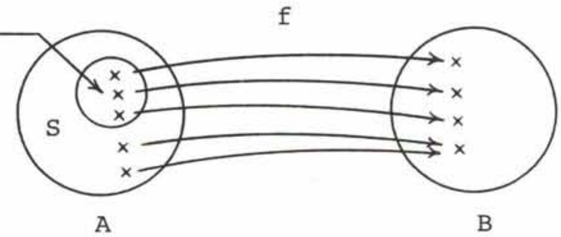
Unit 5: The Chain Rule, Part 1

3.5.1(L)

The main aim of this exercise is to review the basic principles taught in this unit in terms of a rather simple computational exercise.

To begin with, we should observe that the definition of f makes it a function of three independent variables (since its domain is E^3). We next note that P is a subset of E^3 and that if the domain of f is restricted to P, then f is no longer a function of the three independent variables x, y, and z but rather a function of the two independent variables r and s. In other words, once we focus our attention on P, we are no longer dealing with the function, f. [Recall that a function is defined not only by the set of images but also by the domain.] Once we limit f to the domain, P, we have a new function, say g, such that dom g = P. Since P is a proper subset of E^3, $P \neq E^3$; hence, the fact that dom $f = E^3 \neq$ dom g means that f and g are different functions, even though $f(\underline{x}) = g(\underline{x})$ for each $\underline{x} \varepsilon P$. Quite in general, if $f{:}A \to B$ and S is a subset of A, then the function g defined by $g{:}S \to B$ such that $g(\underline{s}) = f(\underline{s})$ for each $s \varepsilon S$ is called the restriction of f to S. Pictorially,

g is this "portion" of f and is called the restriction of f to S.



In this exercise, if we let g denote the restriction of f to P, then g is, in effect, a mapping from E^2 into E. That is, from the equations that determine P, z is determined once x and y are known (or, for that matter, once any two of the three variables x, y, and z are known, the other is uniquely determined). Since P involves linear equations, it is not too difficult to show how we may solve for z in terms of x and y. In fact, the technique which

3.5.1(L) continued

we shall employ is very powerful and is described in great detail in Chapter 6 of our supplementary notes. Thus, the following details are a forerunner of the notes.

We have that

$$\left.\begin{aligned} x &= r + s \\ y &= 2r + 3s \\ z &= 4r + 7s \end{aligned}\right\} \qquad (1)$$

We may easily eliminate r from the second and third equations in (1). That is, we may replace the second equation by the second minus twice the first (equals added to equals are equal, etc.) and we may replace the third equation by the third minus four times the first. This yields

$$x = r + s$$

$$\left.\begin{aligned} y - 2x &= s \;[= (2r + 3s) - 2(r + s)] \\ z - 4x &= 3s \;[= (4r + 7s) - 4(r + s)] \end{aligned}\right\} \qquad (2)$$

In (2), if we subtract three times the second equation from the third, we obtain

$$(z - 4x) - 3(y - 2x) = 3s - 3(s),$$

or

$$z = 3y - 2x. \qquad (3)$$

[As a check, notice that $z = 3y - 2x$, from (1), implies that $z = 3(2r + 3s) - 2(r + s) = 4r + 7s$, which checks with the third equation in (1).]

In other words, if we restrict f to P [i.e., (x,y,z) is restricted to P rather than E^3], we have

$$f(x,y,z) = f(x,y,3y - 2x) = g(x,y).$$

3.5.1(L) continued

In terms of our specific exercise, we see that if f is restricted to P, then

$$
\begin{aligned}
f(x,y,z) &= x^2 + y^2 + z^2 \\
&= x^2 + y^2 + (3y - 2x)^2 \\
&= 5x^2 + 10y^2 - 12xy \\
&= g(x,y).
\end{aligned}
$$

While it may seem contradictory to write that $f(x,y,z)$ is equal to $g(x,y)$ because it then seems as if we are saying that a function of three independent variables is the same as a function of two independent variables, the fact is that we are not saying this at all. In terms of more standard mathematical symbolism, we are not saying that $f = g$ but rather that $f|_P$ (i.e., the restriction of f to P) $= g$. This is no contradiction, since f and $f|_P$ are different functions.

From a pictorial point of view, what we are saying is that if $w = x^2 + y^2 + z^2$, then w is defined at each point in 3-space, but that if we then restrict the study of w to the plane P, then w has two, not three, degrees of freedom. This is no contradiction since if (x,y,z) is restricted to being in the plane P, then x, y, and z are no longer independent.

Hopefully, this discussion makes it clear why the choice of symbols for denoting our function is so important. That is, we must keep track, at each stage, of what variables are independent and which ones depend on the others. Remember that our definition of partial derivatives required that the variables be independent so that we could hold all but one at a time constant.

