These are notes for a Calculus I course taught in fall semester 2013. The approach has been slightly different from the usual one in the way that the concepts of Differential and Integral Calculus and the corresponding rules were developed parallel, so not following the usual path where Integral Calculus is covered in the last third of the course. The starting point is the Fundamental Theorem, and differentiation and integration techniques art developed at the same time, emphasizing the derivative - anti-derivative viewpoint. More emphasis than usual has been given on applications of Calculus in Physics. The course is designed to be used by students who learn Calculus at the same time when they take the Calculus based Physics class.

Each Lecture contains links to related videos that can be found on the internet. Most sections are quite independent, so the hope is that these notes can be used if the course follows a different time-line and more standard approach to Calculus I.
Lecture 1: Quantities, Variables, Functions and Graphs

What is a function: basics and key terms.

**Key concepts:** Numbers and functions; *quantities, measurements and units*

In applications and science we care about *quantities* and their *measurement*. By choosing a *unit* a number can be assigned to a measurement. So if length is measured in meters m then the length of let’s say a stick is compared with a standard length of 1 m. If a stick is $\frac{1}{2}$ m long then we need two sticks to get the standard meter length. Time is another example of a quantity and the unit of second has to be fixed. So don’t forget that the numbers associated to a quantity depend on the chosen unit. Other examples of quantities we study are temperature, current, work etc.

The laws of nature are relations between quantities. They are given by formulas involving the quantities as variables. In practice the result of a measurement is a rational number (like 0.5 in the example above). But we will assume that our quantities can take real number values because this is practical. For instance the length of the diagonal of a unit square is $\sqrt{2}$. This is not a fraction but we can measure it. In general we use symbols as placeholders for quantities, like $t$ for time, $\ell$ for length, $T$ for temperature and so on. Sometimes we talk about a constant quantity, which is just a single number, also denoted by a symbol. This often is comfortable because whether a quantity is constant or variable can depend on the situation considered. For example we may set up an experiment where the temperature is fixed. In another experiment we change the temperature between $-10^\circ$ and $150^\circ$ Fahrenheit and study the resulting electric resistance of a cable. So whether temperature $T$ is variable or constant depends on the situation.

A typical application situation involves many quantities. We want to know how they relate to each other. Let us consider gas contained in a volume $V$ at a temperature $T$ (measured in Kelvin), which will lead to a pressure $p$ according the equation

$$pV = kT,$$

(where $k$ is a constant). In a situation like this we call a variable, which we change in an experiment the *independent variable*. For example this can be the temperature changing within some range. Then, if we keep the volume fixed, the resulting pressure is an example of the corresponding *dependent*
variable. For each fixed volume the temperature determines the pressure, or we say that the pressure \( p \) is a function of temperature \( T \). Here the temperature is changing in some interval \([0, T_{\text{max}}]\). So according to the above law

\[
p = \frac{k}{V} \cdot T,
\]

where now \( c = \frac{k}{V} \) is a constant. The graph of \( p = cT \) is a line through the origin in a coordinate system where we identify lengths along the axes with the numbers associated to \( T \) and \( p \) with respect to units.

**Problem:** Suppose we keep the temperature fixed and change the volume instead. What are the independent and the dependent variable in this case? What is the function describing the relation?

In calculus we often call the independent variable \( x \) and the dependent variable \( y \) and denote their relation by \( y = f(x) \). This means that for each possible value of \( x \) (usually in some interval) there is a uniquely determined value of \( y \). The function can be given by a formula like \( y = x^3 \) or a graph (Give a sketch of the graph of \( x^3 \)).

**Problem:** Write up all the functions and classes of functions you recall from your previous courses!

Another typical example is a point moving along a line, on which we choose a fixed point 0 and a direction. In this case the independent variable is time \( t \) and the dependent variable is position \( s \). This is the real number given by the distance to the origin, but with a negative sign for points to the left of 0. Here before/behind are of course determined by the direction of the line. The position graph \( s(t) \) here is not to be confused with the orbit of the motion.

**Problem:** Describe the motion modeled by the functions \( s = 2t \) and \( s = -3t + 2 \) and sketch the corresponding position graphs.
Lecture 2: Change and Derivative, Gradient/Slope, \( \lim \)-notation, Velocity

**Slopes and tangents** and **Point-slope formula**

**Differentiation** is the technique which enables the velocity or more generally the slope or gradient of a graph to be obtained from an equation without drawing the graph. Before learning those techniques let’s try to see what we are calculating. If a tangent is drawn at a point \( P \) on a curve then the slope (or gradient) of this tangent is said to be the slope of the curve at \( P \):

The slope of the line can be found as rise over run and is 8 units for the graph above: If we go 0.5 units to the right along the \( x \)-axis the graph of the line goes up by 4 units. The equation of the graph above is \( y = f(x) = 4x^2 - 1 \).

Next let us calculate the slope of the chord, which joins \( P(1, f(1)) \) with \( Q(1.5, f(1.5)) \). If we extend the chord to a line we call this a **secant line** and its slope a secant slope. In our example we get

\[
\frac{f(1.5) - f(1)}{1.5 - 1} = \frac{8 - 3}{1.5 - 1} = \frac{5}{0.5} = 10.
\]

This slope is also called the **average rate of change** of \( f(x) \) for the interval \([1, 1.5]\).

If we choose values of \( x \) closer and closer to 1 and calculate the corresponding secant slopes we will get secant slopes closer and closer to 8, which is the slope of the tangent line. For example the slope of the secant line through \((1, f(1))\) and \((1.1, f(1.1))\) is 8.4 (do the calculation!). If we call variable distance on the \( x \)-axis \( h \) then the secant slope is determined by the
and when \( h \) approaches zero these numbers will approximate the slope of \( y = f(x) \). This number is denoted the derivative of \( f(x) \) or \( y \) with respect to \( x \) at 1, and we write

\[
f'(1) = \frac{df}{dx}
\]

\[
\lim_{h \to 0} \frac{f(1 + h) - f(1)}{h} = 8.
\]

The \( \text{lim} \) is the mathematical symbol for the limit, the number that the values \( f(1+h) - f(1) \) are approaching when \( h \) gets smaller and smaller.

In general we denote the slope of the graph of \( f(x) \) at some point \( P(a, f(a)) \) on the graph of \( f(x) \) by

\[
f'(a) = \frac{df}{dx} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}
\]

where \( \Delta y \) is the change of the variable \( y \) resulting from a change of the variable \( x \). These symbols of course depend on the point \((a, f(a))\), where we consider the changes. If \( s(t) \) is the position graph of an object moving along a vertical line then

\[
s'(t) = \frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}
\]

is the velocity because \( \Delta s \) is the change of position within a time period \( \Delta t \).

We call \( \Delta s \) the displacement in the time interval \([t, t + \Delta t]\). If \( \Delta t \) is small enough and the changes take place at some time \( t \) we can assume that the velocity of the object is not changing, so that

\[
v(t_0) = \frac{ds}{dt} \bigg|_{t_0} = \lim_{h \to 0} \frac{s(t_0 + h) - s(t_0)}{h} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t} = \frac{\Delta s}{\Delta t}.
\]

We say that \( v(t_0) \) is the instantaneous velocity at time \( t_0 \). It is what we get from the average velocity for the interval
Lecture 3: Integral and Area, Σ-notation

Integration is the technique that allows to calculate easily the area between the $x$-axis and the graph of a positive function over some interval.

Consider the function $f(x) = 3x^2 + 8x - 1$ and the interval $[1, 3]$. We want to know the area between the $x$-axis and the graph of the function over the interval $[1, 3]$ (the blue shaded area on the left). Of course we could build a lake with that shape, fill in water and measure the volume when the water is 1 foot high. Then we know the area in square feet. Fortunately there is are easier (and more practical) methods getting the answer straight from the equation of the function. We have a symbol for the content of the blue area, namely $\int_1^3 3x^2 + 8x - 1 \, dx$. This is called the definite integral of $f(x)$ over the interval $[1, 3]$.

This number is equal to 56 for the blue area in the picture. We will see next week that numbers like this are remarkably easy to calculate. But first let us try a geometric approach just like we did for the derivative. Let us approximate the area by areas of rectangles. For this divide the interval $[1, 3]$ into ten equal intervals $[1, 1.2], [1.2, 1.4], \ldots [2.8, 3]$ and use these as base sides. For each of these intervals consider the value of the function at the left end-point and use this for the heights. If we add up the areas of these rectangles (see picture above) we get for each rectangle

$$0.2 \times f(\text{left endpoint}),$$
and thus the sum of the areas (see picture on the right) is
\[ 0.2 \cdot (f(1) + f(1.2) + f(1.4) + f(1.6) + \ldots + f(2.8)), \]
which can be calculated to 52.04 (do it!), so slightly smaller than the actual value (because the rectangles miss some of the area under the curve). In order to simplify expressions as above we introduce the so called Σ-notation, where Σ stands for sum. In the above example we would write
\[ 0.2 \cdot \sum_{i=0}^{9} f(1 + 0.2 \cdot i). \]
The numbers \( f(1), f(1.2), \ldots f(2.8) \) for a sequence of numbers \( a_0, a_1, \ldots a_9 \), where \( a_i = f(1+0.2\times i) \). In general, given a sequence of numbers \( a_m, a_{m+1}, \ldots, a_n \) we denote
\[ \sum_{i=m}^{n} a_i = a_m + a_{m+1} + \ldots + a_n. \]
Let’s try the definition:
\[ \sum_{i=3}^{7} i^2 = 3^2 + 4^2 + 5^2 + 6^2 + 7^2 = 9 + 16 + 25 + 36 + 49 = 131. \]
The Σ-symbol satisfies some nice obvious properties like \( \sum(a_i+b_i) = \sum a_i + \sum b_i \) or \( \sum c a_i = c \sum a_i \) (see book p. 289). We can now describe \( \int_a^b f(x)dx \) geometrically for each positive function \( f(x) \) and \( a < b \). It is the area \( A \) between the \( x \)-axis and the graph over \([a,b]\). We can compute it approximately by dividing the interval \([a,b]\) into small intervals of length \( \Delta x \), so have points
\[ a, a + \Delta x, a + 2\Delta x, \ldots, a + (N-1)\Delta x, b = a + N\Delta x \]
dividing \([a,b]\) into \( N \) small intervals. Then we approximate the area \( A \) by
\[ \Delta x \cdot f(a) + \Delta x \cdot f(a + \Delta x) + \Delta x \cdot f(a + 2\Delta x) + \ldots + \Delta x \cdot f(a + (N-1)\Delta x). \]
Using Σ-notation we have
\[ \sum_{i=0}^{N-1} f(a + i\Delta x)\Delta x, \]
and when \( \Delta x \) approaches 0 the corresponding numbers above approach \( A \). We write this as
\[ A = \int_a^b f(x)dx = \lim_{\Delta x \to 0} \sum_i f(a + i\Delta x)\Delta x. \]
Lecture 4: Anti-derivatives, Fundamental theorem of Calculus

Indefinite Integrals as anti-derivatives

Let’s consider the situation of CCC 3 in the case where the function \( f(x) \) is replaced by the velocity function \( v(t) \) of an object moving along a line. We know that if \( v(t) \) is constant it is easy to determine the distance driven. Then we find the position if we know where we are at let’s say \( t = 0 \). When we drive with 30 miles per hour for 30 minutes we have driven 15 miles. Sure! Well this is the basic equation for constant velocity motion:

\[
\text{change in position} = \text{velocity} \times \text{time}
\]

or in formulas:

\[
s(t) = s_0 + v \cdot t
\]

where \( v \) is a constant and \( s_0 = s(0) \) is initial position.

Next suppose our velocity \( v(t) \) is changing, let’s say it looks like on the left below. The position graph \( s(t) \) is on the right hand side. In fact these are the graphs for a body [falling in vacuum] (actually in vacuum all bodies fall the same) with \( t \) in seconds and \( s \) in feet.

Now \( s(t) = 16t^2 \) has derivative \( v(t) = 32t \). How does this relate to the constant velocity motion? Well assume that the motion has constant velocity \( v(t_i) \) on small intervals of length \( \Delta t \) with left endpoints \( t_i \). For example divide the interval \([0, 3]\) into ten intervals of length \( 3/10 = 0.3 \). So we will get \( t_0 = 0, t_1 = 0.3, t_2 = 0.6, \ldots t_9 = 2.7, \) or \( t_i = 0.3 \cdot i \) for \( i = 0, \ldots, 9 \). Then using the basic formula above for each of the intervals we get for the total distance covered:

\[ v(0) \cdot \Delta t + v(0.3) \cdot \Delta t + v(0.6) \cdot \Delta t + \ldots + v(2.7) \cdot \Delta t, \] and so
32(0.3 + 0.6 + \ldots + 2.7) \cdot 0.3 = 32 \cdot 0.3^2 \cdot (1 + 2 + \ldots + 9) = 32 \cdot 0.3^2 \cdot 45 = 129.6,
which is not 144 = s(3). This is smaller than s(3) = 144 because the velocity is increasing on each of the intervals. If we make \( \Delta t \) smaller and smaller then the numbers above will approximate \( s(3) \). In fact they do. Of course these numbers also approximate the area under the graph of \( v(t) \). In fact from the area for a triangle we get for the area over \([0, 3]\):

\[
\frac{1}{2} \cdot 3 \cdot v(3) = \frac{3}{2} \cdot 32 \cdot 3 = \frac{9 \cdot 32}{2} = 9 \cdot 16 = 144 = \sum_{i=0}^{9} v(t_i) \Delta t,
\]

where we used \( \Sigma \)-notation from CCC 3. Thus for \( \Delta t \) very small we get the precise area, or using the integral notation from CCC 3

\[
\int_{0}^{3} v(t) dt = \lim_{\Delta t \to 0} \sum_{i} v(\Delta t \cdot i) \Delta t.
\]

In general we have

\[
s(t_2) - s(t_1) = \int_{t_1}^{t_2} v(t) dt
\]

is the displacement for the interval \([t_1, t_2]\) from the velocity.

Now \( v(t) = \frac{ds}{dt} \) by the very definition of velocity (see CCC 2). Just as \( v(t) \) is the derivative of \( s(t) \) we say that \( s(t) \) is an anti-derivative of \( v(t) \). Of course we can’t determine position from velocity because position always depends on an initial position. The formula

\[
s(t_2) = s(t_1) + \int_{t_1}^{t_2} v(t) dt
\]

describes precisely this. Thus the integral of a function over some interval is the difference of values of an anti-derivative at the end points. But what happens if we move backwards. Let’s look at Definite Integrals. The formula for displacement

\[
s(t_2) - s(t_1) = \int_{t_1}^{t_2} v(t) dt
\]

also holds when we are allowed to go backwards. Geometrically then the integral is the net or signed area. If \( v(t) \) is negative the position gets smaller and we have to subtract. Now the integral symbol picks up area below the \( x- \) or \( t- \) axis with a minus sign.
Lecture 5: Derivatives and Anti-derivatives for Polynomials

Take a look at: Derivatives of polynomials, Linearity of the derivative and Anti-derivatives of polynomials

If \( f(x) \) is a function then we denote the gradient function or derivative \( f'(x) \). It measures the steepness of the graph at all points. Let’s calculate the derivative of \( f(x) = x^3 \) using the approach of CCC 2. We want to use the formula from CCC 2

\[
(*) \quad f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}.
\]

Now

\[
f(x + h) - f(x) = (x + h)^3 - x^3 = x^3 + 3x^2h + 3xh^2 + h^3 - x^3 = 3x^2h + 3xh^2 + h^3
\]

and thus

\[
\frac{f(x + h) - f(x)}{h} = 3x^2 + 3xh + h^2.
\]

Now whatever \( x \) is, when \( h \) gets small, \( 3x^2h + h^2 \) gets small (for example take \( x = 10 \). Then we have \( 300h + h^2 \). If \( h = \frac{1}{1000} = .001 \) is number is \( 0.3 + 0.00001 \) so small). So we see

\[
f'(x) = x^3.
\]

This formula generalizes to the famous Power Rule: \( \frac{d}{dx} x^n = nx^{n-1} \) for all numbers \( n \) (mostly we will have \( n \) a positive integer but in fact the formula holds for all \( n \)). So

\[
\frac{d}{dx} x^1 = 1 \cdot x^0 = 1 \cdot 1 = 1
\]

\[
\frac{d}{dx} x^2 = 2 \cdot x^2
\]

\[
\frac{d}{dx} x^3 = 3x^2.
\]

(Maybe you should recall \( \frac{d}{dx} x^0 = \frac{d}{dx} 1 = 0 \), horizontal graph has steepness zero!) Using the power rule we get in fact derivatives of all polynomials. If

\[
f(x) = x^3 + 4x^2 - x + 6
\]
then
\[ f'(x) = 3x^2 + 8x - 1. \]

Do you see how this works? You take the derivative of each term in the sum separately and add them together. The derivative of \( x^3 \) is \( 3x^2 \). The derivative of \( x^2 \) is \( 2x \), we take this and just multiply by the factor 4 to get \( 8x \). The derivative of \( x \) is 1. Finally the derivative of the constant 6 is zero, 6 does not change! Note that if we multiply a function by a positive constant it’s gradient will change by that constant: \( 2x^2 \) is twice as steep as \( x^2 \) at every point. If you add two functions their gradients will add up too. We call these properties linearity.

\[
\frac{d}{dx} cf(x) = c \frac{df}{dx}, \quad \frac{d}{dx}(f(x) + g(x)) = \frac{df}{dx} + \frac{dg}{dx}.
\]

If \( f(x) \) is the position function \( s(t) \) of a moving object the linearity is more obvious: Suppose I cover 10 miles in an hour. Then if I cover 20 miles in an hour I am twice as fast. Also if \( s_1(t) \) is the motion of the back of a train, and a man moves within the train and the position relative to the back is \( s_2(t) \) then the combined motion is \( s_1(t) + s_2(t) \). The velocity of the absolute motion is \( \frac{ds_1}{dt} + \frac{ds_2}{dt} \). Recall if \( f(x) \) is a function then a function \( F(x) \) such that \( F'(x) = f(x) \) is called anti-derivative of \( f(x) \). We use the notation

\[ \int f(x) dx \]

for anti-derivatives of \( f(x) \). The symbol isn’t really a function but stands for all possible functions. For example

\[ \int 2x dx = x^2 + C, \]

where \( C \) stands for a constant. The power rule above tells us formulas for anti-derivatives

\[ \int x^n dx = \frac{1}{n+1}x^{n+1}. \]

Using the anti-derivatives and the Fundamental Theorem of Calculus we can now for example find change in position from a velocity function.
Lecture 6: Computations with Tangent Lines; Linear motion

Take a look at: Slope and Rate of Change, Velocity and Rate of Change, and Position by integration

Consider the graph of

\[ y = x^3 - 3x^2 + 1 \]

below. We would like to answer the following question: At what points on the graph is the slope 9? What are the equations of the tangent lines at those points.

