Cauchy's Functional Equation and Iffy Derivatives

P. Reany

August 12, 2024

Abstract

We had better be careful when we differentiate across an equation about which derivative we use. We all seem to do it too casually at times, because it's so tedius to show all the details. But this is such an interesting case study, I wanted to showcase it from the viewpoint of Structured Differentiation.

Matters of notation play a considerable role in connection with the chain rule. Wide varieties of usage exist in mathematical writing where the chain rule is concerned.

—Taylor & Mann

1 Introduction

I'm basing this write-up on the opening part of a YouTube video found at:

`https://www.youtube.com/watch?v=UskPO3MwcBc`

Titled: Solving **Cauchy's Functional Equation** When f is Differentiable

By Presenter: SyberMath

The problem is to solve for $f(x)$ given the relation

$$f(x+y) = f(x) + f(y)\,, \tag{1}$$

where we take this f to be differentiable over the real line.

2 The Structured Differentiation

> "When I use a word," Humpty Dumpty said in rather a scornful tone, "it means just what I choose it to mean — neither more nor less.'"
> "The question is," said Alice, "whether you can make words mean so many different things."
> — Lewis Carroll, *Through the Looking Glass*

There are two kinds of people in this world who deal with differentiation: those that see no problem with it and those who see it as a total mess. Decades ago I was of the latter group and I decided to sort out and fix the many problems inherent in modern differentiation — if for no other reason, so that I could at least understand what's going on.

Let's begin with what's not problematic: Let $f(x)$ be a real differentiable function of real variable x. Then, taking the derivative of f by x is not problematic, at least not to me, yielding

$$f'(x) = \frac{df}{dx} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h}\,, \tag{2}$$

if the limit exists (this caveat will be assumed from here on). So, how do we generalize this formula to functions f of multiple real variables $x_1, \ldots, x_n$? That is,

$$f = f(x_1, \ldots, x_n)\,. \tag{3}$$

Well, if we have common sense, we'll start by treating the n real variables $x_1, \ldots, x_n$ as mutually independent of each other. But before we get to that, I want to introduce some much needed terminology for the subject of differentiation. Referring to (3), the function f is explicitly dependent on the n variables $x_1, \ldots, x_n$. I refer to these variables as **variants** of the function. If all the variants of a function are mutually independent of each other, the function is said to be in **primitive form**.

If you see something like this

$$f(x) = g(x, y(x))\,, \tag{4}$$

and then write down its derivative by x as

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial x}\,, \tag{5}$$

do you see a problem with this? I do. The problem is that the 'partial derivative' is technically defined only on primitive functions, like this:

Consider f in (3), where f is considered to be primitve. Then its 'partial derivative' with respect to x_i is defined as

$$\frac{\partial f}{\partial x_i} = \lim_{h\to 0} \frac{f(x_1 \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}\,. \tag{6}$$

Now, don't get me wrong. I'm not faulting the last definition per se. But I'm asking what derivative symbol are we to use on a function that is **not** primitive, such as $g(x, y)$ in (4)? Technically speaking, we're breaking the rules to use such symbols as $\dfrac{\partial g}{\partial x}$ and $\dfrac{\partial g}{\partial y}$, but what's a mathematician, physicist, or engineer to do? The answer to that is that **many** solutions have been devised over the last hundred years or so to deal with this situation. When these varied solutions are taken together, they're pretty conflicting and confusing.[1]

One solution is to claim that when using a 'partial derivative by x', say, on a non-primitive function, we have to 'hold all other variables constant'. Indeed, that is forced when using the partial derivative by its definition, because no two of its variants can be functionally dependent on each other.

The SD solution to this problem is simple. We add to the partial derivative, two other derivatives: one explicit and one implicit, say, as applied to the following function

$$f(\mathbf{x}(t), t) = f(x_1(t), \ldots, x_n(t), t)\,, \tag{7}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$. Now, the function f is implicitly dependent on variable t through the n variables $x_1, \ldots, x_n$ and explicitly dependent on variable t, because t is a variant of f (the $(n+1)$st variant in the list). Therefore, we need two derivatives to account for the possible variations of a non-primitive function on a variable: an explicit derivative and an implicit derivative. Thus, the derivative of f by t in (7) can be represented as

$$\frac{\partial f}{\partial t} + \frac{\not{\partial} f}{\not{\partial} t}\,, \tag{8}$$

where the left derivative SD calls the 'partial derivative' (which is explicit) and the right derivative SD calls the 'copartial derivative'. This copartial derivative (or implicit derivative) is nearly always just using some form of the chain rule, tailored to the situation.

In SD, I refer to $\dfrac{\partial f}{\partial t}$ as the 'partial derivative by t', which is defined similarly to the standard partial derivative, except it allows differentiation of non-primitive functions. I also refer to it as an 'explicit derivative'. Thus,

$$\left(\frac{\partial f}{\partial t}\right)_{\text{exp}} + \left(\frac{\partial f}{\partial t}\right)_{\text{imp}}\,. \tag{9}$$

If we abstract the derivatives from the function in (8), what's left? This is:

$$\frac{\partial}{\partial t} + \frac{\not{\partial}}{\not{\partial} t}\,. \tag{10}$$

What is this? In the case of it being applied to (7), it's obviously the total derivative of ordinary differentiation. And in that case we could write

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \frac{\not{\partial}}{\not{\partial} t}\,. \tag{11}$$

[1]In the wild, a person is far more likely to encounter a nonprimitive function than a primitive function.

