My Defence of Structured Differentiation from 1999, 3

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*[1] 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 Patrick's reply (17 November)

Subject: Re: partial derivation
Date: Wed, 17 Nov 1999 02:36:50 -0700
From: Patrick Reany
Newsgroups: sci.math

Defender wrote:

> In article <3830EB84.E52F8E66> Patrick Reany
> writes:
> [snip]
>
> It's true that I have taught calculus, but it's also true that I have
> not found anything difficult about computing partial derivatives since
> I learned them in a vector calculus course, as a second-year undergrad.

I didn't either until I came to advanced calculus where the variables had
functional dependencies. It's at this point the a notation is tested to the max.

---

[1] The *Arizona Journal of Natural Philosphy* is a defunct journal I published long ago.

And it was here that I lost all sense of what was going on. I came to blame the notation used. I still do today.

>> As another example of the problems that can occur by this particular
>> operator overloading, consider the case found in Taylor-Mann (*Advanced*
>> *Calculus*, 2nd ed, p 271): “Consider the function $G(x, y)$ as a function of u
>> and y, with $x = f(u, y)$.
>
> Many students will be (understandably) confused at this point. I was
> slightly confused myself for a couple seconds. But look, we haven’t
> even said anything about calculus, never mind partial derivatives!
>
> The challenge here is the poor phrasing “$G(x, y)$ as a function of u and
> y...” which, without reading ahead, is meaningless. It is not until
> we find the “with $x = f(u, y)$” that it is even possible to parse the
> phrase.
>
> Of course, once the phrase has been correctly parsed, the student
> should at the very least be thinking if not writing $G(f(u, y), y)$. The
> more pedantic student, or the student rigorously following the chain
> rule (until said student spots the computational speed-ups that we all
> use), will write (something like) the following composition:
> $(u, y) \to (x, v)$ where $x = f(u, y)$, $v = y$
> $(x, v) \to G(x, v)$.

In fact, the true difficulties of so-called partial differentiation are in the complex functional dependencies that can arise in a system of variables. But the notation should not be a further burden to students, and SD shows that this can be accomplished.

[snip]

> Perhaps my biggest concern with your “SD” notation is that, to me, it
> seems to be another “computational trick” rather than an aid in
> understanding the theory, although perhaps it is slightly better than
> most in that it at least points out that there is an issue. But this
> saving grace is pretty much cancelled out by the introduction of new
> notation which obscures the underlying reality that you are still
> computing a derivative of a function, the only difference being that
> there might be some functional composition in the middle.

[SD is not a trick. The more you seee it in use, the clearer that should become.]

I trust that my example problems below will prove that SD is superior in dealing with real problems taking derivatives of functions.

> Moreover,
> the “copartial” which you introduce is (I think) not a well-defined
> operation, and therefore inevitably will cause more confusion than the
> existing notion of derivative.

The copartial in merely the implicit part of the total derivative.

> (The “total derivative” is a concept

> which is already in use in, for example, the study of PDEs on jet
> bundles.)
>

[Let's agree that if I present any equation, not being obviously false, that it is true on some domain.]

The total derivative is the ONLY differential operator that can be applied across an equal sign and is therefore the central operator of so-called 'partial differentiation,' and in fact of all calculus. Partial differentiation is really generalized total differentiation (GTD), and is, in principle, no different from ordinary differentiation. The only major difference is that in GTD the independent variable may be vector valued rather than scalar valued as in ordinary differentiation. I would think that this insight would be encouraging to students of advanced calculus, since they can the leverage what they've already learned and apply it directly to GTD as a smooth transition. The feeling I had when going between the two was more of jumping off into a dark cloud and leaving behind everything I had already learned. I think it is fair to say that ordinary differentiation and GTD are two subjects separated by a lousy notation. (BTW, the ordered set of independent variables is referred to as the 'fundamental' in SD. One reason for this is because it is much easier in change-of-variable problems to think of having a new fundamental dependent on the old fundamental than to have a new set of independent variables dependent on a set of old independent variables.)

SD makes it clear that virtually all the common expressions of ordinary differential equations can be generalized by merely upgrading scalar variables to vector variables. I cannot demonstrate all this here as it is a lot of material and text format is not a good format for mathematics. But I will show a few specifics below.

The standard conventions of so-called partial differentiation are so ambiguous that they are always amended to help remove the inherent ambiguities. So, instead of writing $W = W(t, x(t))$ and differentiating this to get

$$\frac{\partial W}{\partial t} = \frac{\partial W}{\partial t} + \frac{\partial W}{\partial x}\frac{\partial x}{\partial t}\,. \tag{1}$$

One solution is to add to the derivative a declaration of which variables should be held constant during the differentiation. Thus we get

$$\frac{\partial W}{\partial t} = \left(\frac{\partial W}{\partial t}\right)_x + \left(\frac{\partial W}{\partial x}\right)_t \frac{\partial x}{\partial t} \tag{2}$$

which is common in thermodynamics. And when you extract out the operators you get

$$\frac{\partial}{\partial t} = \left(\frac{\partial}{\partial t}\right)_x + \left(\frac{\partial}{\partial x}\right)_t \frac{\partial x}{\partial t} \tag{3}$$

which is the same (conceptually) as

$$\left(\frac{\partial}{\partial t}\right)_{\text{total}} = \left(\frac{\partial}{\partial t}\right)_{\text{explicit}} + \left(\frac{\partial}{\partial t}\right)_{\text{implicit}}\,. \tag{4}$$

I was first made aware of this 'obvious' concept, not from my calculus books, but from a book on statistical mechanics by Kerson Huang. On page 124 we find the statement:

“Let f be a solution of the Boltzmann transport equation. It may depend on time explicitly as well as implicitly through the time dependence of local density, velocity, and temperature. Thus we may decompose the time derivative of f into two terms:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial t}_{\text{explicit}} + \frac{\partial f}{\partial t}_{\text{implicit}} .\text{”} \tag{5}$$

Note that with the f removed, we are left with three distinct differential operators related as follows:

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t_{\text{explicit}}} + \frac{\partial}{\partial t_{\text{implicit}}} . \tag{6}$$

I simply decided to use in SD a δ for the total derivative, a ∂ for the explicit derivative and a ‘del with a forward slash through it’ $\not\partial$ for the implicit (copartial) derivative. There’s nothing new in SD except for its abhorrence for the kludgy notations used previously.

