Introduction

I will start by 'deriving' the limit formula defining the constant e as $\lim_{\{n\to\infty\}} \left(1+\frac{1}{n}\right)^n$ with the motivation being finding a base, b, such that the function b^x is its own derivative. Then I will attempt to generalize this for positive and negative powers of e.

Throughout this article, I will try to describe my thought process behind every step. As such, obvious disclaimer: I aim to be very intutive and beginner-friendly, and I do not pretend to be rigorous.

Finally, before diving into the article, readers may check out the Appendix at the end that covers elementary properties of exponentiation for a quick refresher.

Shrinking Step Sizes of Difference Quotient

So, for some $f(x)$, if I increment x by Δx, $\Delta y = f(x+\Delta x) - f(x)$. And the difference quotient is $\frac{\Delta y}{\Delta x} = \frac{f(x+\Delta x)-f(x)}{\Delta x}$. As $\Delta x \to 0$, the difference quotient converges to the derivative of $f(x)$.

My goal, then, is to find a function, $f(x)$ where this quotient is itself. So I want $\frac{\Delta y}{\Delta x} = f(x)$ which, solving for Δy necessitates that by incrementing x by Δx, $\Delta y = f(x) * (\Delta x)$. That is, Δy scales Δx by a scaling factor that is $f(x)$ itself.

2^x is a Step Size of 1 Approximation

I will straight up present the function $f(x) = 2^x$ as a function that when I take a $\Delta x = 1$ step of size 1, $\Delta y = f(x+1) - f(x) = 2^{x+1} - 2^x = 2 * 2^x - 2^x = 2^x = f(x)$ Thus, the difference quotient for this step size of 1 is $\frac{\Delta y}{\Delta x} = \frac{f(x)}{1} = f(x)$. And so the function $f(x) = 2^x$ satisfies the desired property of $\frac{\Delta y}{\Delta x} = f(x)$ for $\Delta x = 1$.

From Algorithms

Let me end this section by taking a brief, optional, detour to explain why 2^x was a very plausible function. Having a computer "science" background, I am familiar with properties of powers of 2 that arise when dealing with binary objects. There are several recursive algorithms including binary search, quick select, binary tree traversals, that take a problem of size N and in a single step, reduce it to a problem of size $\frac{N}{2}$. Thus the entire problem takes roughly $\log_2(N)$ steps and is very efficient. If I treat steps taken as x values and remaining problem size as $f(x)$ or y values, this results in exponential decay of $f(x) = N * \left(\frac{1}{2}\right)^x$. When $x = 0$, the problem size, or work left, is N units of work, when $x = 1$, the problem size, or work left, is $\frac{N}{2}$ units of work, and so forth. The first step when there is N work left saves $\frac{N}{2}$ work, the second step when there is $\frac{N}{2}$ work left saves $\frac{N}{4}$ work, and each successive step saves half of the work left. "Saving" per each step size where $\Delta x = 1$ is the negative of Δy. So saving from step x to step $x+1$ is $\frac{f(x)}{2}$ meaning $\Delta y = -\left(\frac{1}{2}\right) * f(x)$. But again, as $\Delta x = 1$, $\frac{\Delta y}{\Delta x} = \Delta y = -\left(\frac{1}{2}\right) * f(x)$ for $f(x) = N * \left(\frac{1}{2}\right)^x$. Note that N does not appear in the difference quotient at all. So I have found a function, $f(x) = k * \left(\frac{1}{2}\right)^x$ whose difference quotient scales itself by a factor of $-\left(\frac{1}{2}\right)$. Scaling itself is a crucial property and this suggests an exponential function is desired, though maybe not decay.

From Difference Quotient

A straight forward way to arrive at 2^x is to first assume I am searching for an exponential function. And then using algebra to set the difference quotient (with step size $\Delta x = 1$) of the function b^x equal to itself and solve for b to get $b = 2$:

$$\frac{\Delta y}{\Delta x} = f(x)$$

$$\frac{b^{x+1} - b^x}{1} = b^x$$

$$(b^x) * b - b^x = b^x$$

$$(b^x) * (b - 1) = b^x$$

$$b - 1 = 1$$

$$b = 2$$

Smaller steps

While $f(x) = 2^x$ satisfies $\frac{\Delta y}{\Delta x} = f(x)$ for $\Delta x = 1$, I want to find a different function that satisfies this constraint for a smaller step size.

Given the function 2^x is exponential and I'm working towards finding e where e^x, an exponential function, is it's own derivative, this new function for a smaller step size is presumably exponential as well in the form of $f(x) = b^x$.

