Today is about a basic concept and some extremely useful notation. First the notation.

1 Big-oh notation

Suppose that we wish to determine $f'(x)$ where $f(x) = x^6$. We can proceed as follows:

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^6 - x^6}{h} \\
&= \lim_{h\to 0} \frac{x^6+6x^5h+15x^4h^2+20x^3h^3+15x^2h^4+6xh^5+h^6-x^6}{h} \\
&= \lim_{h\to 0} 6x^5 + 15x^4h + 20x^3h^2 + 15x^2h^3 + 6xh^4 + h^5 \\
&= 6x^5 + 0 + 0 + 0 + 0 + 0 \\
&= 6x^5.
\end{aligned}$$

Which is the correct answer. But we wrote down a great deal of information that we never actually used. Wouldn't it be better to record only the information that we need?

One way to simply the above calculation is to introduce big-oh notation:

$$\text{Anything of the form } Ah + \text{ any number of terms of the form } Bh^m \text{ with } m > 1 \;=\; O(h)$$

where A and B stand for anything that is independent of h. For example,

$$6x^5h + 15x^4h^2 + 20x^3h^3 + 15x^2h^4 + 6xh^5 + h^6 \;=\; O(h)$$

because $A = 6x^5$ is independent of x, and the remaining five terms all have the form Bh^m where B is independent of h and $m > 1$ (in fact, $B = 15x^4$ with $m = 2$, $B = 20x^3$ with $m = 3$, $B = 15x^2$ with $m = 4$, $B = 6x$ with $m = 5$ and $B = 1$ with $m = 6$). Notice the important point that I did *not* write $O(h) = Ah+$ any number of terms of the form Bh^m with $m > 1$, because that would *not* have been correct; one must write $Ah+$ any number of terms of the form Bh^m with $m > 1 = O(h)$, i.e., $O(h)$ must be on the right-hand side. What's the difference? The easiest way to answer this question is to point out that, e.g.,

$$6x^5h + 15x^4h^2 + 20x^3h^3 \;=\; O(h)$$

as well: $O(h)$ does not define a function. Rather, it is just a grab-all notation for a great many things that fall in a certain category, much as dog is a grab-all term for Alsations, beagles, chihuahuas, dachsunds and ... oh, I'm sorry, I can't think of a dog that begins with "e," but you get the idea. In fact, much as adding a huge pack of dogs to another huge pack of dogs gives you just yet another huge pack of dogs, we have

$$O(h) + O(h) \;=\; O(h).$$

For that matter,

$$AO(h) \pm BO(h) \;=\; O(h)$$

as well (or, if you prefer, it doesn't matter whether A and B are positive or negative). This all makes perfect sense when reading from left to right (the only way we allow it to be read), but it would make no sense at all if we tried to read it from right to left.

Well, might you ask, how can something as imprecise as $O(h)$ be of any use in something as precise as mathematics? Again, the best answer to this question will come in demonstrating the use of big-oh notation below, but for now let's say that $O(h)$ is useful for certain purposes because it allows us to retain only the information that we actually need. And what is that information? Primarily, that

$$\lim_{h\to 0} O(h) \;=\; 0$$

(because $h \to 0 \implies Ah \to 0$ and $Bh^m \to 0$ for $m > 1$). But before we can actually demonstrate the use of $O(h)$, we must first observe that

$$Ah^2 \;+\; \text{any number of terms of the form } Bh^m \text{ with } m > 2 \;=\; hO(h)$$

(on multiplying the definition of $O(h)$ by h).

Now we are ready to proceed. Instead of writing

$$(x+h)^6 \;=\; x^6 + 6x^5h + 15x^4h^2 + 20x^3h^3 + 15x^2h^4 + 6xh^5 + h^6$$

we write

$$(x+h)^6 \;=\; x^6 + 6x^5h + hO(h).$$

Don't you agree that that's a simplification? Yet we still retain all of the information we need to calculate $f'(x)$. Just look:

$$\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \frac{(x+h)^6 - x^6}{h} \\
&= \lim_{h\to 0} \frac{x^6 + 6x^5h + hO(h) - x^6}{h} \\
&= \lim_{h\to 0} \frac{6x^5h + hO(h)}{h} \\
&= \lim_{h\to 0} 6x^5 + O(h) \\
&= 6x^5 + 0 \\
&= 6x^5.
\end{aligned}$$

More generally, when finding derivatives of power functions, it suffices to use the binomial theorem in the simplified form

$$(x+h)^n \;=\; x^n + nx^{n-1}h + hO(h).$$

It then follows readily that

$$
\begin{aligned}
\frac{d}{dx}\left(x^n\right) &= \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} \\
&= \lim_{h\to 0} \frac{x^n + nx^{n-1}h + hO(h) - x^n}{h} \\
&= \lim_{h\to 0} \frac{nx^{n-1}h + hO(h)}{h} \\
&= \lim_{h\to 0} nx^{n-1} + O(h) \\
&= nx^{n-1} + 0 \\
&= nx^{n-1}.
\end{aligned}
$$

Thus, for example, $\frac{d}{dx}\left(x^{10}\right) = 10x^9$ and $\frac{d}{dx}\left(x^{54}\right) = 54x^{53}$.