Returning to the specifics of this exercise, we have

a. If (x,y,z) is restricted to being in the plane P, then

3.5.1(L) continued

$$f(x,y,z) = (r + s)^2 + (2r + 3s)^2 + (4r + 7s)^2$$

$$= 21r^2 + 70rs + 59s^2$$

$$= h(r,s). \tag{4}$$

In terms of our earlier discussion, h is something similar to the restriction of f to P, but it is not exactly the same thing since we are using r and s rather than x and y. [If we wished to put everything in terms of x and y, we could return to equations (2) and replace the first equation by the first minus the second, to obtain

$$x - (y - 2x) = (r + s) - s,$$

or

$$3x - y = r.$$

This, coupled with the second equation of (2) tells us that

$$\left.\begin{aligned} 3x - y &= r \\ -2x + y &= s \end{aligned}\right\}, \tag{5}$$

whereupon, $21r^2 + 70rs + 59s^2$ becomes

$$21(3x - y)^2 + 70(3x - y)(-2x + y) + 59(-2x + y)^2 = g(x,y).$$

In other words,

$$h(r,s) = g(x,y) = +5x^2 - 12xy + 10y^2$$

and it is g, technically speaking, that is called the restriction of f to P, even though g and h are the same except for the choice of variables.]

b. We now notice that when we talk about w as a function of r and s on P, we are referring to w in the form

3.5.1(L) continued

$w = h(r,s)$

not in the form

$w = f(x,y,z)$.

Thus, by w_r, we mean $h_r(r,s)$. Then, since r and s are independent on P, we may use equation (4) to obtain

$$w_r = h_r(r,s) = 42r + 70s.$$

[Recall that the partial derivative of $59s^2$ with respect to r is zero since $59s^2$ depends only on s which in turn is independent of r (i.e., $59s^2$ is a constant if we are changing only r).]

c. Up to now, we see that there is no real need for the chain rule. In fact, in this particular exercise, it seems easier to substitute directly. The point is that in other cases it is more cumbersome to substitute directly; and in more important cases, we are often given $f(x,y,z)$ symbolically, without being told explicitly the rule which determines $f:E^3 \rightarrow E$. In this latter case, we must use the chain rule, and this is what the purpose of part (c) is.

The key idea is that from $f(x,y,z)$, it is natural to talk about f_x, f_y, and f_z; while from the fact that x, y, and z are functions of r and s, it is natural to talk about x_r, y_r, z_r, x_s, y_s, and z_s. In other words, when we look at $f(x,y,z)$, we do not worry about the fact that we may be restricting f to a subset of E^3 on which these variables are not independent. We let the other given information handle that part of the problem.

At any rate, the chain rule tells us that

$$w_r = f_x x_r + f_y y_r + f_z z_r \tag{6}$$

or, in "fraction" notation,

3.5.1(L) continued

$$\left(\frac{\partial w}{\partial r}\right)_s = \left(\frac{\partial w}{\partial x}\right)_{(y,z)}\left(\frac{\partial x}{\partial r}\right)_s + \left(\frac{\partial w}{\partial y}\right)_{(x,z)}\left(\frac{\partial y}{\partial r}\right)_s + \left(\frac{\partial w}{\partial z}\right)_{(x,y)}\left(\frac{\partial z}{\partial r}\right)_s .^*$$

Equation (6) becomes easier to memorize if we realize that $f_x x_r$ is the change in f with respect to r due solely to the change in x (and if we remember it this way, we can cancel the x's and write f_r, where, to be rigorous, we must keep in mind that f_r really refers to the derivative with respect to r of the <u>restriction</u> of f to P, since without this restriction, f is <u>not</u> a function of r and s alone). That is,

$$w_r = \underbrace{\underset{\downarrow\, w_r}{\underbrace{f_x x_r}} + \underset{\downarrow\, w_r}{\underbrace{f_y y_r}} + \underset{\downarrow\, w_r}{\underbrace{f_z z_r}}}_{\text{due solely to } \Delta x \;+\; \text{due solely to } \Delta y \;+\; \text{due solely to } \Delta z}{}^{**}$$

= "total" w_r since x, y, and z are independent.

---

*Notice that we are taking certain liberties here with notation. We have that $w = f(x,y,z)$, but restricted to P, $w = h(r,s)$. Thus, $\left(\frac{\partial w}{\partial r}\right)_s$ refers to $\left(\frac{\partial h}{\partial r}\right)_s$ while $\left(\frac{\partial w}{\partial x}\right)_{(y,z)}$ refers to $\left(\frac{\partial f}{\partial x}\right)_{(y,z)}$.

