June 10, 2024
Justin Bloom (justib8@uw.edu)
University of Washington
Department of Mathematics

---

Dear [redacted],

You are absolutely correct that there is a problem in defining the continuous, or *fractional* derivative. You have correctly identified this as a problem of non-uniqueness of interpolating a function over a discrete parameter into one over a continuous parameter. However, when it comes to derivatives of the exponential, perhaps you should consider the following. I will attempt to apply your mode of reasoning to a much simpler setting, to point out its shortcoming, and then offer what I understand about fractional calculus. I have broken my argument into three sections and only begin to talk about derivatives in my third section.

# 1 On interpolating an operator

When defining even something simple like the laws of arithmetic for real numbers, we are always trying to extend definitions for integers to the real numbers in a way that preserves as many of the essential properties as we can. For example, if we try to define $x^n$ as repeated multiplication of $x$ by itself $n$-times, this does not immediately tell you how to interpret the operation $x^{1/2}$ or even $x^0 = 1 \quad (x \neq 0)$. The reason we decide, for $x \geq 0$, that $x^{1/2} = \sqrt{x}$ is due to the essential properties that

$$x^{n+m} = x^n x^m, \quad x^{nm} = (x^n)^m = (x^m)^n, \tag{1}$$

whenever $n, m$ are positive integers. These properties follow from the definition of powers as repeated multiplication, but are meaningless if $n, m$ are not positive integers, as there is no formal meaning to multiplying $x$ by itself even 0 times, let alone 1/2 times. Most math students in, say, highschool or above are probably satisfied with the intuition that an empty product should be 1, but we are trying to be formal in how we define things. So really, the only reason we decide $x^0 = 1$ (for $x \neq 0$) is so that we can maintain the properties that we've decided are essential, namely

$$x^n = x^{0+n} = x^0 x^n,$$

and similarly

$$x = x^1 = x^{2*1/2} = (x^{1/2})^2.$$

Even still, this argument breaks apart when trying to define an irrational power of $x$ such as $x^\pi$. No repeated application of the essential properties listed so far will uniquely determine the value of, for instance $2^\pi$. Yet any scientific calculator will output something like

$$2^\pi \approx 8.824977827.$$

Instead we may observe that $\pi$ can be approximated by a sequence of rational numbers $3, 3.1, 3.14, 3.141, 3.1415, ...$ and defining the fractional powers e.g. $x^{31415/1000}$ as the 1000th root of $x^{31415}$, gives a sequence

$$x^3, x^{31/10}, x^{314/100}, ...$$

which approaches a limit for any positive $x$, and importantly that this limit is actually independent of which approximation of $\pi$ by a sequence of rational numbers is chosen.

Perhaps it is that we have assumed another property to be essential, for (nonnegative) rational numbers $a, b$, that if $a < b$ and $x > 1$, now that we've defined $x^a$ and $x^b$, we can conclude that

$$x^a < x^b.$$

Thus, knowing that $3.1415 < \pi < 3.1416$ means that we should assume for $x > 1$ that

$$8.82441108248 \approx x^{31415/10000} < x^\pi < x^{31416/10000} \approx 8.82502276524.$$

So a sequence of progressively finer upper and lower bounds for the value of $\pi$ squeezes the value of $x^\pi$ into a progressively narrower window.

Ideally, all of the essential properties of taking powers work nicely enough with one another in order to completely determine how to extend the idea beyond integer powers. But it's difficult to know this without really knowing what makes a property essential.

## 2 On essential properties

For better or for worse there is no authority in mathematics to tell us what makes a property essential. In the spirit of your interpolations of derivatives towards a continuous domain, consider the property, whenever $n, m$ are positive integers, of taking powers.

$$x^{n+m} = x^n x^m + \sin(nm\pi). \tag{2}$$

Indeed, we are looking at the same property as before, only I have added a function $f(n, m) = \sin(nm\pi)$, which obvioulsy has a continuous domain, and yet happens to have a value of 0 whenever $n, m$ are both integers.

