Example 2. $f(x) = x^n$ where $n = 1, 2, 3...$

In this example we answer the question "What is $\frac{d}{dx}x^n$?" Once we know the answer we can use it to, for example, find the derivative of $f(x) = x^4$ by replacing n by 4.

At this point in our studies, we only know one tool for finding derivatives – the difference quotient. So we plug $y = f(x)$ into the definition of the difference quotient:

$$\frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = \frac{(x_0 + \Delta x)^n - x_0^n}{\Delta x}$$

Because writing little zeros under all our x's is a nuisance and a waste of chalk (or of photons?), and because there's no other variable named x to get confused with, from here on we'll replace x_0 with x.

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^n - x^n}{\Delta x}$$

Remember that when we use the difference quotient, we're thinking of x as fixed and of Δx as getting closer to zero. We want to simplify this fraction so that we can plug in 0 for Δx without any danger of dividing by zero. To do this we must expand the expression $(x + \Delta x)^n$.

A famous formula called the binomial theorem tells us that:

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times}$$

We can rewrite this as

$$x^n + n(\Delta x)x^{n-1} + O\left((\Delta x)^2\right)$$

where $O(\Delta x)^2$ is shorthand for "all of the terms with $(\Delta x)^2$, $(\Delta x)^3$, and so on up to $(\Delta x)^n$."

One way to begin to understand this is to think about multiplying all the x's together from

$$(x + \Delta x)^n = (x + \Delta x)(x + \Delta x)...(x + \Delta x) \quad \text{n times.}$$

There are n of these x's, so multiplying them together gives you one term of x^n. What if you only multiply together $n - 1$ of the x's? Then you have one $(x+\Delta x)$ left that you haven't taken an x from, and you can multiply your x^{n-1} by Δx. (If you multiplied by x, you'd just have the x^n that you already got.) There were n different Δx's that you could have chosen to use, so you can get this result n different ways. That's where the $n(\Delta x)x^{n-1}$ comes from.

We could keep going, and figure out how many different ways there are to multiply $n - 2$ x's by two Δx's, and so on, but it turns out we don't need to. Every other way of multiplying together one thing from each $(x + \Delta x)$ gives you at least two Δx's, and $\Delta x \cdot \Delta x$ is going to be too small to matter to us as $\Delta x \to 0$.

Now that we have some idea of what $(x + \Delta x)^n$ is, let's go back to our difference quotient.

$$\frac{\Delta y}{\Delta x} = \frac{(x + \Delta x)^n - x^n}{\Delta x} = \frac{(x^n + n(\Delta x)(x^{n-1}) + O(\Delta x)^2) - x^n}{\Delta x} = nx^{n-1} + O(\Delta x)$$

As it turns out, we *can* simplify the quotient by canceling a Δx in all of the terms in the numerator. When we divide a term that contains Δx^2 by Δx, the Δx^2 becomes Δx and so our $O(\Delta x^2)$ becomes $O(\Delta x)$.

When we take the limit as x approaches 0 we get:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = nx^{n-1}$$

and therefore,

$$\frac{d}{dx} x^n = nx^{n-1}$$

This result is sometimes called the "power rule". We will use it often to find derivatives of polynomials; for example,

$$\frac{d}{dx}(x^2 + 3x^{10}) = 2x + 30x^9$$