**In fractional notation, we would have, <u>omitting subscripts</u>,

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r}$$

from which it appears that by "cancellation"

$$\frac{\partial w}{\partial r} = 3\,\frac{\partial w}{\partial r}.$$

The proper use of subscripts shows us, however, that

$$\left(\frac{\partial w}{\partial r}\right)_s = \left(\frac{\partial w}{\partial x}\right)_{(y,z)}\left(\frac{\partial x}{\partial r}\right)_s + \left(\frac{\partial w}{\partial y}\right)_{(x,z)}\left(\frac{\partial y}{\partial r}\right)_s + \left(\frac{\partial w}{\partial z}\right)_{(x,y)}\left(\frac{\partial z}{\partial r}\right)_s .$$

Hopefully, the notation $\left(\frac{\partial w}{\partial x}\right)_{(y,z)}\left(\frac{\partial x}{\partial r}\right)_s$ warns us not to blindly cancel ∂x since the subscripts on each set of parentheses are different.

In the <u>special case</u> of one real variable, no harm is done when we "cancel" dx to obtain

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

since the contribution to $\frac{dy}{dt}$ due to the change in x alone is the total $\frac{dy}{dt}$ since x is the only variable in this case.

3.5.1(L) continued

At any rate, in our particular exercise, we have $f_x = 2x$, $f_y = 2y$, $f_z = 2z$, and $x_r = 1$, $y_r = 2$, $z_r = 4$. Hence,

$$w_r = 2x(1) + 2y(2) + 2z(4)$$

$$= 2x + 4y + 8z. \tag{7}$$

Again, one word of caution about (7); namely, w_r refers to the restriction of f to P. As a check, we have that on P, $x = r + s$, $y = 2r + 3s$, and $z = 4r + 7s$, whereupon (7) yields

$$w_r = 2(r + s) + 4(2r + 3s) + 8(4r + 7s)$$

$$= 42r + 70s \tag{8}$$

and equation (8) checks with the answer we obtained for part (b).

Notice that even in this rather simple case, the use of the chain rule was computationally simpler than direct substitution, at least in the sense that when we used the chain rule, we never had to square the terms $(r + s)$, $(2r + 3s)$, $(4r + 7s)$ and then take the derivative.

As a final word, let us observe that in this example, our choice of specific functions guaranteed that f was a continuously differentiable function of x, y, and z (recall that by continously differentiable, we mean that not only do all the partial derivatives exist but that they are also continuous) and that x, y, and z were each differentiable (but not necessarily continuously differentiable) functions of r and s. Had f not been continuously differentiable, we could have still computed $f_x x_r + f_y y_r + f_z z_r$, but in this case, it might not have been equal to w_r, since the proof that the error in the linear approximation was a second order infinitesimal used the fact that f was continuously differentiable.

[That is, we needed the result that

$$\Delta w = f_x \Delta x + f_y \Delta y + f_z \Delta z + k_1 \Delta x + k_2 \Delta y + k_3 \Delta z,$$

3.5.1(L) continued

where k_1, k_2, $k_3 \to 0$ as Δx, Δy, $\Delta z \to 0$ in order to conclude that

$$\frac{\Delta w}{\Delta r} = f_x \frac{\Delta x}{\Delta r} + f_y \frac{\Delta y}{\Delta r} + f_z \frac{\Delta z}{\Delta r} + k_1 \frac{\Delta x}{\Delta r} + k_2 \frac{\Delta y}{\Delta r} + k_3 \frac{\Delta z}{\Delta r}.$$

Then, as $\Delta r \to 0$, so did Δx, Δy, and Δz since x, y, and z were differentiable with respect to r. Thus,

$$\frac{\partial w}{\partial r} = \lim_{\Delta r \to 0} \frac{\Delta w}{\Delta r} = f_x x_r + f_y y_r + f_z z_r + 0x_r + 0y_r + 0z_r$$

$$= f_x x_r + f_y y_r + f_z z_r .]$$

In summary, if $w = f(x_1, \ldots, x_n)$ and $x_1, \ldots, x_n$ are differentiable functions of $y_1, \ldots, y_m$, then we may conclude, for example, that

$$\frac{\partial w}{\partial y_1} = \frac{\partial w}{\partial x_1} \frac{\partial x_1}{\partial y_1} + \ldots + \frac{\partial w}{\partial x_n} \frac{\partial x_n}{\partial y_1}$$

provided that f is a continuously differentiable function of $x_1, \ldots, x_n$. In most "practical" examples, this is usually the case, but we should always check to be sure.