Now, taking $n = m = \frac{1}{2}$ we should get

$$x = x^1 = x^{1/2+1/2} = x^{1/2}x^{1/2} + \sin(\pi/4).$$

So we may try to solve $y = x^{1/2}$ by noticing

$$x = y^2 + \frac{\sqrt{2}}{2}$$

and therefore $x^{1/2} = y = \pm\sqrt{x - \frac{\sqrt{2}}{2}}$.

Now there are two possible values of $y$ given $x$ (even if $y$ ends up being imaginary for $x < \frac{\sqrt{2}}{2}$), and both of these values are equally absurd to one another, for anyone familiar with basic arithmetic. But the only thing we needed to get to this absurdity was a lack of authority telling us that the property 1 is essential while the property 2 was not. So why is it that everyone seems to agree on 1 but not 2? This is, from what I can tell, the same problem you have raised by seeing different ways to 'fill in the graph' for derivatives of the exponential.

I may have you convinced at this point, but there is in fact a saving grace which at the very least rules property 2 out as not only non-essential, but actually self-contradictory. Because of the associative property

$$(n + m) + \ell = n + (m + \ell),$$

we get a contradiction when extending $n, m, \ell$ to a continuous domain as follows. Assuming property 2, we see that

$$x^{(n+m)+\ell} = x^{n+m}x^{\ell} + \sin((n + m)\ell\pi),$$

and applying 2 again expands this further as

$$\begin{aligned} x^{(n+m)+\ell} &= (x^n x^m + \sin(nm\pi))x^{\ell} + \sin((n + m)\ell\pi) \\ &= x^n x^m x^{\ell} + \sin(nm\pi)x^{\ell} + \sin((n + m)\ell\pi). \end{aligned}$$

But, using the other association, we do the same thing to get

$$\begin{aligned} x^{n+(m+\ell)} &= x^n(x^m x^{\ell} + \sin(m\ell\pi)) + \sin(n(m + \ell)\pi) \\ &= x^n x^m x^{\ell} + x^n \sin(m\ell\pi) + \sin(n(m + \ell)\pi). \end{aligned}$$

Now clearly we need $x^{(n+m)+\ell} = x^{n+(m+\ell)}$ for our continuous extension of raising $x$ to a power, and the two expressions seem to be incompatible. But to be rigorous, we take $n = m = 1$ and try to see what this determines about $x^{\ell}$, for some non-integer value of $\ell$. Now we have

$$\begin{aligned} x^{(1+1)+\ell} &= x^2 x^{\ell} + \sin(\pi)x^{\ell} + \sin(2\ell\pi) \\ &= x^2 x^{\ell} + \sin(2\ell\pi), \end{aligned}$$

while

$$x^{1+(1+\ell)} = x^2\ell + x\sin(\ell\pi) + \sin((1 + \ell)\pi).$$

We may fix $x = 1$ now and study two functions of the continuous value $\ell$. One being $\sin(\ell\pi) + \sin((1 + \ell)\pi)$, we can actually see that this function is identically 0, as

$$\sin((1 + \ell)\pi) = \sin(\pi + \ell\pi) = -\sin(\ell\pi)$$

by symmetry of the sinusoid.

The other function $\sin(2\ell\pi)$ is not identically 0, and actually is only 0 for integer values of $\ell$. Therefore property 2 can only possibly make sense for integer values of $n, m$.

Let's generalize what we've done so far. We started with observing that property 1 is adaptable to a strange looking property

$$x^{n+m} = x^n x^m + f(x, n, m), \tag{3}$$

for a continuous function $f(x, n, m)$, and this can only make sense as an extension of the usual rule for positive integer powers if $f(x, n, m) = 0$ for any positive real number $x$, whenever $n, m$ are positive integers. If $f(x, n, m)$ is identically 0, then we recover property 1, and the previous section explored how this successfully defines $x^a$ for positive rational numbers $a$. The current section showed that in order for property 3 to serve as any sort of defining property for fractional powers, it must at least be compatible with the associative property in order not to contradict itself. That is, property 3 implies that