2 Linear operators

Let's suppose that you own a zoo in a little-known (and far away) country where the total amount of meat in a zoo must be reported daily to the government. Suppose you have three elephants, four giraffes and five zebras, and that the weight of an elephant, giraffe or zebra in your zoo is precisely e, g or z pounds, respectively—absolutely no variation, because you are an amazing scientist who has perfected a process called genetico-environmental cloning that guarantees that all of your animals in a given species have identical weight throughout life. So yesterday you reported to the government that you had $3e + 4g + 5z$ pounds of exotic animal meat in your zoo.

Today, however, the new Department of Zooland Security has decided that your nation will be safer if the weight of exotic animal meat is reported in ounces instead of in pounds, and has introduced the necessary legislation. So you now have the task of converting all your pounds to ounces. How do you do it? You are well aware that w pounds can be converted into ounces by applying the conversion operator C defined by

$$C(w) = 16w.$$

But should you first of all calculate the weight of your meat in pounds, and then convert it to ounces? If so, then you are calculating

$$16\{3e + 4g + 5z\} = C(3e + 4g + 5z).$$

Or should you first of all convert the weight of an elephant, a giraffe and a zebra into ounces and then multiply by as many of each animal as you have before adding? If so, then you are calculating

$$3 \cdot \{16e\} + 4 \cdot \{16g\} + 5 \cdot \{16z\} = 3C(e) + 4C(g) + 5C(z).$$

Well, which? The answer, of course, is that it doesn't matter in the least because

$$C(3e + 4g + 5z) = 3C(e) + 4C(g) + 5C(z).$$

And it wouldn't matter in the least even if you had n_1 elephants, n_2 giraffes and n_3 zebras, because

$$C(n_1 e + n_2 g + n_3 z) \;=\; n_1 C(e) + n_2 C(g) + n_3 C(z)$$

as well. It wouldn't matter because the conversion operator has the special property that you can apply it to a sum (or difference) of individuals either by applying it first to each individual and then summing (or taking a difference), or else by summing (or subtracting) over all individuals and then applying the operator, and the answer is always exactly the same. Operators with this property are known as *linear* operators. For example, C is a linear operator. More importantly—and the only reason for visiting your zoo today—the differential operator $D = \frac{d}{dx}$ is a linear operator. What does that mean in practice?

Once again, an example provides the best answer. Suppose that you wish to determine $g'(x)$ where $g(x) = 3x^{10} - 4x^6 + 5x^4$. You could proceed as follows:

$$\begin{aligned}
g'(x) &= \lim_{h\to 0} \frac{g(x+h) - g(x)}{h} \\
&= \lim_{h\to 0} \frac{3(x+h)^{10} - 4(x+h)^6 + 5(x+h)^4 - 3x^{10} + 4x^6 - 5x^4}{h} \\
&= \lim_{h\to 0} \frac{3\{x^{10}+10x^9h+hO(h)\}-4\{x^6+6x^5h+hO(h)\}+5\{x^4+4x^3h+hO(h)\}-3x^{10}+4x^6-5x^4}{h} \\
&= \lim_{h\to 0} \frac{30x^9h - 24x^5h + 20x^3h + h\{3O(h) - 4O(h) + 5O(h)\}}{h} \\
&= \lim_{h\to 0} \frac{30x^9h - 24x^5h + 20x^3h + hO(h)}{h} \\
&= \lim_{h\to 0} 30x^9 - 24x^5 + 20x^3 + O(h) \\
&= 30x^9 - 24x^5 + 20x^3 + 0 \\
&= 30x^9 - 24x^5 + 20x^3.
\end{aligned}$$

Which is the correct answer. But it is much more efficient to proceed as follows:

$$\begin{aligned}
\frac{d}{dx}\left(3x^{10} - 4x^6 + 5x^4\right) &= 3\frac{d}{dx}\left(x^{10}\right) - 4\frac{d}{dx}\left(x^6\right) + 5\frac{d}{dx}\left(x^4\right) \\
&= 3\cdot 10x^9 - 4\cdot 6x^5 + 5\cdot 4x^3 \\
&= 30x^9 - 24x^5 + 20x^3,
\end{aligned}$$

as before. Thus big-oh notation and linearity have in common that they both enable us to simplify the calculation of derivatives.