3.5.2

On R,

$$w = f(x_1, x_2, x_3, x_4)$$

and

$$x_1 = h_1(t),\ x_2 = h_2(t),\ x_3 = h_3(t),\ x_4 = h_4(t).$$

Thus, if f is continuously differentiable, we have

3.5.2 continued

$$\frac{dw}{dt}^* = \frac{\partial w}{\partial x_1}\frac{dx_1}{dt} + \dots + \frac{\partial w}{\partial x_4}\frac{dx_4}{dt} = \sum_{n=1}^{4}\left(\frac{\partial w}{\partial x_n}\right)\left(\frac{dx_n}{dt}\right). \tag{1}$$

a. In this exercise,

$$w = x_1x_4 + x_1x_2x_3 + x_2x_3x_4.$$

Therefore,

$$\frac{\partial w}{\partial x_1} = x_4 + x_2x_3$$

$$\frac{\partial w}{\partial x_2} = x_1x_3 + x_3x_4$$

$$\frac{\partial w}{\partial x_3} = x_1x_2 + x_2x_4$$

$$\frac{\partial w}{\partial x_4} = x_1 + x_2x_3.$$

Hence, (1) becomes

$$\frac{dw}{dt} = (x_4+x_2x_3)\frac{dx_1}{dt} + (x_1x_3+x_3x_4)\frac{dx_2}{dt} + (x_1x_2+x_2x_4)\frac{dx_3}{dt} + (x_1+x_2x_3)\frac{dx_4}{dt}$$

$$= (x_4+x_2x_3)h_1'(t) + (x_1x_3+x_3x_4)h_2'(t) + (x_1x_2+x_2x_4)h_3'(t)$$

$$+ (x_1+x_2x_3)h_4'(t). \tag{2}$$

[Unless we are told specifically the functions $h_1(t)$, $h_2(t)$, $h_3(t)$ and $h_4(t)$, we cannot "refine" equation (2) further.]

*Since, on R, w depends on the single variable t, we write $\frac{dw}{dt}$ rather than $\frac{\partial w}{\partial t}$. In other words, with only one independent variable, the "ordinary" derivative and the partial derivative are the same, since the "ordinary" derivative is obtained by holding <u>all but one</u> of the independent variables constant. (Here, there is only one independent variable.)

3.5.2 continued

b. Now we are told explicitly that

$h_1(t) = t$, $h_2(t) = 2t$, $h_3(t) = 3t$, and $h_4(t) = 4t$.

Hence,

$h_1'(t) = 1$, $h_2'(t) = 2$, $h_3'(t) = 3$, and $h_4'(t) = 4$.

In this case, (2) becomes

$$\frac{dw}{dt} = \underline{(x_4+x_2x_3) + 2(x_1x_3+x_3x_4) + 3(x_1x_2+x_2x_4) + 4(x_1+x_2x_3)},$$

or, in terms of t,

$$\frac{dw}{dt} = (4t+2t[3t]) + 2(t[3t]+3t[4t]) + 3(t[2t]+2t[4t]) + 4(t+2t[3t])$$

$$= 4t + 6t^2 + 6t^2 + 24t^2 + 6t^2 + 24t^2 + 4t + 24t^2$$

$$= \underline{8t + 90t^2}.$$

By direct substitution, we obtain that on R,

$$w = t(4t) + t(2t)(3t) + 2t(3t)(4t)$$

$$= 4t^2 + 6t^3 + 24t^3$$

$$= 4t^2 + 30t^3.$$

Therefore,

$$\frac{dw}{dt} = \underline{8t + 90t^2}.$$

[The answers check but in this particular case, the direct substitution is more convenient than the chain rule.]

3.5.3(L)

a. To emphasize the chain rule, we write the given equations in the form

$$w = f(x,y) \tag{1}$$

and

$$\left.\begin{aligned} x &= u + v \\ y &= u - v \end{aligned}\right\} . \tag{2}$$

From (2), we have that $x_u = x_v = y_u = -y_v = 1$, so that applying the chain rule to (1), we obtain

$$w_u = f_x x_u + f_y y_u$$

or

$$w_u = f_x + f_y. \tag{3}$$

In a similar way,

$$w_v = f_x x_v + f_y y_v$$

or

$$w_v = f_x - f_y. \tag{4}$$

If we now multiply equations (3) and (4), we obtain

$$w_u w_v = f_x^{\ 2} - f_y^{\ 2} \tag{5}$$

which is the desired result. If we have any quarrel with equation (5), it is only that w_u and w_v suggest that we are viewing w as a function of u and v; say, $w = g(u,v)$. In this case, it would perhaps be less ambiguous to rewrite (5) as

$$g_u g_v = f_x^{\ 2} - f_y^{\ 2}.$$

3.5.3(L) continued

b. What makes this a learning exercise from our point of view is that we are now in a rather good position to show how the chain rule is used in "real-life" situations to solve partial differential equations.

For example, suppose we are looking for a continuously differentiable function $f(x,y)$ which satisfies the partial differential equation (where "partial" is used to describe the fact that we have a partial derivative rather than in the sense of meaning a "partial" equation)

$$f_x{}^2 - f_y{}^2 = 0. \tag{6}$$

From equation (6), we see that $f_y = \pm f_x$, which certainly is a sensible result but perhaps not too obvious as far as guessing what f must look like.