$$\begin{aligned} x^{(n+m)+\ell} &= x^{n+m}x^\ell + f(x, n+m, \ell) \\ &= (x^n x^m + f(x, n, m))x^\ell + f(x, n+m, \ell) \\ &= x^n x^m x^\ell + f(x, n, m)x^\ell + f(x, n+m, \ell), \end{aligned}$$

as well as

$$\begin{aligned} x^{n+(m+\ell)} &= x^n x^{m+\ell} + f(x, n, m+\ell) \\ &= x^n(x^m x^\ell + f(x, m, \ell)) + f(x, n, m+\ell) \\ &= x^n x^m x^\ell + x^n f(x, m, \ell) + f(x, n, m+\ell). \end{aligned}$$

To avoid contradiction, we must have equality

$$f(x, n, m)x^\ell + f(x, n+m, \ell) = x^n f(x, m, \ell) + f(x, n, m+\ell)$$

for all $x, n, m$, and so not just any function $f(x, n, m)$ will do.

So far I have failed to mention that the commutative property must also be respected. That is, because $n + m = m + n$, in order to avoid contradiction, we must also have

$$f(x, n, m) = f(x, m, n)$$

for all $x, n, m$.

So we have that the properties we are looking to preserve when extending to fractional powers come down to a system of functional equations. It may be the case that the only continuous solution to these equations is $f(x, n, m) = 0$, meaning there is only one definitive way to define fractional powers in such a way that extends the property of addition in the exponent. It is beyond my capabilities to try to think of another such $f(x, n, m)$.

But to reiterate the point, given such a rule as property 3, we can try to define the simple fractional power $1^{1/p}$. To do this we first note that by induction, we get a formula for $x^{nm}$ whenever $n$ is an integer, and it's quite a bit more complicated than $(x^m)^n$. Instead we get

$$x^{nm} = x^{m+(n-1)m} = (x^m)^n + f(x,m,(n-1)m) + \sum_{k=1}^{n-2} (x^m)^k f(x,m,(n-k-1)m).$$

Substituting $x = 1, n = p$, and $m = 1/p$, we can deduce a polynomial equation that $y = x^{1/p}$ must satisfy, namely that

$$1 = 1^{p(1/p)} = y^p + f(1,1/p,(p-1)/p) + \sum_{k=1}^{p-2} y^k f(1,1/p,(p-k-1)/p.$$

For $p = 2$, this is simple to solve. It is a quadratic equation

$$1 = y^2 + f(1,1/2,1/2),$$

and therefore the solutions are

$$1^{1/2} = y = \pm\sqrt{1 - f(1,1/2,1/2)}.$$

This is the case we already saw at the beginning of this section.

But already for $p = 3$ this becomes difficult, as we are solving a cubic equation

$$1 = y^3 + f(1,1/3,2/3) + yf(1,1/3,1/3).$$

There is a cubic formula that works to solve equations of this kind, but it is rarely taught because it is quite complicated compared to quadratic equations. I used a computer algebra system to get at least this one solution

$$1^{1/3} = y = \frac{\sqrt[3]{\sqrt{3}\sqrt{27A^2-54A+4B^3+27}-9(A-1)}}{\sqrt[3]{18}} - \frac{\sqrt[3]{\frac{2}{3}}B}{\sqrt[3]{\sqrt{3}\sqrt{27A^2-54A+4B^3+27}-9(A-1)}},$$

where $A = f(1,1/3,2/3)$ and $B = f(1,1/3,1/3)$ are the constants appearing in the original polynomial in the variable $y$. It is clearly in our best interest to assume that $A = B = 0$, so that $1^{1/3} = 1$, for what other sensible values would $1^{1/3}$ ever take on? For any function $f(x,n,m)$ other than 0, what hope is there to ever even conclude how to define $x^{1/p}$, if it involves solving a polynomial equation of degree $p$? No highschool math student is ever going to consider any function $f(x,n,m)$ other than 0 when considering the most natural properties of taking powers of $x$, and we hardly need to go to these depths to see why.

