The Derivative of x^n is nx^{n-1}

Calculus 11, Veritas Prep.

In all of the slopes we've drawn and computed so far, you've probably noticed a pattern. Visually, the slope of a polynomial appears to be one degree less than the polynomial itself. Algebraically, you've seen that:

$$\frac{d}{dx}(x^2) = 2x \qquad \frac{d}{dx}(x^3) = 3x^2 \qquad \text{etc.}$$

Is this pattern more than just a coincidence? Is it true, in the general case, that $\frac{d}{dx}(x^n) = nx^{n-1}$, for any value of n?

Let's find out. In principle, we don't have to do anything particularly different than what we did to differentiate x^2 and x^3. All we need to do is take x^n, plug it into Fermat's difference quotient, fiddle a bit, send h to 0, and voila! we should get nx^{n-1}:

$$\frac{(x+h)^n - x^n}{h}$$

In practice, of course, we have the substantial complication that our exponent is not 2 or 3 or some definite number. Our exponent is n. But how do we simplify this? When we had $(x+h)^2$, we were able to multiply it out; when we had $(x+h)^3$, we were able to multiply that out, too. But what if I want to multiply $(x+h)$ by itself not once or twice but n times? Clearly, there is a pattern, but what is it?

You might recall that we discussed this earlier in the year. We can distill that pattern and write it as an iterated sum (the things with the giant Σ), in a formula known as the *binomial theorem*:

$$\begin{aligned}
(a+b)^0 &= 1 \\
(a+b)^1 &= a+b \\
(a+b)^2 &= a^2 + 2ab + b^2 \\
(a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3 \\
&\vdots \\
(a+b)^n &= \sum_{k=0}^{k=n} \binom{n}{n-k} a^{n-k} b^k
\end{aligned}$$

Now, this looks very scary. There's a giant Greek letter in the middle of the page. But it's not scary. All we're going to do in this proof is do *the same thing we did with x^2 and x^3*. The only difference is that we're going to use a slightly fancier tool—a giant Σ rather than normal arithmetic.

It is the difference between using a hammer to crush a small stone, and using a jackhammer to tear apart asphalt. The jackhammer looks huge and expensive and scary, and it runs on gasoline, not adenosine triphosphate, and you have to borrow it from your cousin in the demolition business. But in principle, it is no different than the hammer. It operates by exactly the same mechanism: hitting things with a blunt object. It is the same with the binomial theorem: it looks different, but it's really just the same thing. There are some extra things we need to do as a result of its presence (get a permit, fill it with gas, put on safety glasses, etc.) but the core idea of the proof is identical. We are going to expand $(x+h)^n$, simplify it, cancel x's and h's out, and then send h to zero.

Now. As to the binomial theorem itself. It is only true, of course, for n being some nonnegative integer, like $0, 1, 2, 3\ldots$ And remember that the $\binom{n}{n-k}$ is a combination. Informally, we can think of a combination like $\binom{a}{b}$ as being the number of different ways we can select a things from a collection of b things, with the order being irrelevant. Formally, we define it as

$$\binom{a}{b} = \frac{a!}{b!(a-b)!}$$

So if we apply this to our problem, we have

$$\frac{(x+h)^n - x^n}{h} = \frac{\sum_{k=0}^{k=n} \binom{n}{n-k} x^{n-k}h^k \quad - \quad x^n}{h}$$

or just

$$= \frac{1}{h}\left(\sum_{k=0}^{k=n}\left[\binom{n}{n-k}x^{n-k}h^k\right] - x^n\right)$$

Which is not immediately more helpful. In fact, it looks worse. But what you may have noticed from working out the derivatives of x^2, x^3, and so forth, is that after you multiply out $(x+h)^2$ (or to the 3), you get an x^2 (or x^3 term), which then cancels out with the $-x^2$ (or $-x^3$) term later on. We want to do the same thing here. We want to somehow extract an x^n term from this sum, such that we can get rid of the $-x^n$. We can do it in this way: we can partially break up the sum. Meaning: we have this sum, consisting of n terms from $k = 0$ to $k = n$. What if we extract the first term (the term where $k = 0$?) Then we would have:

$$\sum_{k=0}^{k=n}\binom{n}{n-k}x^{n-k}h^k = \underbrace{\binom{n}{n-0}x^{n-0}h^0}_{\text{the } k=0 \text{ term}} + \underbrace{\sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k}_{\text{now our sum needs to start at } k=1}$$

$$= \binom{n}{n}x^{n-0}h^0 + \sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k$$

$$= x^n + \sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k$$

I've condensed a whole lot of work into two steps, so stare at this for a while (and try to do the intermediate steps yourself on scrap paper, if you need to) until you see it. Note that I worked out the combination: $\binom{n}{n} = 1$. You could either do this from the formula, or by realizing that if you have a collection of n objects, there's only one way you can pick n of them (if the order doesn't matter).