The point is that what we did in part (a) of this exercise is a fairly representative example of how one uses a change of variables to reduce a differential equation to a different, and hopefully simpler, differential equation. In particular, in equation (5) we saw that $f_x{}^2 - f_y{}^2$ was the same as $g_u g_v$ (or $w_u w_v$). If we put this information into equation (6), we see that

$$g_u g_v = 0. \tag{7}$$

From (7), we see that either $g_u = 0$ or else $g_v = 0$. But to say that $g_u = 0$ is equivalent to saying that g is a function of v alone (since if v is held constant, $g_u = 0$ says that g does not vary with respect to u).

In a similar vein, if $g_v = 0$, then g depends only on u. In other words, we have that either

$$g(u,v) = h(v) \tag{8}$$

or

$$g(u,v) = k(u). \tag{9}$$

3.5.3(L) continued

Since $x = u + v$ and $y = u - v$, it is not hard to see that we may express u and v in terms of x and y as

$u = (x + y)/2$ and $v = (x - y)/2$.*

Putting this into equations (8) and (9), we have that

$$g(u,v) = h\left(\frac{x - y}{2}\right)$$

or

$$g(u,v) = k\left(\frac{x + y}{2}\right).$$

[Notice that while x and y are two variables, $\frac{x + y}{2}$ and $\frac{x - y}{2}$ are each single variables. That is, the sum of two numbers is a number.]

Since

$$w = f(x,y) = g(u,v),$$

we see that our only possible types of solution are:

$$\left.\begin{array}{l} f(x,y) = h\left(\frac{x - y}{2}\right) \\ \text{or} \\ f(x,y) = k\left(\frac{x + y}{2}\right) \end{array}\right\}. \tag{10}$$

To see (10) more clearly in terms of a concrete example, let h and k be specific differentiable functions of t. Say

$$h(t) = t^2$$

*In this particular exercise, it was easy to solve for u and v explicitly in terms of x and y. In many cases, it is more difficult (and sometimes impossible). The general problem of expressing u and v in terms of x and y, knowing that $x = f(u,v)$ and $y = g(u,v)$, is quite difficult to handle.

3.5.3(L) continued

$k(t) = \sin t + e^t.$

Then

$$h\left(\frac{x-y}{2}\right) = \left(\frac{x-y}{2}\right)^2$$

and

$$k\left(\frac{x+y}{2}\right) = \sin\left(\frac{x+y}{2}\right) + e^{\frac{x+y}{2}}.$$

If we let $f(x,y) = \left(\frac{x-y}{2}\right)^2 = \frac{x^2 - 2xy + y^2}{4}$, we have that

$$f_x = \frac{2x - 2y}{4} = \frac{x-y}{2}.$$

Therefore,

$$f_x{}^2 = \left(\frac{x-y}{2}\right)^2$$

while

$$f_y = \frac{-2x + 2y}{4} = \frac{y-x}{2}.$$

Therefore,

$$f_y{}^2 = \left(\frac{y-x}{2}\right)^2 = \left(\frac{x-y}{2}\right)^2.$$

Therefore,

$$f_x{}^2 - f_y{}^2 = 0.$$

Similarly, if we let

$$f(x,y) = \sin\left(\frac{x+y}{2}\right) + e^{\frac{x+y}{2}},$$

3.5.3(L) continued

then

$$f_x = \frac{1}{2}\cos\left(\frac{x+y}{2}\right) + \frac{1}{2}e^{\frac{x+y}{2}}.$$

Therefore,

$$f_x{}^2 = \frac{1}{4}\cos^2\left(\frac{x+y}{2}\right) + \frac{1}{2}e^{\frac{x+y}{2}}\cos\left(\frac{x+y}{2}\right) + \frac{1}{4}e^{x+y}$$

while

$$f_y = \frac{1}{2}\cos\left(\frac{x+y}{2}\right) + \frac{1}{2}e^{\frac{x+y}{2}}.$$

Therefore,

$$f_y{}^2 = \frac{1}{4}\cos^2\left(\frac{x+y}{2}\right) + \frac{1}{2}e^{\frac{x+y}{2}}\cos\left(\frac{x+y}{2}\right) + \frac{1}{4}e^{x+y}.$$

Therefore,

$$f_x{}^2 - f_y{}^2 = 0.$$

In summary, all differentiable functions of the single variable $\frac{x+y}{2}$ or $\frac{x-y}{2}$ are solutions of $f_x{}^2 - f_y{}^2 = 0$.