It seems to be difficult to get precise answers from the graph. So we proceed from the equation. From CCC 5 we get

\[ y' = 3x^2 - 6x \]

for the slope at \((x, y)\) on the graph. So the slope is 9 gives the equation

\[ 3x^2 - 6x = 9, \]

or

\[ 3x^2 - 6x - 9 = 3(x^2 - 2x - 3) = 3(x - 3)(x + 1) \]

So the first tangent line is through the point \((-1, -3)\) on the graph and has the equation:

\[ y - (-3) = 9(x - (-1)) \quad \text{or} \quad y = 9x + 6. \]
The second tangent line is through the point \((3, 1)\) on the graph and has the equation:

\[ y - 1 = 9(x - 3) \quad \text{or} \quad y = 9x - 26 \]

Now let’s state the **general** case. If \(y = f(x)\) is a function then the tangent line at the point \((x_0, y_0)\) on the graph has the equation:

\[ y - y_0 = f'(x_0) \cdot (x - x_0). \]

So for example the tangent line to \(f(x) = -x^4 + 6x^3 + 1\) for \(x = 1\) can be calculated as follows: \(f'(x) = -4x^3 + 18x^2\) so \(f'(1) = 14\), and \(f(1) = -1 + 6 + 1 = 6\). The equation of the tangent line is:

\[ y - 6 = 14(x - 1) \quad \text{or} \quad y = 14x - 8. \]

Now let’s recall our formula for computing position from velocity:

\[ s(t) = s(t_0) + \int_{t_0}^{t} v(u) du \]

to calculate the position from some initial position \(s(t_0)\) and the velocity function using integration. Using the Fundamental Theorem of Calculus we know we find the integral from the anti-derivative of \(v\). Thus there is a way to do the calculation avoiding the integral symbol. Let’s just do an example: Suppose that a motion is described by \(v(t) = 10t - 6t^2\). We have the graph of \(v(t)\) on the left and the graphs of two corresponding position functions (anti-derivatives) on the right, with \(s(0) = 0\) respectively \(s(0) = -3\).
Using our knowledge about anti-derivatives of polynomials it is easy to find formulas for the two position functions. We know that

\[ s(t) = 5t^2 - 2t^3 + C, \]

where \( C \) is a constant. If \( s_1(0) = 0 \) then we calculate \( 0 = C \) and get \( s_1(t) = 5t^2 - 2t^3 \), the top graph. If \( s_2(0) = -3 \) we calculate \(-3 = C\) and so \( s_2(t) = 5t^2 - 2t^3 - 3\). Note that \( s_1(1) = 5 - 2 = 3 \) but \( s_2(1) = 5 - 2 - 3 = 0 \). We can set the initial position for any point in time. Then the velocity function determines the position. Note that if the velocity function is constant \( v \) then \( s(t) = vt + C \) is a special case of the power formula for anti-derivatives. Since \( s(t_0) = vt_0 + C \) we can calculate \( C = s(t_0) - vt_0 \) and get the general formula for constant velocity motion:

\[ s(t) = s(t_0) + v(t - t_0) \]
Lecture 7: The Derivative of Exponential functions; More Linear Motion

Take a look at: Position, velocity and acceleration

Last time we deduced the formula for constant velocity motion. A motion where the velocity is changing in such a way that it changes by equal amounts in equal time intervals is called a constant acceleration motion. In general, the instantaneous change of the velocity is called acceleration and can change with time. It is again a function of time and we have

\[ a(t) = \frac{dv}{dt}. \]

By applying integration as before we get the general formula

\[ v(t) = v(t_0) + \int_{t_0}^{t} a(u)du. \]

Let’s call \( v(t_0) = v_0 \) and \( s(t_0) = s_0 \) because then we see better that it is a constant. We can avoid the integral by just using that \( v(t) \) is an anti-derivative of \( a(t) \). If we can find the anti-derivative of \( a(t) \) then we can determine the constant by some initial condition \( v(t_0) \). Then knowing \( v(t) \) we can find \( s(t) \) as before. For the constant acceleration motion the anti-derivative of the constant \( a \) is \( v(t) = at + C \) and so from \( v(t_0) = at_0 + C \) we get

\[ v(t) = v_0 + a(t - t_0) \]

The anti-derivative of \( v(t) \) is

\[ s(t) = v_0t + \frac{a}{2}(t - t_0)^2 + C'. \]

To see this just differentiate

\[ \frac{d}{dt} \left( v_0t + \frac{a}{2}(t^2 - 2t_0 \cdot t + t_0^2) + C' \right) = v_0 + \frac{a}{2} \cdot (2t - 2t_0) = v_0 + a(t - t_0) \]

Finally from \( s(t_0) = s_0 = v_0t_0 + C' \) we get \( C' = s_0 - v_0t_0 \) and the law for the general constant acceleration motion:

\[ s(t) = s_0 + v_0(t - t_0) + \frac{a}{2}(t - t_0)^2. \]
For the motion in constant gravity \( a = -g \) and \( t_0 = 0 \) we get the simpler formula:

\[
s(t) = s_0 + v_0 t - \frac{g}{2} t^2.
\]

Suppose that \( a = 5 \), \( v_0 = 1 \) and \( s_0 = 2 \) (units could be for example here feet/sec\(^2\), feet/sec). Then for \( t_0 = 0 \) we get

\[
s(t) = 2 + t + \frac{5}{2} t^2.
\]

The formulas above are sometimes comfortable. But you do not have to remember them. Just recall that \( s(t) \rightarrow v(t) \rightarrow a(t) \) is given by taking derivatives and thus going backwards we take anti-derivatives. Here is an example: Suppose that \( a(t) = t \), and at time \( t = 0 \) we have initial velocity \( v_0 = 1 \) and initial position \( s_0 = 0 \). Then \( v(t) = \frac{1}{2} t^2 + C \) is the velocity, and \( 1 = v(0) = C \) shows that \( v(t) = 1 + \frac{1}{2} t^2 \) is the formula for velocity (Check from this formula again that \( v(0) = 1 \) and \( \frac{dv}{dt} = t \)). By taking the anti-derivative again we get \( s(t) = t + \frac{1}{6} t^3 \) and from \( s(0) = 0 \) we get \( C' = 0 \). Thus the solution is

\[
s(t) = t + \frac{1}{6} t^3.
\]

Take a look at: Derivatives of exponential equations

In many applications we need to know how the exponential function is changing. This turns out to be more difficult that our discussion of polynomials. We study first the slope of \( e^t \) at \( t = 0 \) when \( e^0 = 1 \). So imagine a motion with \( s(t) = e^t \) and we want to know the velocity at \( t = 0 \). So let’s calculate the numbers \( e^{\Delta t} - 1 \) and \( \frac{e^{\Delta t} - 1}{\Delta t} \) for \( \Delta t = 0, 0.1, 0.01, 0.001 \) and \( 0.0001 \). Guess the value of \( \frac{e^{\Delta t} - 1}{\Delta t} \) for infinitesimal \( \Delta t \), so answer the question: What value do we approach when \( \Delta t \) gets smaller and smaller?

<table>
<thead>
<tr>
<th>( \Delta t )</th>
<th>( e^{\Delta t} - 1 )</th>
<th>( \frac{(e^{\Delta t} - 1)}{\Delta t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000</td>
<td>0.00000</td>
<td>not defined</td>
</tr>
<tr>
<td>0.1000</td>
<td>0.10517</td>
<td>1.05171</td>
</tr>
<tr>
<td>0.0100</td>
<td>0.01005</td>
<td>1.00502</td>
</tr>
<tr>
<td>0.0010</td>
<td>0.00100</td>
<td>1.00050</td>
</tr>
<tr>
<td>0.0001</td>
<td>0.00010</td>
<td>1.00005</td>
</tr>
</tbody>
</table>

We see: When \( \Delta t \) gets smaller then the value of \( \frac{e^{\Delta t} - 1}{\Delta t} \) approaches 1.
Now we find for the function \( s(t) = e^t \) the displacement \( \Delta s \) in terms of \( \Delta t \), and thus \( \frac{\Delta s}{\Delta t} \). The displacement is (using the law of exponents)

\[
\Delta s = s(t + \Delta t) - s(t) = e^{t + \Delta t} - e^t = e^t \cdot (e^{\Delta t} - 1),
\]

which will approach 0 when \( \Delta t \) approaches 0. The average velocities

\[
\frac{\Delta s}{\Delta t} = \frac{s(t + \Delta t) - s(t)}{\Delta t} = e^t \cdot \frac{e^{\Delta t} - 1}{\Delta t},
\]

which will approach \( e^t \cdot 1 = e^t \) when \( \Delta t \) approaches 0. Thus we get for the instantaneous velocity

\[
v(t) = \frac{ds}{dt} = e^t \cdot 1 = e^t.
\]

Thus we know that \( \frac{d}{dt} e^t = e^t \) or \( \frac{d}{dx} e^x = e^x \). A small algebraic trick actually allows to get the general formula

\[
\frac{d}{dx} e^{kx} = k e^{kx}
\]

for any constant \( k \). In fact

\[
\frac{d}{dx} e^{kx} = \lim_{\Delta x \to 0} \frac{e^{k(x+\Delta x)} - e^{kx}}{\Delta x} = \lim_{\Delta x \to 0} \frac{e^{kx} e^{k\Delta x} - 1}{\Delta x} = \lim_{\Delta x \to 0} (k e^{kx}) \frac{e^{k\Delta x} - 1}{k \Delta x}
\]

Now when we calculate the \( \lim_{\Delta x \to 0} \) we only consider the change of \( \Delta x \), not any change in \( x \), here \( x \) is fixed but arbitrary. But \( \Delta x \to 0 \) certainly precisely when \( k\Delta x \to 0 \) (except maybe when \( k = 0 \) but that will give the constant function \( e^0 = 1 \)). So we can call \( k\Delta x = h \) and get

\[
\frac{d}{dx} e^{kx} = ke^{kx} \lim_{h \to 0} \frac{e^h - 1}{h} = ke^{kx}
\]

So for example \( \frac{d}{dx} e^{5x} = 5e^{5x} \).
Lecture 8: Product Rule and Quotient Rule

Take a look at:

Differentiating factored polynomials (This actually assumes that you know that \( \frac{d}{dx} \sin(x) = \cos(x) \). We will see this soon!) and When to use the quotient rule (well the quotient rule is actually not our last rule, so here you have to know that \( \frac{d}{dx} \sin(4x) = 4\cos(4x) \). We will see this also soon!).

Derivatives can be calculated from few basic examples like \( \frac{d}{dt} (t) = 1 \) and \( \frac{d}{dt} e^t = e^t \) using a collection of Rules. Here we will justify the Product rule, Reciprocal rule, Quotient rule and Power rule from the very definition that \( \frac{ds}{dt} \) is calculated from \( \frac{\Delta s}{\Delta t} \) by taking the limit. So

\[
\frac{ds}{dt} = \lim_{\Delta t \to 0} \frac{\Delta s}{\Delta t}.
\]

Recall that the limit symbol means that \( \frac{ds}{dt} \) is the value that is approached by considering the average velocities \( \frac{\Delta s}{\Delta t} \) when \( \Delta t \) gets smaller and smaller.

Throughout for all functions \( f(t) \) (or \( g(t), s(t), \ldots \)) of \( t \) let us use the symbol

\[ \Delta f = f(t + \Delta t) - f(t), \]

which stands for a change of the value of the function originating from a change from \( t \) to \( t + \Delta t \). Note that this also implies:

\[ f(t + \Delta t) = f(t) + \Delta f. \]

We will assume throughout that \( \Delta f \) approaches 0 when \( \Delta t \) approaches 0 and that the "limits" always exist.

**Product Rule**

Let \( s(t) = f(t)g(t) \) be a product of the functions \( f(t) \) and \( g(t) \). Then

\[
\begin{align*}
  f(t + \Delta t)g(t + \Delta t) &= (f(t) + \Delta f)(g(t) + \Delta g) \\
  &= f(t)g(t) + (\Delta f) \cdot g(t) + f(t) \cdot (\Delta g) + (\Delta f) \cdot (\Delta g)
\end{align*}
\]

Thus

\[
\frac{\Delta s}{\Delta t} = \frac{s(t + h) - s(t)}{\Delta t} = \frac{\Delta f}{\Delta t} g(t) + f(t) \frac{\Delta g}{\Delta t} + (\Delta f) \cdot \frac{\Delta g}{\Delta t}
\]
and it follows for infinitesimal $\Delta t$ (and thus $\frac{\Delta g}{\Delta t}$ just approaching a fixed number while $\Delta f$ is approaching 0):

$$\frac{d}{dt}(f(t)g(t)) = \frac{df}{dt}g(t) + f(t)\frac{dg}{dt}$$

**Reciprocal Rule**

Let $s(t) = \frac{1}{g(t)}$. Then

$$\Delta s = s(t + \Delta t) - s(t) = \frac{1}{g(t + \Delta t)} - \frac{1}{g(t)} = \frac{1}{g(t)} - \frac{1}{g(t) + \Delta g} = \frac{g(t) - (g(t) + \Delta g)}{g(t)(g(t) + \Delta g)} = \frac{-\Delta g}{g(t)(g(t) + \Delta g)}$$

and so

$$\frac{\Delta s}{\Delta t} = \frac{-\frac{\Delta g}{\Delta t}}{g(t)(g(t) + \Delta g)} = \frac{\Delta g}{\Delta t} \left( \frac{-1}{g(t)(g(t) + \Delta g)} \right).$$

If $\Delta t$ becomes infinitesimal also $\Delta g$ becomes infinitesimal (we assume this!) and thus

$$\frac{d}{dt} \left( \frac{1}{g(t)} \right) = \frac{dg}{dt} \left( \frac{-1}{g(t)^2} \right)$$

**Quotient Rule**

**Exercise:** (a) Show that the quotient rule:

$$\frac{d}{dt} \left( \frac{f(t)}{g(t)} \right) = \frac{df}{dt}g(t) - f(t)\frac{dg}{dt}$$

holds as follows:

Let $s(t) = \frac{f(t)}{g(t)} = f(t) \cdot \frac{1}{g(t)}$. Then apply the product rule and the reciprocal rule, and simplify algebraically.

(b) Apply the quotient rule to the special case $f(t) = 1$ and deduce the reciprocal rule from the quotient rule.

**Remarks:** Note that in the formulas above we usually write $\frac{df}{dt}$ for the derivative function when we consider the function $f(t)$ of $t$. Actually this
should be \( \frac{df}{dt}(t) \), but looks clumsy. You may want to rewrite the rules using the alternative notation \( \frac{df}{dt} = f'(t) \). Note that also all rules can be expressed using the independent variable \( x \):

\[
\frac{d}{dx} (f(x)g(x)) = f'(x)g(x) + f(x)g'(x)
\]

and

\[
\frac{d}{dx} \left( \frac{f(x)}{g(x)} \right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}
\]

Write the rules for a function \( P \) of a variable \( T \). How do the rules look like?

The rest is to do examples, as many as you can. Here are a couple.

\[
\frac{d}{dx} \left( \frac{x}{1-x} \right) = \frac{1 \cdot (1-x) - x \cdot (-1)}{(1-x)^2} = \frac{1}{(1-x)^2}
\]

or

\[
\frac{d}{dx} (xe^x) = \left( \frac{dx}{dx} \right) e^x + x \frac{de^x}{dx} = e^x + xe^x = (x+1)e^x
\]

or

\[
\frac{d}{du} \left( \frac{1}{u^2 + 4u + 2} \right) = - \frac{d}{du} \left( u^2 + 4u + 2 \right) = - \frac{2u + 4}{(u^2 + 4u + 2)^2}
\]

What are the rules I used in the three cases above?
Lecture 9: Derivatives and Anti-derivatives of Trigonometric Functions

Take a look at:

Calculating derivatives of trigonometric functions (Their graphs of the trigonometric functions are not correct. The graphs are not built from half-circles.)

A simple harmonic motion is the projection of the motion of a point on a circle rotating with constant angle velocity onto the x- or y-axis. Recall that the coordinates of a point moving in a plane are given by \((x(t), y(t))\). If the point moves on a circle let’s say of radius 1 then

\[ x(t)^2 + y(t)^2 = 1. \]

Thus we can write for each time \(t\), \(x(t) = \sin \alpha(t)\) and \(y(t) = \cos \alpha(t)\), where \(\alpha(t)\) is the angle at time \(t\). The angle velocity is the rate of change of the angle (which we usually measure in radians) with respect to time. If we assume that \(\frac{d}{dt} \alpha(t) = \omega\) is constant then the \(y\)-coordinate of a simple harmonic motion is described by

\[ y(t) = A \sin \omega t, \]

where \(A\) is the amplitude (just the radius of the circle). Note that the motion is periodic. After a time \(T = \frac{2\pi}{\omega}\) the motion repeats because \(\omega(t + \frac{2\pi}{\omega}) = \omega t + 2\pi\) and the sinus function is \(2\pi\)-periodic. If we look at how the velocity is changing it looks like it follows the same pattern. The graph to the left shows the position \(y(t) = \sin(t)\) in blue and the graph of \(\cos(t)\) in the same coordinate system (so we just put \(\omega = 1\) for simplicity). Note that where \(\sin(t)\) has horizontal tangents (velocity 0) the function \(\cos(t)\) is zero. The velocity is \(\frac{\pi}{2}\) behind the position corresponding to the trigonometric identity:

\[ \sin(\alpha) = \cos(\alpha - \frac{\pi}{2}). \]

In fact in the interval \([0, \frac{\pi}{2}]\) the sinus function has positive slope (positive velocity) and the cosine function is positive there. Also when the harmonic motions goes through its center the velocity is maximum.
Let us use our definition of the derivative to find the velocity of the simple harmonic motion from the limit of average velocities definition. So let’s put \( y(t) = \sin(t) \) and calculate

\[
\frac{y(t + \Delta t) - y(t)}{\Delta t} = \frac{\sin(t + \Delta t) - \sin(t)}{\Delta t} = \frac{\sin(t) \cos(\Delta t) + \sin(\Delta t) \cos(t) - \sin(t)}{\Delta t}.
\]

Here we used the angle addition formula

\[
\sin(\alpha + \beta) = \sin(\alpha) \cos(\beta) + \sin(\beta) \cos(\alpha)
\]

There’s also a corresponding formula for the cosine function:

\[
\cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta).
\]

We will also need this formula a lot, that’s why it’s mentioned here. Let’s look back at our calculation and break things apart into factors containing only \( \Delta t \) and others containing only \( t \):

\[
\frac{y(t + \Delta t) - y(t)}{\Delta t} = \cos(t) \frac{\sin(\Delta t)}{\Delta t} + \sin(t) \frac{\cos(\Delta t) - 1}{\Delta t}
\]

Let’s calculate how the expression \( \frac{\sin(\Delta t)}{\Delta t} \) behaves when \( \Delta t \) gets small:

<table>
<thead>
<tr>
<th>( \Delta t )</th>
<th>( \frac{\sin(\Delta t)}{\Delta t} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>0.95885</td>
</tr>
<tr>
<td>0.10</td>
<td>0.9983</td>
</tr>
<tr>
<td>0.01</td>
<td>0.9998</td>
</tr>
<tr>
<td>-0.1</td>
<td>0.9983</td>
</tr>
</tbody>
</table>

Note that actually the values for negative \( \Delta t \) don’t have to be computed (Why?). Looks like \( \lim_{\Delta t \to 0} \frac{\sin(\Delta t)}{\Delta t} = 1 \). We will give a more precise argument for this later on when we understand better how to find limits. Thus we see that \( \cos(t) \frac{\sin(\Delta t)}{\Delta t} \) will approach \( \cos(t) \). Let’s study the second term:

\[
\frac{\cos(\Delta t) - 1}{\Delta t} = \frac{(\cos(\Delta t) - 1)(\cos(\Delta t) + 1)}{\Delta t(\cos(\Delta t) + 1)} = \frac{\sin(\Delta t)}{\Delta t} \frac{\cos(\Delta t) + 1}{\cos(\Delta t) + 1}.
\]