# 3 Fractional derivatives

The purpose of the last two sections was to establish that extending an integer parameterized operator into a continuous parameterized operator can be an extremely subtle task, and that we don't have to look at calculus to find examples of this. I tried to illustrate this with the simple operation of taking powers for a good reason: taking the $n$th derivative itself may be intepretted as an $n$th power of an operator.

When we write the expression $D = \frac{d}{dx}$, we are defining an operator which inputs a smooth function of a single variable $f(x)$, and outputs the derivative $D\{f\}$, another smooth function, and this may also be notated as

$$D\{f\}(x) = \frac{d}{dx} f(x) = f'(x).$$

If you have seen the notation $\frac{d^n}{dx^n} f(x) = f^{(n)}(x)$, this notation is already showing how the $n$th derivative is like an $n$th power. For example at $n = 2$ we are interpretting this as an associative 'multiplication' between operators and the function inputs, i.e.

$$\frac{d^2}{dx^2} f(x) = \left(\frac{d}{dx}\right)^2 f(x) = \frac{d}{dx}\left(\frac{d}{dx} f(x)\right).$$

The notation $D^n = \frac{d^n}{dx^n} = \left(\frac{d}{dx}\right)^n$ is convenient, so that we are not writing the derivative operator as a fraction every time. To be clear, since we know how to take a derivative once, we are representing by the operator $D^n$, taking a derivative $n$ times in succession, just as it makes sense to write $x^n$ to represent multiplying $x$ by itself $n$ times.

Hopefully from the previous two sections you can now be convinced that the only reasonable way to express how one adds powers of a derivative together is as

$$D^{n+m} = D^n D^m,$$

i.e. taking the derivative first $m$ times and then another $n$ times. Thus, however we try to define $D^s$ for fractional powers $s$, it should obey

$$D^{s+t} = D^s D^t,$$

and not a crazy identity dependent on $s, t$ which happens to be 0 whenever $s, t$ are positive integers.

This is hardly enough to even define what is meant by $D^{1/2}$. What we need is an operator $T$, which inputs and outputs a smooth function, such that

$$T\{T\{f\}\}(x) = D\{f\}(x) = f'(x).$$

but it should still highlight a problem in your approach. It is not quite enough to look at the $n$th derivatives of a single function, evaluated at a point like $x = 0$, and then simply guess how to fill it in. We must define what $D^s f$ means for every smooth function $f$.

After chatting today with an expert in analysis and geometry, I have learned a few things about the theory of fractional calculus that I didn't know even yesterday. As I understand, it revolves around finding which properties of an $n$th derivative, beyond how they compose together as operators, interact with one another nicely enough to define what is meant by $D^s$ for fractional values of $s$. But perhaps as you suspected, there is not just one way to do this, as even when refining one's approach, we find not all properties behave nicely together, and some of them contradict each other when trying to define the fractional derivative.

For example, given a constant $a$, if $g(x) = x + a$, and $f \circ g$ denotes the composition $f \circ g(x) = f(g(x)) = f(x + a)$, then the chain rule tells us

$$D\{f \circ g\} = D\{f\} \circ g.$$

I like to interpret this as, if you shift a graph to left by a factor of $a$, then the derivative is also shifted to the left by a factor of $a$.

We also have linearity: that for any constants $a, b$, and any smooth functions $f_1, f_2$, we have

$$D\{af_1 + bf_2\} = aD\{f_1\} + bD\{f_2\}.$$

From these properties of the derivative $D$, we get some properties of the $n$th derivative, namely that

$$D^n\{f \circ g\} = D^n\{f\} \circ g$$

and

$$D^n\{af_1 + bf_2\} = aD^n\{f_1\} + bD^n\{f_2\}.$$

Now, these properties of $D$ combined, actually determine how to take the derivative of an exponential! Consider the function $f(x) = a^x$ for some constant $a > 0$, and let $b$ be some other constant, with $g(x) = x + b$. Then $f(x + b) = f \circ g = a^b f(x)$. Therefore, we have