3.5.4(L)

We have

$$w = f(u) \qquad (1)$$

and

$$u = ax + by. \qquad (2)$$

3.5.4(L) continued

a. From (2),

$$u_x = a \text{ and } u_y = b. \tag{3}$$

Applying the chain rule to (1), we obtain

$w_x = f_u u_x = f'(u)u_x$* (since f depends only on u)

$$w_y = f_u u_y = f'(u)u_y$$

and putting in the results in (3), we obtain

$$\left.\begin{aligned} w_x &= a\ f'(u) \\ w_y &= b\ f'(u) \end{aligned}\right\} . \tag{4}$$

b. From (4), we see that

$$\left.\begin{aligned} b\ w_x &= ab\ f'(u) \\ -a\ w_y &= -ab\ f'(u). \end{aligned}\right\}$$

Therefore,

$$b\ w_x - a\ w_y = 0.$$

[Notice that $\dfrac{\partial f(ax + by)}{\partial x} = a\ f'(u)$ <u>not</u> merely <u>a</u>. For example if $f(ax + by) = (ax + by)^2$, then

$$f_x = 2(ax + by)a = a[2(ax + by)].$$

That is, in this case, $f(u) = u^2$. Therefore, $f'(u) = 2u = 2(ax + by)$, which agrees with what f_x should be. The point is that in this exercise you can get the correct answer by accident if you assume that $f_x = a$ and $f_y = b$.]

*See note at the end of this exercise.

3.5.4(L) continued

c. The trick here is to see that $w = g(3x + 4y)$. This is part (b) with $a = 3$ and $b = 4$. Hence, we may conclude that

$$4w_x - 3w_y = 0,$$

whereupon,

$$4w_x - 3w_y + 2 = 2.$$

[The "glutton for punishment" is invited to solve $4w_x - 3w_y + 2$ directly from

$$w = e^{3x+4y} \tan^{-1}(3x + 4y)^2 \cosh (3x + 4y) \sin h^4(3x + 4y).]$$

d. If we let $a = 1$ and $b = -1$ (or equivalently, $a = -1$, $b = 1$) our previous results indicate that if $w = f(x - y)$ then

$$(-1)w_x - (1)w_y = 0$$

or

$$w_x + w_y = 0.$$

Therefore, if f is any function such that $f'(t)$ exists (i.e. f is any differentiable function of a single real variable), $w = f(x - y)$ is a solution of $w_x + w_y = 0$.

e. Without any boundary condition (that is, a prescribed behavior along a particular path [boundary]), we know that any function of the form

$$w = f(x - y)$$

is a solution to our equation.

We are now told that along the x-axis (i.e., $y = 0$) $w = e^x$. The point is that along the x-axis, since $y = 0$, $w = f(x - 0) = f(x)$. Equating these two values of w yields that $f(x) = e^x$.

3.5.4(L) continued

We now need only keep in mind that $f(x) = e^x$ is defined without recourse to the symbol x. That is, f is defined by

$$f([\ ]) = e^{[\ ]}.$$

In particular, then,

$$f(x - y) = e^{x-y}.$$

Thus, subject to the given boundary condition, w must be defined by

$$w = e^{x-y}.$$

A trivial check shows that this choice of w is indeed a solution of the equation; and clearly, letting $y = 0$ yields that $w(x,0) = e^{x-0}$, or $w = e^x$, as required.

Note

In the statement of the chain rule, nothing limits the number of variables. Thus, we might have that w is a continuously differentiable function of $x_1,\ldots,x_n$ while $x_1,\ldots,x_n$ are differentiable functions of $u_1,\ldots,u_m$; where m can be 1 or n can be 1.

The proof works the same as always. Namely, in our present exercise, from the fact that $w = f(u)$ is differentiable, we have that

$$\Delta w = f'(u)\Delta u + k\Delta u, \text{ where } \lim_{\Delta u \to 0} k = 0.$$

Therefore,

$$\frac{\Delta w}{\Delta x} = f'(u)\,\frac{\Delta u}{\Delta x} + k\,\frac{\Delta u}{\Delta x}$$

Therefore,

$$\frac{\partial w}{\partial x} = f'(u)\,\frac{\partial u}{\partial x} + 0\,\frac{\partial u}{\partial x}$$

$$= f'(u)\,\frac{\partial u}{\partial x}.$$

3.5.4(L) continued

Since in our exercise $u = ax + by$, we have $\frac{\partial u}{\partial x} = a$, so that $\frac{\partial w}{\partial x} = a\,f'(u)$, which checks with our previous result.