But in the last expression the factor \( \frac{\sin(\Delta t)}{\Delta t} \) approaches 1 as before, \( \cos(\Delta t) + 1 \) approaches 2 and \( \sin(\Delta t) \) approaches 0. Thus the whole will approach 0 and thus also if we multiply it by \( \sin(t) \). Now by checking the graphs it should be obvious that \( \frac{d}{dt} \cos(t) = -\sin(t) \) (notice that \( \cos(t) \) is descending in \( [0, \frac{\pi}{2}] \) but \( \sin(t) \) is non-negative there).
So we can add to our collection of derivatives (for fun using the $x$ variable):

\[
\frac{d}{dx} \sin(x) = \cos(x), \quad \frac{d}{dx} (\cos(x)) = -\sin(x)
\]

That’s the point to start using our other rules:

\[
\frac{d}{dx} \tan(x) = \frac{d}{dx} \sin(x) \cos(x) = \frac{\cos(x) \cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \sec^2(x)
\]

and all the other formulas for derivatives of trigonometric functions. Also, as usual we get the formulas for anti-derivatives at the same time:

\[
\int \sin(x) \,dx = -\cos(x) + C, \quad \int \cos(x) \,dx = \sin(x) + C
\]

Now we can use our other rules to calculate more derivatives, for example:

\[
\frac{d}{d\alpha} \left( \alpha^2 \sin(\alpha) \right) = 2\alpha \sin(\alpha) + \alpha^2 \cos(\alpha)
\]

or

\[
\frac{d}{du} \left( \tan(u) \right) = \frac{\tan(u) - u \sec^2(u)}{u^2}
\]

and also anti-derivatives:

\[
\int (2 \sin(x) + \cos(x) + x^3) \,dx = -2 \cos(x) + \sin(x) + \frac{x^4}{4} + C
\]

and we also can calculate areas with this:

\[
\int_0^\pi \sin(x) \,dx = -\cos(x)|_0^\pi = -\cos(\pi) - (-\cos(0)) = 2
\]

Finally let’s just at this point that

\[
\frac{d}{dx} (\sin(kx)) = k \cos(kx), \quad \frac{d}{dx} (\cos(kx)) = -k \sin(kx)
\]

for constants $k$. This is not surprising. For large $k \sin(kx)$ will oscillate faster so the slopes will have bigger magnitude.
Lecture 10: Integration by Parts; Higher Derivatives

Take a look at:

Integration by parts

If we want to calculate anti-derivatives of products we have a problem because there is no rule that applies. In fact the product rule tells that that in general \( \frac{d}{dx} (f(x)g(x)) \neq f'(x)g'(x) \). But we can still use the product rule to calculate anti-derivatives. Let’s look at:

\[
\int (f'(x)g(x) + f(x)g'(x)) \, dx = f(x)g(x) + C
\]

this is true because if we take the derivative of the right hand side by the product rule we get the integrand, or we are just saying that \( f(x)g(x) \) is anti-derivative of \( f'(x)g(x) + f(x)g'(x) \). Usually integrals do not come in the form above and the rule is used as follows:

\[
\int f'(x)g(x) \, dx = f(x)g(x) - \int f(x)g'(x) \, dx
\]

We omit the constant because each indefinite integral sign contains constants. The idea is \( \int f(x)g'(x) \) might be easier than \( \int f'(x)g(x) \, dx \). Here is an example with \( f'(x) = e^x \) and \( g(x) = x \):

\[
\int e^x \, x \, dx = e^x \cdot x - \int x e^x \, dx = e^x - e^x + C = e^x (x - 1) + C
\]

Here we used that \( \int e^x \, dx = e^x + C \). Similarly we can take \( f'(x) = \cos(x) \) and thus \( f(x) = \sin(x) \) is an anti-derivative, and \( g(x) = x \) and get:

\[
\int \cos(x) \cdot x \, dx = \sin(x) \cdot x - \int \sin(x) \cdot 1 \, dx = x \sin(x) + \cos(x) + C
\]

Don’t forget that there are many ways to write down the formula for integration by parts, depending on what names we give to functions and variables. Also the order in which the functions appear can change:

\[
\int u \frac{dv}{dt} \, dt = u(t)v(t) - \int \frac{du}{dt}v \, dt
\]

Here \( u, v \) are functions of \( t \). But we did not write this explicitly under the integral sum to avoid clumsy notation like \( u(t) \frac{dv}{dt}(t) \). Of course the
Fundamental Theorem of Calculus tells us that definite integrals are just given by evaluating anti-derivatives and the formula becomes

\[ \int_a^b u(x)v'(x) \, dx = u(x)v(x)|_a^b - \int_a^b u'(x)v(x) \, dx \]

Here is an example with \( u(x) = x \) and \( v'(x) = e^{4x} \) so that \( \frac{1}{4} e^{4x} \) is an anti-derivative:

\[ \int_0^3 xe^{4x} \, dx = x \cdot \frac{1}{4} e^{4x} \bigg|_0^3 - \int_0^3 \frac{1}{4} e^{4x} \, dx. \]

This can now be evaluated further:

\[ \frac{1}{4} (3e^{12} - 0 - e^0) - \frac{1}{4} \int_0^3 e^{4x} \, dx = \frac{3}{4} e^{12} - \frac{1}{4} \left( \frac{1}{4} e^{4x} \right)_0^3 = \frac{3}{4} e^{12} - \frac{1}{16} (e^{12} - 1) = \frac{11}{16} e^{12} + \frac{1}{16} = \frac{1}{16} (11e^{12} + 1) \]

This looks like a huge number, is that possible. Well look at the area:
Take a look at:

**Higher Derivatives**

In applications higher derivatives often are important. The meaning of these higher derivatives, both geometrically and in physics, is of its own interest. At this point we just want to introduce the corresponding notation. Of course we can take the derivative of a derivative. If we write \( y = f(x) \) then the notation is \( y' \) for the first (usual) and \( y'' = (y')' \) for the second derivative. In Leibniz notation it is \( \frac{df}{dx} \) for the first and \( \frac{d^2f}{dx^2} \) for the second derivative. Here is an example: If \( f(x) = \sin(x) \) then \( f'(x) = \cos(x) \) and \( f''(x) = -\sin(x) \). A typical example is when we start with a position function, let’s say \( x(t) \). Then \( x'(t) \) is the velocity and \( x''(t) \) is the acceleration. So for a simple harmonic motion the acceleration is just the negative of the position at each point in time. This explains why we end up with a periodic motion as we see. If the position gets large positive the acceleration will be negative and get the velocity to become negative and the motion turn around.
Take a look at: Using the chain rule to differentiate (But don’t believe that hanging ropes hang in the form of parabolas!) and at Chain rule

Suppose that you are climbing a hill with constant slope (see left picture above), let’s say 0.3 meters rising vertically per one meter horizontally, and the $x$-coordinate of your position is described by

$$x(t) = 0.5t + 1$$

with $x(t)$ is horizontal meters and $t$ in seconds. It is now easy to calculate the rate of change of your altitude (your $y$-coordinate) with respect to time by plugging in:

$$y(t) = 0.3(0.5t + 1) = 0.3 \cdot 0.5t + 0.3$$

and we see that $y'(t) = 0.15$ vertical meters per second. Note how the units fit: We multiply 0.3 vertical meters/horizontal meters with 0.5 horizontal meters/second to get 0.15 vertical meters/second. This seems to be obvious. A little less obvious example is given when we slide down a hill described by the graph of a function $f(x) = 50 - 0.1x^2$ with both $x$, $y$ measured in feet and $x(t) = 1 + 1.5t^2$ describes our horizontal motion (see right picture). (Note that the unit of 0.1 here actually has to be in vertical feet/square of horizontal feet to get the right units for $y$.) Then of course we can again calculate our vertical motion:

$$y(t) = 50 - 0.1(1 + 1.5t^2)^2 = 50 - 0.1(1 + 3t^2 + 2.25t^4) = 49.9 - 0.3t^2 - 0.225t^4.$$
Thus our vertical velocity at time time \(t\) is \(v(t) = -0.6t - 0.9t^3\). Now our horizontal velocity is \(x'(t) = 3t\) and the change of our vertical position relative to our horizontal position is \(-0.2x\) vertical feet/horizontal feet. Note that we have the right units. The number 0.2 has the above units of vertical feet/square of horizontal feet and is multiplied by \(x\) with units horizontal feet. How can both these numbers together give directly the answer above. Well the answer is that:

\[-0.2 \cdot (1 + 1.5t^2) \cdot 3t\]

The general case is that of a function \(y(t) = f(x(t))\) with derivative

\[y'(t) = \frac{dy}{dt} = f'(x(t))x'(t).\]

The prime notation has to be taken carefully here because \(f'(x(t)) = \frac{df}{dx}\big|_{x(t)}\) while \(x'(t) = \frac{dx}{dt}\), so we are actually taking derivatives with respect to different variables. The easiest way to memorize (and deduce by cheating) the chain rule is:

\[\frac{df}{dt} = \frac{df}{dx} \cdot \frac{dx}{dt}.\]

In the mathematics literature the above is seen as a composition \((f \circ g)(t)\) and written as:

\[\frac{d}{dt} (f \circ g) = (f \circ g)'(t) = f'(g(t)) \cdot g'(t)\]

Note that on the right hand side the \(\cdot\) is a product and not composition, so is often omitted. Instead of the variable \(t\) also often the variable \(x\) is used and then

\[\frac{d}{dx} (f \circ g) = f'(g(x))g'(x)\]

Here is a first example: Take \(f(x) = \sin(x)\) and \(g(x) = x^2\). Then \(f(g(x)) = \sin(x^2)\). We have \(f'(x) = \cos(x)\) and \(g'(x) = 2x\). Thus

\[\frac{d}{dx} \left(\sin(x^2)\right) = \cos(x^2) \cdot 2x = 2x \cos(x^2).\]
Lecture 12: Taking Difficult Derivatives; Derivatives of Inverse Functions

Take a look at:

Applying the rules of differentiation, Derivatives of inverse functions and Derivatives of inverse trigonometric functions

When taking derivatives things often get more difficult in the sense that we have to apply many rules at the same time and also apply the same rule multiple times. An easy case is application of the product rule to a product of multiple factors. Of course this can always be done by two applications of the product rule in a row. But it is worth to find the general pattern:

\[
\frac{d}{dx} (f(x)g(x)h(x)) = \frac{d}{dx} (f(x)(g(x)h(x))) = f'(x)g(x)h(x) + f(x)\frac{d}{dx} (g(x)h(x)) = f'(x)g(x)h(x) + f(x)g'(x)h(x) + f(x)g(x)h'(x)
\]

So you just work through all factors differentiating one factor at a time and summing up the results.

Example 1: For \( f(x) = x^2e^{8x} \sin(7x) \) we get

\[
f'(x) = \frac{d}{dx} x^2e^{8x} \sin(7x) = 2xe^{8x} \sin(7x) + x^2e^{8x} \sin(7x) + 8x^2e^{8x} \cdot 7\cos(7x) = xe^{8x}(2\sin(7x) + 8x\sin(7x) + 7x\cos(7x))
\]

The chain rule gives more subtle situations. Let’s look at \( F(x) = f(g(h(x))) \), which is a composition of three functions. In establishing the general principle we think of this as \( f((g \circ h)(x)) \) and apply the chain rule to the inner function \( g \circ h \) and the outer function \( f \) to get

\[
F'(x) = f'(g(h(x))) \cdot (g \circ h)'(x),
\]

so a further application gives

\[
F'(x) = f'(g(h(x))) \cdot g'(h(x)) \cdot h'(x).
\]

So we work from the outside to the inside. Another way to think of this is to introduce variables \( u = g(v) \) and \( v = h(x) \) and write:

\[
\frac{dF}{dx} = \frac{dF}{du} \cdot \frac{du}{dv} \cdot \frac{dv}{dx}.
\]
The best is not to use the resulting rule here but to do \textit{explicitly} what we showed for the general case:

**Example 2:** Let $f(x) = \cos(\sin(\cos(x)))$. Then

$$f'(x) = -\sin(\sin(\cos(x))) \cdot \frac{d}{dx}(\cos(x))$$

$$= \sin(\sin(\cos(x))) \cdot \cos(\cos(x)) \cdot (-\sin(x))$$

$$= -\sin(\sin(\cos(x))) \cdot \cos(\cos(x)) \cdot \sin(x)$$

The chain rule is the main tool in finding derivatives of inverse functions. This comes from the identity

$$f \circ f^{-1}(x) = f(f^{-1}(x)) = x$$

**Example 3:** Recall that the natural logarithm $\ln(x)$ is the inverse function of the natural exponential function. So we have

$$e^{\ln x} = x$$

Now take derivatives on both sides, using the chain rule on the left side and $\frac{dx}{dx} = 1$ (This is not algebraic cancellation but the fact that a line of slope 1 has derivative 1 because derivative=slope).

$$e^{\ln x} \frac{d}{dx}(\ln x) = 1,$$

so because $e^{\ln(x)} = x$ we get

$$x \frac{d}{dx}(\ln x) = 1$$

and thus

$$\frac{d}{dx}(\ln x) = \frac{1}{x}$$

Of course by using the chain rule this gives us:

$$\frac{d}{dx} \ln(kx) = \frac{1}{k} \ln x + C$$

and by taking the anti-derivatives we get

$$\int \frac{1}{x} dx = \ln(x) + C.$$
Of course since the domain of $\ln(x)$ is only $(0, \infty)$ the formula will only work for $x > 0$.

**Example 4:** The inverse function of the function $\sin x$ is often denoted $\sin^{-1}(x)$, which has to be taken very carefully. Note that we write $\sin^2 x$ to mean $(\sin(x))^2$. This will be that way for all $\sin^k(x)$ and positive integers $k$. But for $k = -1$, $\sin^{-1} x \neq \frac{1}{\sin x}$ but denotes the inverse function. This is standard in most text books. Sometimes people use $\arcsin x$ instead to denote the inverse function of $\sin x$ to avoid this problem. We start with

$$\sin(\sin^{-1}(x)) = x$$

and apply derivatives on both sides:

$$\frac{d}{dx} \sin(\sin^{-1}(x)) = \frac{d}{dx} x$$

Now proceed as in Example 1.

$$\cos(\sin^{-1}(x)) \cdot \frac{d}{dx} \sin^{-1}(x) = 1$$

or

$$\frac{d}{dx} \sin^{-1}(x) = \frac{1}{\cos(\sin^{-1}(x))}$$

Now we draw an angle $\theta$ in a right triangle with hypotenuse length 1, opposite length $x$ and thus adjacent length $\sqrt{1 - x^2}$ (**Do it!**). Then $\sin(\theta) = x$ so $\theta = \sin^{-1}(x)$ and $\cos(\theta) = \sqrt{1 - x^2}$. Thus we get the formula

$$\frac{d}{dx} \sin^{-1}(x) = \frac{1}{\sqrt{1 - x^2}}.$$ 

Again by taking anti-derivatives we get

$$\int \frac{dx}{\sqrt{1 - x^2}} = \sin^{-1}(x) + C$$
Lecture 13: Calculating Integrals; Substitution

Take a look at:

Definite Integrals, Addition property of definite integrals and How to solve integrals using substitution

Recall that for $a < b$ the symbol

$$\int_a^b f(x) dx$$

just represents the signed area for $f(x)$ over the interval $[a, b]$. This is the sum of the signed areas over intervals where $f(x) > 0$ and $f(x) < 0$. Here signed means that we count the area content positively if the graph of $f(x)$ is above the $x$-axis, and negative when the graph of $f(x)$ is below the $x$-axis. Not surprisingly the following rules are compatible with the above interpretation:

$$\int_a^a f(x) dx = 0$$

for all $a$.

$$\int_a^b f(x) dx = -\int_b^a f(x) dx$$

Note that here if $a < b$ then $b < a$, so this extends our definition.

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$$

This is easy to understand geometrically when $a < c < b$ but follows using the previous formula for all cases. So if $a > b$ and $f(x) > 0$ we will have $\int_a^b f(x) dx < 0$, so we always should remember that the interpretation of the integral as signed area refers to a direction of the $x$-axis. If we have $a > b$ we are integrating in the direction of the negative $x$-axis and that reverses the signs of the areas (in fact in our approximation formulas $\Delta x$ are negative quantities). Note that all of the above is also compatible with the Fundamental Theorem of Calculus. If $F(x)$ is an anti-derivative of $f(x)$ then

$$\int_a^a f(x) dx = F(x)|_a^a = F(a) - F(a) = 0$$

Also

$$\int_a^b f(x) dx = F(b) - F(a) = -\left(F(a) - F(b)\right) = -\int_b^a f(x) dx$$
and

$$\int_a^b f(x)dx = F(b) - F(a) = (F(b) - F(c)) + (F(c) - F(a))$$

$$= (F(c) - F(a)) + (F(b) - F(c)) = \int_a^c f(x)dx + \int_c^b f(x)dx$$

So here is an example:

$$\int_{-1}^1 x^2dx = \int_{-1}^0 x^2dx + \int_0^1 x^2dx = 2 \int_0^1 x^2dx = \frac{2}{3}x^3\bigg|_0^1 = \frac{2}{3}$$

where because of the symmetry of $x^2$ we know that the signed areas over $[-1, 0]$ and $[0, 1]$ are the same. The addition property also is important to find integrals of piecewise defined functions. For example consider $f(x) = |x - 2|$. Then

$$f(x) = \begin{cases} 2 - x & x \leq 2 \\ x - 2 & x > 2 \end{cases}$$

So

$$\int_0^3 f(x)dx = \int_0^2 f(x)dx + \int_2^3 f(x)dx$$

$$= \int_0^2 (2 - x)dx + \int_2^3 (x - 2)dx$$

$$= 2x - \frac{x^2}{2}\bigg|_0^2 + \frac{x^2}{2} - 2x\bigg|_2^3$$

$$= (4 - 2) + \left(\frac{9}{2} - 6 - (2 - 4)\right)$$

$$= \frac{9}{2} - 2 = \frac{5}{2}$$
Calculating integrals can be difficult. Let’s consider

$$\int 2x \cos(x^2) \, dx$$

We know that by chain rule

$$\frac{d}{dx} \sin(x^2) = \cos(x^2) \cdot 2x$$

and so

$$\int 2x \cos(x^2) \, dx = \sin(x^2) + C$$

Of course, in this way we can find many anti-derivatives. But usually we have
given a specific integrand. So we want to see whether we can deduce anti-
derivative in a systematic way by thinking the chain rule backwards. This
leads to a new integration technique, which is called substitution. Just begin
with the chain rule in the form:

$$\frac{d}{dx} f(g(x)) = f'(g(x)) \cdot g'(x)$$

and take anti-derivatives on both sides. Note that the anti-derivative of a
derivative gives back the function up to a constant. So we get

$$f(g(x)) + C = \int f'(g(x)) g'(x) \, dx.$$  

But how do we spot that an integrand has this form? The way to do this
systematically is to go and use a substitution $$u = g(x)$$. Then $$g'(x) = \frac{du}{dx}$$.
Using the formal notation $$du = g'(x) \, dx$$ we write

$$\int f'(g(x)) g'(x) \, dx = \int f'(u) \, du = f(u) + C = f(g(x)) + C$$

Of course $$f(u)$$ is just an anti-derivative of $$f'(u)$$ so if $$F(u)$$ is an anti-
derivative of $$f(u)$$ we usually apply the technique as follows:

$$\int f(g(x)) g'(x) \, dx = \int f(u) \, du = F(u) + C = F(g(x)) + C.$$  

Let’s see how this formalism works in a more difficult situation:

$$\int \frac{5x^2}{7 + x^3} \, dx.$$  

34
We write $u = 7 + x^3$. Then we get $du = 3x^2dx$. We can solve for $dx = \frac{du}{3x^2}$ and plug this back in to get

$$\int \frac{5x^2}{7 + x^3}dx = \int \frac{5}{u} \frac{du}{3}$$

We canceled the $x^2$. Make sure that when you take the anti-derivative with respect to $u$ no $x$-variable anymore appears in the integrand. Note that $x$ is not a constant with respect to $u$. So we get

$$\frac{5}{3} \int \frac{du}{u} = \frac{5}{3} \ln(u) + C = \frac{5}{3} \ln(7 + x^3) + C$$

This shows that we found the anti-derivative correctly. You can finally check that the derivative of the result is the original integrand.
Lecture 14: More Substitution

Take a look at:

**More solving Integrals using substitution**

Let’s first look at a case where the substitution has to be set-up slightly differently to be successful:

$$\int \frac{x}{1 + x^4} dx$$

Here the first guess $u = 1 + x^4$ will not be successful. We substitute instead $u = x^2$ so that $2x dx = du$ or $x dx = \frac{du}{2}$. Then

$$\int \frac{x}{1 + x^4} dx = \frac{1}{2} \int \frac{du}{1 + u^2} = \tan^{-1} u + C = \tan^{-1}(1 + x^4) + C$$

There are cases where the integrand does not obviously have the form $f(g(x))g'(x)$ but the substitution method is still working. Let’s consider:

$$\int \frac{x^2}{\sqrt{1 - x}} dx$$

We use the substitution $u = \sqrt{1 - x}$. Then $\frac{du}{dx} = \frac{1}{2\sqrt{1 - x}}(-1) = \frac{-1}{2u}$ or $dx = -2udu$.

