p159,#4: Verify the chain rule for \( \frac{\partial h}{\partial x} \), where \( h(x, y) = f(u(x, y), v(x, y)) \) and

\[
f(u, v) = \frac{u^2 + v^2}{u^2 - v^2}, \quad u(x, y) = e^{-x-y}, \quad v(x, y) = e^{xy}.
\]

Solution: First we will work out what the chain rule looks like in this context. The mapping \( h: (x, y) \mapsto h(x, y) \) is defined to be the composition of two mappings, \( h = f \circ g \):

\[
g: \langle x, y \rangle \mapsto \left( \begin{array}{c} u(x, y) \\ v(x, y) \end{array} \right), \quad \text{and} \quad f: \langle u, v \rangle \mapsto f(u, v).
\]

We agree to write \( f(u, v) \) or \( f\langle u, v \rangle \) interchangeably. We only use the second notation to match with describing the first mapping as \( g: \mathbb{R}^2 \to \mathbb{R}^2 \) and the second mapping as \( f: \mathbb{R}^2 \to \mathbb{R} \), since we have stated that \( \mathbb{R}^n \) is a set of column vectors (not row vectors). The chain rule states that \( D(f \circ g)(x, y) = Df(g(x, y))Dg(x, y) \). Of course \( Dh = (\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}) \), whereas the problem only asks for the first component of this. We have

\[
Dg = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}, \quad Df = \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \\ \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix}.
\]

\( Dg \) gets evaluated at \( \langle x, y \rangle \) whereas \( Df \) gets evaluated at \( \langle u, v \rangle = g(x, y) \). The chain rule says then that

\[
Dh = \begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \end{pmatrix} = Df \cdot Dg = \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} \\ \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} \end{pmatrix}.
\]

Comparing the first components of these row vectors we see that the chain rule claims that

\[
\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x}.
\]

evaluated at the appropriate arguments. So to verify the chain rule in this case we should compute both sides of this equation and show that they are equal. We will start with the left side of the equation.

\[
h(x, y) = f(u(x, y), v(x, y)) = \frac{u(x, y)^2 + v(x, y)^2}{u(x, y)^2 - v(x, y)^2} = \frac{(e^{-x-y})^2 + (e^{xy})^2}{(e^{-x-y})^2 - (e^{xy})^2}
\]
\[ h(x, y) = \frac{e^{-2x-2y} + e^{2xy}}{e^{-2x-2y} - e^{2xy}} \]

\[ \frac{\partial h}{\partial x} = \frac{\frac{\partial}{\partial x} (e^{-2x-2y} + e^{2xy})}{(e^{-2x-2y} - e^{2xy})} \cdot \frac{\frac{\partial}{\partial x} (e^{-2x-2y} - e^{2xy})}{(e^{-2x-2y} - e^{2xy})} \]

\[ = \frac{1}{(e^{-2x-2y} - e^{2xy})^2} \left[ (-2e^{-2x-2y} + 2ye^{2xy})(e^{-2x-2y} - e^{2xy}) \right. \]

\[ \left. - (e^{-2x-2y} + e^{2xy})(-2e^{-2x-2y} - 2ye^{2xy}) \right] \]

Now we will work on the right side of the equation.

\[ f(u, v) = \frac{u^2 + v^2}{u^2 - v^2} \]

\[ \frac{\partial f}{\partial u} = \frac{2u(u^2 - v^2) - (u^2 + v^2)2u}{(u^2 - v^2)^2} \]

\[ \frac{\partial f}{\partial v} = \frac{2v(u^2 - v^2) - (u^2 + v^2)(-2v)}{(u^2 - v^2)^2} \]

\[ \frac{\partial f}{\partial u} (e^{-x-y}, e^{xy}) = \frac{2e^{-x-y}(e^{-2x-2y} - e^{2xy}) - (e^{-2x-2y} + e^{2xy})2e^{-x-y}}{(e^{-2x-2y} - e^{2xy})^2} \]

\[ \frac{\partial f}{\partial v} (e^{-x-y}, e^{xy}) = \frac{2e^{xy}(e^{-2x-2y} - e^{2xy}) - (e^{-2x-2y} + e^{2xy})(-2e^{xy})}{(e^{-2x-2y} - e^{2xy})^2} \]

\[ u(x, y) = e^{-x-y} \]

\[ \frac{\partial u}{\partial x} = -e^{-x-y} \]

\[ v(x, y) = e^{xy} \]

\[ \frac{\partial v}{\partial x} = ye^{xy} \]

So combining everything on the right side we get

\[ \frac{\partial f}{\partial u} (e^{-x-y}, e^{xy}) \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} (e^{-x-y}, e^{xy}) \frac{\partial v}{\partial x} \]

\[ = \frac{2e^{-x-y}(e^{-2x-2y} - e^{2xy}) - (e^{-2x-2y} + e^{2xy})2e^{-x-y}}{(e^{-2x-2y} - e^{2xy})^2} \cdot \left[ -e^{-x-y} \right] \]

\[ + \frac{2e^{xy}(e^{-2x-2y} - e^{2xy}) - (e^{-2x-2y} + e^{2xy})(-2e^{xy})}{(e^{-2x-2y} - e^{2xy})^2} \cdot \left[ ye^{xy} \right] \]

If we apply the distributive law to this, it can be seen to be equal to the left side \( \frac{\partial h}{\partial x} \), as desired.

