My Defence of Structured Differentiation from 1999, 4

P. Reany

January 20, 2024

Abstract

Here I review my defence of Structured Differentiation which I had made in 1999 on sci.math.

1 Introduction

In 1999, I made a defence on sci.math of my notation in Structured Differentiation (SD), which is a notational system I invented to deal with the many confusing (and well-recognized) features that commonly arise in multi-variable calculus. A mathematician on the newsgroup thought he should counter my claims and I'll present his arguments, and my counterarguments. The reader can decide the merits of my system for him or herself.

I think it will become obvious to the reader that the reason partial derivatives is a confusing subject is simply because it employs too few symbols to chase too many concepts. All SD does is to add in a couple more symbols to better distribute the cognitive workload.

Of all the mathematics subjects I've published on in the AJNP the most controversial one is what I call *Structured Differentiation* (SD), which reorganizes and reformulates the so-called theory of "partial differentiation." "Defender" is an alias for a mathematician that defended the status quo for doing so-called partial differentiation [as it was commonly accepted at that time] against my presentation of SD (I have interjected "editorial" comments within square brackets.):

2 Defender's Reply (18 November)

Subject: Re: partial derivation
Date: 18 Nov 1999 15:21:21
From: Defender
Newsgroups: sci.math
In article <3833420F.2390EO37> Patrick Reany writes:

> [snip]
>> Many times I have converted transport-like equations (or similar
>> things) into technically accurate notation. The meaning was not
>> changed, but it was certainly clarified, to great benefit both
>> conceptually and computationally. That you apparently are unable to
>> do so without "changing the meaning of the equation" suggests to me
>> (again) that your problems are with understanding functions, not with
>> computing partial derivatives.
>
> If you had any physicists siding with you up to this post, you just
> lost them. Since you say you have no difficulty *using* 'partial

> derivatives,' suppose you define what one is for us.

Let $f : R^n \to R$ be a function of the n variables $(x_1, \ldots, x_n)$. The
partial derivative of f wrt x_1 (if it exists) at the point $(x_1, \ldots, x_n)$ is

$$\lim_{h \to 0} \frac{f(x_1 + h, x_2 \ldots, x_n) - f(x_1, x_2, \ldots, x_n)}{h} \,. \tag{1}$$

I'll denote this as $D_1 f(x)$. (Of course this presupposes understanding of the limit.)

> [snip]
> However, you felt compelled to introduce the function $h(x)$ (which
> stands for the variant of G) which is OK, but my solution did not
> need to introduce any new functions.

I didn't *introduce* the function h, I *labelled* it. It was (implicitly) introduced when we went from $(G(x_1, x_2, y)$, $f(x_1, x_2)$ to $(y = f(x_1, x_2), G(x_1, x_2, f(x_1, x_2))$.
(We went from $R^3 = \{(x_1, x_2, y)\}$ to the graph $\{(x_1, x_2, f(x_1, x_2)\}$ of the function f.)

The computation could be done without labelling this function h.
However, in many cases it is beneficial to label (if not introduce) such functions,
simply for the sake of explicitly clarifying dependencies and such.

> Why didn't you define what this D of yours is. (Not for me, for the
> beginner students.) Since you didn't, I will. The D operator (in
> this context) means to differentiate the function with respect to
> the function's argument.

I assumed that anyone who was still paying attention is familiar with the definition of
the derivative. Given a function $f : R^m \to R^n$, the derivative (if it exists)
of f at a point x (in R^m) is a linear operator $Df(x)$ satisfying

$$\lim_{|h| \to 0} \frac{I f(x + h) - f(x) - Df(x)hl}{|h|} = 0 \,. \tag{2}$$

Theorem: If the derivative exists, it is unique. In the "standard coordinate system"
$(x_1, ..., x_m)$ on R^m and $(y_1, \ldots, y_n)$ on R^n, $Df(x)$ can be represented
in the obvious induced bases as the $(m \times n)$ matrix with entries $(D_i f_j(x)$, where
$f(x) = (f_1(x), \ldots, f_n(x))$.

> And while we're in the spirit of presenting logical justifications of the
> processes of differentiation, what differential operator did you
> apply to both sides of
>
> $g(x) = G(h(x))$
>
> to get
>
> $Dg(x) = DG(h(x))Dh(x)$?

This is a direct quote of part of the statement of the chain rule.

[What if $g(x) = G(x, h(x))$?]

Theorem: If h is differentiable, G is differentiable, and the composition $(G \circ h)$ (defined by $(G \circ h)(x) = G(h(x)))$ is defined, then it is differentiable, and $D(G \circ h)(x) = DG(h(x))Dh(x)$. Given the definition of g, namely, $g(x) = G(h(x))$, it is immediate that g is differentiable, with its derivative given by the above formula.

> This is one of my biggest irritations against conventional ‘partial
> differentiation’ legerdomain. Either the operator is never
> specifically identified or it’s identified as a ‘partial derivative’
> which is supposed to be interpreted in context.

“The Operator” which I applied was the derivative. I then expressed
the various derivatives in terms of the matrices as represented in the
standard coordinate systems on R^2 and R^3.

[But if, by the defintiton you gave of a partial derivative, that it can only be applied to a primitive function, then what ‘derivative’ is to be applied to $G(x, h(x))$ and how is it to split? What does it look like? What is it called?]