$$D\{f \circ g\} = D\{f\} \circ g,$$

as well as

$$D\{f \circ g\} = a^b D\{f\},$$

and therefore

$$D\{f\} \circ g = a^b D\{f\}.$$

If we plug in 0 we get

$$D\{f\} \circ g(0) = D\{f\}(b) = a^b D\{f\}(0),$$

but this holds for any $b$. So we may let $b = x$ be the variable, which tells us the entire function $D\{f\}(x) = Ca^x = Cf(x)$ for the constant $C = D\{f\}(0)$. We only need the actual definition of the derivative, beyond its properties, to know conclusively that $C = \ln(a)$. In fact, we may repeat this argument, using instead the properties of $D^n$, to know that $D^n\{f\} = Cf$ for some constant $C$ (which again, we should know is $\ln(a)^n$, So when defining the fractional derivative, we should already expect it to behave nicely with exponentials. That is, provided that the properties

$$D^s\{f \circ g\} = D^s\{f\} \circ g, \qquad D^s\{af_1 + bf_2\} = aD^s\{f_1\} + bD^s\{f_2\}$$

hold for fractional values of $s$, we already know that

$$D^s\{a^x\} = C(a, s)a^x$$

for some $C(a, s)$ depending only on $a$ and $s$.

Notice, since $f(0) = 1$, we have that $C(a, s)$ is precisely $D^s\{f\}(0)$. What we know about $C(a, s)$ is that whenever $s = n$ is a positive integer, that $C(a, n) = \ln(a)^n$. What you've shown in the document you sent me, is that there are many continuous functions $C(e, s)$ that have the property that $C(a, n) = \ln(e)^n = 1$. We may find any function $F(s)$ such that $F(n) = 0$ whenever $s = n$ is a positive integer, and then $C(e, s) = 1 + F(s)$ is an 'extension' of the $n$th derivative of $e^x$ to the $s$th derivative for fractional $s$. You found $F(s) = \sin(s\pi)$ as well as $F(s) = \sin(2\pi(s + 0.25)) = \cos(2\pi s) - 1$ as two different options, which may be added to 1 to get some value of $C(e, s)$ which is 1 at each positive integer $s$. It's true now that defining $D^s\{f\} = C(e, s)$ for $f(x) = e^x$ this way does satisfy the shifting and linearity properties. But unfortunately this does not satisfy the first property I mentioned in this section: We still need $D^{s+t} = D^s D^t$. This is fundamentally why it is so important to consider how we are defining the entire operator $D^s$ for any function, rather than the $s$th derivative of a single function at a time.

By this, what I mean is that, we have already concluded, for $f(x) = a^x$, that $D^s\{f\} = C(a, s)f$, by simply assuming the shifting and linearity properties of $D^s$.

Therefore

$$D^s D^t\{f\} = D^s\{C(a, t)f\} = C(a, s)D^s\{f\} = C(a, t)C(a, s)f,$$

again, using only linearity. But if we are to assume $D^{s+t} = D^s D^t$, we must also have

$$D^s D^t = D^{s+t}\{f\} = C(a, s + t)f.$$

We now conclude that the function $C(a, s)$ must satisfy

$$C(a, t)C(a, s) = C(a, s + t).$$

Given $C(a, n) = \ln(a)^n$ for positive integers $s = n$, there is **only one continuous function** which has this property, and it is $C(a, s) = \ln(a)^s$. In particular $C(e, s) = 1$ for all values $s$. In conclusion, $D^s\{f\}(x) = e^x$ for $f(x) = e^x$, i.e. the $s$th derivative of $e^x$ is always $e^x$, even for fractional values of $s$, provided we assume how the $s$th derivative should behave.

---

It turns out, in my chat with an expert, that not all of these properties are always assumed in the theory of fractional calculus. Linearity is practically always assumed, as well as the composition of operators property $D^{s+t} = D^s D^t$, but the shifting property I mentioned is not. So for some other definitions of the fractional derivative, it may be the case that the $s$th derivative of $e^x$ is something other than $e^x$ for non-positive integers $s$. I would not be able to elaborate why. But he said to understand any of this better, you should learn real analysis, linear algebra, and in particular the theory of Hilbert spaces and the spectral theorem. I hope this helps!

Best wishes,

-Justin Bloom