In the calculus of one variable the chain rule is introduced as a method of computing derivatives of complex functions in terms of
derivatives of simpler functions. In our context of several variables it is less about computing the derivative in the easiest or shortest manner, and more about an important structural relationship satisfied by derivatives of compositions of functions.

\textbf{p160,\#8:} Let $f: \mathbb{R}^3 \to \mathbb{R}$ be differentiable. Making the substitution

$$x = \rho \sin \phi \cos \theta, \quad y = \rho \sin \phi \sin \theta, \quad z = \rho \cos \phi$$

(spherical coordinates) into $f(x, y, z)$, compute $\frac{\partial f}{\partial \rho}, \frac{\partial f}{\partial \phi}, \frac{\partial f}{\partial \theta}$ in terms of $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$.

\textbf{Solution:} The statement of this problem (straight out of the text) contains a common, but unfortunate, abuse of notation, one which can cause fairly severe confusion in applications. $f$ is initially defined as a function of $\langle x, y, z \rangle \in \mathbb{R}^3$, so the partial derivatives $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ are unambiguous. This abuse of notation attempts to also think of $f$ as a function of $\langle \rho, \phi, \theta \rangle \in (0, \infty) \times (0, \pi) \times (-\pi, \pi) \subset \mathbb{R}^3$, but this is overloading the symbol $f$ and introducing an ambiguity in the notation for partial derivatives $\frac{\partial f}{\partial \rho}, \frac{\partial f}{\partial \phi}, \frac{\partial f}{\partial \theta}$. Warning about partial derivative notation: Be very careful when interpreting $\frac{\partial f}{\partial u}$ to identify exactly what function $f$ stands for, what all its independent variables are, and to check that $u$ is one of them. When changing variables in applications sometimes (out of laziness or to avoid introducing yet another symbol) function symbols are given multiple conflicting meanings.

So we introduce the mapping $g: (0, \infty) \times (0, \pi) \times (-\pi, \pi) \to \mathbb{R}^3$ by the rule

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = g(\rho, \phi, \theta) = \begin{pmatrix} g^1(\rho, \phi, \theta) \\ g^2(\rho, \phi, \theta) \\ g^3(\rho, \phi, \theta) \end{pmatrix} = \begin{pmatrix} \rho \sin \phi \cos \theta \\ \rho \sin \phi \sin \theta \\ \rho \cos \phi \end{pmatrix}.$$

Here we are using $x, y, z$ as dependant variable names for the three component functions $g^1, g^2, g^3$ of the vector-valued function $g$. Define the mapping $h: (0, \infty) \times (0, \pi) \times (-\pi, \pi) \to \mathbb{R}$ to be $h = f \circ g$. $h(\rho, \phi, \theta) = f(\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$, so when the problem asks for $\frac{\partial f}{\partial \rho}, \frac{\partial f}{\partial \phi}, \frac{\partial f}{\partial \theta}$ it really means $\frac{\partial h}{\partial \rho}, \frac{\partial h}{\partial \phi}, \frac{\partial h}{\partial \theta}$, which are the components of the row vector $Dh$. We use the chain rule to compute it:

$$Dh(\rho, \phi, \theta) = Df(g(\rho, \phi, \theta))Dg(\rho, \phi, \theta),$$

i.e.

$$\begin{pmatrix} \frac{\partial h}{\partial \rho} \\ \frac{\partial h}{\partial \phi} \\ \frac{\partial h}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} \begin{pmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta} \end{pmatrix}.$$
It is well-known that \( \frac{\partial y}{\partial \theta} \) and \( \frac{\partial g}{\partial \theta} \) are synonymous, since
\[ y = g^2(\rho, \phi, \theta). \]
(Traditional notation is loath to introduce the symbol \( g_i \) for the \( i \)th component function of the vector-valued function \( g \).) So now we compute:

\[
\begin{pmatrix}
\frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \phi} & \frac{\partial x}{\partial \theta} \\
\frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \phi} & \frac{\partial y}{\partial \theta} \\
\frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \phi} & \frac{\partial z}{\partial \theta}
\end{pmatrix}
= \begin{pmatrix}
\sin \phi \cos \theta & \rho \cos \phi \cos \theta & -\rho \sin \phi \sin \theta \\
\sin \phi \sin \theta & \rho \cos \phi \sin \theta & \rho \sin \phi \cos \theta \\
\cos \phi & -\rho \sin \phi & 0
\end{pmatrix}
\]

Hence

\[
\left. \begin{pmatrix}
\frac{\partial h}{\partial \rho} & \frac{\partial h}{\partial \phi} & \frac{\partial h}{\partial \theta}
\end{pmatrix} \right|_{(\rho, \phi, \theta)} = \left. \begin{pmatrix}
\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z}
\end{pmatrix} \right|_{g(\rho, \phi, \theta)} \begin{pmatrix}
\sin \phi \cos \theta & \rho \cos \phi \cos \theta & -\rho \sin \phi \sin \theta \\
\sin \phi \sin \theta & \rho \cos \phi \sin \theta & \rho \sin \phi \cos \theta \\
\cos \phi & -\rho \sin \phi & 0
\end{pmatrix}
\]

I will assume you can do this row vector–matrix multiplication, and will leave the answer in this (shorter) form.

