My Defence of Structured Differentiation from 1999, 7

P. Reany

January 21, 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 My Reply (19 November)

Subject: Re: partial derivation

Defender wrote: Like I said, I'm happy to drop it if you are. :-)

Please feel free to not reply to this post, but I wish that you do reply one last time. I don't think that we've reach closure on the subject quite yet, though all the hard work is done. It shouldn't take much of your time to tie up a couple loose ends from your last post. Either way I will not post again in this thread.

You said two things that have me *really* confused.

First I will reproduce your definition of a partial derivative: Let $f : R^n \to R$ be a function of the n variables $\{x_1, ..., x_n\}$. The partial derivative of f wrt x_1 (if it exists) at the point $(x_1, \dots, x_n)$ is

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

(Of course the variables are real.) Then, after I said that your definition ignores any functional dependence of one of the variants of f on any other one, you said: By definition of "$f : R^n \to R$", the only possible dependence of the value

of f on x_1 is explicit. An expression such as "$f(x, y, g(x, y))$" is NOT
a function $R^3 \to R$.
> So, even if f should vary by x_1 through x_2, say,

This violates the hypothesis that f is a function with domain R^n.

I assume you're saying that all the variants (arguments) of f are mutually independent of each other, right? (SD refers to such a function as 'primitive.') I went back to Buck (*Advanced Calculus*, 3rd ed, p. 23) and he defines a function F from R^n to R^m by $y = F(x)$, where $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_m)$, and $y_i = f_i(x_1, \ldots, x_n)$ where $x_1, \ldots, x_n$ are n real variables. Buck doesn't explicitly say that these n variables are mutually independent, so if they are, I should be forgiven this transgression. It's hard to second guess a definition.

Defender, you said that the real confusion lies in this functional dependence stuff, so here's an opportunity to help a confused soul out. I'll come back to this shortly.

>>> "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.
>
> What Defender is saying is that conventional partial differentiation
> has some invisible operator, called 'the derivative,' which has been
> applied, out of sight of the beginner, and then we rely on the
> result of a theorem to write the derivative.

No, what I am saying is that "the derivative" can be (and is) defined
without reference to a choice of coordinate system. I seriously
question the quality of any vector-calculus type of course which does
otherwise (except that it's not necessary to actually talk about "R^n
without a coordinate system"). Of course a lot of things are most
easily computed once a coordinate system is chosen; by choosing
coordinates (or using the standard ones) an almost trivial theorem
tells us how to write down the derivative.

>> Oh, there are so many questions I'd still like to ask Defender, like:

> What symbol represents this invisible 'derivative'?

Usually the derivative of f is represented by Df. What I call delta you call 'D'. Now we're finally getting somewhere.
>> And what is its definition?

I already gave it, but here it is again: if it exists at x, it is a linear operator $Df(x)$ such that

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

(I may have put an extra $|.|$ in the definition I gave before.) I take it that in $Df(x)$, the variants of f are also mutually independent of each other, right? I.e., f is primitive? This is an implicit definition, but an almost trivial theorem tells
you it is unique, and (as above) an almost trivial theorem tells you
how to write it down.

> And how does it relate to the total derivative of ordinary
> differential calculus when there is only one independent variable?

It is the same thing: let $f : R \to R$, and suppose $Df(x)$ satisfies the above definition. Then
$\lim_{h\to 0} |(f(x+h) - f(x) - Df(x)h)/h| = 0$
$\lim_{h\to 0}(f(x+h) - f(x) - Df(x)h)/h = 0$
$\lim_{h\to 0}[(f(x+h) - f(x))/h - Df(x)] = 0$
$\lim_{h\to 0}(f(x+h) - f(x))/h = Df(x)$;
the converse is equally easy. This is the kind of stuff that puts your 'D' in relation to the ordinary derivative. It is useful to students. It is useful for computation too.

>> How does it relate to our explicit derivative defined above?

Since I claim that the "explicit derivative" which you use is unnecessary, this last question is irrelevant if not meaningless.
If by "explicit derivative" you mean the definition of partial derivative which I gave, $Df(x)$ is represented by $[D_i f_j(x)]$.
Where f and f_j are primitive??
> .... But how
> committed are you to maintaining that a partial derivative is an
> explicit derivative?

Since I have no need for what you call an "explicit derivative", I am very happy to discard it entirely and work only with what I call "the derivative" and (given a coordinate system) "a partial derivative". (Though to be truthful, I often work with differentials, as in differential forms on manifolds; these can be interpreted as a generalization of "the derivative".)

[Except that Defender used an explicit derivative by writing $D_1F(x, y, g(x, y))$, which is the same as $\frac{\dot\partial}{\partial x}F(\dot x, y, g(x, y))$, where the overdot indicates how precisely the differentiation is to be carried out. The explicit derivative of F by x is what you get by differentiating F by x, while holding $g(x, y)$ constant. You can instead insist that what Defender has written down is merely the derivative of F by the first slot, but that amounts to the same thing. It amounts to an explicit derivative: All other variants are being held constant. Physicists have understood this from since forever.]