Appendix: Exponentiation Basics

I'll explore basics of exponentiation here using integers and motivate some properties of exponents, especially the property: $\left(b^k\right)^x = b^{k*x}$. Firstly, what does b^x mean? b^x evaluates to $\overbrace{b * b * ... * b}^{\text{x times}}$.

Symbolically, this is a product of x factors of b. Visually, I like to use trees that with branching factor b. For example, below shows a complete binary tree to represent 2^h, the case where $b = 2$.

TODO draw complete binary tree, a "2-Tree" here

The levels of these trees are 0-indexed, meaning at the 0th level, there is 1 node (the root), at the 1st level, there are b nodes, at the second level, there are $b * b$ nodes. Each successive level introduces another factor of b, since every node at the previous level splits into b more nodes. 1 node introduces b child nodes, 2 nodes introduce $2 * b$ nodes, all k nodes introduce $k * b$ children. Thus, at some level, l, there are b^l nodes, and the relation between successive levels is: $b^{l+1} = b^l * b$. And this relation naturally extends to $b^{l+k} = b^l * b^k$, that is, adding k to the exponent introduces k more factors of b that act on b^l.

Different bases

Let me add another base for consideration: 8^x. Below are 2 trees side-by-side that terminate with 64 leaves.

TODO draw these trees and make them line up, so distance between levels of the 8^x would be 3x that of 2^x tree

Observe that these two trees are quite closely related. Let me state the relation exactly as follows: every 3 levels of doubling for the 2-Tree produces the same effect of a single level of the 8-Tree.

So the 8-Tree is a 'compressed' version of the 2-Tree, by a factor of 3, based on the following equivalancy.

TODO draw another side by side picture of 3 levels of the 2-Tree and 1 level of the 8-tree, again lined up

Because $8 = 2^3 = 2 * 2 * 2$, 3 levels of doubling results in 1 level of multiplying by 8.

Let h be the height of the tree where if the lowest, leaf, level is indexed at l, $h = l - 1$. So the 8^h tree has 8^h leaves. When $h = 1$, there are 8 leaves. And when $h = 2$, there are 64 leaves. Now for the 2^h tree, when $h = 3$ there at 8 leaves. And when $h = 6$, there are 64 leaves. So, more generally, this shows that $8^h = 2^{3*h}$. But $8 = 2^3$, so $8^h = \left(2^3\right)^h$, and this proves $\left(2^3\right)^h = 2^{3*h}$ More generally, if X is some number as a power of b, say, $X = b^h$, then X^k multiplies the height of the b-tree representation of X by k. Note that this is only for integer values of k. I will, very soon, motivate this for rational powers as well (namely, $k = \frac{1}{3}$).

Finally, and this is, I suspect how most people including myself learned exponents, I can readily see all this when writing out factors: $8^2 = (8) * (8) = (2 * 2 * 2) * (2 * 2 * 2) = 2^6$. The number of factors is h, the argument of $f(h) = b^h$ and it is evident that the number of factors in the 8-expansion gets multiplied by 3 to get the number of factors in the 2-expansion. Like it takes 2 8's to write out 64 but it takes $6 = 2 * 3$ 2's to write out 64 using factors of all 2's. (If you are familiar with hexademical and binary numberings a similar compression by a factor of 4 happens where every hexadecimal digit valued from 0-15 can be converted into 4 binary digits)

OK, but what about instead of multiplying by 3, dividing by 3. Consider $8^{\frac{1}{3}}$. For the function $f(h) = 8^h$, the input h is the height. But a fractional height doesn't make sense? But if use the relation I just derived, where every 1 level of the 8-Tree is equivalent to 3 levels of 2-Tree, every 2 levels of the 8-Tree is equivalent to 6 levels of the 2-Tree, it follows that 1/3 level of the 8-Tree is equivalent to 1 level of the 2-Tree. That is, I'm assuming the ratio of 1 level 8-Tree : 3 levels 2-Tree,

$$X \text{ 8-level} = X\cancel{\text{8-level}} * \left(\frac{3 \text{ 2-level}}{1 * \cancel{\text{8-level}}}\right) = 3X \text{ 2-level}$$

or, equivalently,

$$X \text{ 2-level} = X\cancel{\text{2-level}} * \left(\frac{1 \text{ 8-level}}{3 * \cancel{\text{2-level}}}\right) = \left(\frac{1}{3}\right) X \text{ 8-level}$$

And so $8^{\frac{1}{3}} = 2^1 = 2$ and more generally, $b^{\frac{1}{k}} = x$ where $x^k = b$. And symbollically, this is readily displayed by $\left(b^{\frac{1}{k}}\right)^k = b$.