If we plug back into the integrand we get

$$\frac{x^2(-2u)}{u} = -2x^2$$

How do we get rid of the $x^2$? We solve $u = \sqrt{1 - x}$ for $x$: $u^2 = 1 - x$ and thus $x = 1 - u^2$. So we finally end with the integral:

$$\int -2(1-u^2)^2 du = -2 \int 1 - 2u^2 + u^4 du = 2 \int 2u^2 - u^4 - 1 du = 2(\frac{2}{3}u^3 - \frac{1}{5}u^5 - u) + C$$

Of course this is not the answer. We always have to substitute back to get the answer:

$$\int \frac{x^2}{\sqrt{1 - x}} dx = \frac{4}{3}(1 - x)^{3/2} - \frac{2}{5}(1 - x)^{5/2} - 2(1 - x)^{1/2} + C$$

It would have been quite difficult to guess the answer in this case. Of course we can now calculate definite integrals by first calculating anti-derivatives and apply the Fundamental Theorem of Calculus. In many situations it is
more comfortable to avoid the back-substitution and instead transform the limits. The idea is to note that

\[
\int_a^b f'(g(x))g'(x)dx = \int_{g(a)}^{g(b)} f(u)du
\]

To see this just use the anti-derivative \( F(u) \) of \( f(u) \) to see that the right hand side gives \( F(g(b)) - F(g(a)) \). But this is also the answer we get by first finding the anti-derivative \( F(u) \), plugging back \( u = g(x) \) and then calculate \( F(g(x)) \bigg|_a^b \).

**Example:** We want to calculate

\[
\int_0^7 \sqrt{4+3x}dx
\]

We use the substitution \( u = g(x) = 4 + 3x \) and so \( du = 3dx \) or \( dx = \frac{du}{3} \). Then \( g(0) = 4 \) and \( g(7) = 25 \). So

\[
\int_0^7 \sqrt{4+3x}dx = \int_4^{25} \sqrt{\frac{du}{3}} = \frac{2}{9} \left(5^3 - 2^3\right) = 26
\]

In applications the integral appears as the mathematical tool to accumulate instantaneous information to calculate the total change of a quantity. We have seen this for change of position over a time interval as a result of the function of instantaneous velocities. Suppose we study any quantity \( f(t) \) of time and consider

\[
\int_a^b f'(t)dt.
\]

The independent variable could be time or any other quantity, for example the change in the length of a stick caused by a change in temperature. The Fundamental Theorem of Calculus tells us that

\[
f(b) - f(a) = \int_a^b f'(t)dt
\]

Here is how we may understand in a different way why this is the case referring back to the subdivide-approximate-sum up idea.

\[
\int_a^b f'(t)dt \approx \sum f'(t_i)\Delta t \approx \sum \frac{\Delta f}{\Delta t} \cdot \Delta t = \sum \Delta f,
\]

37
where each $\Delta f$ is the change of the variable $f$ over the interval corresponding to a rectangle of the subdivide procedure of base length $\Delta t$. Now suppose we are in the situation of the chain rule, so the integral we consider has the form

$$\int_a^b f'(g(t))g'(t)dt$$

With $u = g(t)$ and $g'(t) = \frac{\Delta u}{\Delta t}$ we get similarly to the above:

$$\int_a^b f'(g(t))g'(t)dt \approx \sum \Delta f \cdot \frac{\Delta u}{\Delta t} \cdot \Delta t = \sum \Delta f,$$

The change in the variable $y = f(u)$ is happening from $u = g(a)$ to $u = g(b)$.

**Application:** Suppose that a person exhales $\frac{1}{6}(t - 1.5)^{-2/3}$ liter/sec for 1 second. How much air has exited his lungs. The total amount of air will sum up from the amounts exited over small intervals of time $\Delta t$. For those time intervals the value of $\frac{1}{6}(t - 1.5)^{-2/3}$ evaluated at some point in the interval multiplied by $\Delta t$ is an approximation, which becomes the exact value as $\Delta t$ gets smaller and smaller. Thus what we are looking for is

$$\int_0^1 \frac{1}{6}(t - 1.5)^{-2/3}dt,$$

which we can calculate by the substitution method using the substitution $u = t - 1.5$ so $du = dt$.

$$\int_0^1 \frac{1}{6}(t - 1.5)^{-2/3}dt = \int_{-1.5}^{-0.5} u^{-2/3}du$$

$$= 3u^{1/3}|_{-0.5}^{-1.5}$$

$$= 3(1.5^{1/3} - 0.5^{1/3}) \approx 1.05 \text{ liter}$$
Lecture 15: Hyperbolic Functions, Logarithmic Differentiation

Take a look at:

- Logarithmic differentiation
- Derivatives of hyperbolic functions
- Derivatives of inverse hyperbolic functions

First note that the derivative \( \frac{d}{dx} \ln x = \frac{1}{x} \) only makes sense for \( x > 0 \) since the domain of \( \ln x \) is \([0, \infty)\). There is an easy way to extend this to a very important result using an easy application of the chain rule. In fact the function \( \ln|x| \) is defined for all \( x \neq 0 \). \( x > 0 \) it is the function \( \ln x \) considered before. But for \( x < 0 \) \( \ln|x| = \ln(-x) \). From the chain rule we get

\[
\frac{d}{dx} \ln(-x) = \frac{1}{-x} \cdot (-1) = \frac{1}{x}.
\]

In the following we assume that the values of the independent variables are always chosen such that a corresponding argument of \( \ln \) is not 0. So we can summarize:

\[
\frac{d}{dx} \ln|x| = \frac{1}{x}.
\]

In logarithmic differentiation we use the property that

\[
\frac{d}{dx} \ln|f(x)| = \frac{f'(x)}{f(x)}.
\]

If we use \( y = f(x) \) then the basic formula reduces to \( \frac{d}{dx} \ln|y| = \frac{y'}{y} \). Note that \( \frac{d}{dy} \ln|y| = \frac{1}{y} \). This shows that we always have to pay attention to the variable with respect to which we differentiate. Recall the laws of logarithms. Since we will mostly use them for the natural logarithm we only state them here in this case:

\[
\ln\left(\frac{A}{B}\right) = \ln A - \ln B, \quad \ln(A \cdot B) = \ln A + \ln B, \quad \ln(A^r) = r \ln A.
\]

The best is to show the method with an example. In the following we start with an equation, then take absolute values on both sides, then we expand the right hand side using the laws of logarithms. Then we take derivatives using necessary rules, finally we get the answer for the derivatives of the function.
\[ y = f(x) = \frac{\sqrt{|x|}(x^2 - 1)^5}{(x + 2)(x - 4)^3} \]
\[ |y| = \frac{\sqrt{|x|}|x^2 - 1|^5}{|x + 2| \cdot |x - 4|^3} \]
\[ \ln |y| = \ln \left( \frac{x^{1/2} \cdot |x^2 - 1|^5}{|x + 2| \cdot |x - 4|^3} \right) \]
\[ = \frac{1}{2} \ln |x| + 5 \ln |x^2 - 1| - \ln |x + 2| - 3 \ln |x - 4| \]

Now we take the derivatives with respect to \( x \) on both sides:
\[
\frac{y'}{y} = \frac{d}{dx} \left( \frac{1}{2} \ln |x| + 5 \ln |x^2 - 1| - \ln |x + 2| - 3 \ln |x - 4| \right)
\]
\[ = \frac{1}{2x} + \frac{10x}{x^2 - 1} - \frac{1}{x + 2} - \frac{3}{x - 4} \]
So the final answer is:
\[ y' = y \left( \frac{1}{2x} + \frac{10x}{x^2 - 1} - \frac{1}{x + 2} - \frac{3}{x - 4} \right) \]

This could be multiplied out after substituting back \( y = f(x) \). This would be still quite some work and you might wonder whether logarithmic differentiation is easier than using quotient, product and power rule (Note that the laws of logarithm are precisely replacing the use of those rules by corresponding laws of logarithms.) But suppose we want to know \( f'(2) \). Then \( f(2) = \frac{\sqrt{2} \cdot 3^5}{4 \cdot (-2)^7} = -\frac{243\sqrt{2}}{32} \) and we get
\[ f'(2) = -\frac{243\sqrt{2}}{32} \cdot \left( \frac{1}{4} + \frac{20}{3} - \frac{1}{4} \cdot \frac{3}{2} \right) = \frac{3969\sqrt{2}}{64} \approx 87.7 \]
It would have been much harder to get this result without logarithmic differentiation.

Note that the original function \( y = f(x) \) is defined for all \( x \neq -2, 4 \). But we should expect some problems also for \( x = 0 \) because the derivatives of \( \sqrt{x} \) is \( \frac{1}{2\sqrt{x}} \) and thus is not defined for \( x = 0 \). When we apply the rules of logarithms we get terms, which are only defined for \( x \) not \( 0, -1, -2, 1, 4 \).
Just like trigonometric functions parametrize the points on the unit circle $x^2 + y^2 = 1$ the points on the hyperbola $x^2 - y^2 = 1$ are parametrized by the hyperbolic functions

$$
\cosh(t) = \frac{1}{2}(e^t + e^{-t}), \quad \sinh(t) = \frac{1}{2}(e^t - e^{-t})
$$

(see the graph in CCC 16). In fact, an easy calculation gives $\cosh^2 t - \sinh^2 t = 1$. From the definitions and $\frac{d}{dt} e^{kt} = ke^{kt}$ for constants $k$ we get immediately:

$$
\frac{d}{dt} \sinh(t) = \cosh(t), \quad \frac{d}{dt} \cosh(t) = \sinh(t)
$$

and of course the corresponding anti-derivative formulas:

$$
\int \cosh t \, dt = \sinh t + C, \quad \int \sinh t \, dt = \cosh t + C
$$

We can use this to practice how to find derivatives of inverse functions. Starting from (switching to variable $x$ for fun!)

$$
cosh(\cosh^{-1}(x)) = x
$$

by taking derivatives on both sides and applying the chain rule on the left side we get

$$
\sinh(\cosh^{-1}(x)) \cdot \frac{d}{dx} \cosh^{-1}(x) = 1
$$

or

$$
\frac{d}{dx} \cosh^{-1}(x) = \frac{1}{\sinh(\cosh^{-1}(x))}
$$

From $\cosh^2 x - \sinh^2 x = 1$ we get $\sinh^2 x = \cosh^2 x - 1$ and so $\sinh x = \sqrt{\cosh^2 x - 1}$. (Note that $\cosh^2 x > 1$ for all $x$. So the domain of $\cosh^{-1} x$ is $[1, \infty]$. Moreover, in order to actually have $\cosh(x)$ an invertible function we need to restrict the domain of $\cosh(x)$ and so the range of $\cosh^{-1} x$ to $x \geq 0$. Since we have $x \geq 0$ the positive root suffices above!). Using $\cosh^2(\cosh^{-1}(x)) = (\cosh(\cosh^{-1}(x)))^2 = x^2$ it follows:

$$
\frac{d}{dx} \cosh^{-1} x = \frac{1}{\sqrt{x^2 - 1}}
$$
Lecture 16: Implicit Differentiation

Take a look at:

How to find derivatives of implicit functions

Often curves are not given as the graph of a function but by an equation in for example $x, y$. So the equation $x^2 - y^2 = 1$ describes a hyperbola, which is the set of all points with coordinates $(x, y)$ satisfying the equation. Below we see the graph and the tangent line at the point $(2, \sqrt{3})$. How do we find the equation of the tangent line?

In this case we can solve for $y = \pm \sqrt{x^2 - 1}$ and analyze the derivatives for each branch separately. Implicit differentiation just replaces $y = f(x)$ and uses the chain rule when taking the derivative of the equation:

$$x^2 - f(x)^2 = 1$$

so

$$\frac{d}{dx}(x^2 - f(x)^2) = \frac{d}{dx}1$$

and we get

$$2x - 2f(x) \cdot f'(x) = 0$$

We can solve for $f'(x)$ and get

$$f'(x) = y' = \frac{x}{y}$$
For $y = \sqrt{x^2 - 1}$ this gives $y' = \frac{x}{\sqrt{x^2 - 1}}$, and for $y = -\sqrt{x^2 - 1}$ we get $y' = -\frac{x}{\sqrt{x^2 - 1}}$. These are just the answers we would have gotten directly. We can easily find the equation of a tangent line let’s say at $(2, \sqrt{3})$ by using $y' = \frac{2}{\sqrt{3}}$ and get $y = \sqrt{3} + \frac{2}{\sqrt{3}}(x - 2)$ or

$$y = \frac{2}{\sqrt{3}}x - \frac{1}{\sqrt{3}}$$

There are more difficult curves when it is impossible or impractical to solve for $y$ as a function of $x$, like

$$x \sin x + y \sin y = 1$$

Some idea how difficult this should be you can get from looking at some picture of it

For example we might be interested in finding the points where the tangent line is horizontal.

Implicit differentiation, which essentially is just application of the chain rule, allows us to do so.

We begin with the equation $x \sin x + y \sin (y) = 1$ and take derivatives with respect to $x$ on both sides.

$$\frac{d}{dx} x \sin x + \frac{d}{dx} y \sin y = \frac{d}{dx} (1)$$

43
and get by pretending that \( y = f(x) \)

\[
\sin x + x \cos x + \frac{d}{dx} f(x) \sin f(x) = 0,
\]

so by using the chain rule

\[
\sin x + x \cos x + f'(x) \sin f(x) + f(x) \cos f(x) \cdot f'(x) = 0
\]

To simplify notation we usually don’t even write \( f(x) \) but write \( y \) for it:

\[
\sin x + x \cos x + y' \sin(y) + y \cos y \cdot y' = 0.
\]

We can solve this for \( y' \) to get

\[
y' = -\frac{x \cos x + \sin x}{\sin y + y \cos y}
\]

A fraction is zero when the numerator is zero and the denominator is \( \neq 0 \).

So we have a horizontal tangent whenever \( x = 0 \). Here is the graph of \( y = x \cos x + \sin x \):

![Graph of y = x cos x + sin x](image)

Note that is a non-trivial task to find \( y \)-values even corresponding to \( x = 0 \). We see from the graph that there are many values of \( y \) satisfying

\[
y \sin y = 1 \quad \text{or} \quad \sin y = \frac{1}{y}
\]

Note that this is also quite clear from the formula: If \( y \) gets large then \( \frac{1}{y} \) gets small (in fact also for negative \( y \)), and \( \sin y \) runs through all numbers
in $[-1,1]$ again and again. So no wonder that $y \sin y = 1$ occurs again and again. But it is difficult to find the $y$-values exactly. $(0,1.1111)$ is almost on the graph: You can check that $1.1111 \cdot \sin(1.1111) \approx 0.9957536773$. We will get back to the problem of finding approximate solutions, also using Calculus tools, later in the course. If we try to find other values of $x$ with $x \cos x + \sin x = 0$ we can plot the graph and see that there should be an approximate solution for $x = 2$. By asking a computer we get $x = 2.02876$ for an approximate solution.
Lecture 17: Related Rates

Take a look at:
A point moving on a graph  |  Hot air balloon problem  
Check out more youtube videos on related rates!

Suppose a particle is moving along the ellipse $x^2 + 4y^2 = 1$ and it’s horizontal velocity (the $x$-component of its velocity vector) is $\frac{dx}{dt} = 12$ when $x = 0.5$ and $y = 0.25\sqrt{3}$. Assume that we measure lengths in meters and time in seconds. We want to know the vertical velocity at that point. We know that the velocity vector will point in the direction of the tangent line and the $x$-component of the vector is positive, so the vector will point to the right hand side. So we know some $\Delta x$ for the tangent line, and after calculating the slope of the tangent line, we can find the corresponding $\Delta y$. This is one way to think about it and we will follow this way to check our thinking below. But let’s look at the motion of the particle in the plane given by $(x(t), y(t))$, where $x, y$ are related by the equation above, but we actually don’t have any further information about $(x(t), y(t))$. But the relation between the position variables will lead to a relation between the corresponding velocities, or rates of change. Problems asking how rates of change are related for quantities related by equations are called Related rates problems. So the problem above asks to find

$$\frac{dy}{dt} \Big|_P$$

if we assume that

$$\frac{dx}{dt} \Big|_P = 12,$$

where $P(0.5, 0.25\sqrt{3})$. By implicit differentiation we get

$$2xx' + 8yy' = 0$$

and so

$$y' = -\frac{1}{4} \frac{xx'}{y}.$$

Here the prime is differentiation with respect to $t$. So at the point $P$ we get

$$y' = -\frac{6}{\sqrt{3}}$$

46
Note that if we had first calculated the slope of the tangent line by implicit differentiation

\[ 2x + 8y \frac{dy}{dx} = 0 \]

and solved:

\[ \frac{dy}{dx} = -\frac{x}{4y}, \]

so we get

\[ \frac{dy}{dx} \bigg|_P = \frac{-1}{2\sqrt{3}}. \]

We know if there are no forces anymore acting on the moving particle to hold it on the curve it will keep on moving along the line with constant velocity. Then, within \( \Delta t = 1 \) second it will move \( \Delta x = 12 \) meter. so, using that the slope of the tangent line is \( \frac{dy}{dx} \bigg|_P = \frac{\Delta y}{\Delta x} \) above we get that the particle moves \( \Delta y = 12 \cdot \left(-\frac{1}{2\sqrt{3}}\right) = -\frac{6}{\sqrt{3}}, \) as expected. It is important to note that in the situation above the curve \( x^2 + 4y^2 = 1 \) is not the position graph of the function (note that this would mean the particle have to be at two points at the same time). The horizontal \( x \)-axis is \textbf{not time} but a space coordinate. Thus it is important to differentiate carefully the meaning of \( \frac{dy}{dx} \) from the velocity of the moving particle.

Another typical Related Rates problem is the so called \textbf{Ladder problem}:

Here is a typical question: A 12- m ladder is leaning against a wall (long ladder \( \Theta \), maybe somebody cleaning a window on a 4th floor). The ladder
slips and begins to fall. If the foot of the ladder moves away from the wall at a constant speed of 2 m/sec, how fast is the top of the ladder approaching the ground when the base is 7 m away from the wall.