But we should anticipate more general situations, such as when $f : \mathbb{R}^n \to \mathbb{R}^m$, which means that f takes in n variables and renders m scalar functions f_j $(j = 1 \ldots m)$. Thus, we'll denote the general total derivative (also referred to as the *deltal derivative*) as

$$\frac{\delta}{\delta t} = \frac{\partial}{\partial t} + \frac{\not\partial}{\not\partial t}\,. \tag{12}$$

3 A Second Look

So here we are again, back at Cauchy's functional equation, where we are asked to solve for $f(x)$ of the given relation

$$f(x + y) = f(x) + f(y)\,, \tag{13}$$

where we take this f to be differentiable over the real line.

Okay, before we worry about the differentability of f, let's look at it algebraically. There are solutions of the Cauchy's functional equation from the rationals to the rationals which do not employ differentiation. The solutions can be cavalier about the domains and ranges because it won't matter. But when differentiating, we really need to be careful about domains and ranges.

So, since we are looking for $f(x)$, we have to assume the obvious: that

$$f : \mathbb{R} \to \mathbb{R}\,. \tag{14}$$

But what do we make of the function $f(x + y)$? Remember that x and y are treated as independent variables. This is an abuse of notation, which in algebra is tolerable, but in calculus is not. When we go to use the chain rule on $f(x+y)$, we're going to have to use it correctly. So, let's set it up correctly.

$$f(x + y) \to f(u(x, y))\,, \tag{15}$$

where the variable u was chosen arbitrarily, and

$$u : \mathbb{R}^2 \to \mathbb{R}\,, \quad u(x, y) = x + y\,. \tag{16}$$

Now we know what we're dealing with.

So, here's the fundamental question: Say we begin with a true equation built of differentiable functions. Which derivative can be taken across such an equation and (practically) guarantee the resulting equation to be true? Your choices are:

a) a total derivative, b) an explicit derivative,
c) an implicit derivative, c) a partial derivative.

And, no, this is **not** a trick question! In fact, it's a very practical question. But I feel that for most poeple, the derivative to use looks like a partial derivative, but it can function in any way that the user wants it to in any given situation.

I call that approach to differentiation "unstructured." But I have learned that a lot of people are quite comfortable with that approach.

If you watch someone at a board talk about taking a derivative without telling you exactly what derivative it is or without writing down a symbolic representation of the derivative, then you might not be able to be sure what is going on. Now, I'm not intensionally picking on SyberMath for doing this, because practically everyone does this at some point — including myself if I'm not careful, because stopping the flow to fill in all the details for the functional dependencies can be so tedius. We sort of hand-wave our way through the differentiation.

Now, it turns out that in the case of SyberMath and this problem, the means to the answer are correct. It's my self-appointed job right now to show the details of why that method works by supplying in some omitted details.

SyberMath said that when going from (1) to the pair of equations he needs, he started with taking the derivative of (1) by x while holding y constant. In the common parlance of physics, that would be called taking an explicit derivative.

So, let's employ the procedure used in the YouTube presentation (more or less)

$$\frac{\partial}{\partial x} f(x+y) = \frac{\partial}{\partial x}[f(x) + f(y)\,]\,, \tag{17a}$$

$$\frac{\partial}{\partial y} f(x+y) = \frac{\partial}{\partial y}[f(x) + f(y)\,]\,. \tag{17b}$$

But x and y are completely independent of each other, so, setting $u \equiv x + y$, we get derivative by x and the LHS of (17a) becomes:

$$\big[\frac{\partial}{\partial u} f(u)\big]\frac{\partial u}{\partial x} = f'(u)\frac{\partial}{\partial x}(x+y) = f'(u)\,, \tag{18}$$

and the RHS simplifies to

$$\frac{\partial}{\partial x}[f(x) + f(y)\,] = \frac{\partial}{\partial x} f(x) = f'(x)\,, \tag{19}$$

where the prime will always denote differentiation by the function's argument. (From the viewpoint of SD, it's a strange thing that a 'partial' derivative will end up being used in a chain-rule expansion.) Now, if we repeat this process by y instead of x, we get the couple

$$f'(u) = f'(x)\,, \tag{20a}$$

$$f'(u) = f'(y)\,. \tag{20b}$$

Therefore, from these last two equations, we have that

$$f'(x) = f'(y)\,. \tag{21}$$

4 Once More, but with SD

This time we'll use SD. Again, the given relation is

$$f(x+y) = f(x) + f(y)\,, \tag{22}$$

but it is now replaced by the more tedius form

$$f(u(x,y)) = f(x) + f(y)\,. \tag{23}$$

We need to differentiate twice, once by x and once by y, yielding,

$$\frac{\delta}{\delta x} f(u(x,y)) = \frac{\delta}{\delta x}[f(x) + f(y)]\,, \tag{24a}$$

$$\frac{\delta}{\delta y} f(u(x,y)) = \frac{\delta}{\delta y}[f(x) + f(y)]\,. \tag{24b}$$