If we take what we just learned, and apply it to our derivative formula, we'll have:

$$= \frac{1}{h}\left(\left(x^n + \sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k\right) - x^n\right)$$

And THEN we can cancel out the lonely x^n's on the right side and get:

$$= \frac{1}{h}\left(\sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k\right)$$

or just

$$= \frac{1}{h}\sum_{k=1}^{k=n}\binom{n}{n-k}x^{n-k}h^k$$

Then we can bring the $1/h$ inside of the sum. We can do this because a sum is just a bunch of things added together, and a giant Σ is like a pair of parentheses around them—we're just distributing the $1/h$ to the inside.

$$= \sum_{k=1}^{k=n}\binom{n}{n-k}\frac{x^{n-k}h^k}{h}$$

But using laws of exponents we can simplify this to:

$$= \sum_{k=1}^{k=n} \binom{n}{n-k} x^{n-k} h^{k-1}$$

And it's not entirely clear where we go from here... but wait! What if we try that same fancy trick we tried before, where we pull out the first term of the sum? We can extract the $k = 1$ term, and rewrite this as:

$$= \binom{n}{n-1} x^{n-1} h^{1-1} \; + \; \sum_{k=2}^{k=n} \binom{n}{n-k} x^{n-k} h^{k-1}$$

If we use the formula for $\binom{n}{n-1}$, we can discover that it's just n.

$$\binom{n}{n-1} = \frac{n!}{(n-1)!(n-(n-1))!} = \frac{n!}{(n-1)! \cdot 1!} = \frac{n!}{(n-1)!} = \frac{n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdots}{(n-1) \cdot (n-2) \cdot (n-3) \cdots} = n$$

Which should make sense—if I have a collection of n objects, there are n different ways I can choose $n-1$ objects from it—I just systematically exclude one of the objects. So then this becomes:

$$= nx^{n-1} h^0 \; + \; \sum_{k=2}^{k=n} \binom{n}{n-k} x^{n-k} h^{k-1}$$

or just:

$$= nx^{n-1} \; + \; \sum_{k=2}^{k=n} \binom{n}{n-k} x^{n-k} h^{k-1}$$

ALMOST DONE. See the sum? The sum starts at $k = 2$, which means that the greatest-powered h inside the sum will be $h^{2-1} = h^1 = h$. Meaning that if we were to write this out, we'd get something like:

$$= nx^{n-1} + (\text{blah})h + (\text{blah blah})h^2 + (\text{blah blah blah})h^3 + \cdots$$

And so then when h gets really small, all of that other stuff will go away! As h goes to 0, it will drag the rest of the sum with it, down into the depths of nothingness! And all that will be left, when h goes to 0, is the exact slope of x^n, the derivative, or just

$$= nx^{n-1}$$

Problem

Strictly speaking, we've only proven that $(x^n)' = nx^{n-1}$ when n is a positive integer (or zero). Our procedure was based on combinatorics and the binomial theorem, which are all based on the natural numbers $(0, 1, 2, 3, \dots)$. (What would it mean to "choose -2 objects from a group of π objects?" What is 7.2 factorial?) There is nothing wrong with this approach; it's just that we'd like to be able to prove that $(x^n)' = nx^{n-1}$ for ALL real-numbered values of n. But what if n is a negative integer? What if we want to find the derivative of, say, x^{-1} (i.e., $1/x$)? Or x^{-2} (i.e., $1/x^2$)? what if n is a fraction? what if n is an irrational number? We'd like to prove that:

$$\frac{d}{dx}(x^\pi) = \pi x^{\pi - 1} \qquad \frac{d}{dx}(x^{3/2}) = \frac{3}{2} x^{1/2}$$

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

and so forth.

There is a very beautiful proof for the case of n being any real number that uses the derivatives of a logarithm. It only takes two or three lines and is very elegant. But you don't know the derivative of a logarithm yet, and it takes a bit of work to get there.

On the other hand, it is not too difficult to generalize our binomial-theorem-based proof to negative values of n. Imagine that we want to prove that $\frac{d}{dx}(x^n) = nx^{n-1}$ for n being some negative integer. This is the same as proving that $\frac{d}{dx}\left(\frac{1}{x^n}\right) = \frac{-n}{x^{n+1}} = -nx^{-n-1}$ for some positive integer n. So then, if we use our definition of a derivative, we'll have

$$\frac{\frac{1}{(x+h)^n} - \frac{1}{x^n}}{h}$$

which is

$$\frac{1}{h}\left(\frac{1}{(x+h)^n} - \frac{1}{x^n}\right)$$

which, if we substitute in the binomial theorem, will become

$$\frac{1}{h}\left(\frac{1}{\sum_{k=0}^{k=n}\binom{n}{n-k}x^{n-k}h^k} - \frac{1}{x^n}\right)$$

For extra credit, finish this proof, write it up, and turn it in on the Monday after Thanksgiving. (The next step, in parallel to our derivation of the derivatives of $1/x$ and $1/x^2$, is to combine the two fractions atop a common denominator. This will get messy, but you certainly are able to do it.) When your obnoxious relatives ask during Thanksgiving dinner what you're doing in school ("What grade are you in now? 9th?"), you can pause for a moment, as if to wonder how to explain your work to the laity, and then say, "I'm trying to use the binomial theorem to differentiate x^n for negative n." And you will sound like a graduate-student-in-the-making.

If you do write it up, please take the time to write it up *nicely*—don't hand in a paper full of scratch work. Use sentences, explain your methodology, and justify each of your steps in English and in math. Somewhere on the internet there's an excellent four-page article entitled "How to Write Math in Paragraph Style," by Tim Hsu. Read it if you want an idea of what I mean. Either Google it, or find it at `http://www.math.sjsu.edu/~hsu/`