Above you see two pictures to visualize the situation. The right hand side shows some snapshot in time. It is important to understand that the 12 meters are fixed while the distance from the top of the ladder to the ground is actually changing. Call this distance $y$ and call the corresponding horizontal distance from the base of the ladder to the wall $x$. Note that we actually have $x(t)$ and $y(t)$. We have the basic relation:

$$x^2 + y^2 = 12^2$$

The question is to find $y' = \frac{dy}{dt}$ when $x = 7$ and we know that $x' = \frac{dx}{dt} = 2$. By taking derivatives we get

$$2xx' + 2yy' = 0$$

and solving for $y'$ we have

$$y' = -\frac{xx'}{y}$$

Now when $x = 7$ m then $y^2 = 144 - 49 = 95$ m$^2$ and so $y = \sqrt{95}$ m, and we get that for this moment in time:

$$y' = -\frac{14}{\sqrt{95}} \text{ m/sec.}$$

Of course $y'$ is negative because the top of the ladder is moving downwards. Thus by choosing the coordinate system with $y$-axis upwards the distance to the ground is shrinking. So our final answer is: The top of the ladder does approach the ground with a speed of

$$\frac{14}{\sqrt{95}} \approx 1.44 \text{ m/sec}$$
Lecture 18: Limits

Take a look at:
- Using a graph to define limits
- Understanding limits - using notation
- How to determine limits of functions
- Understanding the properties of limits

Intuitively, a limit is what the output of a function approaches as the input of that function approaches some value. We have been using limits in our study of slope. The derivative at $a$ is determined as:

$$f'(a) = \frac{df}{dx}_{x=a} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(a + \Delta x) - f(a)}{\Delta x}$$

Here when we take the limit we consider a function of $\Delta x$ (the input) and study how the values of the function $\frac{\Delta f}{\Delta x}$ are changing when we change the input by making it getting closer and closer to 0.

There are many other situations where limits appear naturally. For example in Special Relativity the velocity of two objects moving where the first object moves with velocity $u$ to the lab and the second one moves with velocity $v$ with respect to the first object is given by:

$$w = \frac{u + v}{1 - \frac{uv}{c^2}}$$

Here $c$ is the speed of light. Note that if we add two velocities, each equal to $c$ we get

$$w = \frac{c + c}{1 + \frac{c^2}{c^2}} = c.$$ 

It is even more interesting to study a case when we consider objects where one is moving with velocity $c$ and the other one with velocity $-c$. Classically the answer would be 0. But the special relativity formula does not give a well-defined answer either because both numerator and denominator become 0 in this case. But mathematics can still be used to give an interesting prediction. In fact we may first set $u = c$, which is not a problem to get

$$f(v) = w = \frac{c - v}{1 - \frac{v}{c}}.$$ 

Note that if we just substitute $v = c$ we get an expression of the form $\frac{0}{0}$, called an indeterminate form. Note that for all $v \neq c$ the output of this expression
is $c$ because we can actually cancel. so even though $f(c)$ is not defined the limit

$$\lim_{v \to c} f(v) = c$$

is defined. This is like turning on a flashlight in a train moving with the speed of light, opposite to the direction of the train. This is like a standard situation in Calculus when limits are important. In most cases until we have been using tables to find limits. But there are nice algebraic methods to calculate limits. Consider the function

$$y = \frac{x^2 - 4}{x - 2}$$

This function is not defined for $x = 2$. On the other hand for $x \neq 2$ we can just divide by $x - 2$ and see that for $x \neq 2$ we have $y = x + 2$. Here is the graph of this function on the left:

Here is the formal Definition: A number $L$ is the limit of a function $f(x)$ as $x$ approaches a value $c$ if the outputs $f(x)$ are all as close as we want to $L$ if the values of $x$ are sufficient close to $L$. But we never consider actually $f(c)$, it might not even be defined as for the function $f(x)$ above. The notation is:

$$\lim_{x \to c} f(x) = L.$$  

Next we consider

$$g(x) = \begin{cases} x + 2 & x < 2 \\ x + 4 & x \geq 2 \end{cases}$$
where you see the graph on the right side above. Notice that in this case the values of $f(x)$ are all close to 4 for $x$ close to 2 and $x < 2$. For this we write

$$\lim_{x \to 2^-} g(x) = 4.$$ 

This is called the left-limit of $g(x)$ as $x$ approaches 2. On the other hand for $x > 2$ and close to 2 the values of $f(x)$ are close to 6. This we write:

$$\lim_{x \to 2^+} g(x) = 6$$

This is called the right-limit of $g(x)$ as $x$ approaches 2. Of course the definitions for left- and right-limits apply for general functions. Note that for the function $g(x)$ there is no number $L$ satisfying the properties of the Definition of limit. We say that the limit does not exits, and abbreviate DNE.

Now if the values of a function $f(x)$ are getting larger and larger (and are positive) when $x$ approaches $c$, we write

$$\lim_{x \to c} f(x) = +\infty.$$ 

If they are negative and getting larger and larger in size when $x$ approaches $c$ (by abuse of notation we just say the values are getting smaller and smaller) we write

$$\lim_{x \to c} f(x) = -\infty.$$ 

Examples would be $f(x) = \frac{1}{|x-3|}$ for the first case and $f(x) = -\frac{1}{(x-3)^2}$ for the second case. The example $f(x) = \frac{1}{x-3}$ shows that it also is of use to have notation when the left and right-hand values differ by sign but get larger in size.
Lecture 19: Limits and Asymptotes

Take a look at:

Graphs and limits defining asymptotes and infinity and Horizontal asymptotes and limits and infinity

The mathematical definition of \( \lim_{t \to \infty} f(t) \) for a function \( f(t) \) is very similar to the definitions in CCC 18. We will say that \( \lim_{t \to \infty} f(t) = L \) for a number \( L \) if all the values \( f(t) \) for very large \( t \) will be as close to \( L \) as we want.

Of course we will also introduce notation \( \lim_{t \to \infty} f(t) = \infty \) meaning that the values are all very large and get be made arbitrarily large for sufficiently large \( x \). But only if we have a limit \( L \), which is a number, we will say that the limit exists. The main starting point for understanding limits at infinity

A typical example is the quotient of two polynomials \( \frac{P(t)}{Q(t)} \). You probably have seen these results before. If the degree of \( P(t) \) is larger than the degree of \( Q(t) \) then \( \lim_{t \to \infty} \frac{P(t)}{Q(t)} = \pm \infty \) with the sign depending on whether the quotient of the leading coefficients is a positive or negative number. If the degree of \( P(t) \) is smaller than the degree of \( Q(t) \) then \( \lim_{t \to \infty} \frac{P(t)}{Q(t)} = 0 \). If both polynomials have the same degree then the limit is the quotient of the leading coefficients. This can be understood by writing with \( a_n, b_m \neq 0 \):

\[
\frac{P(t)}{Q(t)} = \frac{a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0}{b_m t^m + b_{m-1} t^{m-1} + \ldots + b_1 t + b_0} \approx \frac{a_n t^n}{b_m t^m} = \frac{a_n}{b_m} t^{n-m}
\]

and

\[
\lim_{t \to \infty} t^i = \begin{cases} 
+\infty & i > 0 \\
-\infty & i < 0 \\
1 & i = 0 
\end{cases}
\]

The \( \approx \) method has to be used carefully. In this case we are using that \( t^n \) is much greater than \( t^m \) for \( n > m \) and large \( t \). More precise arguments will use the limit laws described below. But let’s look at an application first:

If a function \( f(t) \) represents a system that varies in time, the existence of \( \lim_{t \to \infty} f(t) \) means that the system reaches a steady state state (or equilibrium). For example suppose that the amplitude of an oscillator is given by

\[
a(t) = 2 \frac{t + \sin t}{t}.
\]

Let’s look at the graph to guess the limit:
We see that the values are oscillating around the $y$-value of 1 and are all getting closer and closer to this value when $t$ gets larger and larger, so our conjecture is that the limit is 1.

We first note that

$$\lim_{t \to \infty} \frac{t + \sin t}{t} = \lim_{t \to \infty} \left(1 + \frac{\sin t}{t}\right)$$

This uses one of the so-called limit laws, which apply to limits at $c$ but also to limits at $\infty$ (let’s use the notation $c$ for both cases). In fact, if $\lim f(t)$ and $\lim g(t)$ exist then

$$\lim_{t \to c} (f(t) + g(t)) = \lim_{t \to c} f(t) + \lim_{t \to c} g(t)$$

We also have for $k$ a constant:

$$\lim_{t \to c} kf(t) = k \lim_{t \to c} f(t)$$

Compare these with the linearity properties for derivatives and anti-derivatives (and integrals). This is the case because for functions $f(x)$, $g(x)$ also $\Delta(f + g) = \Delta f + \Delta g$ and $\Delta(kf) = k\Delta f$. But note that $\Delta(f \cdot g) \neq \Delta f \cdot \Delta g$ (just
\[ \Delta f \cdot \Delta g = (f(t + \Delta t) - f(t)) \cdot (g(t + \Delta t) - g(t)) \]
\[ = f(t + \Delta t)g(t + \Delta t) - f(t)g(t) - f(t)g(t + \Delta t) + f(t + \Delta t)g(t) \]
\[ = \Delta f \cdot g - f(t)g(t + \Delta t) - f(t + \Delta t)g(t) \]

This is the origin of the product rule. A proof actually requires the following product rule for limits:

\[ \lim_{t \to c} (f(t)g(t)) = \lim_{t \to c} f(t) \cdot \lim_{t \to c} g(t) \]

and for \( \lim_{t \to c} g(t) \neq 0 \) we have

\[ \lim_{t \to c} \frac{f(t)}{g(t)} = \frac{\lim_{t \to c} f(t)}{\lim_{t \to c} g(t)} \]

Now if we want to argue the limit \( \lim_{t \to \infty} \frac{\sin t}{t} \) for \( t \) to \( \infty \) we use the so called Sandwich Theorem: Since \(-1 \leq \sin t \leq 1\) and \( t > 0 \) we get

\[ \frac{-1}{t} \leq \frac{\sin t}{t} \leq \frac{1}{t} \]

This means that \( \frac{\sin t}{t} \) is squeezed like the meat/cheese in your sandwich between the functions \(-\frac{1}{t}\) and \( \frac{1}{t}\). Since both approach 0 for \( t \to \infty \) we conclude that also

\[ \lim_{t \to \infty} \frac{\sin t}{t} = 0. \]

This is the algebraic proof of what we saw in the graph. The limit laws also can be applied to give a more precise argument for the limit of a quotient of two polynomials. Suppose for example that \( n = m \):

\[ \frac{P(t)}{Q(t)} = \frac{a_nt^n + a_{n-1}t^{n-1} + \ldots + a_1t + a_0}{b_nt^n + b_{n-1}t^{n-1} + \ldots + b_1t + b_0} = \frac{a_nt + a_{n-1}t^{n-1} + \ldots + a_0t^n}{b_n + b_{n-1}t^{n-1} + \ldots + b_0t^n} \]

Here we are dividing both numerator and denominator by \( t^n \), which is allowed because we can assume \( t \neq 0 \). Then since \( \lim_{t \to \infty} t^i = 0 \) for \( i < 0 \) we conclude by first applying the rule for quotients and then for sums in the numerator and denominator:

\[ \lim_{t \to \infty} \frac{P(t)}{Q(t)} = \frac{\lim_{t \to \infty} f(t)}{\lim_{t \to \infty} g(t)} = \frac{a_n}{b_n}. \]
Lecture 20: Continuity and Differentiability

Take a look at:

One-sided limits and continuity and What it means to be differentiable

In may situations we have \( \lim_{x \to c} f(x) = f(c) \), so that the values \( f(x) \) of a function for \( x \) close to \( c \) are close to \( f(c) \), in particular \( f(c) \) is defined. In a situation like this we say that \( f(x) \) is continuous at \( c \). A function is continuous on an interval if it is continuous at all \( x \) in the interval, and we say the function is continuous if it is continuous on its domain. Examples are polynomials, which are continuous for all real numbers, or rational functions are continuous on their domain (the set of points where the denominator is \( \neq 0 \)). For the graph the continuity condition means that we can draw the graph without suddenly lifting our pencil. There is no jump or gap in the graph at points of the domain. Take a look at the two functions whose graphs are shown in CCC 18. Both functions are not continuous at 2. In the first case \( f(2) \) actually is not defined. In the second case, \( \lim_{x \to 2} f(x) \) does not exist. If we extend the definition in the first case by defining \( f(2) = 2 \) the function becomes continuous, while if we choose any other value we get \( \lim_{x \to 2} f(x) \neq f(2) \). Continuity at \( c \) should be seen really as three conditions: \( f(c) \) is defined, \( \lim_{x \to c} f(x) \) exists, and \( \lim_{x \to c} f(x) = f(c) \). If continuity fails at \( c \) we see that \( f(x) \) is discontinuous at \( c \). Failure of any of the three previous conditions will tell us that \( f(x) \) is discontinuous at \( c \). Note that for the second function in CCC 18 we have \( \lim_{x \to c^+} f(x) = f(c) \). So the values of the function for \( x > c \) close to \( c \) are close to \( f(c) \). In this situation we say that \( f(x) \) is right continuous. Similarly left-continuous is defined. Of course \( f(x) \) is continuous at \( c \) if it is both left-continuous and right-continuous at \( c \). If one of the two conditions holds but \( f(x) \) is not continuous at \( c \) then we will say that \( f(x) \) has a jump discontinuity at \( c \). We will say that \( f(x) \) has an infinite discontinuity at \( c \) if one or both of the two limits \( \lim_{x \to c^\pm} f(x) \) is infinite. (Note that we can still have \( f(c) \) defined).

It is not hard to check that we can conclude from the limit laws that sums of continuous functions are continuous or products etc. So, for example because trigonometric functions are continuous, a function like \( \sin x + \cos x + x^5 \) is continuous. Continuity is used all the time when we evaluate limits by substituting \( c \) into the function. Then we are using the basic definition \( \lim_{x \to c} f(x) = f(c) \) to conclude from the value \( f(c) \) the value of the limit. But be careful, this is only legitimate if you know that the function is con-
tinuous at \( c \). Compositions of continuous functions are continuous. Here is an application: Since \( e^x \) and \( \sin x \) are continuous \( \lim_{x \to 0} e^{\sin x} = e^{\sin 0} = e^0 = 1 \). In fact this kind of behavior with respect to limit is a natural property of continuous functions. If \( f(x) \) and \( g(x) \) are continuous then:

\[
f(\lim_{x \to c} g(x)) = \lim_{x \to c} f(g(x)),
\]

so you can switch limits with continuous functions.

One of the main properties of continuous functions is the Intermediate Value Property, which is just the fact that a continuous function can’t jump, so in particular you cannot jump over values at a point where a function is continuous (like it is happening for the second function in CCC 18). The typical jump function is

\[
f(x) = \begin{cases} 
-1 & x < 0 \\
0 & x = 0 \\
+1 & x > 0 
\end{cases}
\]

Note that \( f(-1) = -1, f(1) = +1 \) but not all values in \((-1, 1)\) are values of \( f(x) \) for \( x \) in \((-1, 1)\). This behavior is not possible for continuous functions on intervals. We have the:

**Intermediate Value Theorem:** Suppose that \( f(x) \) is continuous on an interval \([a, b]\) and \( f(a) \neq f(b) \). Then \( f(x) \) takes any value between \( f(a) \) and \( f(b) \) on the interval \((a, b)\). In other words for each value \( M \) between \( f(a) \) and \( f(b) \) we can find at least one \( c \) in \((a, b)\) with \( f(c) = M \).

Finally consider the function \( f(x) = |x| \). This is a continuous function but there is no slope defined at \( x = 0 \). We say that this function is not differentiable at \( x = 0 \). In fact, at \( x = 0 \) the slope jumps from \(-1\) to \(+1\):

56
The slope at \( x = 0 \) is not defined. This is obvious from the picture. We can also justify it using limits: \( \lim_{h \to 0^+} \frac{|h|}{h} = \lim_{h \to 0^-} \frac{|h|}{h} = 1 \) while \( \lim_{h \to 0^-} \frac{|h|}{h} = -1 \). Since the left- and right-limits are different the limit for \( t \) approaching 0, which is the derivative at 0, does not exist.
Lecture 21: L’Hospital’s Rule

Take a look at: 

- What is l’Hospital’s rule
- Applying l’Hospital’s rule in simple cases
- Applying l’Hospital’s rule in complex cases

If you want to know a little historic background go to de l’Hospital. Also check out Real-world applications, background and examples.

The rule of l’Hospital is a comfortable method to compute limits for indeterminate forms \( \frac{0}{0} \) and \( \frac{\infty}{\infty} \). It actually isn’t as important as you may assume from Calculus text books. In Calculus II when you will learn how to calculate improper integrals you will find some nice applications. At this point it is a nice way to use our differentiation tool box to calculate limits. Kind of not surprising that we can pay back to the definition. Last but not least derivatives are limits themselves.

Roughly the rule says that

\[
\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}
\]

if we have that the right hand side exists, and the limits of both numerator and denominator are 0, or both \( \pm \infty \). Here \( c \) can be a number or \( \pm \infty \), the rule applies in both cases. If we assume that the derivatives are continuous one can actually give a quite simple justification (see your text book).

It gives nice justifications in cases where our feeling tells us that’s how it should work. For example we know that exponential functions grow much faster than any polynomials. Well note that differentiating a polynomial sufficiently often it will become a constant: For example if we differentiate \( f(x) = 5x^7 + 6x^5 - 3x + 8 \) seven times we get 7! \( \cdot \) 5 with 7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1. Similarly if \( P(x) = a_n x^n + \ldots \) with \( a_n \neq 0 \) is a degree \( n \) polynomial by differentiating \( n \) times we get:

\[
\lim_{x \to \infty} \frac{P(x)}{e^x} = \lim_{x \to \infty} \frac{n! \cdot a_n}{e^x} = 0.
\]

Each time the application is justified because the limits of numerator and denominator are \( \pm \infty \), with the sign depending on the sign of \( a_n \). You may wonder why we don’t differentiate once more. Well l’Hospital’s rule does not apply once more because the limit of the constant numerator is the constant and not \( \infty \).
Another nice application is
\[
\lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{\cos x}{1} = 1.
\]
Of course we used this identity to prove that \( \frac{d}{dx} \sin x = \cos x \). So don’t fool yourself, the application of l’Hospital’s rule is a nice memo to remember the limit, but not a proof.

Also don’t use l’Hospital’s rule in simple situations like for the asymptotics of rational functions, where elementary methods easily give the answer.

Here is a useful application of the chain rule in combination with l’Hospital’s rule (\( p, q \) constants):
\[
\lim_{x \to 0} \frac{\sin px}{\sin qx} = \lim_{x \to 0} \frac{p \cos(px)}{q \cos(qx)} = \frac{p}{q}.
\]
There is an elementary argument in this case too but l’Hospital’s rule is easier in this case.