3.5.5(L)

The definition of a homogeneous function is modeled after the idea of what it means analytically when we talk about the degree of a polynomial term. For example, if f is a function of x and/or y such that

$$f(x,y) = x^3$$

or

$$f(x,y) = x^2y$$

or

$$f(x,y) = xy^2$$

or

$$f(x,y) = y^3$$

we seem to sense that in each case, f is a <u>third degree</u> polynomial, or in a different choice of words, that the dimension of x^3, x^2y, xy^2, and y^3 are all <u>three</u>. [In this case, there is a rather simple geometric interpretation in the sense that all four expressions represent the volumes of parallelepipeds. That is, x^3 is the volume if the dimensions are x by x by x, x^2y is the volume if the dimensions are x by x by y, etc.]

The point is that the definition of a homogeneous function of degree k as stated in this exercise generalizes our above discussion for terms of degree greater than three, as well as to functions which are not polynomials [as in part (b) of this exercise]. For example, in our above illustration of f, let us in each case

3.5.5(L) continued

replace x and y respectively by tx and ty where t is an arbitrary independent variable (which may be thought of as simply being some "symbol"). We then see that if $f(x,y) = x^2y$, for example, $f(tx,ty) = (tx)^2(ty)* = t^3(x^2y)$. Then, since $x^2y = f(x,y)$, we have that

$$f(tx,ty) = t^3f(x,y),$$

so that f is homogeneous of degree 3 according to the definition in this exercise, and this agrees with our earlier remarks.

Notice that the sum of homogeneous functions of the same degree is also a homogeneous function of that same degree. For example, if we let

$$g(x,y) = x^3 + 5x^2y + y^3$$

it follows that

$$\begin{aligned} g(tx,ty) &= (tx)^3 + 5(tx)^2(ty) + (ty)^3 \\ &= t^3x^3 + 5t^3x^2y + t^3y^3 \\ &= t^3(x^3 + 5x^2y + y^3) \\ &= t^3(x,y) \end{aligned}$$

so that by our definition, g is homogeneous of degree 3.

It should not be believed that all functions are homogeneous. For example, if we let

---

*Think of $f(x,y) = x^2y$ as a mapping of an ordered pair into a number. x denotes the first member of the pair and y the second. Just as in the single-variable case, the names of the variables are not important. For example, we might think of $f(x,y) = x^2y$ "pictorially" as

$f(\square,\triangle) = \square^2\triangle$ so that

$f(tx,ty) = f(\boxed{tx}, ty) = \boxed{tx}^2\, ty = (tx)^2(ty) = t^3x^2y.$

3.5.5(L) continued

$$h(x,y) = x^2 + y^3$$

then

$$h(tx,ty) = (tx)^2 + (ty)^3$$

$$= t^2x^2 + t^3y^3$$

and it should now be apparent that there is no power of t that can be factored out of the right hand side of the equation so that the remaining factor will be $x^2 + y^3$. What is true is that x^2 is homogeneous of degree 2 and that y^3 is homogeneous of degree 3. The sum of these two homogeneous functions, however, is not homogeneous. In other words, the sum of homogeneous functions is also homogeneous if each function in the sum has the same degree.

A general mathematical result that we shall not discuss here but simply state is that any function can be represented as a sum of homogeneous functions.

With this as background, we proceed with the exercise.

a. Given that

$$f(x,y,z) = x^3yz \tag{1}$$

we see that

$$f(tx,ty,tz) = (tx)^3(ty)(tz)$$

$$= t^5(x^3yz),$$

or, from (1),

$$f(tx,ty,tz) = t^5f(x,y,z). \tag{2}$$

According to our definition, then, equation (2) tells us that f is homogeneous of degree 5.

b. Given that

3.5.5(L) continued

$$f(x,y) = \sin\left(\frac{y}{x}\right),$$

we see that

$$f(tx,ty) = \sin\left(\frac{ty}{tx}\right) = \sin\left(\frac{y}{x}\right) = f(x,y) = t^0 f(x,y)^*$$

so that again by our definition, f is homogeneous of degree zero (and this is usually referred to as a dimensionless function).

c. $f(x_1,x_2,x_3,x_4) = x_1^{\,3}x_4 + x_2^{\,2}x_3^{\,2}$

Therefore,

$$\begin{aligned} f(tx_1,tx_2,tx_3,tx_4) &= (tx_1)^3(tx_4) + (tx_2)^2(tx_3)^2 \\ &= t^4x_1^{\,3}x_4 + t^4x_2^{\,2}x_3^{\,2} \\ &= t^4(x_1^{\,3}x_4 + x_2^{\,2}x_3^{\,2}) \\ &= t^4 f(x_1,x_2,x_3,x_4). \end{aligned}$$

Therefore, f is homogeneous of degree 4.

3.5.6(L)

a. We have that

$$w = f(tx_1,\ldots,tx_n)$$

where f is a continuously differentiable function from E^n to E. To emphasize the chain rule, we may rewrite w in the form

*Notice that both $\frac{ty}{tx} = \frac{y}{x}$ and $t^\circ = 1$ require that $t \neq 0$, but the fact that $t \neq 0$ is ensured by our definition of homogeneous.