More common with mathematicians is to ‘reduce to primitive form,’ as I call it. That is to replace a function having implicit dependence on a variable with a function that has only explicit dependence on the variable. For example, the equation $W = W(t, x(t))$ would have the function W reduced to primitive form by writing $f(t) = W(t, x(t))$, say. On differentiating this by t we get

$$\frac{\partial f}{\partial t} = \frac{\partial W}{\partial t} + \frac{\partial W}{\partial x}\frac{\partial x}{\partial t} \tag{7}$$

which has the advantage of dropping the variable subscripts which indicate which variables are held fixed during the differentiation. But it suffers from the redundant introduction of extra variables. But this isn’t the only problem with this reduction. In cases where in physics the implicit vs explicit dependence of a variable have very different meaning and physical origins (such as in the case of the Boltzmann transport equation above and also for Lagrangians); such equations do not like being contorted into primitive form which, though allowable formally, would change the meaning of the equations.

Let me demonstrate just how clean and minimalist SD is on a real problem.

Consider a problem from Taylor/Mann, *Advanced Calculus* (2nd ed) pg (?)

If $G_1(x_1, x_2, y)$, $G_2(x_1, x_2, y)$, and $f(x_1, x_2)$ are given, and if

$$g_i(x_1, x_2) = G_i(x_1, x_2, f(x_1, x_2)) \quad (i = 1, 2) , \tag{8}$$

then show that[2]

$$\det\left[\frac{\partial(g_1, g_2)}{\partial(x_1, x_2)}\right] = \det\left[\frac{\partial(G_1, G_2)}{\partial(x_1, x_2)}\right] + \frac{\partial f}{\partial x_1}\det\left[\frac{\partial(G_1, G_2)}{\partial(y, x_2)}\right] + \frac{\partial f}{\partial x_2}\det\left[\frac{\partial(G_1, G_2)}{\partial(x_1, y)}\right] \tag{9}$$

where an expression such as $\partial(g_1, g_2)/\partial(x_1, x_2)$ depicts a Jacobian (or in this case, a Jacobian-like) matrix, but not its determinant, as is commonly done.

Proof: First we note that g_i is the reduction of G_i to primitive form. (This reduction is NOT necessary in SD though.) Let $g = (g_1, g_2)$, $G = (G_1, G_2)$,

---

[2] That is, g_i is the reduction of G_i to primitive form to the intent that $\frac{\delta g}{\delta x} = \frac{\partial g}{\partial x}$.

and $x = (x_1, x_2)$. Now, since $g(x) = G(x, z(x))$, we can differentiate both sides by x to get

$$\frac{\delta g}{\delta x} = \frac{\delta G}{\delta x}\,. \tag{10}$$

(Again: The total derivative is the central derivative of calculus because it's the only one that can be taken across an equal sign with impunity [in the general case].)

Then we get

$$\begin{aligned}\det\left[\frac{\partial(g_1, g_2)}{\partial(x_1, x_2)}\right] &= \det\left[\frac{\partial g}{\partial x}\right] = \det\left[\frac{\delta g}{\delta x}\right] \\ &= \det\left[\frac{\delta G}{\delta x}\right] = \det\left[\frac{\delta(G_1, G_2)}{\delta(x_1, x_2)}\right] \\ &= \frac{\delta G_1}{\delta x_1}\frac{\delta G_2}{\delta x_2} - \frac{\delta G_1}{\delta x_2}\frac{\delta G_2}{\delta x_1}\,,\end{aligned} \tag{11}$$

and the result follows by expanding the delta (total) derivatives and regrouping. (After I presented this problem/solution to a group of upper-division math majors I received an ovation for its simplicity and conciseness. I recount this event solely to point out that if SD were just another hohum notational system then why the excitement in it from a group of cynical math majors?)

Now I will address the accusation that the copartial is not well defined in SD. I will demonstrate that it is well defined in a special case and trust that it is obvious how to generalize.

Let $f = f(x, y, z)$. Consider the infinitesimal δf. By expanding this through its variant (i.e., using the chain rule) we get

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z\,. \tag{12}$$

Now we divide through by δx, say.

$$\begin{aligned}\frac{\delta f}{\delta x} &= \frac{\partial f}{\partial x}\frac{\delta x}{\delta x} + \frac{\partial f}{\partial y}\frac{\delta y}{\delta x} + \frac{\partial f}{\partial z}\frac{\delta z}{\delta x} \\ &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\delta y}{\delta x} + \frac{\partial f}{\partial z}\frac{\delta z}{\delta x}\,,\end{aligned} \tag{13}$$

where the first term on the RHS is the partial derivative and all the rest is the copartial derivative. Of course it's well defined.

Note that

$$\delta f = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y + \frac{\partial f}{\partial z}\delta z = (\nabla f)\cdot(\delta x, \delta y, \delta z) \tag{14}$$

which is another connection to vector calculus.

cheers,

Patrick