$. As I make the time intervals smaller, decreasing ‘d’, I also make more subintervals, increasing ‘n’. Does ‘d’ go to zero faster than ‘n’ goes to infinity? It is not immediately clear.

For example, if the given time interval is [0 , 1] and I make the length of the subintervals, 1/n, then it is possible that the error on each of the n subintervals is $1/\sqrt{n}$, and the aggregate error is $n \times 1/\sqrt{n} = \sqrt{n}$. The length of the subintervals goes to zero and the error on each of the subintervals goes to zero, but the total error goes to infinity. I hope that this is not possible.

If my idea is going to work, the error on each of the subintervals must go to zero faster than the number of intervals gets large.

I am taking it for granted that the error, ‘d’, on the subinterval, goes to zero as the length of the subinterval goes to zero. I can not bear the thought of it being otherwise so I won’t think about it.

My decision to try for an approximation necessitates some thought about what approximation means. I suppose that a guess is an approximation and an educated guess is a better one. What makes an approximation ‘good’? An approximation is good if it satisfies some practical requirement, like being within .0001 of the actual distance. Its degree of goodness follows from how close it is to the ideal. I think I’ve heard that song before.

I am interested in an approximation <u>process</u>. Approximations in an approximation process depend on some ‘quantity’. As the ‘quantity’ approaches something, say 0 or infinity, the approximations approach the actual number being approximated. I decide how small I want the error to be and then determine a value of this ‘quantity’ that makes the approximation closer to the actual value than the allowable error.

In the approximation of 1/3 by 0.333...3, the ‘quantity’, which goes to infinity, is the number of 3’s. If I want the decimal to be within 0.0001 of 1/3, I make the number of 3’s greater than 4. In the approximation of speed by the average speed, the ‘quantity’ is the length of the averaging interval, and the ‘quantity’ goes to zero.

In the approximation of the distance traveled, the ‘quantity’ will be the length of the largest subinterval and it goes to zero. The general idea is that I make the approximation better by making the subintervals smaller. Given an allowable error, I would like to be able to find a positive number, δ, such that if the length of all the subintervals is less than δ, then the difference between the approximation and the actual distance is less than the allowable error.
There is a price to be paid for good approximations. Good approximations require smaller subintervals, which means more subintervals, which means more computations. Good approximations are paid for in the ultimate commodity, time.

The difference between Ideal World approximation and Real World approximation is that Ideal World approximation can be made as close to the ideal as I want it to be, and Real World approximation is limited by the accuracy of measuring instruments and the very nature of the Real World. In the Ideal World I can make the time interval that I am averaging over as small as I need to get whatever accuracy I want, in say, finding speed. I can make it 10^{-100} without batting an eye.

In the Real World, I would suppose that 10^{-9} seconds is about as small as I can make the interval. This is a tremendous advantage for the Ideal World.

LECTURE 8-2

My car slowly took shape out of the darkness as I walked toward it. Low and sleek, it was built for speed. I stood for a moment and just looked at it before I opened the door...

I am prepared to approximate distance traveled on small intervals. I'll call the function that models speed g, I'll suppose that the object was observed at t = 2, and I'll consider the interval from t = 2 to t = 2 + h where h is positive. I am going to use the interval [2,2+h] because I hope that the error of the approximation will depend on the length of the interval, h, and I want it involved in the action.

It occurs to me that if I knew the average speed over the interval, I would then know the exact distance traveled. The average speed is computed using the exact distance so

exact distance traveled = average speed $\times$ h.
in the interval [2,2+h] over the interval

As h goes to zero, it's the average speed that changes as it gets close to the speed, g(2), so I tend think of the average speed approximating speed. But I can look at it the other way around. The speed, g(2), is close to the average speed over the interval, [2,2+h], and I can think of g(2) as approximating average speed.

exact distance = (average speed over the interval) $\times$ h,

is approximated by

speed $\times$ h = g(2) $\times$ h.

I don't see anything really special about the time, t = 2, and it seems to me that the speed at any time between 2 and 2 + h should be an approximation of the average speed. I wonder if the speed at some other time in the interval would give a better approximation of average speed than g(2).

Wondering doesn't put any hay in the barn, so I think I'll address this possibility with some degree of seriousness. I need a toehold in this problem and I am going to look at an example to see if I can find one. I know that if the function relating distance and time is $f(t) = 16t^2$, then the function that relates speed and time is g(t) = 32 t. This gives me a speed to work with that I know quite a bit about.

The average speed over the interval [t,t+h] is

$$f(t + h) - f(t) / h = 16\{(t + h)^2 - t^2 / h\} = 32 t + 16h,$$

so the average speed over the interval [2,2.1] is 65.60.

I can now compare the average speed, 65.60, with the speeds at 0.01 intervals between 2 and 2.1 to see if there is a time when the speed is closer to 65.60 than g(2).

time	speed	
2.00	64.00	
2.01	64.32	
2.02	64.64	
2.03	64.96	
2.04	65.28	
2.05	65.60	Hmmmmmm
2.06	65.92	
2.07	66.24	
2.08	66.56	
2.09	66.88	
2.10	67.20	

This is spectacular enough to make me do a little algebra. Could it be possible that I can find a value of time, c, in any interval, [t , t+h], where the 'speed at time = c' <u>equals</u> the average speed,?

I have an expression for average speed over [t , t+h], 32t + 16h, and I have an expression for speed at t = c, g(c) = 32c. I can use this information to write an equation that can be solved for c,

(speed at t = c) = average speed over [t,t+h],

$$32c = 32t + 16h.$$

The solution of this equation is

$$c = t + .5h.$$

It appears that the 'magic time', c, is the midpoint of the interval, and the distance traveled in the interval, [t , t+h], is $g(t + .5h) \times h$.

This is amazing. The interval doesn't have to be small, it can be of any length, and I am not approximating distance, the answer is exact. I better tie this down with the example.

I know the speed of a dropped rock is given by g(t) = 32t. I want to find how far it falls between times a = 2 and b = 5. In this case, the value of t in the 'magic number' formula is 2 and the value of h is 3. The speed at '$c = 2 + .5 \times 3 = 3.5$ sec.' equals the average speed over the interval, so the average speed is $g(c) = 32 \times 3.5 = 112$ ft./sec.. The distance equals the average speed times the length of the time interval, which is 3 sec., so, the distance is
$112 \times 3 = 336$ ft.. This is about the height of a 33 story building, which is pretty tall.

If I use this technique on the interval [0,t], I get

(The distance the rock falls between 0 and t)

equals

(the average speed over [0,t]) $\times$ (the length of the interval = t)

equals

(speed at the 'magic time = t/2') $\times$ (the length of the interval = t).

equals

$g(t/2) \times t$

equals

$(32 \times t/2) \times t$

equals

$16t^2$.

(The distance the rock falls between 0 and t) = $f(t) = 16t^2$, and I am back to the rule of the function that relates distance and time.

This seems so slick. If I want to find a distance, I do a 'little algebra', find the 'magic time', multiply it by the length of the time interval, and there it is, the distance.

distance at time, $t = g(c) \times t$,

where c is a number such that g(c) = average speed over [0,t] .

The lynch pin of this method is the 'magic time', the time when speed equals average speed. If I know how to find 'magic times' for any interval of time, I can go from the function that models speed to the function that models distance. This doesn't seem so hard.

My smug self-satisfaction is not long lived. The 'little algebra' step was setting up and solving an equation that involved the function, f. The equation, for the interval, [0,t], was

$$f(t) - f(0) / t - 0 = g(c),$$

and I solved it for c. Unfortunately, I now remember that I was just pretending that I knew f, in an effort to see what was going on. I don't know f, I only know g. The whole point of the process is to find f.

Unless by some miracle the 'magic time' is always the midpoint of the time interval for every object in motion, I need to know the function I'm looking for in order to find the function I'm looking for. Since I don't think a miracle is going to happen, I'm going to look else where for distance.

It seems ironic that if I know the function that relates distance and time, then I know average speeds, but if I know the speed at every time in an interval, average speed seems inaccessible. This doesn't seem reasonable and I rather expect that the way to get to the average speeds from speed is through the function that relates time and distance. I need to find that elusive function.

I have been sidetracked from my original plan with a target of opportunity and it didn't work out. You win some, you lose some.

LECTURE 8-3

There was a small dent that I had put in the right front fender last weekend and I felt a twinge of annoyance at my carelessness. I opened the door and got in...

I don't at the moment see how 'magic times' are going to lead me to the function that relates time and distance, but I am finding it hard to put the idea down. I find the possibility of a time when the speed equals the average speed intriguing.

I think that it was a fluke that I could solve so easily for the 'magic time' in the example. I think that equations rarely have explicit solutions, so that even if I have the function that models distance, I could only approximate the 'magic time'. Worse yet, there are lots of equations that don't have any solutions at all, $x^2 + 1 = 0$, for example, and the equation for the 'magic time' could be one of those. There could be functions that don't have 'magic times'.

I can eliminate the no 'magic time' possibility and will argue that 'magic times' exist for any object in motion on any interval of time. There is always a time in any interval when the speed equals the average speed.

My first clue that this is true comes from the table I constructed earlier.

time	speed
2.00	64.00
2.01	64.32
2.05	65.60
2.09	66.88
2.10	67.20

The speed at $t = 2$ is less than the average speed and the speed at $t = 2.1$ is greater. The 'principle of continuity' tells me that the rock must attain all the speeds in between 64 and 67.2, so the rock must attain the average speed, which lies between them.

If the speed of an object is always greater than some number, 'a', then it seems to me that the average speed would be greater than 'a'. If I keep the speedometer over 60 mph all the way to Santa Fe, I have no trouble believing that my average speed would be over 60 mph. If this is the case, then the speed can't always be greater than the average speed because if it were, the average speed would be bigger than itself. How absurd.

The same argument implies that the speed can't always be less than the average speed.

This leaves two possibilities, either the speed is always equal to the average speed, or it is sometimes greater and sometimes less than the average speed. Since the speed changes continuously, there must be at least one ‘magic time’ when the speed equals the average speed.

I am going to express these ideas in terms of functions.

If g is the function that models speed, then for every interval, [a, b], in the domain of g, there is a ‘magic time’, c, $a < c < b$, where g(c) = (average speed over the [a ,b]).

In terms of the functions f and f `, there is a time, c, $a < c < b$, where

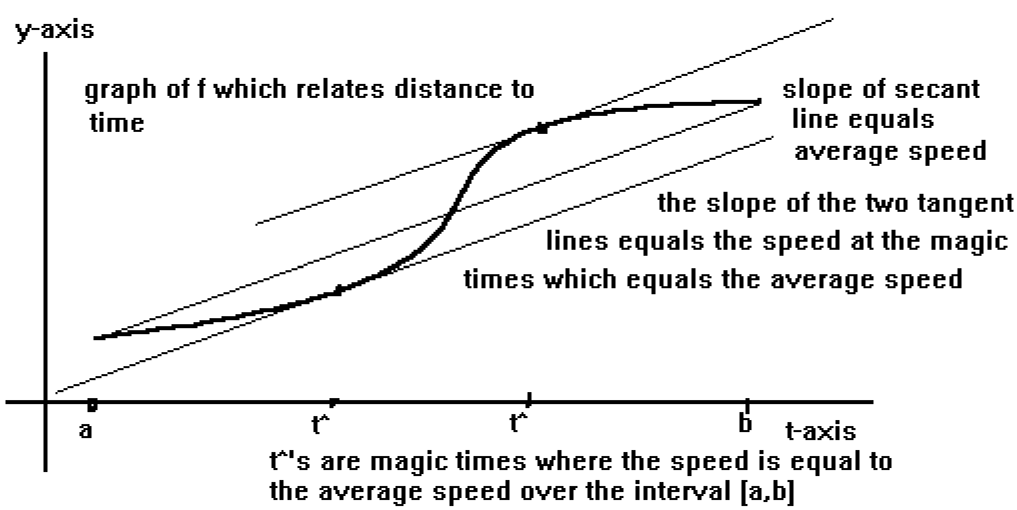
$$f(b) - f(a) / b - a = f`(c)$$

or

$$(\text{distance traveled between a and b}) = f(b) - f(a) = f`(c) \times (b - a).$$

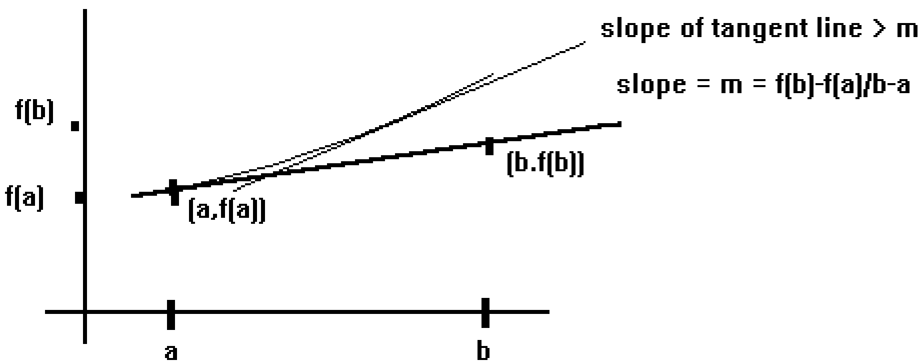
This looks like such a promising way to find distance that I am tempted to keep after it, but a person has to know when to fish and when to cut bait. I choose the latter but unlike Richard the Lionhearted, who turned his back on Jerusalem when he realized he couldn’t conquer it, I shall have a last look at ‘magic times’ before I leave.

