Math Diversion Problem 469: Partial Derivatives Using SD

P. Reany

March 17, 2025

> 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 [5]

1 Introduction

This is the seventh in a series of articles on how to perform the computations of partial differentiation by use of a formalism call 'Structured Differentiation'. Partial differentiation is tricky enough when we keep the set of independent variables fixed, but it can get much trickier when we allow for a change in the set of independent variables. In SD, we are permitted to refer to an independent variable as a 'fundamental' variable, and to refer to an ordered list of fundamental variables as the 'fundamental'.

I have been criticized for introducing the term 'fundamental' variable into SD, because it seems redundant. Maybe technically it is, but psychologically it isn't. For the longest time I got confused when tackling problems in which the independent variables could change — the problem below is a case in point. I'd ask myself, "How can an independent variable change? It's either independent or it isn't, right?"

Well, actually, one has to be more flexible than that. The following problem helps to reveal the subtleties involved.

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

$$\frac{\delta \eta_i}{\delta \eta_j} = \delta_{ij} \,, \tag{1}$$

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.[1]

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).[2]

2 An Example from Thermodynamics

A certain thermodynamic state can be represented by the differential equation of state

$$\frac{\partial U}{\partial V} - T\frac{\partial P}{\partial T} + P = 0\,, \tag{2}$$

where V and T are the fundamental variables. What, then, is the appropriate form that (2) takes when U and P replace V and T as the fundamental variables? The current problem is from Buck [6], prob. 18, pg. 146.

Let the old fundamental be represented by

$$\boldsymbol{\eta} = (V, T)^t\,, \tag{3}$$

and the new fundamental by

$$\boldsymbol{\eta}' = (U, P)^t\,. \tag{4}$$

For convenience, we also define the state vector[3]

$$\psi = (U, P, V, T)^t\,. \tag{5}$$

Then

$$\psi'(\boldsymbol{\eta}') = \psi(\boldsymbol{\eta}(\boldsymbol{\eta}'))\,. \tag{6}$$

Differentiating this by $\boldsymbol{\eta}'$, we get[4]

$$\frac{\delta\psi'}{\delta\boldsymbol{\eta}'} = \frac{\delta\psi}{\delta\boldsymbol{\eta}}\frac{\delta\boldsymbol{\eta}}{\delta\boldsymbol{\eta}'}\,, \tag{8}$$

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

[2]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.

[3]This so-called 'state vector' ψ is not a state variable in any sense that aligns with thermodynamics. Instead, it refers to the space of all 'fundamental' variables of the problem. I use it because I find it convenient.

[4]The usual form for the chain rule is

$$\frac{\delta\psi'}{\delta\boldsymbol{\eta}'} = \frac{\partial\psi}{\partial\boldsymbol{\eta}}\frac{\delta\boldsymbol{\eta}}{\delta\boldsymbol{\eta}'}\,, \tag{7}$$

but this won't do in this case, because we need total derivatives to guarantee that the total derivative of a variable by itself is 1 and that the total derivative of a variable by its cofundamental (other than itself) is zero. If we just used partial derivatives (in SD), it would be ambiguous.

and simplifying, we have

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} = \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{J} \tag{9}$$

where $\mathbf{J} = \partial\boldsymbol{\eta}/\partial\boldsymbol{\eta}'$. (See Appendix 1 to see the important steps left out here.)

The relations we want can be extracted from (9) by inspection (taking determinants of 2×2 submatrices), yielding

$$\frac{\partial T}{\partial P} = \frac{\partial U}{\partial V}\, J \tag{10a}$$

and (dropping the minus signs), and where $J \equiv \det(\mathbf{J})$. Similarly,

$$\frac{\partial V}{\partial U} = \frac{\partial P}{\partial T}\, J\,. \tag{10b}$$

Solving these for $\partial U/\partial V$, $\partial P/\partial T$ and then substituting the results into (2) yields

$$\frac{\partial T}{\partial P} - T\frac{\partial V}{\partial U} + PJ = 0\,. \tag{11}$$

And we're done.

The technique used to derive (10a) and (10b) from (9) we refer to as *decomposition*, which we describe in brief. From the first and fourth rows of (9), we may write the 2×2 matrix equation[5]

$$\begin{bmatrix} 1 & 0 \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} = \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ 0 & 1 \end{bmatrix} \mathbf{J}\,, \tag{12}$$

and from the second and third rows, we may similarly write

$$\begin{bmatrix} 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \end{bmatrix} = \begin{bmatrix} \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \end{bmatrix} \mathbf{J}\,. \tag{13}$$

Equations (10a) and (10b) follow by taking the determinant of these last two matrix equations, employing the fact that the determinant of a product is the product of the determinants.