OK, I need to rephrase my question. How committed are you to maintaining that a partial derivative acts only on primitive functions? (Be patient with me. Remember, I'm just a confused soul.)

> My guess is that your notion of "invariantly defined" and mine of
> "complete equational operator" is the same thing. An operator is
> said to be "complete equational" if, when applied to any equation,
> it *always* maintains the equality.

As you pointed out, this was a poor choice of words on my part. Now that you have defined what you mean by "invariantly defined," I agree with you all the way. We start with the "invariantly defined" derivative (what I call 'total derivative'). Then we extract an explicit representation of it suitable for computation in specific coordinate systems. SD was invented partly to facilitate making those computations, and you admitted that it has at least some merit in that. What I want is to complete the task of making it rigorous. I don't see any fundamental difference between what you do and what I do when you perform computations using the implicit function theorem, though some of your notation still confuses me.

Anyway, I'll define an operator to be "complete equational" if, when applied to any true equation A (that it's allowed to be applied to) returns a true equation B which is implied by A. I agree with you that the (total) DERIVATIVE is a complete equational operator because it is invariantly defined. (If you still don't like this definition it can be ignored.)

>>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: $D_1F(x, y, g(x, y)) + D_3F(x, y, g(x, y))D_1g(x, y) = 0$.

I do apologize for the misunderstanding here. I'm not sure how it occurred, but I meant no criticism of the equation you derived or of your use of the implicit function theorem. I was questioning calling the derivative 'partial,' and I still do but for different reasons now.

> In SD this ghostly 'derivative' that Defender refers to comes out of
> the shadows.

This "ghostly derivative" is THE DERIVATIVE. It is the fundamental
object in all of differential calculus. OK, I am really confused here. First you say that you're applying the'partial derivative' to $F(x, y, g(x, y)) = 0$, then you say it's THE DERIVATIVE you applied. In SD I'd write this operator as δ_x (or δ_1 if you prefer), but I don't see anyway to write this in your notation. If I use D_1, I get into trouble because you already have $D_1F(x, y, g(x, y))$ in the expansion. (I'm assuming that D_1 is the same as ∂_1 in your notation, right?) In other words, I fear to write

$$D_1F(x, y, g(x, y)) = D_1F(x, y, g(x, y)) + D_3F(x, y, g(x, y))D_1g(x, y)\,. \tag{3}$$

Again I'm stuck with an expanded derivative of $F(x, y, g(x, y))$, but no clear operator (i.e., a specific typographical depiction of the operator) that acted on it. In any case you certainly didn't use the DERIVATIVE because that would have given you the three-component vector, $\delta F = (\delta_1F, \delta_2F, \delta_3F) = (\delta_1F, \delta_2F, \partial_3F)$. (You'll have to forgive me for now that the only notation I can use to unambiguously represent what I'm saying is in SD.) In SD I would write the last equation as

$$\frac{\delta}{\delta x}F(x, y, g(x, y)) = \frac{\partial}{\partial x}F(x, y, g(x, y)) + \frac{\partial}{\partial z}F(x, y, g(x, y))\frac{\partial g(x, y)}{\partial x}\,. \tag{4}$$

And reducing F to primitive form by introducing $G(x, y) = F(x, y, g(x, y))$ doesn't solve the problem of what form the operator takes on the RHS. For comparison, in SD I can write $\frac{\delta}{\delta x}G(x, y) = \delta_xF(x, y, g(x, y)) = 0$ and get a completely unambiguous equation and its expansion. But what form does this take in your notation?

My second confusion comes from your insistence that the 'derivative' and/or 'partial derivative' must ONLY act on primitive functions (because those are the only objects you have defined them to act on). I find the following confusing for two reasons: The term $D_1F(x, y, g(x, y))$ is the *value* of D_1F (where $F : R^3 \to R$) evaluated at the point $(x, y, g(x, y))$ in R^3.

Obviously in '$D_1F(x, y, g(x, y))$' F is not primitive, nevertheless has D_1 (a partial derivative) acting on it, but you defined the partial derivative ONLY to act on primitive functions, right? Furthermore, you said previously that this F is not a function $F : R^3 \to R$, but in this last paragraph you say that it is. I'm lost here.

(You previously said:

By definition of "$f : R^n \to R$", the only possible dependence of the value of f on x_1 is explicit. An expression such as "$f(x, y, g(x, y))$" is NOT a function $R^3 \to R$.)

Please explain.

cheers,

Patrick

P.S. It's ironic to me that if you really believe that THE DERIVATIVE D is the fundamental object in all of differential calculus, it took nearly half a dozen posts and quite a bit of prodding from me before you manifested it in the light of day. With that, my posting to this thread is over.