There is geometric meaning in the expression $f(b) - f(a) / b - a = f`( c )$. It says that there is a point c, $a < c < b$, where the slope of the secant line joining the points (a,f(a)) and (b, f(b)) on the graph of f, equals f `(c). Actually there could be several such points.



‘Magic times’ owe their existence to the existence of the tangent lines at each point and the ‘principle of continuity’. Looking at this picture, it seems to me that if f is any function that has a derivative for values of t satisfying $a \leq t \leq b$, and <u>the derivative satisfies the ‘principle of continuity’</u> then there is a value, c, lying strictly between a and b, $a < c < b$, such that [f(b) - f(a)] / b-a = f(c). This is not only true, it has a name: The Mean Value Theorem.

Since f doesn't model anything, I can't use a speed argument so I'll try a geometric argument.



If the slope of the tangent line is always greater than the slope of the secant line, the graph of f rises faster than the secant line and the graph can't get back down to pass through the point (b,f(b)). The same idea shows that the slope of the tangent line can't always be less than m. That means that the slope of the tangent line must either always equal m or, be sometimes larger and sometimes smaller, than m. The derivative satisfies the 'principle of continuity' so the slope of the tangent line satisfies the 'principle of continuity', and there must be a value of x, c, where the slope of the tangent line equals m = f(b)- f(a) / b -a. Using the fact that the instantaneous rate of change is the slope of the tangent line, I can say that there is a value of x, c, where the instantaneous rate of change of f equals the average rate of change.

$$f`(c) = f(b) - f(a) / b - a.$$

This is all very nice, but I just don't see where to go from here with 'magic times'.

LECTURE 8-4

I closed the door and started the engine. Something didn't feel quite right. I turned on the radio and caught the end of "Cowgirl in the Sand". It has an excellent stereo...

In my first attempt to find distance, I fixed the interval and tried to find a time when the speed was as close to the average speed as possible. My thought was to multiply this speed by the time and use this product to approximate the product of the average speed and time, which is the actual distance traveled. In one sense I was quite successful because I convinced myself that there was a time when the speed actually equaled the average speed. This speed multiplied by time would give me distance. I was less successful in the sense that I could find no way to compute that time and I was stuck.

Actually, I had strayed from my original intent of adding up the distances traveled over small intervals of time to find the total distance traveled. I was beguiled by the possibility of a quick, slick way to find distance over the entire interval and, like most quick, slick fixes, it didn't work out. I should have carried out my first idea and not gotten side tracked, but what's done is done. "The moving finger writes and having writ moves on, nor all your wit nor piety can lure it back to cancel half a line, nor all your tears wash out a word of it." I do not regret that I strayed but I think I should get back to the original plan.

It is my contention that I can approximate the distance traveled on a subinterval by multiplying the speed at <u>any</u> time in the subinterval by the length of the subinterval and then, approximate the 'distance traveled on any interval of time' by summing the approximations over the subintervals.

This approach will require a lot of talk about sums of 'approximations over subintervals' and I need a name for these sums. With a clear lack of imagination, I'm going to call them 'approximations' or 'approximating sums'.

The first step is to say exactly what I mean by 'approximate the distance traveled in an interval of time'. It is always so much easier to do something if you know what it is you are trying to do.

When I say that I can 'approximate the distance traveled' I mean that if I am given any error tolerance, $\varepsilon > 0$, I can make the difference between the 'distance traveled' and the 'approximation' less than ε, if I make the subintervals small enough.

The approximation of distance has two parts:

1. The approximations get close to some number as the subintervals get small.

2. The number the approximations get close to is equal to the 'distance traveled'.

This is what I mean by approximating the distance traveled.
It is my contention that I can take any function, g, that models speed, and make

| distance traveled - approximation |

as small as I please if I make the subintervals small enough. This would show that I can approximate the distance traveled.

Just to see if I am in the ballpark, I am going to try this process with the function that I have come to know and love, the function that models the speed of the falling rock. The speed of the rock is given by the function whose rule is $g(t) = 32t$. In this case I know that the function that relates distance and time is $f(t) = 16t^2$ and, in particular, that the object moves $f(2) - f(0) = 64$ ft. in the interval of time, [0,2].

First I'm going to show that the approximations get close to the actual distance traveled during the interval of time, [0,2]. If this effort is successful, I'm going to find the number that the approximations get close to. If this number is 64, I will know that I have led a good life and that Providence is smiling upon my efforts.

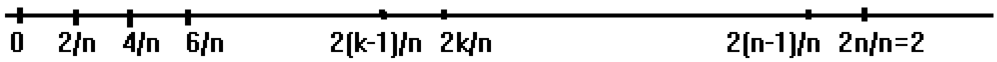
The first step is to break up the interval [0,2] into subintervals and approximate the distance on each subinterval. I think that I will make all the subintervals have the same length. If all the subintervals are small, it shouldn't make any difference if they are the same size or not, but I think the arithmetic will be a lot easier if they are all the same size. I'm doing this to see if the idea is viable and I see no advantage of doing it with weights on my legs. If this works, I'll see if I can do it without this restriction.

Ultimately, I want as few restrictions on the approximations as possible and this is for two reasons. If I have a real problem and am getting paid to find distance, I want as much freedom as possible in making the approximations. I don't want to have to worry about whether the subintervals all have the same length, just whether they are small enough. Likewise, the only restriction on g should be that it model speed. I am starting with a very particular function, g, that models the speed of a falling rock. I am going to try to make g as general as I can and will remove restrictions whenever and wherever I can.

The second reason has to do with esthetics. The esthetics of the Ideal World require that truth be expressed in the most general terms possible. If my hypothesis requires that a function is continuous and increasing but the result is true if the function is just continuous, I have made an inelegant statement. Even though all the functions that model the Real World process I'm interested in are increasing, the Ideal World sees it as lack of art. Real World utility is of little interest to Ideal World taste.

The hypothesis should be that g is any function that models speed, and anything short of this is inelegant. I will strive for that perfection.

I’ll break the interval into n equal subintervals, each of length of 2/n.



The kth subinterval is [2(k-1)/n , 2k/n].

The error on each subinterval is the key and I want to examine it very carefully. I might as well start with the first one, [0,2/n].

The function g is increasing, so $g(0) < g(2/n)$, and $g(0) \leq g(t`) \leq g(2/n)$, if t\` is <u>any</u> value in [0,2/n], $0 \leq t` \leq 2$. When I multiply these three speeds by the length of the interval, 2/n, I get three approximations of the distance traveled over [0,2],

$$g(0)\ 2/n \leq g(t`)\ 2/n \leq g(2\backslash n)\ 2/n.$$

The approximation ‘g(t\`) 2/n’ lies between the extreme approximations, ‘g(0) 2/n’ and ‘g(2\n) 2/n’, and this is because g is increasing.

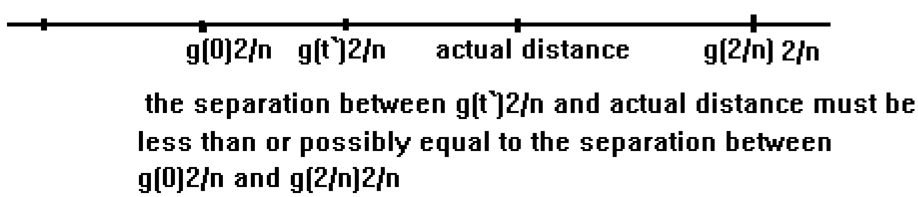
The speed, g(0), is the least speed in the interval, [0,2/n], so ‘g(0) 2/n’ is the least possible distance the rock could fall . Likewise, g(2) is the largest speed in the interval and ‘g(2/n) 2/n’ is the greatest possible distance the rock could fall. This means that the actual distance traveled also lies between g(0) 2/n and g(2\n) 2/n.

$$g(0)\ 2/n \leq \text{actual distance} \leq g(2/n)\ 2/n.$$

Both the actual distance and the estimate g(t\`) 2/n are trapped between g(0) 2/n and g(2\n) 2/n. Again I get a chance to smile.

$$|\ \text{actual distance} - g(t`)\ 2/n\ | \leq g(2/n)2/n - g(0)2/n$$

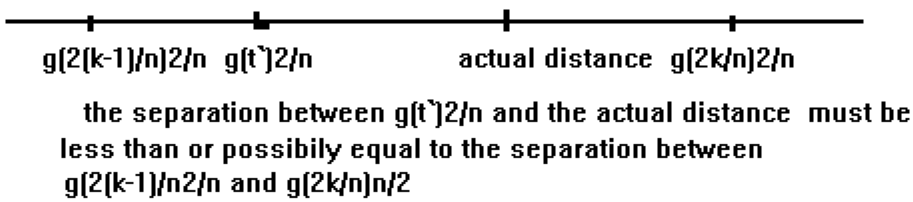
$$\leq [g(2/n) - g(0)] \times 2/n\ .$$



There is nothing special about the first interval in the estimate. If $t`_k$ is <u>any</u> point in the kth subinterval, [2(k-1)/n , 2k/n], then

$$g(2(k-1)/n) \le g(t`_k)) \le g(2k/n).$$

As before, the ‘actual distance’ and the estimate, ‘$g(t`_k)$ 2/n’, are trapped between the least possible distance, g(2(k-1)/n) 2/n and the greatest possible distance, g(2k/n) 2/n.



$$| \text{actual distance} - g(t`_k)\, 2/n | \le [g(2k/n) \times 2/n] - [g(2(k-1)/n)] \times 2/n$$

$$\le [g(2k/n) - g(2(k-1)/n)] \times 2/n.$$

I have been using g(t) instead of 32t because I hope someday to go through this process for any function, g, that models speed and this helps me see if it’s possible. So far I have had no need to use 32t, but now I do. I’ll evaluate g at the endpoints of the subinterval and get an explicit estimate of the error on a subinterval.

$$\text{Error on a subinterval} = |\text{actual distance} - g(t`_k)\, 2/n| \le [32 \times 2k/n - 32 \times 2(k-1)/n]\, 2/n$$

$$\le 64[k/n - (k-1)/n]\, 2/n$$

$$\le 128\; 1/n^2 .$$

The error estimate doesn’t depend on which interval I’m looking at, just on the size of n which determines the interval length. I love it when things like this happen. I think this is going very well.

If I denote the distance traveled in the kth interval by S(k), and the distance traveled in [0,2] by S, then

$$S = S(1) + S(2) + ... + S(k) + ... + S(n-1) + S(n) = \text{distance traveled}$$

and

$$A = g(t`_1)\, 2/n + g(t`_2)\, 2/n + ... + g(t`_k)\, 2/n + ...+ g(t`_{n-1})\, 2/n + g(t`_n)\, 2/n$$

is an approximation of the distance traveled.
Total error = | S - A |

$$= | [S(1) - g(t`_1) 2/n] + [S(2) - g(t`_2) 2/n] + ... + [S(n) - g(t`_n) 2/n] |$$

$$\leq | S(1) - g(t`_1) 2/n| + |S(2) - g(t`_2) 2/n| + ... + |S(n) - g(t`_n) 2/n |$$

$$\leq 128\ 1/n^2 + 128\ 1/n^2 + ... + 128\ 1/n^2 .$$

There are n equal terms in the last sum so

$$\text{Total error} \leq n \times 128\ 1/n^2 = 128/n.$$

The fact that the error is the same on each subinterval has made a substantial contribution to the quality of my life.

Making the lengths of the subintervals, 2/n, close to zero, makes n large, and *vice versa*. This means that letting the length of the subintervals go to zero is equivalent to letting n go to infinity. As the interval length goes to zero, n gets large and the error, which is bounded by 128/n, goes to zero.

If I want the error in my approximation to be less than 0.0001, I make n so large that 128/n < 0.0001, or n > 1280000, or the length of the subintervals less than 1/640000. I don't know what the number for distance is, but I know some numbers that are within 0.0001 of it. Up through the first three decimals, I can't tell the distance apart from the approximation.

I have shown that if the subintervals have equal length, then the approximations approach the distance traveled, which I think is quite an accomplishment, but I am not done. I have shown that the approximations approach the distance but I haven't found the number itself.

In the Real World everything is approximation anyway and there is neither need nor capability of continuing and finding the number exactly. I can't let n go to infinity and I can't let the subintervals shrink to zero.

In the Ideal World, numbers are usually given by specifying ways to approximate them and I suspect that the Ideal World is satisfied that these approximations deliver the distance. But in the Ideal World I can let n go to infinity, I can let the subintervals shrink to zero, and I have a shot at finding the actual number, whether the Ideal World thinks it necessary or not.