> Context may make it total, explicit, or implicit. Good luck! (Too
> bad for beginner students though.)

Too bad for the student who blindly starts computing partial
derivatives without thinking about what is a function of what.

[This has been Defender’s position from the beginning: *The only thing that’s tricky is functional dependence*. Defender knows quite well that I have not advocated blind differentiation! What I have argued is that functional dependence is such a tricky concept for students that they could do without the Byzantine notational conventions commonly in use in textbooks!]

> [snip]
> Remember that SD was invented to be a learner’s formalism [...].

I would *very* much prefer that the learner learn and understand
things such as functional notation, composition of functions, etc.,
rather than learning formalism (without theory) which (perhaps) aids
computation in some set of problems.

[I assure you that in my many SD papers is plenty of theory. But I didn’t think that that theory was appropriate on this thread. But why? Because I was not concerned about the techicalities about when a differentiation can be applied to a functional equation, but when it can be applied, what forms do the derivatives takes and what names should we give them.]

>> Compute the copartial wrt t of $f(t, x) = t + x$. I get $x'(t)$.
>
> In SD:
> $\not{\partial}_t f(t, x) = \partial_x f(t, x)\delta_t x = \partial_x(t + x)\delta_t x = \delta_t x$
>
> and this is as far as I am allowed to go. To say that $\delta_t x$ is
> equal to $x'(t)$ requires you to make some tacit assumption about the
> functional dependence of x on t, which is unjustified by the
> information you gave. Are you assuming that $x = x(t)$? No, you can’t
> because you don’t like that notation.

OK, you caught me. (I said it was bad, I didn’t say I never do it.) :-)

Let me rephrase: suppose $x = X(t)$. Compute the copartial of $f(t, x) = t + x$.
Then the answer is $X'(t)$. Similarly, $y = Y(t)$; the copartial of $g(t, y) = 2t + y$ is $Y'(t)$.
Now let $X(t) = 2t$, $Y(t) = t$, and evaluate the above copartials; the answers

are respectively 2 and 1. However, $f(t, x) = g(t, y)$, and this discrepancy (2 vs 1) **is what I mean** by the copartial not being well-defined; more precisely, it is not invariantly defined. The exact same statement is true of the partial derivative: its value depends on the coordinate system.

I have no problem with things of this nature; the problem is with your that replacing something standard by something non-standard, which has the same problems but buried one level deeper, is an improvement.

[I have already shown Defender that the parametric split of the "derivative" in SD into an explicit and implicit part is both well-defined and simpler than his method of invoking the chain rule (which is really just the implicit part of the derivative anyway). Furthermore, Defender himself uses the parametric split frequently within this thread. The parametric split is one of the most common features of "partial differentiation" to be found all over the literature.]

> [snip]
> You aren't allowed to apply an implicit derivative operator to two
> sides of an equation! In my first post I showed how disobeying this
> central rule can get you into trouble when differentiating $F(x, y, z) = 0$
> explicitly by x, say, getting wrongly that $\partial_x F = 0$.
> Remember ?

Yes, I remember. I remember thinking "blatant misunderstanding or intentional deception".

If $F(x, y, z) = 0$ for all x, y, z (in a neighbourhood of a point), then $\partial_x F = 0$. On the other hand, if you are using $F(x, y, z) = 0$ to define a set of points (namely, the set of points which satisfy the equation), then you must recognize that the equation is true only at certain values (x, y, z). Generically, the equation will be false, and if you want to apply the "partial derivative operator" ∂_x to the equation, you will get something which is generically false.

What you can do, though, is appeal to the implicit function theorem, which says (in part, and given additional hypotheses) that the local solution of the equation $F(x, y, z) = 0$ is given by $z = g(x, y)$ .Then, by construction, $F(x, y, g(x, y)) = 0$ will be true for all x, y (in a neighbourhood), and then applying the "partial derivative operator" ∂_x to this always-true equation will give another true equation:

$$D_1F(x, y, g(x, y)) + D_3F(x, y, g(x, y))D_1g(x, y) = 0. \quad \text{[Bingo. The "parameteric split"!]} \qquad (3)$$

I have seen many students "doing implicit differentiation" without stopping to think about what they're doing, specifically, without thinking about what is a function of what. (I'm sure you've seen students who write down, as the result of an implicit differentiation, what "dy/dx" is, without realizing that for some reason y is a function of x.)

[I cannot by fiat change the bad habits of human beings, but by fiat I can change the lousy habits of the conventional formalism for so-called "partial differentiation." The very name "partial differentiation" contributes to the subject's power to mislead and confuse students. The subject is really **generalized total differentiation**!]

> [snip]
> It has nothing to do with the copartial being well defined or not,
> and everything to do with the copartial derivative being an

> incomplete equational operator!

I'm not familiar with this terminology, "incomplete equational operator". I'm not sure I want to be. **I agree with your point about confusions arising from standard notation; your method works computationally** but in my opinion does little towards aiding understanding and what it doesn't aid with it; buries deeper. Obviously you disagree. Perhaps we come from – different perspectives (theoretical understanding vs computational efficiency or something). As this discussion seems to be going nowhere, I'm happy to drop it if you are.

[Obviously, Defender believes that the subject had been thoroughly covered at this point, but I disagreed.]