3 Appendix 1

Here we fill in some of the missing steps from above problem. The old fundamental is

$$\boldsymbol{\eta} = (V, T)^t\,, \tag{14}$$

[5]Proof given in Appendix 1.

and the new fundamental is

$$\boldsymbol{\eta}' = (U, P)^t \,. \tag{15}$$

We also define the state vector

$$\psi = (U, P, V, T)^t \,. \tag{16}$$

Then

$$\psi'(\boldsymbol{\eta}') = \psi(\boldsymbol{\eta}(\boldsymbol{\eta}')) \,. \tag{17}$$

Differentiating this by $\boldsymbol{\eta}'$, we get

$$\frac{\delta\psi'}{\delta\boldsymbol{\eta}'} = \frac{\delta\psi}{\delta\boldsymbol{\eta}}\frac{\delta\boldsymbol{\eta}}{\delta\boldsymbol{\eta}'} \,, \tag{18}$$

which looks like this

$$\begin{bmatrix} \delta U/\delta U & \delta U/\delta P \\ \delta P/\delta U & \delta P/\delta P \\ \delta V/\delta U & \delta V/\delta P \\ \delta T/\delta U & \delta T/\delta P \end{bmatrix} = \begin{bmatrix} \delta U/\delta V & \delta U/\delta T \\ \delta P/\delta V & \delta P/\delta T \\ \delta V/\delta V & \delta V/\delta T \\ \delta T/\delta V & \delta T/\delta T \end{bmatrix} \mathbf{J} \,, \tag{19}$$

where $\mathbf{J} = \partial\boldsymbol{\eta}/\partial\boldsymbol{\eta}'$ is a 2×2 matrix. Now, let's analyze this last equation in depth.

First, since the total derivative of any variable by itself is unity; then for the derivatives in the 4×2 matrix on the LHS, we have that

$$\delta U/\delta U = 1, \quad \delta P/\delta P = 1 \,, \tag{20}$$

and since U and P are independent of each other,

$$\delta U/\delta P = \delta P/\delta U = 0 \,. \tag{21}$$

Therefore, upon substitution and simplification, we have

$$\begin{bmatrix} \delta U/\delta U & \delta U/\delta P \\ \delta P/\delta U & \delta P/\delta P \\ \delta V/\delta U & \delta V/\delta P \\ \delta T/\delta U & \delta T/\delta P \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} \,, \tag{22}$$

where we used Rules 1 and 2 on page (2). Arguing similarly for the 4×2 matrix on the RHS, we have that

$$\delta V/\delta V = 1, \quad \delta T/\delta T = 1 \,, \tag{23}$$

and since V and T are independent of each other,

$$\delta V/\delta T = \delta T/\delta V = 0 \,. \tag{24}$$

Therefore, upon substitution and simplification, we have

$$\begin{bmatrix} \delta U/\delta V & \delta U/\delta T \\ \delta P/\delta V & \delta P/\delta T \\ \delta V/\delta V & \delta V/\delta T \\ \delta T/\delta V & \delta T/\delta T \end{bmatrix} \longrightarrow \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{25}$$

where, again, we used Rules 1 and 2 on page (2). Now we substitute these simpler expressions back into (19) to get

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} = \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{J} , \tag{26}$$

where $\mathbf{J}$ is equal to

$$\mathbf{J} = \frac{\delta \boldsymbol{\eta}}{\delta \boldsymbol{\eta}'} = \frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\eta}'} = \begin{bmatrix} \partial V/\partial U & \partial V/\partial P \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} . \tag{27}$$

Now, the fast way to get the information we want out of these matrices is to take the determinants of 2×2 submatrices. Well, what is the exact information we need? Remember that the equation we started with is

$$\frac{\partial U}{\partial V} - T \frac{\partial P}{\partial T} + P = 0 . \tag{28}$$

But the whole point of this exercise is to revamp this equation by replacing the partials $\frac{\partial U}{\partial V}$ and $\frac{\partial P}{\partial T}$ by their equivalent partials after we change the fundamental variables.

Let's look first at $\frac{\partial U}{\partial V}$. If we could form the matrix equation made of submatrices formed by using only the first and fourth rows, we'd get

$$\begin{bmatrix} 1 & 0 \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} = \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ 0 & 1 \end{bmatrix} \mathbf{J} . \tag{29}$$

Before I show the reader how to justify this step, I want to show its value. By taking the determinant across this equation, we get

$$\frac{\partial T}{\partial P} = \frac{\partial U}{\partial V} J . \tag{30}$$

From this we get that

$$\frac{\partial U}{\partial V} = \frac{\partial T}{\partial P} J^{-1} . \tag{31}$$

That leaves the replacement for $\frac{\partial P}{\partial T}$. For this we take the second and third rows, to get

$$\begin{bmatrix} 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \end{bmatrix} = \begin{bmatrix} \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \end{bmatrix} \mathbf{J} . \tag{32}$$

On taking the determinant across this equation, we get

$$-\frac{\partial V}{\partial U} = -\frac{\partial P}{\partial T}\, J\,. \tag{33}$$

From this we get that

$$\frac{\partial P}{\partial T} = \frac{\partial V}{\partial U}\, J^{-1}\,. \tag{34}$$

So, on substituting this result and the result from (31) into (28), we have

$$\frac{\partial T}{\partial P}\, J^{-1} - T\frac{\partial V}{\partial U}\, J^{-1} + P = 0\,, \tag{35}$$

which simplifies to

$$\frac{\partial T}{\partial P} - T\frac{\partial V}{\partial U} + PJ = 0\,. \tag{36}$$