I would like a number. Maybe it is true that for a general function, g, that models speed, I would have to settle for the approximations, but this isn't a general function, it's g(t) = 32t. I want a number.

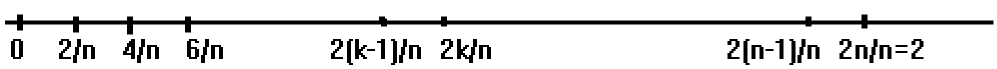
LECTURE 8-5

It was dark as I left the parking lot and headed for the 2nd Street on ramp. This car looked like my car and drove like my car, but it wasn't my car...

I am going to do some algebra on the approximations that used $g(t) = 32t$ for speed and [0,2] for the interval, and find the number they approach as the length of the subintervals goes to zero. Since I showed that approximations with equal length subintervals approach the distance, I will use equal length subintervals. Because I know that the approximations approach the distance, I know that the number I find is, indeed, the distance. I am trying to save my arm for the big game, so I am going to write 'equal subinterval' for 'equal length subinterval'.

Because I know that all the approximations with equal subintervals approach the distance as the length of all the subintervals goes to zero, any sequence of approximations, whose equal subintervals all go to zero, should approach the distance.

I will follow the procedure that I have already used and break up the interval, [0,2], into n equal subintervals, each of length of 2/n, and approximate the distance on each subinterval.



The kth subinterval is [2(k-1)/n , 2k/n].

The fact that I can approximate the distance using the speed at any time in a subinterval, means that I have to pick a time. The function g is increasing so the time at the left end of the subinterval gives the least speed in the subinterval and the approximation would give an under estimate of the distance. Using the right end of the interval would give an over estimate of the distance.

The mid-point of the interval would probably give the best estimate but I think the algebra would be harder. I'm going to opt for easier and evaluate the speed at the right end point of each subinterval. It isn't supposed to make any difference which time I pick it and it would seem perverse to choose a 'hard' time.

The approximation of distance traveled in the first subinterval is

$$g(2/n) \times 2/n = (32 \times 2/n) \times 2/n.$$

The approximation in the 2nd subinterval is

$$g(4/n) \times 2/n = (32 \times 4/n) \times 2/n.$$

The approximation in the kth subinterval is

$$g(2k/n) \times 2/n = (32 \times 2k/n) \times 2/n.$$

In this approximation, g(2k/n) is the speed and 2/n is the length of time. Their product is my approximation of the distance traveled during the interval of time [2(k-1)/n , 2k/n].

The approximation in the next to last subinterval is

$$g(2(n-1)/n) \times 2/n = (32 \times 2(n-1)/n) \times 2/n$$

and the approximation in the last subinterval is

$$g(2n/n) \times 2/n = (32 \times 2n/n) \times 2/n.$$

I add all these up to get the approximation of the total distance traveled in the interval, [0,2].

$$32(2/n)(2/n) + 32(4/n)(2/n) + ...+ 32(2k/n)(2/n) + ...+ 32(2(n-1)/n)(2/n) + 32(2n/n)(2/n)$$

I wrote 2n/n instead of 2 in the last term because I am looking for patterns. Every term has a 32 and twice an integer in the numerator and there is an n^2 in every denominator. I want the last term to reflect this regularity.

I set sail to capture the great white whale,

$$32(2/n)(2/n) + 32(4/n)(2/n) + ...+ 32(2k/n)(2/n) + ...+ 32(2(n-1)/n)(2/n) + 32(2n/n)(2/n).$$

The first thing I do is ‘kick the can’ and factor out everything in sight.

$$(32 \times 2 \times 2 / n^2)(1 + 2 + ... + k + ... + (n-1) + n)$$

$$128 / n^2 \ (1 + 2 + ... + k + ... + (n-1) + n).$$

Maintaining a perpetual state of “want” , I want

$$128 / n^2 \ (1 + 2 + ... + k + ... + (n-1) + n)$$

to approach 64 as n gets large. I don't find this obvious because as n gets large, $128/n^2$ goes to zero, $(1 + 2 + ... + k + ... + (n-1) + n)$ goes to infinity, and the destination of the product is in doubt. There is a race between $128/n^2$, going to zero, and $(1 + 2 + ... + k + ... + (n-1) + n)$, going to infinity. If I want to know the outcome of the race, I must deal with the sum

$$(1 + 2 + ... + k + ... + (n-1) + n).$$

In the Real World I might just leave the sum, since it is grist for the computer mill, but I am thinking in the Ideal World. I want to see what happens as n goes to infinity.

LECTURE 8-6

I slid into the sparse, early evening traffic. The green freeway signs ticked by and I thought that I could have been in any of a hundred cities, driving any fast, expensive car. I wondered whose this one was...

There are a couple of ways that I can deal with the sum of consecutive integers and the first is algebraic. My approach is motivated by an example with n = 4.

$$\begin{array}{c} 1 + 2 + 3 + 4 \\ 4 + 3 + 2 + 1 \\ \hline 5 + 5 + 5 + 5 \end{array}$$

$$2 \times (1+2+3+4) = 4 \times 5 \quad \text{and}$$

$$(1+2+3+4) = 1/2\ (4 \times 5) = 10.$$

I start by naming the sum; $S = (1 + 2 + ... + k + ... + (n-1) + n)$. Since the order in which I add the terms doesn't matter, I can write them in reverse order and still have the same sum.

$$S = 1 + 2 + 3 + ... + k + ... + (n-2) + (n-1) + n.$$

$$S = n + (n-1) + (n-2) + ... + (n-k+1) + ... + 3 + 2 + 1.$$

$$2S = (n+1)+(n+1)+(n+1) + ... + (n+1) + ... + (n+1)+(n+1)+(n+1).$$

The remarkable fact is that if I pair the terms vertically, the sum of each pair is (n + 1). Since there are n pairs,

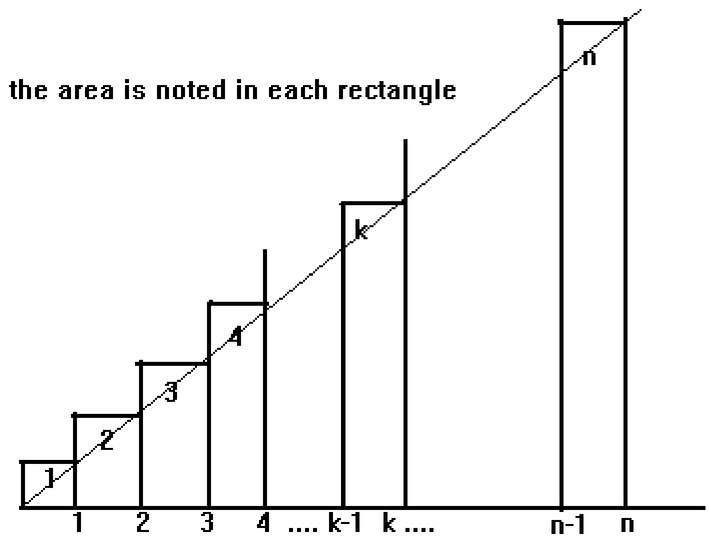
$$2S = n(n+1)$$

and

$$S = n(n+1) / 2.$$

That's one way to add up the first n consecutive integers.

The second way is geometric. An integer, k, is numerically equal to the area of a rectangle of height k and width 1, so (1 + 2 + ... + k + ... + (n-1) + n) equals the combined area of the rectangles in the picture.



(1 + 2 + ... + k + ... + (n-1) + n) = area of triangle whose base is n and whose height is n

+ the area of the n little triangles, each of whose area is 1/2

$= (1/2 \times n \times n) + (n \times 1/2)$

$= 1/2\ n(n+1)$.

1/2 n(n+1) is called the **closed form** of (1 + 2 + ... + k + ... + (n-1) + n).

This step in the process of finding distance from speed is not always quite so straight forward. Often the sum of terms in the approximation of distance can not be put into closed form, and computation of a Real World estimate is the best that can be done.

But in the case at hand, I am after the exact Ideal World answer, nothing less.

I can now continue my effort to find the number the approximations approach. My approximation of the distance traveled is

$$32(2/n)(2/n) + 32(4/n)(2/n) + ...+ 32(2k/n)(2/n) + ...+ 32(2(n-1)/n)(2/n) + 32(2n/n)(2/n)$$

$$(128 / n^2) (1 + 2 + ... + k + ... + (n-1) + n)$$

$$(128 / n^2) 1/2\ n(n+1)$$

$$(64 /n^2) (n^2 + n).$$

$$64 (1 + 1/n)$$

$$64 + 64 / n.$$

I must stop for a moment and gaze in amazement at this result. My ungainly 'great white whale' has boiled down to a minnow.

$$32(2/n)(2/n) + 32(4/n)(2/n) + ...+ 32(2k/n)(2/n) + ...+ 32(2(n-1)/n)(2/n) + 32(2n/n)(2/n)$$

$$= 64 + 64 / n.$$

As n goes to infinity, 64/n goes to zero, and the approximation goes to 64.

I used $g(t) = 32t$, because I knew that the function that relates distance and speed is $f(t) = 16t^2$ and that, in particular, the object moves $f(2) - f(0) = 64$ ft. in the interval of time, [0,2]. My answer is correct and my method is still flying.

I think success is within my grasp. For one function at least, I have computed distance from speed. I don't say that it is the easiest process in the world, but the point is not that finding distance may be computationally hard, the point is how amazing it is that I can do it at all.

LECTURE 8-7

I pulled off the freeway and stopped at a convenience store. The registration was in the glove box and I got out and checked the license. As far as I could tell, this car was exactly like my car...

I have used particular examples so that I could gain insight into what might be true generally and I think the time has come to see what insight I have gained.

I think that the error in each subinterval had the same bound because the length of each subinterval was the same and because g had a specific, easy rule. I doubt that this happens often and I certainly won't count on it.

The equal subinterval requirement seems artificial and the error in the estimate should depend only on the maximum of the lengths of the subintervals. If the subintervals are all small, the approximation should be close to the ideal whether the subintervals are all the same length or not. I'm going to drop the requirement that the subintervals have equal length and accept some loss of control over the error. I think I can still make the error go to zero as the subinterval length goes to zero.

The function, g, models something that happens in the Real World, namely speed, and so I can suppose that it is continuous. There is another condition on g that has been hidden in plain sight. Speed is an inherently non-negative quantity. It is, at heart, the ratio of distance to time and both distance and time are non-negative. This forces, perhaps unfairly, the ratio to be non-negative. Included in my statement that a function models speed is the fact that the function is non-negative.

The speed of the rock is always increasing, which is the same as saying that the function, $g(t) = 32t$, that models the speed, is an increasing function. I am going to assume that the function that models speed is increasing. This condition also seems artificial but I really used it a lot in the example. Most of the error estimates in the approximation of distance on the subintervals depended on the fact that g was an increasing function, and I don't want to think about dropping that condition now. I would like to eventually get rid of it, but I've got enough balls in the air and I don't need another one right now. Later, maybe.

My next step is to approximate distance traveled where g is any continuous, increasing function that models speed.

I don't say, "My next step is to find distance.... .", I say, "My next step is to approximate distance.... ." This is hard for me to say and I think it's hard because of my own ante-calculus experience. In my years before calculus, it seemed to me that mathematics was finding numbers. In arithmetic I found numbers by adding, multiplying, subtracting and dividing. In algebra I solved equations which had explicit answers and found the width of a sidewalk, the price of a bus ticket, the roots of an equation, and so on.

It was my perception that mathematics found numbers explicitly and that the last line of a problem's solution should read, "x = 4" or whatever.

I wanted a number and there was an explicit formula that gave that number. It was the responsibility of a productive member of society to get a number and I felt that if I didn't get a number, I was remiss in my duties as a citizen.

My introduction to calculus seemed to imply that it was more of the same with explicit formulas that gave derivatives and what-not. I could find explicit numbers that were the volumes of strange solids, the speeds of moving objects, centers of mass, moments of inertia, largest areas, smallest volumes, ... I was caught in a maelstrom of numbers and formulae.

People of good will told me that many of the things I was using were limits of approximation processes, but it didn't seem to really matter. It didn't seem to matter because the limits of the approximation processes were usually explicit formulas and I could use the formulas and forget about where they came from. Actually, the fact that calculus can be used without understanding its foundation in approximation is what makes calculus such a popular applied tool. It's like my car, I can use it knowing absolutely nothing about the internal combustion engine. Until it breaks, that is.

And then one day I had an epiphany and I saw that calculus was all approximation and that large parts of ante-calculus mathematics were as well. The approximation process was not introduced by calculus but was there all along. Calculus just used it to the max.

If I am waiting for the bus and someone comes up to me with a circle whose radius is 2 and asks for its area, I would answer, "4π". It might seem that I have given them a number explicitly, but I have not. I have given them an approximation process. Specifically, I told them to inscribe a regular polygon of n sides in a circle of radius 2, let n get large, and notice that $4 \times$ (ratio of the circumference of the n-gon to the radius of the circle) approaches a number, which is the area of their circle. The number approached by (ratio of the circumference of the n-gon to the radius of the circle) appears so often in mathematics that it has been given a special name, π.

