The chain rule

In Leibniz's notation that features **variables** that are related by functions behind the scenes, the chain rule says that if y depends on u and u depends of x, then

$$\frac{dy}{dx} = \frac{dy}{du}\,\frac{du}{dx}$$

This looks right, but it hides the fact that the derivative of y with respect to u needs to be expressed in terms of x when it is part of the derivative of y with respect to x.

The chain rule for functions

If $h = f \circ g$, and $y = h(x) = f(u)$ with $u = g(x)$, then

$$h'(x) = f'\big(g(x)\big)g'(x).$$

You should probably think of this using the notation of variables with the understanding that every variable has a fixed meaning in terms of a single **independent variable** that is x in this discussion.

Proving the chain rule

The chain rule can **almost** be proved just by writing dy/dx as a limit of $\Delta y/\Delta x$ and sticking $(1/\Delta u)\Delta u$ after the first factor. Unfortunately, this doesn't work if Δu can be zero. This can be patched, but it makes the proof more technical.

A better approach, though still technical, is to make the **tangent line** the main consideration, and to change the definition of the derivative to be the slope of the tangent line.

Defining the tangent line

The tangent line to the graph of $y = f(x)$ at the point where $x = a$ is a line with equation $y = A + M(x - a)$ where, for all $\epsilon > 0$, the **vertical distance**

$$|f(x) - A - M(x - a)| < \epsilon \, |x - a|$$

as long as x is **close enough** to a. It is easy to see that there is **at most** one line satisfying this definition, and that $A = f(a)$ and $M = f'(a)$. Since this avoids fractions, there is no danger of dividing by zero, and the proof of the chain rule is direct.

Tangent lines are often very good approximations

If the graph of $y = f(x)$ is **smooth**, then the **vertical distance** between the curve and the tangent line is bounded by a fixed multiple of $|x - a|^2$ on some interval around $x = a$. This will be made precise later (probably in the second semester).

Using the chain rule

To use the chain rule, all you need is the ability to recognize compositions of functions, and to write a correct expression for the derivative based on the composition that you have. In practice, either of the functions being composed may require further rules of calculus — including the chain rule — in the computation of the derivative.

In writing an expression for dy/dx, the steps from x to y are usually unraveled starting with the **last** step.

Exercises 3.5

#1 $y = \sin 4x$

#3 $y = (1 - x^2)^{10}$

#4 $y = \tan(\sin x)$

#11 $y = 1/(t^4 + 1)^3$

#17 $y = (1 + 4x)^5(3 + x - x^2)^8$

#35 $y = \sec^2 x + \tan^2 x$

#39 $y = \sqrt{x + \sqrt{x}}$

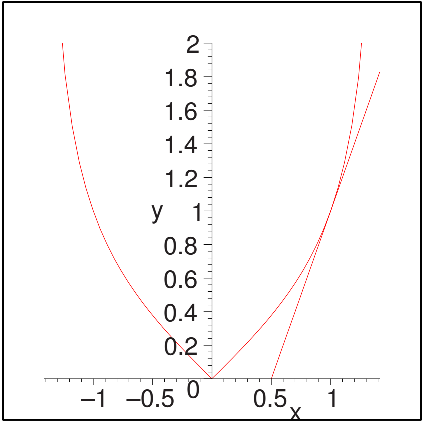
Exercise 3.5 #48

The next slide shows a picture of the **bulletnose curve**

$$y = \frac{|x|}{\sqrt{2 - x^2}}$$

and its tangent line at (1, 1).

The graph



Implicit differentiation

Suppose we want to find the tangent to the circle

$$x^2 + y^2 = 1.$$

We don't have y as a function of x, although it is not difficult **in this case** to obtain **two** such functions: one for the upper semicircle and one for the lower semicircle. However, these functions are much more complicated than the given equation for the whole circle. It would be nice if we could use that equation directly.

Variables are better than functions

If we have any way to get dy/dx when $x^2 + y^2$ simplifies to the constant function 1, then

$$\frac{d}{dx}\left(x^2 + y^2\right) = \frac{d}{dx}\left(1\right)$$
$$2x + 2y\frac{dy}{dx} = 0.$$

This equation can be solved for dy/dx

Very much better!

The solution is

$$\frac{dy}{dx} = -\frac{x}{y}.$$

This is a **direct** expression of the fact that the tangent line at a point of a circle is perpendicular to the radius to that point. The coordinates of the point can be used in this formula without first finding a suitable function.

Better yet!!

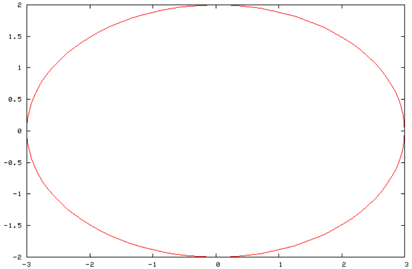
If you have a curve defined **any** equation of the form $F(x, y) = 0$, the process of **assuming** y to be a function of x and differentiating the given **identity** $F(x, y) = 0$ will lead to a **linear** equation for dy/dx in terms of x and y. Solving this is **easy**. At any point P where the solution makes sense, this gives a tangent line that is not vertical.

The implicit function theorem

The tangent line at P is a good enough approximation to the curve that it can be used to start a process that leads to y as a function of x on a small interval around P satisfying $F(x, y) = 0$. The general definition of "function" allows the existence of functions to be proved where they are needed — even if no explicit formula for the function is known. The existence of such functions is known as **the implicit function theorem**.

Exercise 3.6 #2

$$4x^2 + 9y^2 = 36$$



Exercise 3.6 #11

$$x^2 + y^2 + x \sin y = 4$$