Another application is to compare the growth of logarithms and polynomials. If we assume that \( P'(x) \neq 0 \) then
\[
\lim_{x \to \infty} \frac{\ln x}{P(x)} = \lim_{x \to \infty} \frac{1}{xP'(x)} = 0
\]
Of course, \( P'(x) = 0 \) means that \( P(x) \) is constant and \( \lim_{x \to \infty} \ln x = \infty \). Here is another application: Note that \( \lim_{x \to 0^+} \ln x = -\infty \) and \( \lim_{x \to 0^+} x^k = 0 \) for each constant \( k > 0 \). So \( x^k \ln(x) \) is an indeterminate form \( 0 \cdot \infty \). In this case we can think like \( \frac{1}{0} = \infty \) and write
\[
\lim_{x \to 0^+} x^k \ln x = \lim_{x \to 0^+} \frac{\ln x}{\frac{1}{x^k}} = \lim_{x \to 0^+} \frac{\ln x}{x^{-k}} = \lim_{x \to 0^+} x^{-1} - k x^{-k-1} = \lim_{x \to 0^+} x^k - k = 0.
\]
Roughly this means that even \( \ln x \) approaches it’s vertical asymptote much more slowly as \( x^k \) approaches 0 when \( x \) approaches 0 from the right.

Here are [Drill problems] with solutions.
Lecture 22: Linear Approximation

Take a look at:

- Linearization of functions
- How to estimate function values using linearization

Consider the graphs of the functions $\sin x$ and $x$ on the interval $[0, 1]$ on the left below, and look close to 0:

If we zoom in and consider only on the interval $[0, 0.1]$ the graphs are getting even closer. Now note that the derivative of $\sin x$ is $\cos x$ and $y = x$ actually is the equation of the tangent line. Such approximations often are important in applications. For example the mathematics of the pendulum is difficult because one has to solve a differential equation involving both the derivative of the angle $\theta$ and $\sin(\theta)$. So in a first approximation one considers the so called mathematical pendulum, see [Pendulum] and replaces $y = \sin x$ by its tangent line $y = x$.

This appears all over in physics. Let’s consider the gravity exerted by the earth on an object of height $h$ over the surface of the earth. [Newton’s law of gravitation] tells that the force is

$$F = -\gamma \frac{mM}{(R + h)^2} = \gamma m M \cdot \frac{1}{(R + h)^2},$$

where $R$ is the radius of the earth, $M$ is the mass of the earth and $m$ is the mass of the object, $\gamma$ is a constant. Usually we assume the force does not depend on $h$ and just put $h = 0$ to get $F_0 = -\frac{2M}{R^2} = mg$ where $g = \cdots$
\(-\frac{\gamma M}{R^2} \approx -9.8 \text{ m/s}^2\) is the acceleration due to gravity. But of course we know that the force of gravity will get smaller in large heights. A first order approximation will replace the force not by a constant force but replace the function by its tangent line at \(h = 0\). We can calculate:

\[
\frac{d}{dh} \left( \frac{1}{(R + h)^2} \right)_{h=0} = -2(R + h)^{-3}|_{h=0} = -\frac{2}{R^3}
\]

The equation of the tangent line is

\[
F_1 = -\gamma m M \left( \frac{1}{R^2} - \frac{2}{R^3} \cdot h \right) = mg(1 - 2\frac{h}{R})
\]

Note that the force is decreasing.

In a Calculus text book the above problem probably would appear as follows. Consider the function \(f(x) = \frac{1}{(x+1)^2}\). Find the linear approximation, i.e. tangent line approximation for \(x = 0\). Then

\[
\frac{d}{dx} \frac{1}{(x + 1)^2} \bigg|_{x=0} = -2(x + 1)^{-3}|_{x=0} = -2
\]

and we get

\[
\frac{1}{(x + 1)^2} \approx 1 - 2x
\]

Look at the graphs of the function and its linear approximation:
Of course the linear approximation is still not telling the truth when we go far away from the point of tangency. In fact for the gravitation example above it would tell us that for $h = \frac{R}{2}$ the force due to gravity vanishes, which is of course not true. Just like in the graphs above the actual force will always be there and never vanish. You can’t turn off gravity!

Linear approximation can be used to approximate specific values. Like in the example above, we would get for $f(x) = \frac{1}{(x+1)^2}$

$$f(0.1) \approx 0.8.$$ 

Of course in this case we know that actually $f(0.1) = \frac{1}{1.1^2} \approx 0.826446$. But if you think about more difficult functions, just like $\sin x$ above the linear approximation can be helpful:

$$\sin(0.5) \approx 0.4794255$$

is the calculator approximation. The linear approximation gives 0.5, well, not too bad. Note that your calculator also has to use an algorithm to approximate numbers like this. Linear approximation is a rough first order method.

Linear approximation is often also done relative to the value of the function. This is just using a different notation, we are actually not doing anything different. If $y = f(x)$ then

$$\frac{\Delta y}{\Delta x} \approx \frac{dy}{dx},$$

so

$$\Delta y \approx dy = \frac{dy}{dx} \cdot \Delta x = f'(x) \Delta x,$$

so the actual difference in $y$-values $\Delta y$ is approximated by $dy$. Here we set $\Delta x \equiv \Delta x$. Why is this the same as before. Well $\Delta y$ is a difference of two values of the function. So lets say $\Delta y = f(x) - f(a)$. Then $\Delta y = dy$ means

$$f(x) - f(a) = dy = f'(a) \Delta x = f'(a)(x-a)$$

or

$$f(x) = f(a) + f'(a)(x-a),$$

so we recognize the equation of the tangent line for $f(x)$ at $x = a$. We call $dx, dy$ the differentials.
Finally linear approximation and differentials are used in error propagation. Suppose you measure a quantity with some specific error (usually predicted by the laboratory equipment). If we calculate a quantity depending on the measured quantity the error will propagate to the new quantity. Of course we will be interested in how big that error will be. Here is a simple example:

Suppose the radius $r$ of a sphere was measured and found to be 21 cm with a precision of 0.05 cm. We want to give an estimate for error in the calculated resulting value of $V = \frac{4}{3} \pi r^3 \approx 338792$ cm$^3$. The approximation symbol here refers both to the error resulting from using a calculator and the error resulting from the original measurement. Using

$$dV = 4\pi r^2 dr$$

we find for $r = 21$ and $dr = 0.05$ that

$$dV \approx 277 \text{ cm}^3$$

and thus a relative error of about

$$\frac{\Delta V}{V} \approx \frac{dV}{V} = \frac{4\pi r^2 dr}{\frac{4}{3} \pi r^3} = 3\frac{dr}{r}$$

So the relative error in the volume is about three times of the value of the relative error in the radius. Note that the original relative error was

$$\frac{dr}{r} = \frac{0.05}{21} \approx 0.0024 = 0.24\%.$$
Lecture 23: Newton’s Method

Take a look at:

What is Newton’s method and How to use Newton’s method to find roots of equations

Newton’s method uses the idea of linear approximation to find solutions of equations. First note that we can write any equation into the form $f(x) = 0$. Suppose we want to solve the equation

$$2 \cos x = x - 2.$$ 

So we are looking for zeroes of the function:

$$f(x) = 2 \cos x - x + 2.$$ 

You see the graph of $f(x)$ on the left:

On the right hand side you see the graphs of $f(x)$ and the tangent line at $x = 2$ over the interval $[1.6, 2.1]$. We chose $x_0 = 2$ as an easy to calculate starting value. The idea is since the value is not too far away from 0 the value of the $x$-intercept of the tangent line is close to the value of the root we are looking for, well at least closer than $x_0 = 2$. The calculator value for the solution here is $1.714191494$.

In general we do not have to set up the equation of the tangent line but can determine $x_1$ directly. Let’s look at the general case: The equation of the tangent line at $(x_0, f(x_0))$ is

$$y = f(x_0) + f'(x_0)(x - x_0).$$
We define the new approximation \( x_1 \) to be the \( x \)-intercept of this line, so \( x_1 \) satisfies:

\[
0 = f(x_0) + f'(x_0)(x_1 - x_0)
\]

or

\[
x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}
\]

Note that this does not work if we pick a value of \( x_0 \) where \( f'(x_0) = 0 \) because a horizontal tangent line has no \( x \)-intercept. But then we can proceed with \( x_1 \) as we did with \( x_0 \) to get an even better approximation. The value for \( x_1 \) for the above example is

\[
2 + \frac{2 \cos(2)}{2 \sin(2) + 1} \approx 1.7047133
\]

which is much closer to the actual value than the starting value \( x_0 \). In general we will define in this a sequence of numbers \( x_0, x_1, x_2, \ldots, x_n, \ldots \) recursively by

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

Recursively means that after we have calculated \( x_n \) the above formula tells us how to calculate \( x_{n+1} \). In good cases the values of the sequence will approximate the values of the actual root. If \( c \) is the actual value we write:

\[
c = \lim_{n \to \infty} x_n.
\]

The method works great to approximate the values of roots. Suppose we want to find \( \sqrt{7} \). The calculator approximation is 2.6457513. Now \( \sqrt{7} \) is the unique zero of the function \( f(x) = x^2 - 7 \), which has \( f'(x) = 2x \). If we plug this into the above equation we get

\[
x_{n+1} = x_n - \frac{x_n^2 - 7}{2x_n} = \frac{1}{2} (x_n + \frac{7}{x_n})
\]

\[
\begin{array}{|c|c|}
\hline
n & x_n \\
\hline
0 & 3.0000000 \\
1 & 2.6666667 \\
2 & 2.6458333 \\
3 & 2.6457513 \\
4 & 2.6457513 \\
\hline
\end{array}
\]

A good starting value \( x_0 \) should be \( x_0 = 3 \):
After only three approximations the values stabilize and we get the calculator value of $\sqrt{7}$. Astonishing!
Lecture 24: Taylor Polynomials

Take a look at:

Using a polynomial to approximate a function and dance the Taylor Polynomial Tango

Obviously polynomials are much easier to evaluate and work with than transcendental functions like exponential and trigonometric functions and their inverses. Thus it is obvious to try to approximate difficult functions by the simpler polynomial functions. We already did this by approximating the graph of a function by a tangent line at a point. In this case the values and slopes of the function and its tangent line at \( a \) did agree at \( a \), and the values on the tangent line are approximations of the values of the functions close to \( a \). Below you see the graph of the function \( e^x \) (the top one) and graphs of a sequence of polynomial functions

\[
y_0 = 1, \quad y_1 = 1 + x, \quad y_2 = 1 + x + \frac{x^2}{2}, \quad y_3 = 1 + x + \frac{x^2}{2} + \frac{x^3}{6}, \quad y_4 = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}, \quad y_5 = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}
\]

The graphs of the polynomial functions \( y_0, y_1, \ldots, y_5 \) seem to get closer and closer to the graph of \( e^x \). What is the relation between \( e^x \) and those
polynomials. Well for \( y_0(x) = 1 \) the number 1 is just given by the value of \( e^x \) at 0. The function \( y_1(x) \) is the linearization or tangent line at 0. Can we calculate the equation of the blue curve \( y_3 = 1 + x + \frac{x^2}{2} \) from \( e^x \)? Now a polynomial is determined by its coefficients so we only have to find the coefficients.

The linearization \( y_1 \) is a line and completely determined by intercept and slope. The relation with \( e^x \) is that the value and the slope are actually those of \( e^x \) at \( a = 0 \). So we may ask in general: Given a polynomial

\[
P = a_0 + a_1 x + a_2 x^2,
\]

can we find all the coefficients \( a_0, a_1, a_2 \) from the values of the polynomial close to 0? Well \( a_0 = P(0) \) and \( P' = a_1 + 2a_2 x \), so \( a_1 = P'(0) \). Finally \( P'' = 2a_2 \) and so \( a_2 = \frac{P''(0)}{2} \). Actually the second derivative of a degree 2 polynomial is constant. So maybe this is a coincidence. Let’s consider a degree 3 polynomial:

\[
Q = a_0 + a_1 x + a_2 x^2 + a_3 x^3.
\]

Then \( Q' = a_1 + 2a_2 x + 3a_3 x^2 \) and \( Q'' = 2a_2 + 6a_3 x \). In this case \( Q'' \) is not constant, but \( Q''(0) = 2a_2 \). In fact since \( Q'' = 6a_3 \) we can also calculate \( a_3 = \frac{Q''(0)}{6} \). If we compare \( P \) and \( Q \) for small \( x \), their values should be quite close to each other because \( x^3 \) is a small number for small \( x \). Compare with \( 1 + x + \frac{x^2}{2} \) and \( 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \) above.

Let’s look at the general case. It seems that a polynomial \( T_n \) is a best approximation up to order \( n \) close to \( a \) if the first \( n \) derivatives of the polynomial and the function at \( a \) (including the 0-th derivative which we think of as the function itself) coincide at \( a \):

\[
f^{(j)}(a) = T_n^{(j)}(a) \quad \text{for } j = 0, 1, 2, \ldots n
\]

Just as in the cases \( a = 0 \), \( n = 1, 2, 3 \) above the coefficients of the polynomial are determined by this condition. For convenience we write \( T_n \) centered at \( a \)

\[
T_n(x) = c_0 + c_1(x - a) + c_2(x - a)^2 + \ldots + c_n(x - a)^n = \sum_{i=0}^{n} c_i(x - a)^i,
\]

where we use the \( \Sigma \)-notation from CCC 2. Let’s calculate the \( j \)-th derivative of \( T_n(x) \) at \( a \). In principle we apply the chain rule each time we take the
derivative of a power of \((x-a)\). But the inner derivative is 1, so it won’t matter. If we take the derivative we get

\[
T'_n(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2 + \ldots + nc_n(x-a)^{n-1},
\]

so \(T'_n(a) = c_1\). For the second derivative we get

\[
T''_n(x) = 2c_2 + 3 \cdot 2c_3(x-a) + \ldots + n \cdot (n-1)c_n(x-a)^{n-2},
\]

and so \(T''_n(a) = 2c_2\) or \(c_2 = \frac{T''_n(a)}{2}\). Let’s see the next step explicitly:

\[
T''''_n(x) = 3 \cdot 2c_3 + 4 \cdot 3 \cdot 2c_4(x-a) + \ldots + n \cdot (n-1) \cdot (n-2)c_n(x-a)^{n-3},
\]

so \(T''''_n(a) = 6c_3\) or \(c_3 = \frac{T''''_n(a)}{6}\). Note that the number 6 came from applying the power rule to \((x-a)^3\) three times resulting in factors 3, 2, 1. So if we keep on differentiating and substituting \(x = a\) (which sets the higher order terms 0) we see that

\[
T^{(j)}_n(a) = j \cdot (j-1) \ldots 2 \cdot 1 c_j = j! \cdot c_j,
\]

where we have used the factorial notation \(j! = 1 \cdot 2 \cdot 3 \ldots (j-1) \cdot j\) for the product of the first \(j\) natural numbers. Thus, from \(c_j = \frac{T^{(j)}_n(a)}{j!}\) and recalling that we want \(T^{(j)}(a) = f^{(j)}(a)\), we get the definition of the \(n\)-th order Taylor polynomial of the function \(f(x)\) at \(a\):

\[
T_n(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n = \sum_{i=0}^{n} \frac{f^{(i)}(a)}{i!}(x-a)^i.
\]

Of course, for this to make sense the function has to be at least \(n\)-times differentiable at \(a\). For \(a = 0\) the Taylor polynomial is also called the **McLaurin polynomial**.

Let’s calculate the third order Taylor polynomial of \(\sin x\) at \(x = 0\). Since \(f'(x) = \cos x\), \(f'(0) = 1\), and \(f''(x) = -\sin x\), \(f''(0) = 0\), and \(f'''(x) = -\cos(x)\) and \(f'''(0) = -1\) we calculate \(T_2(x) = T_1(x)\), and

\[
T_3(x) = x - \frac{x^3}{6}
\]

Just like the linearization \(T_1(x) = x\) has been a best linear approximation close to 0, \(T_2(x)\) is a best **quadratic approximation** of the function close to 0, and \(T_3(x)\) is the best cubic approximation.
There is a way to use the Fundamental Theorem of Calculus to understand how the Taylor polynomials are approximating functions. Note that

\[ f(x) - f(a) = \int_a^x f'(t)dt \]

and

\[ f'(t) - f'(a) = \int_a^t f''(u)du, \]

so \( f'(t) = f'(a) + \int_a^t f''(u)du \). Let’s see what happens when we substitute this in for \( f'(t) \):

\[ f(x) = f(a) + \int_a^x \left( f'(a) + \int_a^t f''(u)du \right) dt. \]

We can use linearity of the integral to get

\[ f(x) = f(a) + \int_a^x f'(a)dt + \int_a^x \int_a^t f''(u)du = f'(a)(x-a) + \int_a^x \left( \int_a^t f''(u)du \right) dt. \]

Here we used that \( f'(a) \) is a constant to calculate the first integral. Of course the two integrals nested look scary. But we have another application of the Fundamental Theorem of Calculus:

\[ f''(u) - f''(a) = \int_a^u f'''(v)dv \]

and so \( f''(u) = f''(a) + \int_a^u f'''(v)dv \). If we ignore the integral and just plug in \( f''(a) \) for \( f''(u) \) what we get:

\[ \int_a^x \left( \int_a^t f''(a)du \right) dt \]

Let’s calculate the interior integral first:

\[ \int_a^t f''(a)du = f''(a) \bigg|_a^t = f''(a)(t-a) \]

and when we plug in:

\[ \int_a^x f''(a)(t-a)dt = f''(a) \int_a^x (t-a)dt = f''(a) \left. \frac{1}{2}(t-a)^2 \right|_a^x = \frac{f''(a)}{2}(x-a)^2, \]

which is the second term of the Taylor polynomial. We could proceed in this way to get the formula for the Taylor polynomial above. It’s all in the
Fundamental Theorem of Calculus. It also becomes obvious that the error
$f(x) - T_n(x)$ should be related to $f^{(n+1)}(x)$. A more precise analysis following
the lines above or the textbook gives:

**Taylor’s Theorem** The remainder $R_n(x) = f(x) - T_n(x)$ of the Taylor
approximation is

$$R_n(x) = \frac{1}{n!} \int_a^x (x - u)^n f^{(n+1)}(u) du$$

This formula shows that if $x$ is not too far from $a$, and thus $[a, x]$ has
small length, and $f^{(n+1)}(x)$ is small on this interval then the remainder term
is small, and we don’t make too much of an error replacing $f(x)$ by $T_n(x)$.
Lecture 25: Extreme Values

Take a look at:

- How to determine maximum and minimum values of a graph.
- Using differentiation to find maximum and minimum values.

Below you find the graphs of the velocity function \( v(t) \) (in ft/sec) and the corresponding acceleration function \( a(t) \) of a space telescope, from liftoff at \( t = 0 \) until \( t = 126 \) seconds.

In order to design the technical equipment it is very important to understand the maximum values of acceleration in that time period. Let’s consider the problem to determine the maximum and minimum values of acceleration. From the graph we can guess the minimum value at around 25 seconds of about 22 ft/sec\(^2\), the maximum value will be taken at the end of the time interval at 126 sec, and we can guess it will be about 63 ft/sec\(^2\).

Of course it is often important to get precise numerical values based on a model for the motion. Suppose we have given the cubic polynomial

\[
v(t) = 0.001302t^3 - 0.09029t^2 + 23.61t - 3.083
\]

as a model for the motion. Then we can calculate the derivative, which is the quadratic function

\[
a(t) = \frac{dv}{dt} = 0.003906t^2 - 0.18058t + 23.61.
\]

The minimum value is, obviously from the graph, taken at that point in time when the acceleration function has zero slope. We can find the time and the
value much more precisely using Calculus. The acceleration has a horizontal tangent when
\[ a'(t) = 0.007812 t - 0.18058 = 0, \]
from which we get \( t_1 = \frac{0.18058}{0.007812} \approx 23.12 \). In order to be sure that the lowest value of \( a(t) \) on the interval \([0, 126]\) is taken at \( t_1 \) we should check also the value at \( t = 0 \) because the graph might be misleading. We calculate
\[ a(0) = 23.61, \quad a(t_1) = 21.52, \quad a(126) \approx 62.869, \]
so the minimum acceleration on the interval \([0, 126]\) is \( a(t_1) \) and the maximum acceleration is \( a(126) \).

