Math Diversion Problem 449: Partial Derivatives Using SD

P. Reany

March 7, 2025

> Time and again I come across people who, although very
> bright, and having received advanced mathematical
> education, have a completely confused notion of
> what partial differentiation is.
> — andrewkirk

1 Introduction

This is our fifth journey into so-called 'partial differentiation', as seen through the lens of structured differentiation. We've already touched on the change of variables when differentiating, but this time I want to go into it a bit more formally.

However, that doesn't mean that I'll go into delta-epsilon rigor. Although I want to be accurate in what I report, it's basically just to show how to work most practical problems in the subject. I usually don't concern myself with technicalities in these problems, because technical issues almost never arise. For example, the existence and continuity of partial derivatives will almost always be assumed, especially in problems of physical content. And if I'm using the Implicit Function Theorem, concerns about certain determinants not being zero at a certain points, arise naturally in the solution — depending on what's being asked to solve for.

By the way, it may be useful for the reader to read these articles in the order they are presented, as the terminology and the symbolism I present in them are nonstandard, and require some getting used to.

2 Some Definitions

We define the *fundamental vector* (or just *fundamental*) of the function $\mathbf{f}$ to be the ordered tuple of all the "independent" variables of $\mathbf{f}$. The order of the variables in the fundamental, though arbitrary, must be adhered to throughout any given problem. And the function $\mathbf{f}$ need not be vector-valued; it can be scalar-valued as well.

It should be clear that $\delta\mathbf{x}/\delta\mathbf{x} = \mathbf{I}$ is the identity matrix, if and only if $\mathbf{x}$ is the fundamental or a subset of it. In like manner, we say that for some function f, $\delta f/\delta x = 0$ if an only if f is functionally independent of x.

Similarly, we define the *variant vector* (or just *variant*) of a function to be the ordered tuple of all the variables on which the function is explicitly dependent. If we say 'a variant of f', we refer to just one of those variables.

Warning: In problems in which the independent variables can change, what is the 'new' verses the 'old' fundamental can be quite fluid. So, let context be your guide, but in the end, one may just have to declare the matter to remove ambiguity.

3 Theory for the 'Change of Variables'

Up to this point we have encountered functions that are dependent on variable $\mathbf{x}$ maybe both explicitly and implicitly, in which case by differentiation, we get

$$\frac{\delta\mathbf{f}}{\delta\mathbf{x}} = \frac{\partial\mathbf{f}}{\partial\mathbf{y}}\frac{\delta\mathbf{y}}{\delta\mathbf{x}}\,, \tag{1}$$

where $\mathbf{f} = \mathbf{f}(\mathbf{y})$ and $\mathbf{y}$ is the variant of $\mathbf{f}$, say. There are many situations that require us to extend our theory of SD to include problems in which the fundamental itself will change, that is, will be replaced by some other fundamental.

Now, this is not a new concept even to a calculus student. Say we have the integral

$$I = \int A(x)\,dx \tag{2}$$

and we prefer to integrate with respect to some other variable, such as u. Then we recast this integral as

$$I = \int A(u)\,\frac{dx}{du}\,du\,. \tag{3}$$

This expression $\dfrac{dx}{du}$ is called the *Jacobian* of the transformation of variables from the old variable x to the new variable u. We will soon encounter problems where the old variables and the new variables both are treated as vectors in SD.

I will now present a couple definitions of SD on how to deal with derivatives resulting from a change of fundamental, which I cannot promise that all authors will use.

One Definition of Jacobian matrix

The total derivative of the old fundamental with respect to the new fundamental shall be called the *Jacobian matrix* of the transformation of variables.[1] Often in practice, the total derivatives can be reduced to partial derivatives. The determinant of the Jacobian matrix is called the *Jacobian* of the transformation.

So, besides integration, where does this kind of change of variables come up? For one, when one wants to change coordinates, such from rectangular to polar or in 3D spherical polar coordinates, or back the other way. Another is in the transformation of variables in thermodynamics. We will soon see an example where a thermodynamic equation which is known for one pair of fundamental variables is to be converted to its equivalent for when a new couple of fundamental variables is chosen.

Now, a subtlety arises that hasn't to this point. We got by just fine with using $\dfrac{\partial \mathbf{f}}{\partial \mathbf{y}}$ in Eq. (1), but it may be safer to use a parallel to that which is done is ordinary chain rule. So, if $f(x) = f(u(x))$ then

$$\frac{df}{dx} = \frac{df}{du}\frac{du}{dx}\,. \tag{4}$$

Therefore, if $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{u}(\mathbf{x}))$, it may be safer (that is, more accurate) to write

$$\frac{\delta \mathbf{f}}{\delta \mathbf{x}} = \frac{\partial \mathbf{f}}{\partial \mathbf{u}}\frac{\delta \mathbf{u}}{\delta \mathbf{x}}\,. \tag{5}$$

Rule 1) When one fundamental variable η_i is totally differentiated by a cofundamental variable η_j, the result is

$$\frac{\delta \eta_i}{\delta \eta_j} = \delta_{ij}\,, \tag{6}$$

where δ_{ij} is the Kronecker delta, meaning that when a fundamental variable is totally differentiated by itself, the result is unity, but when it is totally differentiated by a different cofundamental variable, the result is zero; hence, the notion of *variable independence* between different elements in a given fundamental is captured by this rule. Note: To remove ambiguity, any variable in a fundamental set is cofundamental to itself.[2]

Rule 2) In all other cases, i.e., when a new fundamental variable differentiates an old fundamental variable, or vice versa, the deltal derivative usually reduces to an explicit derivative. This is the standard thing to do in thermodynamics because, by default, we regard state variables as having no implicit dependence on other variables (at least this is my knowledge regarding them).[3]

[1] Of course, if both fundamentals are just scalars, then the Jacobian 'matrix' will be just a 1×1 matrix, or, a scalar.