The formula for the area of a circle is no more exact than a distance that I might compute from some speed function. Ideal World formulas seem exact because I have given names to a lot of numbers that are defined as the limits of approximation processes. While $\sqrt{2}$ is an Ideal World number in good standing, it is still <u>defined</u> as the limit of an approximation process. If I say that a root of the equation, $x^2 - 2 = 0$, equals $\sqrt{2}$, I mean that the root is the limit of a certain approximation process.

I now see that approximation is what it's all about. I should be able to give my answer as a limit and not feel as though I had somehow failed, but the impressions gained in youth die hard and I still feel that I should explain and justify my failure to come up with something explicit.

When I say that I can ‘find distance’, which I would rather say than ‘approximate distance’, I do not mean that I can give a formula or a number, it means that I can give an approximation process that approaches distance. Specifically I mean the following:

Given <u>any</u> positive tolerance ε, say $\varepsilon = 0.0001$, I am able to find some positive number, δ, such that <u>any</u> approximation, A, all of whose subintervals have length less than δ, will be within the tolerance of the actual distance, S, or the difference between S and E is less than $\varepsilon = 0.0001$.

If I have an approximation, A, and I want know where S is, I think that S lies between A - 0.0001 and A + 0.0001,

$$A - 0.0001 < S < A + 0.0001.$$

I would say that the distance, S, equals $A \pm 0.0001$.

If I am making approximations and I want to know how close I am to S, I think that A lies between S - 0.0001 and S + 0.0001,

$$S - 0.0001 < A < S + 0.0001.$$

These two ways of thinking about it are exactly equivalent and each implies the other. If

$$A - 0.0001 < S < A + 0.0001$$

then

$$S - 0.0001 < A < S + 0.0001$$

and *vice versa*.

This is what I mean by saying that I can find distance. It is not quite like saying, “distance equals speed times time” .

In the Ideal World generally, if I say that I can find a number, I mean that I can approximate it as closely as I please, and this usually means that I can write the number as a limit.

Sometimes I can find the limit number exactly as in

$$\lim_{h \to 0} 32 + 2h = 32.$$

Sometimes my approximations of distance will approach a number that I can write down, like 64, in the example.

Sometimes I can’t find the limit number exactly, as in

$$\lim_{h \to 0} (1 + h)^{1/h} ,$$

and I give it a name, such as ‘e’.

Sometimes I know the limit number exists but since the number doesn’t come up very often, I don’t bother to name it. That’s the way life is. Sometimes.

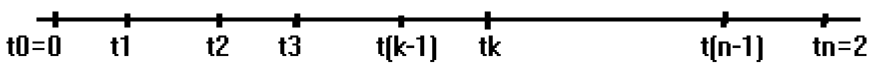
LECTURE 8-8

I did not recognize the convenience store and I felt a need for familiar surroundings. I got back on the freeway...

I am ready for the next step. I'll let g be any increasing, continuous function that models speed on the interval, [0,2]. I, with some confidence, am going to find distance traveled.

I break up the interval into subintervals by designating points in the interval, not necessarily equally spaced.

$$0 = t_0 < t_1 < t_2 < ... < t_k < ... < t_n = 2.$$



The set of points is called a **partition of the interval [0,2].** The kth subinterval is $[t_{k-1}, t_k]$ and I let $t`_k$, $t_{k-1} \leq t`_k \leq t_k$, denote <u>any</u> point in this subinterval. I call $t`_k$ a **sample point** in the subinterval because I am going to sample the speed at that point. I approximate the distance traveled on the subinterval, $[t_{k-1}, t_k]$, by

$$g(t`_k)\,[t_k - t_{k-1}].$$

An approximation of the total distance traveled is

$$A = g(t`_1)(t_1 - t_0) + g(t`_2)(t_2 - t_1) + ... + g(t`_k)(t_k - t_{k-1}) + ... + g(t`_n)(t_n - t_{n-1}).$$

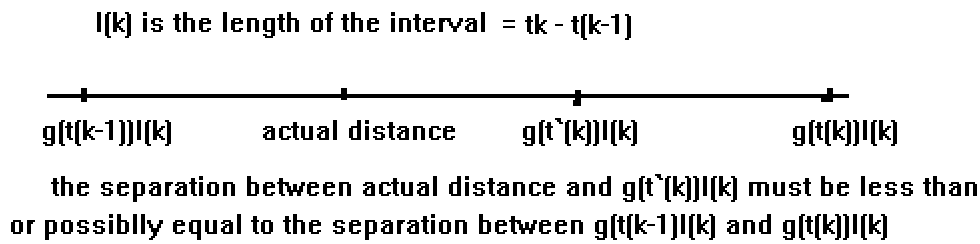
Since g is increasing, the speeds satisfy, $g(t_{k-1}) \leq g(t`_k) \leq g(t_k)$, and

$$g(t_{k-1})\,[t_k - t_{k-1}] \leq g(t`_k)\,[t_k - t_{k-1}] \leq g(t_k)\,[t_k - t_{k-1}].$$

This is why I wanted g to be increasing.

Since $g(t_{k-1})\,[t_k - t_{k-1}]$ is the least possible distance and $g\,(t_k)[\,t_k - t_{k-1}]$ is the greatest, the 'actual distance traveled in the subinterval' is trapped, along with $g(t`_k)[t_{k-1} - t_k]$, between the least and greatest possible distances

$$g(t_{k-1})[\,t_k - t_{k-1}] \leq g(t`_k)\,[t_k - t_{k-1}]\text{ , actual distance } \leq g(t_k)\,[t_k - t_{k-1}].$$



The 'actual distance traveled in a subinterval' can't differ from the approximation, $g(t`_k)[t_k - t_{k-1}]$,by more than the difference between the extremes. If I denote the actual distance traveled in the kth subinterval by S(k),

$$\text{Error on kth subinterval} = |S(k) - g(t`_k)[t_k - t_{k-1}]| \le g(t_k)[t_k - t_{k-1}] - g(t_{k-1}))[t_k - t_{k-1}]$$
$$\le [g(t_k) - g(t_{k-1})][t_k - t_{k-1}]$$

Everything on the right side of the inequality is positive so I don't need absolute value signs. I am certainly glad that g is increasing.

I am going to name the 'maximum of the lengths of the subintervals', **the mesh of the partition,** or usually just **'the mesh'** and call it **l.** Then $t_k - t_{k-1} \le \mathbf{l}$ for all the intervals and I can rewrite the error estimate as

$$|S(k) - g(t`_k)[t_k - t_{k-1}]| \le [g(t_k) - g(t_{k-1})][t_k - t_{k-1}] \le \{g(t_k) - g(t_{k-1})\}\mathbf{l}.$$

The difference between the actual distance, S, and the approximation, A, can now be estimated.

$$\text{Error} = |S - A| = |S(1) + S(2) + ... + S(n-1) + S(n) -$$

$$\{g(t`_1)(t_1 - t_0) + g(t`_2)(t_2 - t_1) + ... + g(t`_k)(t_k - t_{k-1}) + ... + g(t`_n)(t_n - t_{n-1})\}|$$

$$= |S(1) - g(t`_1)(t_1 - t_0) + S(2) - g(t`_2)(t_2 - t_1) + ... + S(k) - g(t`_k)(t_k - t_{k-1}) + ... +$$
$$S(n) - g(t`_n)(t_n - t_{n-1})\}|$$

$$\le |S(1) - g(t`_1)(t_1 - t_0)| + |S(2) - g(t`_2)(t_2 - t_1)| + ... + |S(k) - g(t`_k)(t_k - t_{k-1})| + ... +$$
$$|S(n) - g(t`_n)(t_n - t_{n-1})|$$

$$\le (g(t_1) - g(t_0))\,\mathbf{l} + (g(t_2) - g(t_1))\,\mathbf{l} + ... + (g(t_k) - g(t_{k-1}))\,\mathbf{l} + ... +$$
$$(g(t_n) - g(t_{n-1}))\mathbf{l}$$

Admittedly, the page is dense with symbols, but as I look it over I see that I have just grouped the terms that involve a particular subinterval and repeated one computation n times. That computation is

$$\begin{aligned} | S(k) - g(t`_k)(t_k - t_{k-1}) | &\le g(t_k)(t_k - t_{k-1}) - g(t_{k-1})(t_k - t_{k-1}) \\ &\le \{g(t_k) - g(t_{k-1})\}(t_k - t_{k-1}) \\ &\le \{g(t_k) - g(t_{k-1})\}\mathbf{l}. \end{aligned}$$

I find the error on each subinterval and add them up just as I had planned. It is <u>crucial</u> that g is an increasing function so that $\{g(t_k) - g(t_{k-1})\}$ is positive.

And now a miracle occurs.

$$| S - A | \le (g(t_1) - g(t_0))\,\mathbf{l} + (g(t_2) - g(t_1))\,\mathbf{l} + \ldots + (g(t_k) - g(t_{k-1}))\,\mathbf{l} + \ldots + (g(t_n) - g(t_{n-1}))\mathbf{l}$$

reduces to

$$| S - A | \le [\, g(t_1) - g(t_0) + g(t_2) - g(t_1) + \ldots + g(t_k) - g(t_{k-1}) + \ldots + g(t_n) - g(t_{n-1})]\,\mathbf{l}$$

and then to

$$| S - A | \le (g(t_n) - g(t_0))\,\mathbf{l}.$$

Finally, since $g(t_n) = g(2)$ and $g(t_0) = g(0)$, and I have the moderately pleasing estimate of the error in my approximation,

$$| S - A | \le (g(2) - g(0))\,\mathbf{l}.$$

Personally, I think this is kind of astounding.

The process that collapses all the g's is called **telescoping**. Every term except the first and last has its opposite in the sum.

If I want my approximations to be within 0.0001 of the actual distance, I make the mesh, **l**, so small that $(g(2) - g(0))\,\mathbf{l} < 0.0001$ or $\mathbf{l} < 0.0001 / (g(2) - g(0))$. In the case where $g(t) = 32t$, $\mathbf{l} < 0.0001/64$ is required.

<u>It doesn't make any difference how I pick the subintervals or how I pick the sample points in them as long as I make the mesh less than 0.0001 / (g(2) - g(0)).</u>

$$| S - A | \le \text{tolerance} \quad \text{if} \quad \mathbf{l} < \text{tolerance} / g(2) - g(0).$$

If I have any sequence of approximations that approaches a number, that number must be distance. If I choose my sample points and intervals fortuitously, I may be able to get a sequence of approximations where I can actually see what number they approach; Ideal World number of course. This is what I did in the example using g(t) = 32t. I chose right end points of subintervals as sample points and equal subintervals, and these choices turned out to be fortuitous.

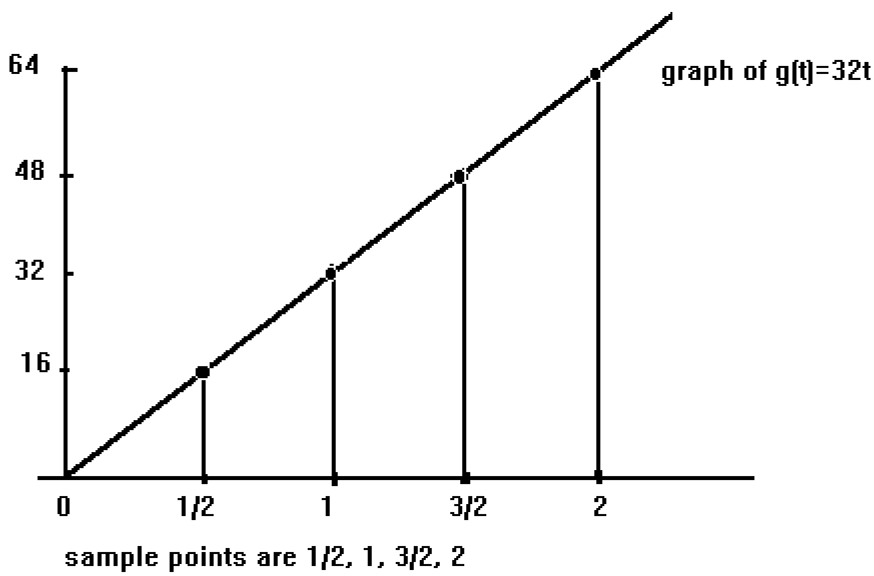
I have effectively solved the problem of finding distance if the speed is increasing. It must surely work about the same way if the speed is decreasing and I assume it does. This little comment lies here quietly but it doubles the number of functions I can use for speed and find distance traveled. If g is any continuous function that models speed which is always increasing or always decreasing, I can find the distance traveled.

LECTURE 8-9

I drove for some time. Even though the exits all seemed familiar, I knew this wasn't my town. Was I the person I thought I was?