This leaves only the justification for my extracting 2×2 submatrices from those 4×2 matrices in Eq. (26). Let's take the case where I extracted the second and third rows. Suppose we multiply on the left Eq. (26) by the 2×4 matrix

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{37}$$

and use the key fact that matrix multiplication is associative to get

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \\ \partial T/\partial U & \partial T/\partial P \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \partial V/\partial U & \partial V/\partial P \end{bmatrix}\,. \tag{38}$$

Now, applying the matrix on the RHS, we get

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{J} = \begin{bmatrix} \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \end{bmatrix} \mathbf{J}\,, \tag{39}$$

It's obvious that by this technique that you can pick out any two rows you want. And this completes the justification.

Now, for an exercise, what information could we get be extracting on rows one and three from Eq. (26)? (Try to do this by inspection only.) Answer:

$$\frac{\partial V}{\partial P} = -\frac{\partial U}{\partial T}\, J\,. \tag{40}$$

Now, a final comment on the SD approach. I trust the reader can see why in Eq. (25) I had to use total derivatives to get 1's and 0's as entries, and thus

$$\begin{bmatrix} \delta U/\delta V & \delta U/\delta T \\ \delta P/\delta V & \delta P/\delta T \\ \delta V/\delta V & \delta V/\delta T \\ \delta T/\delta V & \delta T/\delta T \end{bmatrix} \longrightarrow \begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\,. \tag{41}$$

And this justifies why I had to use the equation

$$\frac{\delta\psi'}{\delta\boldsymbol{\eta}'} = \frac{\delta\psi}{\delta\boldsymbol{\eta}}\frac{\delta\boldsymbol{\eta}}{\delta\boldsymbol{\eta}'}\,, \tag{42}$$

instead of the equation

$$\frac{\delta\psi'}{\delta\boldsymbol{\eta}'} = \frac{\partial\psi}{\partial\boldsymbol{\eta}}\frac{\delta\boldsymbol{\eta}}{\delta\boldsymbol{\eta}'}\,. \tag{43}$$

So, I can just hear a defender of the conventional way of writing these derivatives argue that convention is just fine in how it does this. So, instead of the SD version of (41), one could use the conventional 'partial' derivative the way SD used the deltal derivative and write

$$\begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ \partial V/\partial V & \partial V/\partial T \\ \partial T/\partial V & \partial T/\partial T \end{bmatrix} \longrightarrow \begin{bmatrix} ? & ? \\ ? & ? \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{44}$$

In SD the partial derivative is an explicit derivative, but by convention, it is a total derivative. Therefore, to make an explicit derivative out of a partial derivative comes with a price. One gets

$$\begin{bmatrix} \partial U/\partial V & \partial U/\partial T \\ \partial P/\partial V & \partial P/\partial T \\ \partial V/\partial V & \partial V/\partial T \\ \partial T/\partial V & \partial T/\partial T \end{bmatrix} \longrightarrow \begin{bmatrix} (\partial U/\partial V)_T & (\partial U/\partial T)_V \\ (\partial P/\partial V)_T & (\partial P/\partial T)_V \\ 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{45}$$

Okay, this system works, but I consider it cumbersome and ugly. Anyway, what else must change by use of the standard convention? Ans: First off, we'd probably never see an equation like (42) or like (43), but if we did, it would probably appear as

$$\frac{\partial\psi'}{\partial\boldsymbol{\eta}'} = \frac{\partial\psi}{\partial\boldsymbol{\eta}}\frac{\partial\boldsymbol{\eta}}{\partial\boldsymbol{\eta}'}\,. \tag{46}$$

But then what about Rules 1 and 2? They'll have to be modified to apply to the partial derivative instead of the deltal derivative. Additional changes would involve the two constraint equations. Eq. (28) becomes

$$\Big(\frac{\partial U}{\partial V}\Big)_T - T\Big(\frac{\partial P}{\partial T}\Big)_V + P = 0\,, \tag{47}$$

and (36) becomes

$$\Big(\frac{\partial T}{\partial P}\Big)_U - T\Big(\frac{\partial V}{\partial U}\Big)_P + PJ = 0\,. \tag{48}$$

References

[1] K. Huang. *Statistical Mechanics*. John Wiley & Sons. New York. 126 (1963).

[2] S.N. Rasband. *Dynamics*. John Wiley & Sons. New York. 152 (1983).

[3] C. Page. *Physical Mathematics*. D. Van Nostand Co. Princeton, New Jersey. 44 (1955).

[4] M. Abbott and H. Van Ness. *Theory and Problems of Thermodynamics*. Mc-Graw Hill Book Co. 66-97 (1972).

[5] A. Taylor and W.R. Mann. *Advanced Calculus,* 2ed. John Wiley & Sons. New York (1972).

[6] R.C. Buck. *Advanced Calculus*, 3ed. McGraw-Hill Book Co. (1984).