From these we get

$$\frac{\not\partial}{\not\partial x} f(u(x,y)) = \frac{\partial}{\partial x}[f(x) + f(y)]\,, \tag{25a}$$

$$\frac{\not\partial}{\not\partial y} f(u(x,y)) = \frac{\partial}{\partial y}[f(x) + f(y)]\,. \tag{25b}$$

On simplifying, we have that

$$f'(u)\frac{\partial}{\partial x}(x+y) = \frac{\partial}{\partial x}[f(x)]\,, \tag{26a}$$

$$f'(u)\frac{\partial}{\partial y}(x+y) = \frac{\partial}{\partial y}[f(y)]\,. \tag{26b}$$

And, as before,

$$f'(u) = f'(x)\,, \tag{27a}$$

$$f'(u) = f'(y)\,. \tag{27b}$$

Therefore, from these last two equations, we have that

$$f'(x) = f'(y)\,. \tag{28}$$

5 The Final Lap

I'll complete the proof from here, but the point of this article has already been achieved.

Now, x and y are distinct variables, therefore for this last equation to be true, it must equal a constant, which we'll call k.

$$f'(x) = f'(y) = k\,. \tag{29}$$

On integrating, we get

$$f(x) = kx + c_x \,, \tag{30}$$
$$f(y) = ky + c_y \,. \tag{31}$$

Going back to (1), let $x = y = 0$, and we get[2]

$$f(0) = f(0) + f(0) = 2f(0)\,. \tag{32}$$

From this we conclude that $f(0) = 0$, from which we conclude that $c_x = c_y = 0$. Essentially, Eqs. (30) and (31) are the same equation, so we'll just use x as the independent variable. We end up with the final answer as

$$f(x) = kx \,, \tag{33}$$

where k is an arbitrary real number.

6 Epilogue

I have made the claim that one can always take a total derivative across a true equation and get back a true equation.[3] I can't prove this, but I haven't yet found a contradiction to this claim. However, let's look at an equation that looks to provide us with a contradiction. Let's look at again at our own humble Equation (4)

$$f(x) = g(x, y(x))\,. \tag{34}$$

Normally, one would differentiate this equation by x, but that happens if we take the total derivative by y across this equation? Then we have that

$$\frac{\delta}{\delta y} f(x) = \frac{\delta}{\delta y} g(x, y(x))\,. \tag{35}$$

On expanding, we have that

$$\left(\frac{\partial}{\partial y} + \frac{\not\partial}{\not\partial y}\right) f(x) = \left(\frac{\partial}{\partial y} + \frac{\not\partial}{\not\partial y}\right) g(x, y(x))\,. \tag{36}$$

Now, on the LHS, since f does not have y as a variant, then $\dfrac{\partial f}{\partial y} \equiv 0$. But since x is not dependent on y then $\dfrac{\not\partial f}{\not\partial y} \equiv 0$, so the LHS is identically zero. Is that right? However, on the RHS we have a surviving term, namely, $\dfrac{\partial g}{\partial y} \not\equiv 0$.

[2]Like I said before, in doing algebra, we can treat f as a function on $\mathbb{R}^2$ or on $\mathbb{R}$ as suits us at the moment. Algebra is far more tolerant of being 'unstructured' than is differentiation. That's what makes this particular problem so interesting because we can compare the two usecases.

[3]Again, we assume that the relevant limits exist.

Well, is this a problem? And if so, is there a resolution? On the face of it there definitely **is** a problem. But there is also a resolution, which lies in the **Implicit Function Theorem**, which states, in brief, that under usually very mild restrictions a function of a variable can be reversed for the purpose of differentiation. So, in the case of $y = y(x)$, if y is differentiable by x then on some set (perhaps small) set $x = x(y)$ exists and is differentiable by y. Thus (34) should be better expressed as

$$f(x(y)) = g(x, y(x)) \,, \tag{37}$$

on some domain of applicability. Then, on differentiation by y we get

$$\frac{\not\partial}{\not\partial y} f(x(y)) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial y} = \frac{\partial g}{\partial y} \,, \tag{38}$$

which can be rewritten as

$$\frac{\partial f}{\partial x}\frac{dx}{dy} = \frac{\partial g}{\partial y} \,. \tag{39}$$

And we see that the chain rule should usually (always?) come out of the copartial derivative.

So, if the need for reform of so-called ‘partial differentiation’ is so obviously needed, why is it resisted so much? It’s all in the psychology of it. Those against SD or any other form of reform in differeniation, don’t see any need for it. To them, a derivative is whatever they want it to mean in the moment. It is exactly what they say it is — nothing more and nothing less. They may not admit to that, but that’s the way they treat it when they use it, expecially when they use it in a hurry.

P.S. I have many articles on Structured Differentiation on my website on the math page. Please feel free to read them.