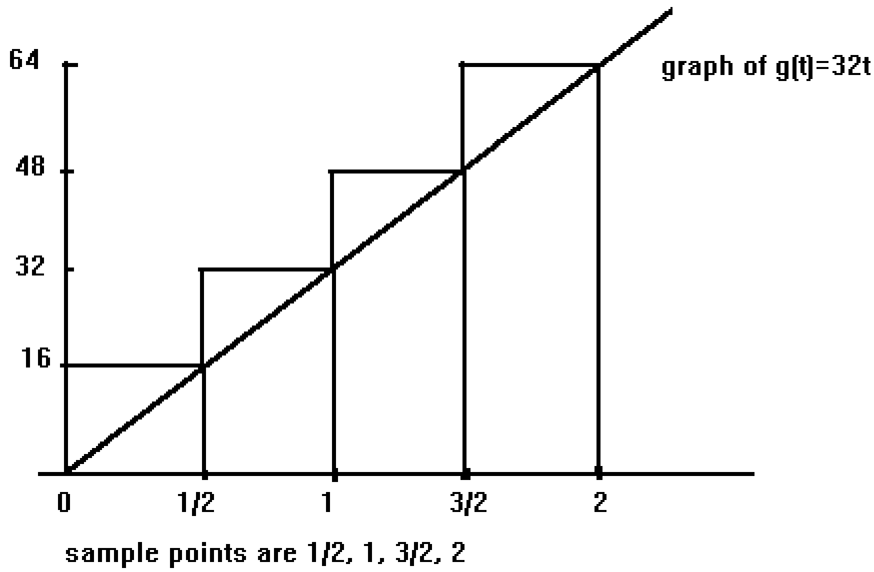
Before I tackle the case where g is any function that models speed, I want see what the picture of the approximations looks like. I would, in the not too distant future, like to have some geometric interpretation of distance as I had for speed. Until now I have been using my mind and its tool, algebra, to think about how speed and distance relate and to compute how the approximations behave. It is time to use my eyes and see if I can picture the approximations.

I'll go back to the example where g(t) = 32t and write an approximation of distance using n = 4 and the right end points of subintervals as sample points. The partition is $0 < 1/2 < 1 < 3/2 < 2$. I can draw the graph of g and designate the speeds at the sample points.



The approximation is g(1/2) 1/2 + g(1) 1/2 + g(3/2) 1/2 + g(2) 1/2. I notice that the 1/2 in each term is the height of a trapezoid in the picture.

But g(2), say, is the vertical distance from t = 2, on the t-axis, to the graph, and remembering the geometric demonstration that 1+2+...+n = n(n+1)/2, I might better think of the '1/2' as the width of a rectangle whose height is g(2). What's good enough for one of the subintervals is good enough for all of them.



Each term in the approximation,

g(1/2) 1/2 + g(1) 1/2 + g(3/2) 1/2 + g(2) 1/2 = 16 × 1/2 + 32 × 1/2 + 48 × 1/2 + 64 × 1/2

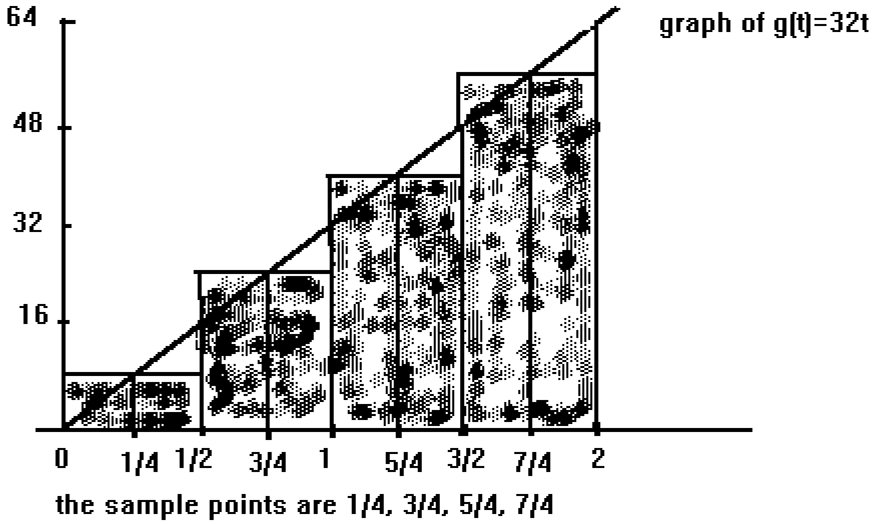
is numerically the same as the area of one of the rectangles. The approximation is numerically equal to the combined areas of the four rectangles.

The geometric interpretation of the derivative, f `(a), is the slope of the line that is tangent to the graph of f at the point (a, f(a)). **The geometric interpretation of an approximating sum is the total area of a collection of rectangles**.

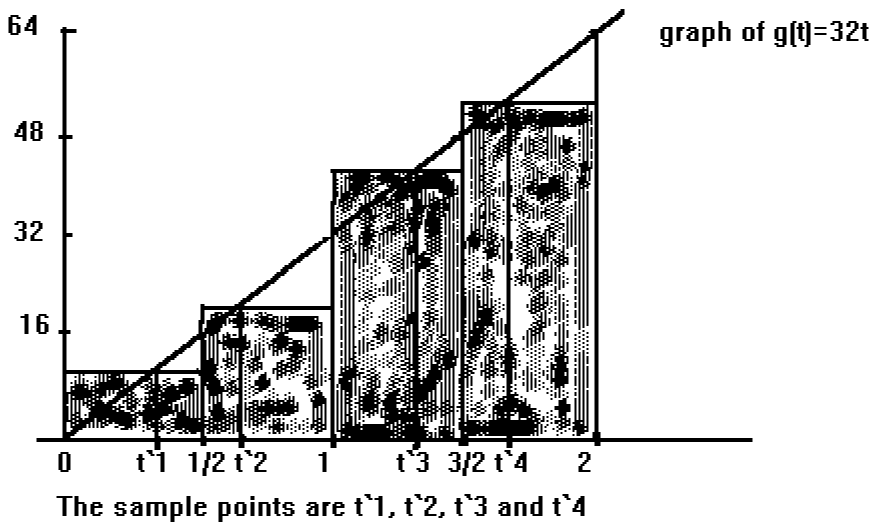
I think I'll look at some approximations using different sample points and intervals. I want to get a feel for these rectangles.

I'll try the midpoints of the intervals as sample points. The approximating sum is

g(1/4) × 1/2 + g(3/4) × 3/2 + g(5/4) × 1/2 + g(7/4) × 1/2.

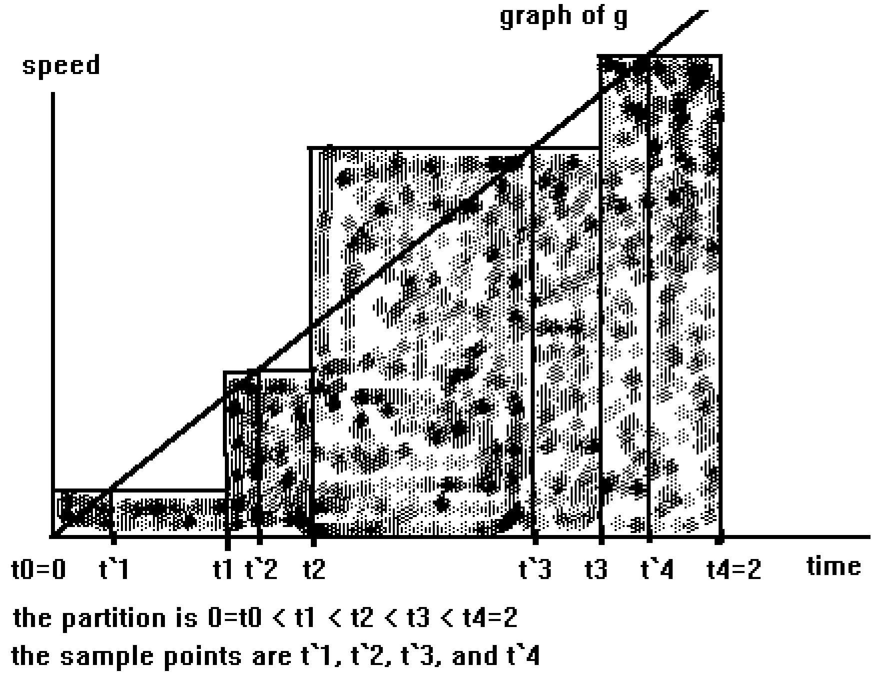


If I chose the sample points randomly in the subintervals, I get something like this:



The approximation is $g(t`_1)1/2 + g(t`_2)1/2 + g(t`_3)1/2 + g(t`_4)1/2$.

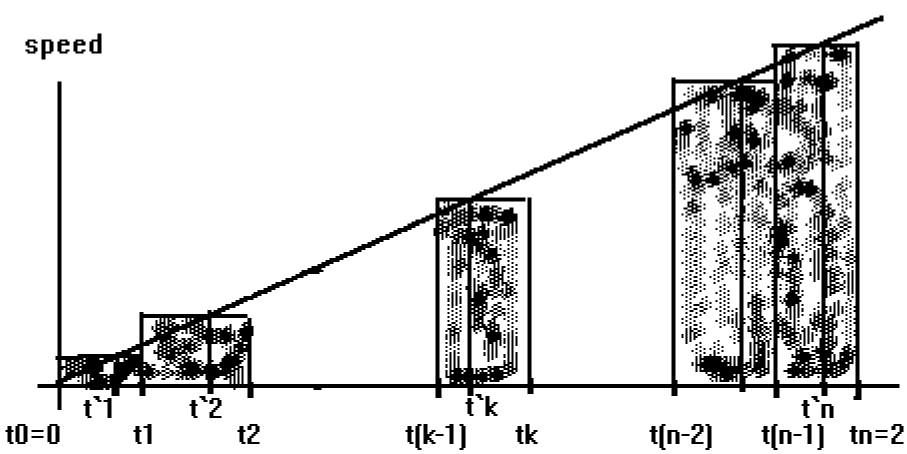
Since I have gone this far, I might as well go all the way. I'll see what it looks like with random sample points in different sized subintervals.



The approximation using n subintervals is

$$g(t`_0)(t_1 - t_0) + g(t`_1)(t_2 - t_1) + ... + g(t`_k)(t_k - t_{k-1}) + ... + g(t`_n)(t_n - t_{n-1}).$$

The kth term is numerically equal to the area of a rectangle whose height is $g(t`_k)$ and whose base is $(t_k - t_{k-1})$. Since n is not specifically given, the possibility of drawing every rectangle seems slight. If n is too large, I can't draw the picture even if I know what n is. None-the-less, I do want to draw something in these cases.



This is the picture that I use to see the general approximation.
The last picture is a sketch of the graph and the approximating rectangles. It is a picture of an idea and is successful if, when I look at it, the correct mathematical idea comes to mind.

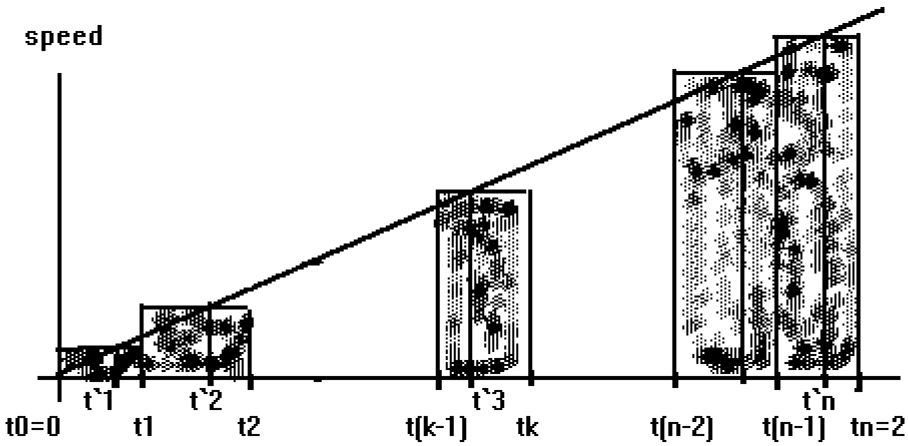
The reason that I went to all the trouble to draw these graphs and rectangles is that now I have a picture that pops into my mind whenever I read, hear, or think about the sums that approximate distance. If I talk and think about sums, I am also talking and thinking about rectangles, and the other way around.

A distance approximation gives rise to a collection of rectangles. The total area of these rectangles is my geometric interpretation of the distance approximations.

I'm going to call the total area of the rectangles determined by a distance approximation, the **'area of the rectangles'**. The phrase, 'area of the rectangles' refers only to rectangles that arise from an approximation.

LECTURE 8-10

How could I tell if I was myself or just a clever copy? I tried to devise a test...



Looking at the picture, I am led to think that the total area of the rectangles gets close to the area of the region between the graph of g and the t-axis as the mesh goes to zero, the 'mesh' being the maximum of the lengths of the subintervals. It, perhaps, sounds a little technical to say, "...as the mesh goes to zero.." but it is so much more convenient than saying, "... as the lengths of the subintervals go to zero..." ,as well as being more precise.