The above example is an application of the Closed Interval Method to find the maximum and minimum values of a continuous function on a closed interval \([a,b]\). To understand how to apply the method precisely we need the following:

**Definition:** (i) The function \( f(x) \) has an absolute minimum on an interval \( I \) at \( x = a \) in \( I \) if \( f(a) \leq f(x) \) for all \( x \) in \( I \). (ii) The function \( f(x) \) has a local minimum at \( x = c \) if \( f(c) \) is the minimum value of \( f(x) \) on some open interval containing \( c \).

Of course there are the corresponding definitions of absolute and local maximum.

Recall that an open interval is an interval of the form \((a,b)\), so consists of all numbers \( x \) with \( a < x < b \). A closed interval is an interval \([a,b]\) including the two endpoints. There is the

**Theorem:** A continuous function \( f(x) \) on a closed and bounded interval \( I = [a,b] \) takes on both minimum and maximum values on this interval.

If a function is not continuous there is not necessarily maximum/minimum. For example consider \( f(x) = \frac{1}{x-2} \) on \([1,3]\) (we may even define \( f(2) \) in an arbitrary way), then the function has no maximum or minimum values on the interval. If the interval is not closed like \((0,1)\) we may even have a continuous function like \( f(x) = \frac{1}{x} \), there is no maximum. This is why the way has nothing to do with the vertical asymptote at 0. If we consider the function \( f(x) = 2 - x \) for example \((0,2)\) it will also not take a maximum value. Note that \( f(x) < 2 \) for all \( x \) in \([0,2]\) but \( f(0) \) is not a value of the function on \((0,2)\). In this case \( 2 - x \) will though on the closed interval \([0,2]\), namely \( f(0) = 2 \).

The Closed Interval Method is based on the observation that for a continuous function on an interval the absolute minimum is either taken at the
boundary, or is a relative local (of course there could be many) in the interior of the interval. To find the local maxima and minima we introduce the definition of a critical number: Given a function $f(x)$ a number $c$ is a critical number if either (i) the function is not differentiable at $c$, or (ii) $f'(c) = 0$ (in which case we have a horizontal tangent). So $c = 0$ is a critical number of $\sqrt{x}$ because the derivative is not defined there. Also $c = 0$ is critical number of $f(x) = x^2$ because $f'(x) = 2x$. Note that $f(x)$ has a minimum there. But the example of $f(x) = x^3$ with $f'(x) = 3x^2$ shows that we can have a horizontal tangent without having a local maximum or minimum there. So care has to be taken when we apply:

**Fermat’s Theorem:** If $f(x)$ has a local maximum and minimum at $c$ then $c$ is a critical number.

Here is the Closed Interval Method in an nutshell: Find the critical numbers and calculate the values of $f(x)$ at those numbers, calculate $f(x)$ at the two endpoints of your interval. Then find the maximum and minimum value among the values you calculated.

Let’s look at a simple Example: Consider the function $f(x) = xe^{-x}$ on the interval $[1, 2]$:

![Graph of $f(x) = xe^{-x}$ on $[1, 2]$](image)

It looks like the maximum is taken close to $c = 1$ but we can check using the Closed Interval Method and calculate $f'(x) = e^{-x} - xe^{-x}$. So $f'(x) = 0$ if $e^{-x}(x - 1) = 0$, which only happens when $x - 1 = 0$ or $x = 1$ (note that
\( e^{-x} \neq 0 \) for all \( x \). Thus is only critical number is \( c = 1 \). Without knowing the graph we can calculate \( f(0) = 0, f(1) = \frac{1}{e} \approx 0.37 \) and \( f(2) = \frac{2}{e^2} \approx 0.27 \). Thus the maximum value is \( \frac{1}{e} \) and the minimum value is 0. We also say that the maximum value is \( f(1) = \frac{1}{e} \) also indicating the \( x \)-value where the maximum occurs.
Lecture 26: Mean value theorem and Monotonicity

Take a look at:

- What is the mean value theorem?
- Rolle’s theorem - a special case of the mean value theorem
- Find and classify critical points

It seems geometrically quite obvious observation that if a function \( f(x) \) satisfies \( f(a) = f(b) \) then there will be some \( c \) between \( a \) and \( b \) such that \( f'(c) = 0 \). This is called Rolle’s theorem and you see an example on the left:

There we can have \( f(-1) = f(1) = 0 \) and find that there is a \( c \) such that \( f'(c) = 0 \). The way to find this \( c \) is of course to solve the equation \( f'(x) = 0 \). Note that differentiability of the function is essential as the example on the right hand side shows. Obviously there is no horizontal tangent line for any \( c \) between 0 and 3. You may think proving Rolle’s theorem by moving the line joining \((a,f(a))\) and \((b,f(b))\) vertically until the two points on the graph vanish, well if they vanish at a differentiable point we have a line touching the graph only at this point at a local minimum or maximum, so there should a critical point. Well, there is.

The Mean Value Theorem is a more general statement but actually based on the same idea. There is Calculus Cop Story exemplifying the idea: You get stopped by a cop telling you that 15 minutes ago you passed a police camera and that was 15 miles away. You have been driving on a road with a 45 miles/hour speed limit. The cop argues that you were speeding and gives you a ticket. Would you start a discussion with him? What is his thinking here: In the 15 minutes you changed your position by 20 miles. So at some
point in time your speed was the average speed, which is $10/\frac{1}{4}$ miles/hr, so 60 miles per hour.

The general situation is that of a differentiable function $f(x)$ and considering the average rate of change of $f(x)$ over the interval $[a, b]$. The Mean Value Theorem states that there is a $c$ between $a$ and $b$ for which the slope is the average rate of change. A typical example is shown with the following graph with $a = -1$ and $b = 2$. The slope of the blue top line is the average rate of change of $f(x)$ over the interval $[-1, 2]$. This rate of change is easily calculated to be

$$\frac{f(2) - f(-1)}{2 - (-1)} = \frac{3 \cdot 4 - 0}{3} = 4.$$

The bottom green line is the tangent line at $c = -\frac{2}{3} + \frac{1}{3}\sqrt{19}$, which is a solution of the equation $f'(x) = 4$.

So in general, the Mean Value Theorem (short MVT) states that if a function is differentiable then there is a $c$ between $a$ and $b$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

To memorize: The average rate of change over some interval is equal to the instantaneous rate of change at some point in the interval. There is one very important consequence of this result. Suppose that $f'(x) = 0$ for all $x$ in some interval. Then it follows from the MVT that $f(b) = f(a)$ for the two endpoints. Since we could choose endpoints anywhere in the interval it follows that $f(x) = c$ is a constant on the interval. We already used this
quite obvious fact before when we were discussing that anti-derivatives of a function are determined by the function up to a constant.

We know that a function is increasing when the slope is positive and decreasing when the slope is negative. In order to make more precise mathematical statements we should define what we mean by increasing and decreasing. We will say that \( f(x) \) is increasing on an interval \((a, b)\) if \( f(x_1) < f(x_2) \) whenever \( x_1 < x_2 \) and \( x_1, x_2 \) are in \((a, b)\). Now suppose that \( f'(x) > 0 \) for all \( x \) in \((a, b)\). The MVT tells us that

\[
f(x_2) - f(x_1) = f'(c)(x_2 - x_1).
\]

So for \( x_2 - x_1 > 0 \) and because we know \( f'(c) > 0 \) we can deduce \( f(x_2) > f(x_1) \).

Of course a similar statement holds for decreasing. The sign of the derivative tells us whether the function is increasing or decreasing. So, in order to change from increasing to decreasing (or vice versa) the sign of the derivative has to change. If \( f'(x) \) is continuous itself, then the sign can only change at a value where the derivative is zero, so at a critical point. Of course if we change from increasing to decreasing there will be a local maximum, and if we change from decreasing to increasing there will be a local minimum. Thus what’s going on near critical points will tell us what kind of local extremum we have (if there is one). The following is always true.

**First Derivative Test:** If \( f(x) \) is differentiable close to \( c \) (but not necessarily at \( c \)), and \( c \) is a critical number for \( f(x) \). Then, if \( f'(x) \) changes sign from + to − at \( c \) we have that \( f(c) \) is a local maximum. If the sign changes from − to + we have a local minimum at \( c \).

Here is the basic example to have in mind: \( f(x) = x^2 \) has \( f'(x) = 2x \) and \( c = 0 \) is a critical number. At this point \( f'(x) \) changes sign from − to +. So \( f(x) \) changes from decreasing to increasing there, which means there is a local minimum. Well we knew that, but that’s why we should keep in mind precisely this situation.
Lecture 27: Concavity and the Shape of a Graph

Take a look at:

Concavity and inflection points on a graph and Understand concavity and inflection points with differentiation

Below on the left you see the graph of the function

\[ y = \frac{e^x}{e^x + 1} \]

over the interval \([-3, 5]\):

If we look at how the shape of the graph is changing something seems to be going on for \(x = 0\). But what is it? The function is increasing over the whole interval. Let’s look at the derivative function. The slopes are first getting larger, so the derivative function is increasing. Then at \(x = 0\) the slopes start to get smaller, so the derivative function is decreasing. What is happening is that at \(x = 0\) the derivative function (do the calculation using the quotient rule):

\[ y' = \frac{e^x}{(e^x + 1)^2} \]

has a local maximum there. Of course we can find this local maximum of the derivative function by taking the derivative again:

\[ y'' = \frac{e^x - e^{2x}}{(e^x + 1)^2} \]

and in fact the graph of \(y''\) looks as follows (graph on the left)
We say that the function \( y = f(x) \) is \textit{concave up} on the interval \([-3,0]\) and is \textit{concave down} on the interval \([0,5]\). You could consider this just as new language to say that the slope function is increasing on \([-3,0]\) and is decreasing on \([0,5]\). But the important point here is the interpretation in terms of the original function \( f(x) \). If \( f'(x) \) is increasing at some point \((c, f(c))\) on the graph then \( f(x) \) grows faster than its linearization at that point, the tangent line graph, because the latter one’s slope is not changing at all. Thus close to \( c \) the graph of \( f(x) \) will be \textit{above} the tangent line. You can see in the picture on the right above the graphs of \( y \) and the tangent lines at \( c = -1 \) and at \( c = 1 \). The graph of \( f(x) \) is curved \textit{upwards} on \([-3,0]\) but is curved downwards \([0,5]\). Note that the choice of interval \([-3,5]\) was quite arbitrary. Of course when we check on \( f'(x) \) we see that the function is increasing on \((\infty,0)\) and is decreasing on \((0, \infty)\). A point where concavity changes is called an \textit{inflection point} of the function \( f(x) \). So \((0, \frac{1}{2})\) is inflection point for our function \( f(x) \). We find the intervals of concavity by checking \( y'' \). If \( f''(x) > 0 \) then \( f'(x) \) is increasing and \( f(x) \) is concave upwards, similarly when \( f''(x) < 0 \) then \( f'(x) \) is decreasing and \( f(x) \) is concave downwards. Note that \( f''(x) = 0 \) means that \( f'(x) \) \textbf{may} change sign there but we do not have guarantee for this. Always have in mind the always concave up graph \( f(x) = x^4 \) which satisfies \( f''(0) = 0 \) but \((0, f(0))\) is \textbf{not} an inflection point. We have \( f'(0) = 0 \) but \( f'(x) \) is nevertheless increasing there. The \textit{thinking} works only \textbf{one-way} here. If \( f'(c) > 0 \) for some \( c \) then \( f(x) \) is increasing close to \( c \). But \( f'(c) = 0 \) does not \textbf{necessarily} mean that \( f'(x) \) changes sign there. So be careful!

If we are close to a local minimum \( f(c) \) of a function \( f(x) \) of course the
function is concave up close to $c$ (the derivative change from decreasing to increasing, the tangent line is horizontal at $c$ with the graph above it). Thus if $c$ is a critical number of the differentiable function $f(x)$ then $f''(c) > 0$ means that $f(c)$ is a local minimum. When $f''(c) < 0$ then $f(c)$ is a local minimum. But if $f''(c) = 0$ we don’t know. We say that the second derivative test is inconclusive.

When you think about concavity and the tests have in mind two important examples: $f(x) = x^3$ and $f(x) = x^4$ (do the examples!).

Just a remark in the end. The word concave describing a specific shape of course is related to the shape of lenses:

Take the concave lens and rotate by 90 degrees. Then the top is shaped concave up and the bottom is shaped concave down.
Lecture 28: Why Graph Sketching?
(or can’t my calculator do that much better)

Take a look at:

Data mining function properties from derivatives and Data mining identifying functions from derivative graphs

This section really is just summarizing what the derivatives and the asymptotic behavior can tell us about the graph of a function. It is not about the technical aspects of graphing. You are correct, your calculator will just evaluate the function at many points and interpolate intermediate values. But a function is telling a story. The local extrema, the inflection points, the intervals of monotonicity and concavity. Think about the position graph of a motion on a line: The local extrema are the turning points, the inflection points are where we switch between speeding up and slowing down. The intervals of monotonicity are intervals in which we move forward (increasing) or backward (decreasing). When the graph is concave upwards the motion is positively accelerated, we are speeding up in the sense that our velocity is increasing (so if our velocity is negative it will change towards the positive and its absolute value, called speed, will actually be decreasing). When the graph is concave down then the motion is negatively accelerated and our velocity is changing towards the negative. At an inflection point the acceleration is zero. Note that this does not necessarily mean that the acceleration changes at each of its zeroes. So we can have acceleration zero without having an inflection point. In this sense graphing functions is like kinematics. The force is giving us the acceleration: then we figure out the basic features of the motion.

Let’s consider a motion of an object of 1 kg along the $y$-axis under a force

$$F(t) = 60t^3 - 30t$$

Newtons with initial velocity $v(0) = 0$ and initial position $y(0) = 3$. By Newton’s law we know that the acceleration is

$$a(t) = 60t^3 - 30t$$

meter/sec$^2$.

This of course is a very non-constant acceleration motion. Of course we know how to find the velocity and position functions by integration:

$$v(t) = 15t^4 - 15t^2$$

meter/sec
and

\[ s(t) = 3t^5 - 5t^3 + 3 \text{ meter}. \]

Then we may go ahead and graph the position function to understand the basic features of the motion. We may choose a standard window \([-10, 10]\) and get the graph on the left. It looks like nothing interesting is happening in the time interval \([-2, 2]\). But is this really the case? Well if we are careful we zoom in and see that there is actually quite some action on the interval \([-2, 2]\). Well Calculus will tell us the story. We know that the turning points are the local extrema: Now \( v(t) = 15t^2(t^2 - 1) = 0 \) at \( t = -1, 0, 1 \). Since \( v(t) \) is not changing sign at \( t = 0 \) there is no local extremum. But at \( t = \pm 1 \) there is a change of sign of the slope function. We can easily get a sign chart (by for example plugging in values \(-2, 0, 2\)) for the intervals of monotonicity: \( y(t) \) is increasing on \( (-\infty, -1) \cup (1, \infty) \) but is decreasing on \( (-1, 1) \). Note that it is impossible to see that from the left-hand graph (well eagle eyes might see it!). The first derivative test tells us that there is a local maximum \( y(-1) = 5 \) and a local minimum \( y(1) = 1 \). Well, the intervals of concavity are the intervals where the force is positive/negative. The acceleration is

\[ a(t) = 30t(1 - 2t^2), \]

which vanishes at \( t = 0 \) and at \( t = 0 \) and at \( t = \pm \sqrt{2} \). Since these are simple roots there are sign changes in all three places. So from \( a(-1) = -90 \) we get that \( y(t) \) is concave downwards on \( (-\infty, -\sqrt{2}) \cup (0, \sqrt{2}) \) and concave downwards on \( (\sqrt{2}/2, 0) \cup (\sqrt{2}, \infty) \). Note that these are just math words for saying
that the force is in the negative $y$-direction in times $(-\infty, -\frac{\sqrt{2}}{2}) \cup (0, \frac{\sqrt{2}}{2})$ and the force is in the positive $y$-direction in times $(0, \frac{\sqrt{2}}{2}) \cup (\frac{\sqrt{2}}{2}, \infty)$. Walk along the graph in time and feel the force. Then you understand concavity.
Lecture 29: Area and Accumulation Functions

Take a look at: Fundamental Theorem of Calculus Part II, The Fundamental Theorem, Part II and The FTC and the Chain Rule.

Recall our very definition of the symbol:

\[ \int_a^x f(t) \, dt, \]

which is the net area of the function \( f(t) \) over the interval \([a, x]\) for \( x > a \). For \( x < a \) we have defined

\[ \int_a^x f(t) \, dt = -\int_x^a f(t) \, dt, \]

so we consider the net area over \([a, x]\), but with a minus sign.

Imagine that \( x \) is moving along the \( t \)-axis. The net area will change correspondingly. So we have a function of \( x \), let’s denote it

\[ A(x) = \int_a^x f(t) \, dt, \]

and call it an area or accumulation function for the function \( f(x) \). Note that there is an in general different area function for each value of \( a \) but by the interval additivity properties of the integral the area functions will only differ by a constant.

Let’s try to understand how \( A(x) \) is changing when \( x \) is moving away from \( a \) in the positive direction along the \( t \)-axis. We are accumulating net area. As long \( f(t) > 0 \) we are accumulating the area with a + sign, and so \( A(x) \) is increasing. When \( f(t) < 0 \) then we are accumulating the area with a – sign, we are subtracting from what we have accumulated up to this point, and \( A(x) \) is decreasing.

Note that the Fundamental Theorem of Calculus is not the definition of the symbol! It tells us how to calculate it. So we know that

\[ F(x) - F(a) = A(x) \]

where \( F(x) \) is anti-derivative of \( f(t) \): \( F'(x) = f(x) \). \( F(a) \) is a constant, so \( F(x) \) is increasing or decreasing precisely when \( A(x) \) is doing the same thing. (In fact, \( A(x) \) is the anti-derivative satisfying \( A(a) = 0 \).) This makes
sense: We know that \( F(x) \) is increasing/decreasing when \( F'(x) = f(x) \) is positive/negative. The Fundamental Theorem of Calculus is compatible with the area/accumulation interpretation.

What we are doing here is an example of rechecking our concepts. This is important. If you do not practice this you will not have understood things conceptually. Thinking about accumulation functions as above also is not new at all. When the function \( f(t) \) is a velocity function \( v(t) \) and \( s_0 \) is a number then

\[
 s_0 + \int_{t_0}^{t} v(u)du
\]

is the position function \( s(t) \) with \( s(t_0) = s_0 \). The position function is just an accumulation function for the velocity function. It accumulates the infinitesimal changes in position described by the velocity function. Similarly the velocity function is an accumulation function of the acceleration function. We talk about area functions when we want to emphasize the geometric meaning of the numbers with respect to the graphs of the integrands.

Since \( A(x) = \int_{a}^{x} f(t)dt \) is an anti-derivative of \( f(x) \) we know that

\[
 A'(x) = \frac{d}{dx} \int_{a}^{x} f(t)dt = f(x)
\]

Again this is not new and we essentially derived in CCC3. It is often called the Fundamental Theorem of Calculus II. It says that the derivative of an area or accumulation function of a function \( f(x) \) gives you back \( f(x) \). It emphasizes the fact that differentiation and integration are inverse processes.