3.5.6(L) continued

$$w = f(u_1,\ldots,u_n) \tag{1}$$

where

$$\left.\begin{array}{l} u_1 = tx_1 \\ \vdots \\ u_n = tx_n \end{array}\right\} . \tag{2}$$

Therefore, by the chain rule

$$\frac{\partial w}{\partial t} = f_{u_1}\frac{\partial u_1}{\partial t} + \ldots + f_{u_n}\frac{\partial u_n}{\partial t}. \tag{3}$$

From (2),

$$\frac{\partial u_1}{\partial t} = x_1,\ldots,\ \frac{\partial u_n}{\partial t} = x_n.$$

Hence, (3) may be written as

$$\frac{\partial w}{\partial t} = x_1 f_{u_1} + \ldots + x_n f_{u_n}. \tag{4}$$

Notice that the result obtained in (4) does not depend on f being homogeneous. If, however, f is homogeneous of degree k, then by definition

$$w = f(tx_1,\ldots,tx_n) \equiv t^k f(x_1,\ldots,x_n). \tag{5}$$

From the right side of equation (5), it follows that

$$\frac{\partial w}{\partial t} = kt^{k-1}f(x_1,\ldots,x_n)$$

[since $f(x_1,\ldots,x_n)$ is independent of t]. Equating this expression for $\frac{\partial w}{\partial t}$ with the expression in (4) yields

$$x_1 f_{u_1} + \ldots + x_n f_{u_n} = kt^{k-1}f(x_1,\ldots,x_n). \tag{6}$$

3.5.6(L) continued

Since equation (6) must hold for all values of t, it holds, in particular, when $t = 1$.

With $t = 1$, it follows that $u_1 = x_1, \ldots, u_n = x_n$ (since $u_k = tx_k$), so that (6) becomes

$$x_1 f_{x_1} + \ldots + x_n f_{x_n} = k\, f(x_1, \ldots, x_n). \tag{7}$$

Equation (7) is the desired result.

b. From part (a), we have just seen that if w is homogeneous of degree k, then

$$x\, w_x + y\, w_y + z\, w_z = kw.$$

In particular, then, if $k = 1$

$$x\, w_x + y\, w_y + z\, w_z = w.$$

Hence, if $w = f(x,y,z)$ is any homogeneous function of degree 1,

$$x\, w_x + y\, w_y + z\, w_z = w.$$

One such (trivial) function is

$$w = x + y + z.$$

A somewhat less trivial function is

$$w = x \sin \frac{y}{z} \quad (z \neq 0).$$

[That $x \sin \frac{y}{z}$ is homogeneous of degree 1, follows from

$$(tx) \sin \left(\frac{ty}{tz}\right) = t\left(x \sin \frac{y}{z}\right).]$$

As a check, notice that $w = x \sin \frac{y}{z}$ implies

3.5.6(L) continued

$$w_x = \sin \frac{y}{z}$$

$$w_y = x\left[\left(\cos \frac{y}{z}\right)\left(\frac{\partial \frac{y}{z}}{\partial y}\right)\right] = x\left[\frac{1}{z} \cos \frac{y}{z}\right] = \frac{x}{z} \cos \frac{y}{z}$$

$$w_z = x\left[\left(\cos \frac{y}{z}\right) \frac{\partial \frac{y}{z}}{\partial z}\right] = x\left[-\frac{y}{z^2} \cos \frac{y}{z}\right] = -\frac{xy}{z^2} \cos \frac{y}{z}.$$

Therefore,

$$\begin{aligned} x\, w_x + y\, w_y + z\, w_z &= x \sin \frac{y}{z} + \frac{xy}{z} \cos \frac{y}{z} - \frac{xy}{z} \cos \frac{y}{z} \\ &= x \sin \frac{y}{z} \\ &= w. \end{aligned}$$

3.5.7

We know from the previous exercise that if w is homogeneous of degree k, then

$$x\, w_x + y\, w_y = kw. \qquad (1)$$

Equation (1) is the present exercise with $k = 3$. Hence, if $f(x,y)$ is any homogeneous function of degree 3, $w = f(x,y)$ is a solution of the equation

$$x\, w_x + y\, w_y = 3w.$$

As a particular example,

$$f(x,y) = x^3 + x^2y + y^3$$

is homogeneous of degree 3. Hence,

$$w = x^3 + x^2y + y^3$$

3.5.7 continued

should satisfy the equation. To check this

$$w_x = 3x^2 + 2xy$$

$$w_y = x^2 + 3y^2.$$

Therefore,

$$x\,w_x + y\,w_y = (3x^3 + 2x^2y) + (x^2y + 3y^3)$$

$$= 3(x^3 + x^2y + y^3) = 3w.$$