The set of numbers determined by the 'area of the rectangles' is exactly the same as the set of numbers determined by the approximations. This means that the 'area of the rectangles' approaches the same number as the approximations when the mesh goes to zero. Since the approximations approach the distance traveled, 64, so must the 'area of the rectangles'.

The question now is, "Does the 'area of the rectangles' approach the area between the graph of g and the t-axis?" The length of a subinterval is the width of a rectangle, so as the mesh goes to zero, so does the width of all the rectangles and, visually at least, the collection of rectangles seems to approach the region between the graph of g and the t-axis. This would seem to imply that the 'area of the rectangles' approaches the area of the region between the graph of g and the t-axis. Since I know that the 'area of the rectangles' approaches 64, the area must be 64.

My argument here is that the <u>region</u> of the rectangles approaches the <u>region</u> under the curve and so the <u>area</u> of the rectangles approaches the <u>area</u> of the region under the curve. My justification of the first part is that it 'looks that way' and my justification of the second part is that 'this is a self-evident property of area'.

I can check this answer. The area between the graph of g and the t-axis is the area of a triangle whose base is 2 and whose height is g(2) = 64. The area of this triangle is 64.

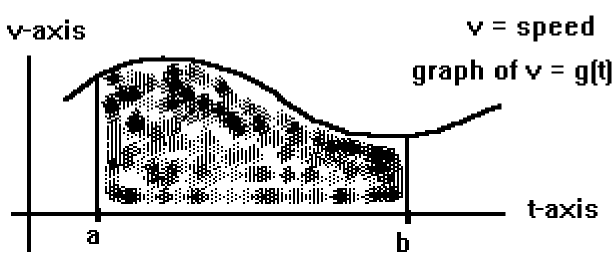
In this case, my intuition is correct and the 'area of the rectangles' approaches 'the area under the graph of g', an expression I'll occasionally use instead of 'the area between the graph of g and the t-axis'. I can do this because g is always non-negative and its graph is always above the t-axis. This automatically makes the t-axis the lower boundary of the region 'under the graph of g'. When I talk about the region, 'under the graph of g', the phrase, 'and above the t-axis' is implied.

If g is any function that models the speed of an object on an interval of time, [a,b], is it true that the 'area of the rectangles' approaches the area under the graph of g?

If g is any function that models the speed of an object on an interval of time, [a,b], is it true that the approximations approach the distance traveled during [a,b]?

So far, the answer to the second question is, "yes", if g is an increasing function or a decreasing function and the answer to the first question is, "yes", if g(t) = 32t. I'm sure I can do better than this.

If the answer is, "yes" to both questions, then, as the mesh goes to zero, the 'area of the rectangles' approaches the area under the graph of g, and the approximations approach the distance traveled. The approximations and the 'area of the rectangles', being identical numerically, must approach the same number as the mesh goes to zero, and this number must be both the area under g and the distance traveled. If the answer is, "yes" , to both questions, then the area 'under the graph of g' equals, numerically, the distance traveled and the 'area between the graph of g and the t-axis' would be the correct geometric interpretation of distance.

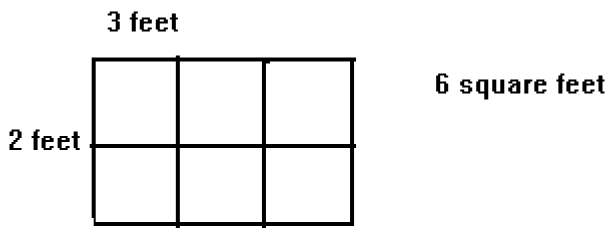


Is the disance traveled in the interval of time [a,b] equal to the shaded area?

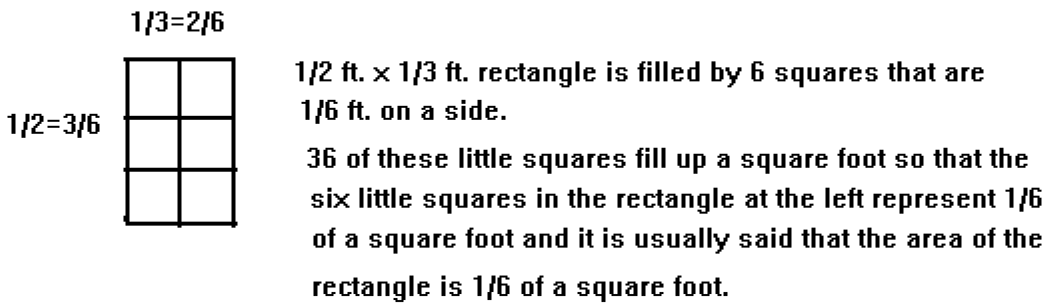
While I greatly desire that area should be the geometric interpretation of distance, I am stymied by the fact that I don't really know what area is in the Ideal World. It is hard to come to grips with justifying that the 'area of the rectangles' approaches the area the under the graph of g when the very idea of area is none too clear. I am first going to look into what I might mean by area and why the 'area of the rectangles' might approach that area.

The idea of area seems to be straight forward in the beginning. I consider area to be a number that is attached to a 'space' in the plane and not the 'space' itself.

A 2 foot by 3 foot rectangle has an area of 6 square feet. I am going to use square feet because I need some kind of unit and I like square feet. My basic concept of area is contained in the units, square feet. I can <u>fill the rectangle with squares</u> that are one foot on a side. There are two rows of three squares or three columns of two squares, and using the basic idea of what multiplication means, the area is $2 \times 3 = 3 \times 2 = 6$ square feet. The familiar formula, area = length times width, holds, which is reassuring.



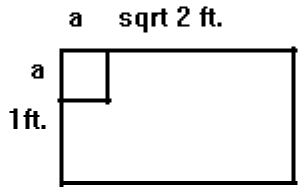
If the rectangle is 1/3 foot by 1/2 foot, the idea of area is a little more subtle. I can't fill this rectangle with an integral number of square feet, but I can fill it with squares.



My naive concept of the area of a rectangle is that I can fill it and a 'square foot' with squares which are all the same size. The ratio of (the number squares it takes to fill the rectangle) to (the number of squares it takes to fill a square foot) is what I choose to call 'the area of the rectangle'.

My naive concept of area comes from my experience in the Real World. In the Real World, the sides of rectangles are given by finite decimals and I can always fill them with squares. For example, if a rectangle is 2.54 ft. by 3.57 ft., I can fill it with 90,678 squares that are 0.01 ft. on a side. I can fill a square foot with 10,000 such squares so the area is 9.0678 square feet. If the sides of the rectangle are given by fractions, then the reciprocal of the common denominator of the fractions gives a side length for squares that will fill the rectangle.

Until they happened upon $\sqrt{2}$, the Greeks thought that all numbers were fractions and this idea of area worked great.
My naive concept of area comes upon hard times in the Ideal World. In the Ideal World I can have a rectangle that is $\sqrt{2}$ ft. by 1 ft. <u>This rectangle can not be filled with squares that are all the same size</u>.



If the axa squares fill the rectangle, there is an integer n such that nxa=sqrt 2. This means that a must be an irrational number. There is also an integer, m, such mxa=1. This means that a is a rational number. There is no such a.

I am at home and happy with the area of the 2ft. by 3ft. rectangle being six square feet. I am happy if I have a rectangle that I can fill with congruent squares. If I have a 1ft. by $\sqrt{2}$ ft. rectangle, I can't fill it with squares and my idea of area comes to a halt. I am less happy about this. How do I give the area of the rectangle in square 'something's' when I can't fill it with squares?

The only collection of squares that might fill up the '1 by $\sqrt{2}$' rectangle would be squares of side zero. The areas of these "very small" squares could add up to the area of the rectangle in the same way that the lengths of the points in the interval [0,1] add up to 1. This is a kind of interesting idea and I'll try to remember it.

Well, I want this '1 by $\sqrt{2}$' rectangle and its ilk to have some kind of area, so I'll give them an area by *fiat*. I'll **define the area of rectangles to be length times width**. If the length and width are measured in 'something's', then the area will be in square 'something's', even though it is not really clear what the units mean in the case of the '1 by $\sqrt{2}$' rectangle. This definition gives the same number as my naive area for the rectangles that can be filled with squares. The rectangles that can't be filled with squares now have a number for area and they should be thankful that they have one and not complain about the fact that it isn't too clear what the number means.

'The area of a rectangle is length times width' and I think of this statement as the genesis of area in general. All other areas follow from the area of rectangles. The area of a rectangle is to area, as the speed of uniform motion is to instantaneous speed. The speed of uniform motion is the genesis of instantaneous speed.

I will say that the area of a '1ft. by $\sqrt{2}$ ft.' rectangle is $\sqrt{2}$ square feet. This number can not be turned into an integral number of little squares that fill up the rectangle and I find this hard to imagine with my Real World mind. What do units like 'square feet' mean if I can't fill the rectangle with squares?

I am in much the same position with area as I was with speed before I had the derivative. I thought that I knew what speed was until I actually sat down and thought about it. I thought I knew what area was until I actually sat down and thought about it.

The units of speed, feet /second, lose their meaning when the motion is non-uniform. There are no feet at a point and there are no seconds at an instant of time; distance and time are both zero at the instant I am considering speed. Speed gets the units, feet/second, from average speed and average speed gets them from uniform motion. Speed gets the units rather like an inherited title; them don't mean much.

The units of area lose meaning when the region can't be filled with squares. I keep the units 'square feet' for sentimental reasons.

I now have the area of the 'region' enclosed by a rectangle. I am going to use the word 'region' for the space enclosed by a curve, the curve in this case being a rectangle.

A circle surrounds a region. I can't fill the region inside a circle with squares because a circle has round sides and squares have straight sides. Be that as it may, I still want to attach a number to this region that describes it in the same way that '6 square feet' describes the region inside the 2 ft. by 3 ft. rectangle.

Well, in what way does '6 square feet' describe the region inside the rectangle? What are the properties of area?

Generally, I want to attach numbers to regions bounded by curves in such a way that the following are true:

1. If two regions are attached to the same number, it takes the same amount of Ideal World paint to cover each of them.

2. If I split a region into two parts, the amount of paint needed to cover one part added to the amount of paint needed to cover the other part equals the amount of paint needed to cover the whole thing.

3. If one region is contained inside another region, it takes less paint to cover the region on the inside.

4. If two regions are about the same size, it takes about the same amount of paint to cover them.

This doesn't seem to be an unreasonable desire. Should I be successful fulfilling this desire, I will call the number, area.

So, in the final analysis, <u>area is a number assigned to certain subsets of the plane</u>. There are subsets of the plane that I can't assign an 'area number' to, for example, a quadrant of the plane is too big to assign a number to.

If I can imagine what a set looks like and can draw a picture of it on Real World paper, then it has an area. Sets that are bounded by curves I can draw in the Real World look as if they have area and they do.

In the nineteenth and twentieth centuries, mathematics became painfully aware of pathology, and bounded subsets of the plane were constructed which can't be given an area. I can't draw pictures of such sets, they do not exist in the Real World, and they never, a dangerous word, come up in applications.

What kinds of sets do have area and how are the numbers assigned?

When I think of a set in the plane, I think of a region bounded by a curve whose picture can be drawn on a piece of paper in the Real World. It is my opinion that all of these sets can be given an area, and for practical purposes, all subsets of the plane that I run into have area. Any subset of the plane that I can imagine visually, has area and I'm not going to worry today about the area of sets I can't imagine. I'll think about them tomorrow. There are sets in the plane that have area and are not bounded by curves, for example, a single point. The set consisting of a single point has an area of zero and is not bounded by a curve.

The way I think about it is that the 'area numbers' are assigned by a function, h, whose domain is the class of sets in the plane that have area, and whose range is in the non-negative numbers. If S is a set with area, then

'the function evaluated at S' = h(S) = the area of S.

I realize that the definition of the domain leaves much to be desired, in particular, how to tell if a region has area or not. My 'down and dirty' definition of sets that have area is that **if a set looks like it has area, then it has area and is in the domain of h**. There are other sets in the plane that have area but they are the fringe elements of the society of sets and, as so often happens to the fringe, I am going to forget about them.

My life has too few hours in it that I can spend them proving things I know are true just for the sake of doing it. Showing that sets that look like they have area, actually do have area, doesn't make the cutoff. The young mathematician, which is still inside me, looks at this last remark with dismay, but there was an eternity of time stretching in front of him that is no longer there for the older version.

The function, h, has to obey some rules if it is going to qualify as an area and in particular,

1. If a set, S, with area, is divided into two sets with area, U and V, then the area of S should equal the area of U plus the area of V; h(S) = h(U) + h(V).

2. If one set with area is contained in another, the area of the contain-ee should be less than or equal to the area of the contain-or. If U is contained in V, then $h(U) \leq h(V)$.

3. If two sets with area seem to be pretty close, their areas should be pretty close.

4. The area of a rectangle is length times width. If R is a rectangle whose length is 'l' and whose width is 'w', then $h(R) = l \times w$.