According to the above we understand that the monotonicity (increasing/decreasing) behavior of an area function \( A(x) \) is just corresponding to the sign of the function \( f(x) \). Let’s discuss the shape of area functions with respect to the second derivative: What is the meaning of \( A''(x) \) in terms of \( f(x) \) and accumulating the area. Of course we know that \( A''(x) = f'(x) \) by the Fundamental Theorem. Thus when \( f(x) \) is increasing then \( A(x) \) is concave upwards. This means that the slope of \( A(x) \) is increasing, or the rate of change of the area is speeding up when moving along the x-axis with constant speed. Does this make sense? Well \( A''(x) \) is looking at the rate of change with respect to \( x \) of things like \( \frac{\Delta A}{\Delta x} \) for small \( \Delta x \). Let’s consider some point \( c \) on the x-axis where \( f'(c) = 0 \), and assume we are at a local maximum of \( f(x) \) and \( f(c) > 0 \). Consider the intervals \([c-\Delta x, c]\) and \([c, c+\Delta x]\). The values \( \frac{\Delta A}{\Delta x} \) are about the same, so there is no change in the rates of change, \( A''(c) = 0 \).
Now consider the same picture but assume that \( f'(c) > 0 \). Then the value of \( \frac{\Delta A}{\Delta x} \) for the interval \([c, c + \Delta x]\) will be larger than the corresponding value for \([c - \Delta x, c]\). The quantities \( \frac{\Delta A}{\Delta x} \) are increasing, or \( A''(c) > 0 \).

Let’s consider an example: \( f(x) = \sin(x) \). Then \( A(x) = -\cos(x) \) is the area function for \( A(0) = -1 \). See the graph on the left:

You can see from the graph: Where \( f(x) \) has positive slope the area function \( A(x) \) is concave upwards. Similarly, where the function \( f(x) \) has negative slope the area function \( A(x) \) is concave downwards. The local extrema of \( f(x) \) are precisely for those \( x \)-values where \( A(x) \) has inflection points. Let’s look at another example. Maybe the above is coincidence? If we have \( f(x) = 2x^3 - 3x^2 - 12x \) then \( A(x) = \frac{1}{2}x^4 - x^3 - 6x^2 \) is an area function. These two functions are graphed on the right above. Do you see it: When the cubic function is increasing the graph of fourth degree polynomial is concave upwards. Inflection points of the area function correspond to local extrema of the function.

Are our observations above new. Of course not. The accumulation function of a velocity function is a corresponding position function. When the velocity has a local extremum that’s where the position function has an inflection point. When you are driving your car, the inflection points of your position function will be the moments in time when you switch from the gas pedal to the brake pedal. Just letting go the gas pedal for a moment is not sufficient. Then the acceleration becomes zero but does not change sign! Then you just go with constant velocity for a moment before you speed up again, you never switch to slowing down. No inflection point here.
Lecture 30: Optimization

Take a look at:
- Optimization and differentiation
- Optimizing simple systems

We have already been discussing the calculus of optimization problems in CCC 25. There a function was given and our task has been to find the absolute maximum and minimum values on some interval, open or closed. In this section we take a closer look at several application problems leading to those situations. Note that in order to apply our calculus methods we need to find a domain and an objective function for which we want to find an extreme value. Let’s start from very simple examples.

**Example 1** Find the points on the hyperbola
\[ y^2 - x^2 = 4 \]

that are closest to the point \((2, 0)\).

The text contains the key word closest, which is a distance that we want to be minimal. This is the objective function. In order to find the points of minimal distance we need to determine their coordinates. But let’s look at the graph of the hyperbola and the point to see whether our question makes sense.

A look at the graph below should convince us that because of symmetry there should be two answers of the form \((x_0, \pm \sqrt{4 + x_0^2})\). But how do we find the precise answer for \(x_0\)? If we zoom in it looks like it could be \(x_0 = 1\) or maybe
the point $x_0 = \frac{5}{4}$? How can we be sure. Let’s compute the distance from a general point $(x, y)$ on the hyperbola to the point $(2, 0)$. We get using the formula for the distance between two points (this is just the Pythagorean theorem):

$$d = \sqrt{(x - 2)^2 + y^2},$$

and since for all points on the hyperbola the equation $y^2 - x^2 = 4$ is satisfied we know $y^2 = x^2 + 4$ and thus

$$d(x) = \sqrt{(x - 2)^2 + x^2 + 4}.$$

We may now use this as objective function. But is is more practical to work with the objective function

$$f(x) = d(x)^2 = (x - 2)^2 + x^2 + 4.$$

This is because we avoid having to take the derivative of the root. Make clear to yourself if you work with the original function and apply the chain rule you will arrive at the same answer. It is clear from the fact that $\sqrt{x}$ is increasing that extrema of $\sqrt{g(x)}$ and $g(x)$ are the same for functions with $g(x) \geq 0$. What is the domain of the objective function? Obviously $(-\infty, \infty)$.

Note that calculus is not necessary to find the minimum in this case because a parabola has only one extreme value, namely at its vertex. But let’s do the calculation to get a feeling for the procedure. We take the derivative:

$$f'(x) = 2(x - 2) + 2x.$$

So we find that the answer is

$$4x_0 = 4 \quad \text{so} \quad x_0 = 1 \quad \text{and} \quad \left(1, \pm \sqrt{5}\right)$$

is the solution. Now it is clear from the theorem of Pythagoras that the point on the hyperbola $(x_0, y_0)$ of minimal distance should be the one where the line through $(2, 0)$ and $(x_0, y_0)$ is perpendicular to the tangent line. What is the equation from this condition? The line through $(2, 0)$ and $(x_0, y_0)$ has the equation

$$y - y_0 = \frac{y_0}{x_0 - 2}(x - x_0)$$

and the tangent line at $(x_0, y_0)$ has the equation:

$$y - y_0 = \frac{x_0}{y_0}(x - x_0)$$

89
(implicit differentiation). From the formula for the slopes of perpendicular lines it follows that

\[- \frac{y_0}{x_0} = \frac{y_0}{x_0 - 2}\]

or \(x_0 - 2 = -x_0\) thus \(x_0 = 1\).

**Example 2:** Find the area of a largest rectangle that can be inscribed in a right triangle with legs of lengths 3 cm and 4 cm if two sides of the rectangle lie along the legs.

The inscribed rectangle is determined by having one of its corners on the hypotenuse.

By placing the right triangle in a coordinate system as above the point is on the line \(y = 4 - \frac{4}{3}x\). The area of the corresponding rectangle is

\[A = xy,\]

but we know that \(y = 4 - \frac{4}{3}x\), and thus our objective function is

\[A(x) = x(4 - \frac{4}{3}x) = 4x - \frac{4}{3}x^2.\]

The domain for \(x\) is \(0 \leq x \leq 3\). We calculate

\[A'(x) = 4 - \frac{8}{3}x,\]
which is 0 when \( x = \frac{12}{8} = \frac{3}{2} \). From \( A(0) = A(3) = 0 \) and \( A(x) > 0 \) for \( x \) in \([0,3]\) we know that this will determine the absolute maximum of the objective function. The corresponding area is

\[
A\left(\frac{3}{2}\right) = \frac{3}{2} \cdot \left(2 - \frac{4}{3} \cdot \frac{3}{2}\right) = \frac{3}{2} \cdot 2 = 3.
\]

So we get the answer: *The maximum area of an inscribed rectangle is \( A = 3 \). It is given when the side of the rectangle (on the leg of the triangle of length \( 3 \)) has length \( \frac{3}{2} \) and the other side has length 2.*

Here are suggested steps in solving an optimization problem.

- Read the problem in detail, possibly several times. Determine: What is the quantity \( Q \) you want to extremize? What are the quantities changing, your variables? What quantities are fixed.
- Draw a diagram showing the variables.
- Introduce notation, so assign a symbol for each variable.
- Express the quantity \( Q \) in terms of the variables and fixed data. This is our objective function.
- If more than one variable is involved in \( Q \) use the geometry of your diagram to eliminate variables in the most efficient way, so express \( Q \) as function of a single variable. Find the domain of this function with respect to the variable (this is often not the largest interval for which the function \( Q \) is defined. It is the largest domain for which the quantity \( Q \) makes sense. If e. g. \( Q \) is a length then we want \( Q \geq 0 \).)
- Apply the calculus methods from CCC 25 to determine absolute maximum or minimum of \( Q \) on the interval.
- In a sentence formulate the answer to the problem.
Lecture 31: More Optimization

Take a look at:

**Optimizing complex systems**

We will discuss a more difficult problem in detail with two quite different solutions. The solutions will be based on **different independent** variables. They both lead to the correct solution. But the routes taken and techniques needed are quite different. It is supposed to show you how versatile calculus techniques can be.

Find the minimum length $\ell$ of a beam that can clear a fence of height $h$ and touch a wall located $b$ ft behind the fence (see picture on back side). *(The answer for this problem is given in our text book and is)*

$$(b^{2/3} + h^{2/3})^{3/2} \text{ ft.}$$

Solution 1 (trigonometry free): Using the notation from the picture
by the theorem of Pythagoras we get for the length $\ell = \ell_1 + \ell_2$ of the ladder

$$\ell^2 = (\ell_1 + \ell_2)^2 = (b + x)^2 + (h + y)^2.$$  

From similarity of triangles we know $\frac{y}{b} = \frac{x}{x}$ and so we get

$$f(x) = \ell^2(x) = (b + x)^2 + (h + \frac{bh}{x})^2.$$  

Minimizing $\ell^2$ will also minimize $\ell$. The interval we consider is $(0, \infty)$. Note that

$$\lim_{x \to 0} f(x) = \lim_{x \to \infty} f(x) = \infty.$$  

Thus a critical number in $(0, \infty)$ will tell us the absolute minimum because $f(x) = \ell^2 \geq 0$. From

$$f'(x) = 2(b + x) + 2(h + \frac{bh}{x})(-\frac{bh}{x^2}) = -\frac{2h^2b}{x^2} \cdot \frac{b + x}{x} + 2(b + x) =$$

$$= 2(b + x)(\frac{bh^2}{x^3} + 1) = 0$$

we get $x^3 = bh^2$ or $x = (bh^2)^{1/3}$. If we plug this back into the formula for $\ell^2$ we get for the minimum length $\ell^2 = (b + (bh^2)^{1/3})^2 + (h + \frac{bh}{(bh^2)^{1/3}})^2 = (b^{1/3}h^{2/3} + b^{1/3}h^{2/3})^2 + (h^{1/3}h^{2/3} + b^{2/3}h^{1/3})^2 = h^{2/3}(b^{2/3} + h^{2/3})^2 + h^{2/3}(b^{2/3} + h^{2/3})^2 = (b^{2/3} + h^{2/3})(b^{2/3} + h^{2/3})^2 = (b^{2/3} + h^{2/3})^3$. Thus

$$[\ell = (b^{2/3} + h^{2/3})^{3/2}]$$

is the minimal length.

**Solution 2 (using trigonometry):** Let $\theta$ be the angle between the ladder and the ground. The fence cuts the length of the ladder $\ell$ into the two parts of lengths $\ell_1$ and $\ell_2$ (see the picture above). Using the two triangles we get

$$\ell(\theta) = h \csc(\theta) + b \sec(\theta).$$

The interval of interest is $(0, \frac{\pi}{2})$, where $\theta \to 0$ obviously corresponds to an almost horizontal ladder, while $\theta \to \frac{\pi}{2}$ corresponds to an almost vertical ladder. Obviously in both cases the length becomes infinite, which we also can find from

$$\lim_{\theta \to 0^+} \ell(\theta) = \lim_{\theta \to \frac{\pi}{2}^-} \ell(\theta) = +\infty$$

93
resulting from
\[
\lim_{\theta \to 0^+} \csc(\theta) = \lim_{\theta \to \frac{\pi}{2}^-} \sec(\theta) = +\infty
\]
and \( h, b > 0 \). Moreover since \( \sin(\theta) > 0 \) and \( \cos(\theta) > 0 \) on \((0, \frac{\pi}{2})\) a local extremum in this interval has to be the absolute minimum. It follows that

\[
\ell' = -h \csc(\theta) \cot(\theta) + b \sec(\theta) \tan(\theta) = -h \frac{\cos(\theta)}{\sin^2(\theta)} + b \frac{\sin(\theta)}{\cos^2(\theta)}
\]

Thus \( \ell' = 0 \) if an only if

\[
\tan^3(\theta) = \frac{h}{b}
\]
or

\[
\theta = \tan^{-1}(x),
\]

where we denote \( x = \left(\frac{h}{b}\right)^{1/2} \). Using a suitable triangle, compare 1.4 and 3.8 from the textbook (use a triangle with , we get

\[
csc(\tan^{-1}(x)) = \frac{\sqrt{1 + x^2}}{x}, \quad \sec(\tan^{-1}(x)) = \sqrt{1 + x^2}.
\]

Thus we get

\[
\ell = h \sqrt{1 + \left(\frac{h}{b}\right)^{2/3}} + b \sqrt{1 + \left(\frac{h}{b}\right)^{2/3}},
\]

which can be calculated to be

\[
h^{2/3} \sqrt{b^{2/3} + h^{2/3}} + b^{2/3} \sqrt{b^{2/3} + h^{2/3}} = (b^{2/3} + h^{2/3})^{3/2}
\]

See the graph for \( h = 2 \) and \( b = 1 \) with \( y = \ell(\theta) \), in which case the solution for the minimal length is is for \( \theta \approx 51.2^\circ \) or \( \approx 0.9 \) rad with \( \ell \approx 4.16 \).
Lecture 32: Examples and Applications

Take a look at: Exponential growth and decay and Exponential growth Calculus version and Displacement and total distance covered

In solving linear motion problems it is important to distinguish between the change in position and distance traveled. If we have given a velocity function on a time interval \([t_0, t_1]\) for a motion along the \(y\)-axis then the change in position is

\[
\Delta y = y(t_2) - y(t_1) = \int_{t_0}^{t_1} v(t) dt.
\]

Of course this quantity can be positive or negative. If \(y(0) = 0\) and the object is moving down the \(y\)-axis on the interval then \(\Delta y\) will be negative. Often we are much more interested in the distance traveled than the change in position. If you drive from Boise to Twin Falls and back to Boise the change in position could be 0 but of course you covered quite some distance, which will be important for the calculation of fuel consumption. The distance traveled will be

\[
\int_{t_1}^{t_2} |v(t)| dt.
\]

The quantity \(|v(t)|\) is the speed and is always positive. So in terms of graphs while displacement is given by net area, the distance traveled is given by actual area.

**Example:** A cyclist is pedals for \(0 \leq t \leq 3\) hours along a straight road with velocity

\[
v(t) = 2t^2 - 8t + 6 \ \text{mi/hr}
\]

The graphs of \(v(t)\) and \(|v(t)|\) are given on the left and right side below. We calculate for the displacement during those three hours:

\[
\Delta y = \int_0^3 2t^2 - 8t + 6 \, dt = \frac{2}{3}t^3 - 4t^2 + 6t \bigg|_0^3 = 0.
\]

If we compare the net area we see that this result looks reasonable. In order to find the distance covered we have to calculate \(|v(t)|\), which is just \(v(t)\) at times of non-negative velocity, and is \(-v(t)\) at times of negative velocity. We need to find the points of possible sign change in \([0, 3]\) and factor

\[
2t^2 - 8t + 6 = 2(t^2 - 4t + 3) = 2(t - 1)(t - 3) = 0
\]
if $t = 1$ or $t = 3$. Because we have an upwards parabola the sign will change from positive to negative at $t = 1$.

So we get

$$v(t) = \begin{cases} 2t^2 - 8t + 6 & 0 \leq t \leq 1 \\ -(2t^2 - 8t + 6) & 1 < t \leq 3 \end{cases}$$

We calculate

$$\int_0^1 2t^2 - 8t + 6dt = \left. \frac{2}{3} t^3 - 4t^2 + 6t \right|_0^1 = \frac{2}{3} - 4 + 6 = \frac{4}{3}.$$ 

Of course we can conclude at this point that the distance covered is $\frac{8}{3}$ because we know that the net area of $v(t)$ over $[1, 3]$ is $-\frac{4}{3}$. But we can easily confirm this:

$$\int_1^3 -2t^2 + 8t - 6dt = -\left. \frac{2}{3} t^3 + 4t^2 - 6t \right|_1^3 = \frac{4}{3}.$$ 

Now let’s study motion problems given in a slightly different way. Suppose that the acceleration of a moving object is given by $a(t) = -kv(t)$, where $k$ is a positive constant and $v$ is the object’s velocity. Assume that the initial velocity is $v(0) = 10$ and the initial position is $s(0) = 0$. We want to find the position of the object as a function of time and it’s velocity as a function of position. First note that we can use $a(t) = v'(t)$ to write

$$\frac{dv}{dt} = -kv$$
This is a first example of a **Differential Equation**. In words the equation says that the rate of change of the velocity is proportional to the velocity itself. So if $v(t)$ is large in value then the velocity will decay fast. Then while the velocity is decreasing the rate with which the decrease happens will also decrease. How do we find the solution? Let’s rewrite the equation using differentials:

$$\frac{dv}{v} = -k dt$$

and then integrate both sides:

$$\int \frac{dv}{v} = - \int k dt$$

Now from our knowledge of anti-derivatives (recall that $\frac{d}{dx} \ln |x| = \frac{1}{x} + C$) we get

$$\ln |v| = -kt + C,$$

where $C$ is a constant. By applying the exponential function on both sides we get

$$|v| = e^{-kt+C} = e^{-kt} e^C.$$

Note that $e^C$ is a positive constant, and the equation only gives us $|v|$. But there is a way out. Note that $|v| = \pm v$ depending on the actual sign. So we get the solution:

$$v(t) = Ae^{-kt},$$

where we allow *any* constants $A$ (corresponding $A = \pm e^C$, note that $e^{-kt}$ is always positive.) In our case we know $v(0) = 10$ so $v(0) = Ae^{-k\cdot0} = A = 10$ and get $A = 10$. Thus the equation for velocity is

$$v(t) = 10e^{-kt}$$

From this we get the position function by integration:

$$s(t) = \int 10e^{-kt} dt + C = -\frac{10}{k} e^{-kt} + C,$$

and the initial condition $s(0) = 0$ gives

$$0 = s(0) = -\frac{10}{k} e^{-k\cdot0} + C = -\frac{10}{k} + C,$$

98
and thus $C = \frac{10}{k}$. Thus the answer for the position function is

$$s(t) = \frac{10}{k}(1 - e^{-kt})$$

For $k = 1$ we get the following graph on the left:

Finally we were interested in knowing how the velocity is changing as a function of the position. We can use the chain rule to get the answer. Note that:

$$\frac{dv}{dt} = \frac{dv}{ds} \cdot \frac{ds}{dt}.$$ 

We already know $\frac{dv}{dt} = -kv(t)$ and so:

$$-kv = \frac{dv}{ds} \cdot v.$$ 

After canceling $v$ on both sides we get

$$\frac{dv}{ds} = -k$$

and so

$$v = -ks + C.$$ 

At time 0 we have position 0 and velocity 10 and so $v(0) = 10$. Note that the variable in the velocity function is now a position! So $10 = C$ and

$$v = 10 - ks$$
is the answer for the velocity as a function of position. The graph is above on the right. But what about the negative velocities for $s > 10$? Well the graph on the left or the calculation $\lim_{t \to \infty} 10(1 - e^{-t})$ show that we approach the position 10 but actually never reach it. Correspondingly the velocities are approaching the values zero, we are slowing down and keeping on doing this, but we never actually stop and reach velocity 0.
Literature


More Videos and online help:

- [10 things I hate about Calculus](#)
- [Calculus Rap](#)
- [Calculus](#)
- [Santa Rosa Calculus](#)
- [Khan Academy](#)