[2] We will consider it axiomatic in SD that the total derivative of any variable with respect to itself is unity.

[3] When a state variable **does** have an implicit dependence on some variable, then we cannot just drop the implicit derivative as we have done in this case.

4 Problem: An Example from Vector Analysis

Let $\mathbf{u}$ be a vector field defined on the open subset $\mathcal{D}$ of R^n, let $\mathbf{x}$ be a vector field defined on the open subset $\mathcal{D}'$ of R^n, and let T be a one-to-one, onto transformation which takes $\mathbf{u}$ to $\mathbf{x}$ by

$$\mathbf{x} = T(\mathbf{u}) = \mathbf{x}(\mathbf{u})\,. \tag{7}$$

Now T^{-1} exists on $\mathcal{D}'$ so that

$$\mathbf{u} = T^{-1}(\mathbf{x}) = \mathbf{u}(\mathbf{x})\,. \tag{8}$$

The *Jacobian matrix* of the transformation is

$$\mathbf{J}_T = \frac{\partial T}{\partial \mathbf{u}} = \frac{\partial \mathbf{x}}{\partial \mathbf{u}}\,, \tag{9}$$

where $\mathbf{x}$ is the 'old' fundamental and $\mathbf{u}$ is the 'new' fundamental. Thus, the *Jacobian* of the transformation to be

$$J_T \equiv |\mathbf{J}_T|\,, \tag{10}$$

where $|\mathbf{J}_T|$ is the determinant of the Jacobian matrix. We will show that

$$\big(\mathbf{J}_T\big)^{-1} = \mathbf{J}_{T^{-1}} \tag{11}$$

and

$$\big(J_T\big)^{-1} = J_{T^{-1}}\,. \tag{12}$$

To prove this we take the convolution of $\mathbf{x}$ through $\mathbf{u}$ and get

$$\mathbf{x} = \mathbf{x}(\mathbf{u}(\mathbf{x}))\,. \tag{13}$$

On taking the total derivative of (13) by $\mathbf{x}$ and simplifying we obtain

$$\mathbf{I} = \frac{\partial \mathbf{x}}{\partial \mathbf{u}}\frac{\partial \mathbf{u}}{\partial \mathbf{x}}\,. \tag{14}$$

But $\partial\mathbf{u}/\partial\mathbf{x} = \mathbf{J}_{T^{-1}}$; therefore, (14) establishes (11). To establish (12) we merely take the determinant of both sides of (14).

5 Transformation on Coordinates: Rectangular to Spherical

Let

$$\left.\begin{array}{l} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{array}\right\} \tag{15}$$

Find ∇r, $\nabla\theta$, and $\nabla\phi$ where $\nabla = \partial/\partial\mathbf{x}$. For example,

$$\nabla r = \left(\frac{\partial r}{\partial x}, \frac{\partial r}{\partial y}, \frac{\partial r}{\partial z}\right)\,. \tag{16}$$

Solution:

One way to solve this problem is to algebraically solve for r, θ, ϕ in terms of x, y, z, then differentiate (since we were given the components of $\mathbf{x}$ by $\mathbf{u}$ in (15)). But that's a lot of 'unnecessary' work. We will take another path. We will instead take the derivative of $\mathbf{x}$ by $\mathbf{u}$, one component at a time (which is easy) and then take the matrix inverse of that result. Of course, this other path will have a significant amount of matrix 'work' to do, but it's more pro forma than the other way. In either case, some form of inversion is necessary here: Either we do the algebraic inversion at the beginning, or we do the matrix inversion at the end.

The gradients we want are just the row vectors of the matrix $\partial\mathbf{u}/\partial\mathbf{x}$ where $\mathbf{u} = (r, \theta, \phi)$. Now, treating $\mathbf{x}$ as the old fundamental and $\mathbf{u}$ as the new fundamental,[4] the Jacobian matrix is

$$\frac{\partial\mathbf{x}}{\partial\mathbf{u}} = \begin{bmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{bmatrix} \tag{17}$$

and the Jacobian is:

$$\left|\frac{\partial\mathbf{x}}{\partial\mathbf{u}}\right| = r^2\sin\theta\,. \tag{18}$$

Where the Jacobian is nonzero,[5] we have from (14)

$$\frac{\partial\mathbf{u}}{\partial\mathbf{x}} = \left(\frac{\partial\mathbf{x}}{\partial\mathbf{u}}\right)^{-1}. \tag{19}$$

Therefore,

$$\frac{\partial\mathbf{u}}{\partial\mathbf{x}} = \begin{bmatrix} \nabla r \\ \nabla\theta \\ \nabla\phi \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \dfrac{1}{r}\cos\theta\cos\phi & \dfrac{1}{r}\cos\theta\sin\phi & -\dfrac{1}{r}\sin\theta \\ -\dfrac{\sin\phi}{r\sin\theta} & \dfrac{\cos\phi}{r\sin\theta} & 0 \end{bmatrix}. \tag{20}$$

[4] Which is consistent with (15).

[5] Here we encounter one of those technicalities I mentioned before. To invert a square matrix, one needs to divide through by the determinant of the original matrix. Hence this inverse will be undefined where the determinant is equal to zero.