I am not going to show that a function with these properties exists, I am going to assume that someone else has . I do not have fun showing h exists nor have I ever particularly enjoyed reading about it. There are tunes that I don't care for and so, I neither play nor listen to them. There are those, however, who do like these area tunes and I have faith that when they tell me h exists, it does. If h didn't exist, area wouldn't be studied in schools and I wouldn't have spent all that time learning formulas to find it.

I am assuming that h exists and has assigned 'area numbers' to all the sets that have area. I don't know, however, what numbers are assigned to particular sets other than rectangles. I know that h exists, not what numbers it assigns to the sets. It's my job is to find these numbers using the properties of h.

I find it easy to accept the fact that curves in general and lines in particular have area and that the their area is zero. I am talking here about the curves themselves and not what they might surround. A circle encloses an area that is non-zero, but the circle itself has an area of zero. Curves are too thin to have a positive area. I enthusiastically embrace the standard formulae for the areas of triangles and trapezoids, which I can get by chopping up rectangles and putting them back together in clever ways. Any finite collection of points has an area of zero.

The circle is something else again. The Greek's solution to this problem was to approximate the region inside the circle with regular polygons. They knew the areas of the polygons and they could compute the number their areas approached as the number of sides got large. This number turned out to be the circumference of the circle times the radius divided by two. They realized the remarkable fact that the ratio of the circumference to the diameter is the same for all circles, gave the ratio a name, π, and turned their formula for area into the more familiar πr^2, where r is the radius of the circle.

They felt that this was the correct number to give as the area inside the circle and, certainly, any other number would have faced criticism. If S is the region inside a circle of radius, r, then $h(S) = \pi r^2$.
I think the Greeks had a good idea about how to find the area of regions bounded by curves and I want to use it. If I have a region in the plane and if

1. I can approximate that region with regions of known area, and

2. The known areas approach a number,

then that number is the area of the region.

There are those who might say that this is a pretty broad interpretation of what the Greeks said but I think it captures the essence of the idea.

When I say that I can approximate a region with other regions, I mean that, eventually, I can't tell the approximating regions, visually, from the region being approximated. The idea works because of the third property of h; If two sets with area seem to be pretty close, their areas should be pretty close.

I like this idea because I am thinking about my rectangles, with known area, approximating the region between the graph of g and its domain on the t-axis.

If g is a function that models speed, I am perfectly willing to believe, on the evidence of my eyes, that the region between the graph of g and the interval, [a,b], has an area. It seems visually evident that the rectangles approximate the region. Indeed, if the mesh is small enough, and the width of all the rectangles is small, I can't distinguish between the collection of rectangles and the region. In the spirit of the Greeks, I believe that the region has an area and that the total area of the rectangles approaches the area of that region as the mesh goes to zero. It certainly looks that way to me.

The rectangles that arise from an approximation approach the area under the graph of g because they look like they do.

I can't believe I've said that, but I love it. I feel as if I have just come out of the closet. I can now answer "yes" to the question of whether the 'area of the rectangles' approaches the area under the graph of g.

LECTURE 8-11

On the other hand, if the copy were perfect, what difference would it make?...

Using the 'keen eyesight' technique, I have shown that, if g is any function that models speed, the 'area of the rectangles', always approaches the area under the graph of g. Now, I'll turn my attention to the approximations.

I know that the approximations approach the distance traveled if the speed is increasing or if the speed is decreasing. Since the speed of most objects in motion both increases and decreases during the trip, I want to allow the function, g, that models speed, to both increase and decrease.

I am going to let g be <u>any</u> function that models speed over an interval of time, [a,b]. I'll form a partition of the interval,

$$a = t_0 < t_1 < \ldots < t_{k-1} < t_k < \ldots < t_{n-1} < t_n = b,$$

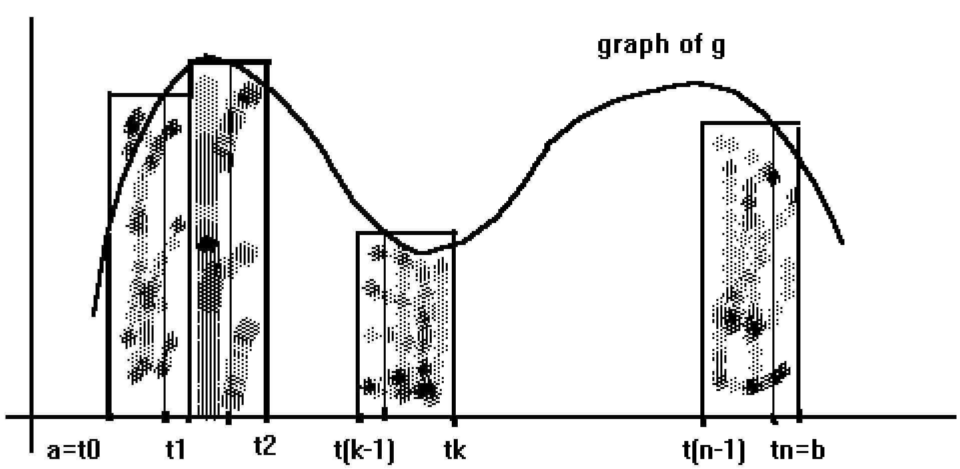
and pick a sample point, $t`_k$, in each interval $[t_{k-1}, t_k]$. The distance traveled over each subinterval is approximated by

$$g(t`_k)(t_k - t_{k-1})$$

The distance traveled during the interval, [a,b], is approximated by

$$A = g(t`_1)(t_1 - t_0) + g(t`_2)(t_2 - t_1) + \ldots + g(t`_k)(t_k - t_{k-1}) + \ldots + g(t`_n)(t_n - t_{n-1}).$$

This is old stuff so far, but now I have something new to add; a picture and area. Each term in the approximation is equal to the area of a rectangle whose height is $g(t`_k)$ and whose width is $(t_k - t_{k-1})$.



The sums,

$$g(t`_1)(t_1 - t_0) + g(t`_2)(t_2 - t_1) + ... + g(t`_k)(t_k - t_{k-1}) + ... + g(t`_n)(t_n - t_{n-1})$$

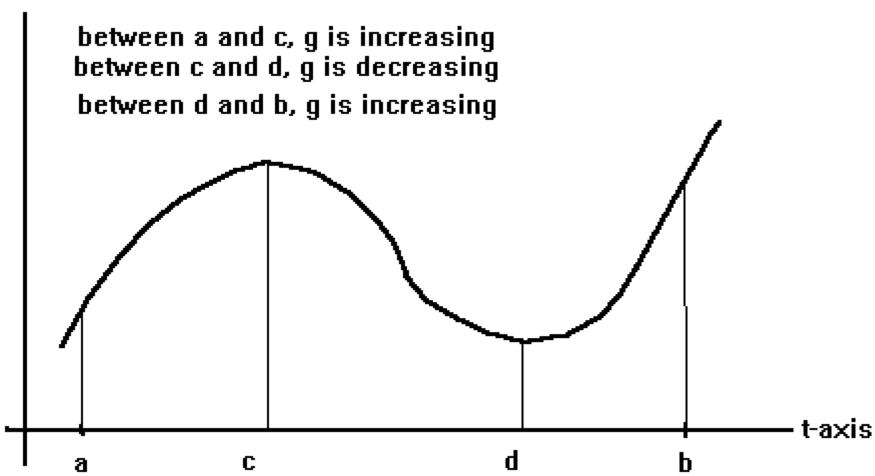
are important enough to have a name. They are called '**Riemann sums**' after the man who had the idea first. For a given value of n, there are an infinite number of ways to choose the partition and there are an infinite number of ways to choose the sample points in each of the subintervals of a partition. There are an awful lot of Riemann sums.

The word 'approximation' when referring to 'sums of distance approximations on subintervals', is going to be replaced by 'Riemann sum'. *Sic transit gloria mundi*.

Using the new name, I can say, **Riemann sums approach the area of the region under the graph of g as the mesh goes to zero**.

Do the Riemann sums approach the distance traveled?

If g is increasing or decreasing, I have decided that the answer is "yes" and it seems to me
that if g increases and decreases in some reasonable way, the answer should be "yes" as well. The idea is to look at the increasing and decreasing parts of g separately, get the distances traveled on each of the parts, and then add them to get the entire distance.



The function is increasing between 'a' and 'c', so, as the mesh goes to zero, the Riemann sums over the interval [a,c] go to the distance traveled over [a,c]. Likewise, the Riemann sums over the intervals [c,d] and [d,b] approach the distance traveled over [c,d] and [d,b].

Now, a Riemann sum over the whole interval, [a,b], is the sum of three 'Riemann sums' over the intervals, [a,c], [c,d], and [d,b], respectively. Each of these three 'Riemann sums' approaches the distance traveled over their respective intervals.

These remarks imply that

Riemann sum [a,b] =

Riemann sum over [a,c] +	$\rightarrow$	distance traveled over [a,c] +	=	distance traveled
Riemann sum over [c,d] +		distance traveled over [c,d] +		over [a,b].
Riemann sum over [d,b]		distance traveled over [d,b]		

The Riemann sums of g, over the interval [a,b], approach the distance traveled over [a,b] as the mesh goes to zero.

This argument is a little funky because it seems to always use the points c and d as points in the partition of the whole interval, [a,b], and they needn't be there. I rise above this minor objection and my belief in the result is not shaken.

This example is typical and I am going to argue from the particular to the general. If the function that models speed increases and decreases, that is, oscillates, continuously a finite number of times in any finite interval, the Riemann sums approach the distance traveled as the mesh goes to zero.

I think this includes all the functions that model speed. I can't conceive of a function that models speed and oscillates an infinite number of times in a finite interval of time, not even in the Ideal World.

Given any function that models the speed of an object in motion, I can find the distance traveled in any interval of time during the motion, in the sense that I can approximate the distance as closely as I please by Riemann sums. A geometric representation of this distance is the area between the graph of g and the interval of time on the horizontal axis.

On the basis of a funky argument used on one function, I have concluded that the Riemann sums approach the distance traveled for any function that models speed. The fact is, the one function and the funky argument make me believe it. I have no need to do more.

I had a friend who was staying with me and she never washed her dishes. When I complained, she said that she didn't mind dirty dishes until there weren't any more dishes, and then she would wash them. It was her feeling that if I was obsessed about clean dishes, it was my own personal problem, not hers, and I should deal with it, probably by washing the dishes. I saw the logic in her point of view and thought about how it related to cleaning up my proofs.

I tend to worry about little gaps in my arguments and I try not to. I want my arguments to give me a ‘deep down’ belief and worry over details can obscure the real reason why something is true. When I was a kid our family was taken out to dinner by a couple who were friends of my parents. The male half of the couple was giving a monologue on the political situation as he parked the car on a hill. We all got out of the car and stood on the sidewalk waiting for him to finish. His wife finally interrupted him, “Jim! Jim! The car’s rolling down the hill!” She was right, it was. He turned to her. “Woman, don’t bother me with minutia.” He continued with his discourse. I realize that he was a little rude, but I have always admired his attitude.

I have found a method to get the distance traveled over any interval of time using only the function that models speed. My first step, after fooling around with ‘magic’ times, was to use the idea of approximating the distance on subintervals and adding them up to get an approximation of the distance traveled on the whole interval. I showed that the idea worked for the particular function, $g(t) = 32t$, and then that it worked if g was increasing. With a wave of my hand I brought decreasing functions into the fold, and finally, I saw that g could be any function that modeled speed.

I have a method to find distance, but what a method! I don’t think the idea is conceptually hard but when I think about actually finding a number, I view the computation with despair, particularly if I have to do it by hand.

For reasons that I can’t recall, I once taught myself to program in the C language. I did this by choosing things to program and my first self imposed exercise was to program Riemann sums. I used the interval [1,3], equal subintervals, the mid-points of the subintervals as sample points, and $g(t) = t^2$. I chose to do Riemann sums first out of respect for the beauty, depth, and utility of the concept, but I only did the one.

Fortunately, there are people who really like these approximations and spend their time trying to find more economical ways of doing them. Riemann sums work but there may be other kinds of sums that work better from a time and money stand point. From the point of view of people who love Riemann sums, it isn’t a drag that the computations are tedious, it is amazing and wonderful that they can find distance at all. Before Riemann sums, there wasn’t a method, and when you don’t have a method, any method looks good.

The fact that distance can only be approximated is hardly a Real World concern, since approximation is all the Real World ever gets anyway. I might also say that using Riemann sums isn’t the worst approximation method in the world, in fact, it’s rather nice. In the Real World, an explicit Ideal World formula for distance would still put me on the computer because I have to evaluate the formula and this requires approximation. It is Ideal World elitism that makes me avoid Riemann sums as a computational method.

Actually, there is a little computational edge in the process. If I set a tolerance for how close I want the approximations to be to the distance, there is a positive number, δ, such that if the mesh, **l**, is less than δ, all of the Riemann sums with mesh **l** are within the tolerance of the distance.

If **l** < δ, then

$$| \text{distance} - \text{Riemann sum} | < \text{tolerance}.$$

Once I get the mesh small enough, I can use <u>any</u> partitions and <u>any</u> sample points I want to. This gives me the freedom to choose the partition and sample points in ways that make my life a little sunnier.

For example, I can make all the subintervals of equal length and use the end points of the subintervals for sample points. In an actual computation there is the problem of finding δ if I am given a tolerance. If the function g is increasing on the interval [a , b], then δ = tolerance / g(b) - g(a).

LECTURE 8-12

A clone could be made while the original slept and the original locked away in some tower. The clone would never know...

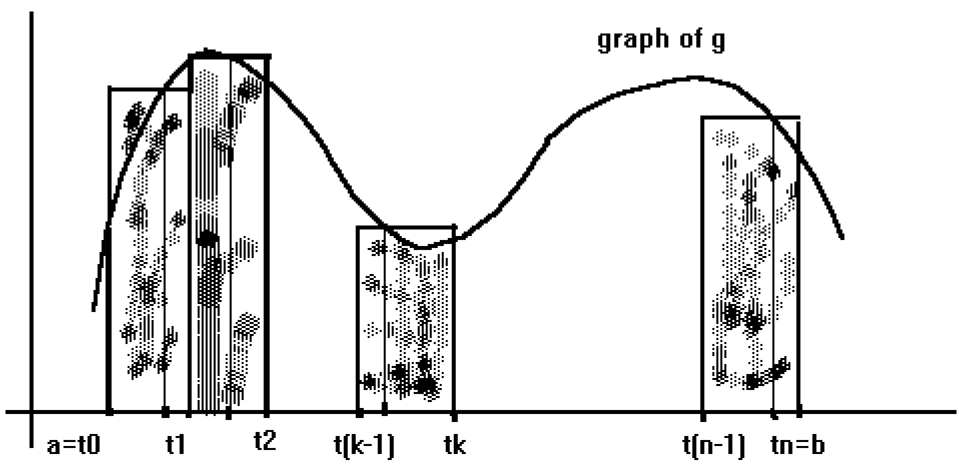
If g models speed, the Riemann sums approach the distance traveled, but the Riemann sums don't know that. All they know is that they are approaching a number and if they were going to interpret the number at all, they would probably interpret it as area.

All the Riemann sums care about is the area under the graph of g, they don't care what g models. I think that if g is non-negative and the region between the graph of g and the horizontal axis has an area, then the Riemann sums approach that area as the mesh goes to zero, regardless of what g may model.

There is a new problem I want to solve.

The problem is to find area between the graph of a non-negative function, g, and an interval [a,b], contained in the domain of g. <u>I am assuming as common knowledge that if regions of known area approximate a region, then the known areas approximate the area of the region.</u> This provides a way to compute an arbitrarily close approximation of area, if not the number itself.

My approach is to find a reasonable way to approximate the region under the graph of g using regions of known area. Rectangles seem the simplest and using Occam's razor, I'll try them. I partition [a,b] to get the bases of the rectangles and sketch them in.



I see that the heights of the rectangles are given by g evaluated at points in the intervals that form their bases and I can write the area of these rectangles as

$$A = g(t`_1)(t_1 - t_0) + g(t`_2)(t_2 - t_1) + \ldots + g(t`_k)(t_k - t_{k-1}) + \ldots + g(t`_n)(t_n - t_{n-1}).$$

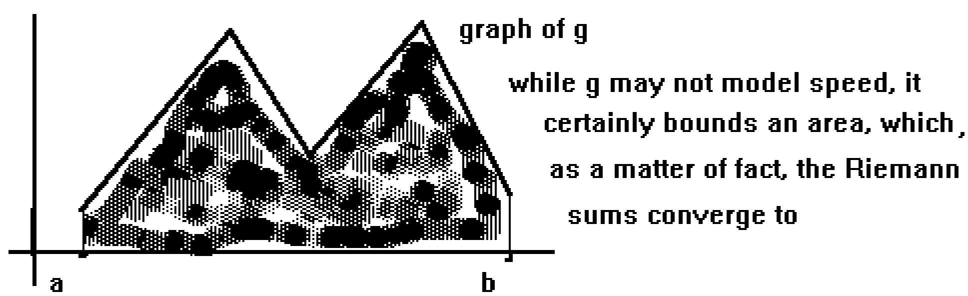
Each term in the approximation is equal to the area of a rectangle whose height is $g(t`_k)$ and whose width is $(t_k - t_{k-1})$. And here they are, the '**Riemann sums**'.
Because of my basic assumption that the areas of approximating regions approach the area of the region, I know that the Riemann sums approach the area under the graph of g. I don't have to worry about whether g is increasing or decreasing, continuous or not, just if it has a graph that looks like it has area under it.

This seems to take a lot less effort than showing the Riemann sums approach the distance traveled if g models speed, and there is a reason. I have assumed that the areas of approximating regions approach the area of the region, and I have done this because of 2000 years of tradition and the fact that it is visually self-evident. I look at the picture and I believe it.

The corresponding statement when I was computing distance was that the sum of distance approximations over the subintervals approached the distance traveled over the interval. This has only about 150 years of tradition behind it and it is not visually self-evident; not to me anyway. When the Riemann sums were approximations of distance, I had to do something to make myself believe they approached the distance traveled, and I did.

I will finish with a couple of examples.

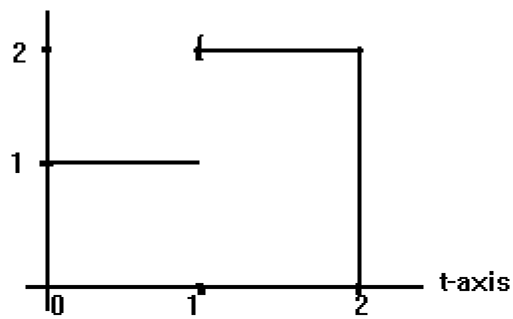
The Riemann sums approach a number, area, for functions that could not possibly model speed, for example, a function whose graph has sharp corners.



I doubt that the above function models speed because then the acceleration would not be continuous. Acceleration is the rate of change of speed and so equals the slope of the line segments that form the graph of g. The acceleration would have to change instantaneously from the slope on the left of a 'point' to the slope on the right of the 'point', and this would violate the 'principle of continuity'. Physical quantities can't jump.

The Riemann sums may converge to a number even if the function is not continuous; the number, of course, is area. Take for example, the function, h, whose domain is [0,2] and whose rule is

$$h(t) = 1 \quad 0 \le t \le 1$$
$$h(t) = 2 \quad 1 < t \le 2.$$



The area under the graph equals 3 and the Riemann sums converge to 3. The fact that the Riemann sums converge to a number even though the function doesn't model speed, may seem at first like one of the curious irrelevancies that mathematics is so famous for. If I look at it again, however, I can conceive of a situation where the speed increases so quickly that the function h may be the best way to model it. The speed of a small electric motor can change very quickly. The actual function that models the speed could be very complicated and h is very simple. The function h is particularly appealing if the change in speed takes place in an interval of time smaller than the degree of accuracy of my clock.

When it comes to modeling, I go for results. There aren't any rules in a knife fight.

While I can find distance traveled using Riemann sums, it seems clear to me that area is the 'thing' in the same way that the 'play' is. What has happened here is something that happens so often in mathematics that it is the rule rather than the exception. While trying to solve a problem, finding distance traveled, I have stumbled across an idea that is far larger that just finding distance. It is the idea of Riemann sums and their relation to area.

Riemann sums originated in the computation of distance traveled using the function that modeled speed. They then found area.

Isaac Asimov has presented a future where the far reaches of the galaxy are populated with humans and they have forgotten that thousands of years ago their species originated on earth. So it is that I see Riemann sums. They have gone to the far reaches of Ideal World space and have become so magnificent that their humble origin is hidden by their brilliance. They have gone to places of beauty and depth that are impossible to imagine from the earth of area and distance. They are the superstars of the Ideal World. They are Wayne Gretzsky, Richard Petty, Elvis, and Muhammad Ali, all rolled into one.

LECTURE 8-13

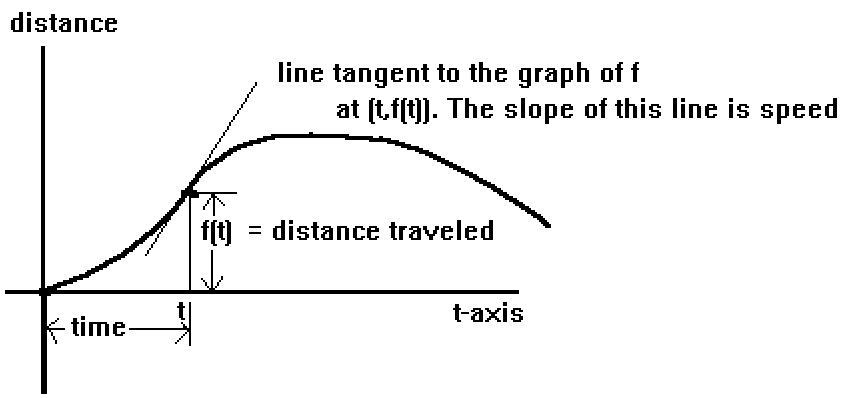
Reality eludes me...

There is a constant interplay between words, symbols, and pictures and I want to particularly consider how I 'see' time, distance, and speed. In my opinion, calculus and physics are joined at the hip, and pictures are the connective tissue.

The first thing I see are the actual physical quantities. I 'see' time by watching a clock. I 'see' the distance an object has moved by seeing the spot it is now, the spot it was then, and the space in between. I 'see' speed as how fast the object passes through my field of vision.

This is how I 'see' in the field of action. How do I see them when I am back home?

On a piece of paper I 'see' these quantities in the graph of the function that relates time and distance. I 'see' time as the scaled length of a line segment on the horizontal axis. I 'see' the distance traveled at time, t as the scaled distance of the point on the graph, (t, f(t)), above the horizontal axis. I 'see' the speed at a time, t, as the 'steepness' or slope of the line that is tangent to the graph at (t, f(t)).

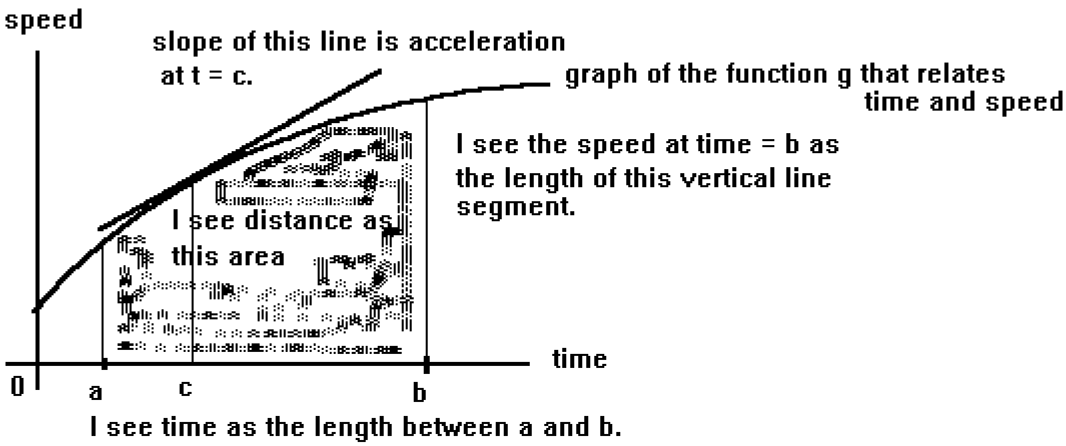


I think that speed is a typical example of the use of the derivative, so I think that the picture is also typical. If f is any differentiable function, I can 'see' the derivative of f as the slope of the tangent line to the graph of f.

My intuitive ability to turn a length into a number is better than my ability to turn 'steepness' into a number. Because of this, I practice by trying to judge the slopes of pitched roofs, open transom windows, anything that tilts.

I can also see time, distance, and speed in the graph of the function that relates speed to time. As before, I see time as the scaled length of a segment on the horizontal axis. Now I see the speed at a time, t_0, as the length of a vertical line segment from t_0 on the horizontal axis to the point, $(t_0, g(t_0))$, on the graph. Speed is now seen as a 'length' instead of 'steepness'.

In this graph, I see the distance traveled between 'time = a' and 'time = b' as the area between the segment [a, b] and the graph.



The slope of a tangent line to this graph is the instantaneous rate of change of speed, which is acceleration.

I express speed analytically as the derivative, that is, the number approached by the average speeds of the function, f, that models distance. I express speed visually as the steepness of a line tangent to the graph of f.

I express distance analytically as the number approached by the Riemann sums of the function, g, that models speed, and I express distance visually as an area under the graph of g.

I began, so long ago, with the goal of describing the motion of a falling rock. I have done this and more. I can use speed, defined as the derivative, to describe the motion of any moving object. I have defended speed as the correct way to describe motion by showing that I can find the distance an object travels from the knowledge of speed.

If the theory stopped here, it would be a valuable theory and would be studied in schools around the world, but it doesn't stop here. There is a connection between the derivative and Riemann sums that is simply remarkable and makes the computation of many of the numbers that Riemann sums approach easy. Maybe not 'easy' exactly, but 'easier'.