**p160, #10:** Let \( f(u, v, w) = (e^{u-w}, \cos(v + u) + \sin(u + v + w)) \) and
\( g(x, y) = (e^x, \cos(y - x), e^{-y}) \). Calculate \( f \circ g \) and \( D(f \circ g)(0, 0) \).

**Solution:**

\[
(f \circ g)(x, y) = f(g(x, y)) = f(e^x, \cos(y - x), e^{-y})
= \begin{pmatrix}
e^x \\
\cos(y - x) \\
e^{-y}
\end{pmatrix}
\]

By the chain rule \( D(f \circ g)(0, 0) = Df(g(0, 0)) Dg(0, 0) \). Computing:

\[
g(x, y) = \begin{pmatrix}
e^x \\
\cos(y - x) \\
e^{-y}
\end{pmatrix}
\]

\[
g(0, 0) = \begin{pmatrix}
e^0 \\
\cos(0 - 0) \\
e^{-0}
\end{pmatrix} = \begin{pmatrix}
1 \\
1 \\
1
\end{pmatrix}
\]

\[
Dg(x, y) = \begin{pmatrix}
e^x & 0 & 0 \\
\sin(y - x) & -\sin(y - x) & 0 \\
0 & -e^{-y} & 0
\end{pmatrix}
\]

\[
Dg(0, 0) = \begin{pmatrix}
e^0 & 0 & 0 \\
\sin(0 - 0) & -\sin(0 - 0) & 0 \\
0 & -e^{-0} & 0
\end{pmatrix} = \begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}
\]

\[
f(u, v, w) = \begin{pmatrix}
e^{u-w} \\
\cos(v + u) + \sin(u + v + w)
\end{pmatrix}
\]
\[
Df(u, v, w) = \begin{pmatrix}
e^u - w & 0 & -e^u - w \\
- \sin v + u & - \sin v + u & \cos u + v + w \\
+ \cos u + v + w & + \cos u + v + w & \\
0 & -1 & \\
\end{pmatrix}
\]

\[
Df(1, 1, 1) = \begin{pmatrix}
e^1 - 1 & 0 & -e^{1-1} \\
- \sin 1 + 1 & - \sin 1 + 1 & \cos 1 + 1 + 1 \\
+ \cos 1 + 1 + 1 & + \cos 1 + 1 + 1 & \\
0 & 1 & \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & 0 & -1 \\
- \sin 2 + \cos 3 & - \sin 2 + \cos 3 & \cos 3 \\
\end{pmatrix}
\]

\[
D(f \circ g)(0, 0) = \begin{pmatrix}
1 & 0 & -1 \\
- \sin 2 + \cos 3 & - \sin 2 + \cos 3 & \cos 3 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 \\
0 & 0 \\
0 & -1 \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
1 & - \cos 3 \\
- \sin 2 + \cos 3 & - \cos 3 \\
\end{pmatrix}
\]

\[
Df(u, v, w) \text{ is a } 2 \times 3 \text{ matrix; what appears above to be a 3rd row is a vertical continuation of the 2nd row.}
\]

**p160,#14:** Let \( f : \mathbb{R}^n \to \mathbb{R}^m \) be a linear mapping so that by a previous homework problem \( Df(y) = f \) for all \( y \in \mathbb{R}^n \). (We can think of \( Df(y) \) as the \( m \times n \) matrix of partial derivatives or as the linear mapping \( \mathbb{R}^n \to \mathbb{R}^m \) it induces.) Check the validity of the chain rule directly for linear mappings.

**Solution:** Let \( g : \mathbb{R}^p \to \mathbb{R}^n \) be another linear mapping. We claim that \( f \circ g : \mathbb{R}^p \to \mathbb{R}^m \) is linear. To check this suppose \( x, y \in \mathbb{R}^p \) and \( a \in \mathbb{R} \). Then

\[
(f \circ g)(xa + y) = f(g(xa + y)) = f(g(x)a + g(y))
\]

\[
= f(g(x))a + f(g(y)) = (f \circ g)(x)a + (f \circ g)(y).
\]

These equalities are because of the definition of \( f \circ g \), the linearity of \( g \), and the linearity of \( f \). Thus \( f \circ g \) is linear. By the earlier homework problem \( Dg(x) = g \) and \( D(f \circ g)(x) = f \circ g \). Hence (for any \( y \in \mathbb{R}^n \))

\[
D(f \circ g)(x) = f \circ g = Df(y) \circ Dg(x) = Df(g(x))Dg(x)
\]

for all \( x \in \mathbb{R}^p \), as the chain rule claims. Note that as linear mappings we could write \( Df(g(x)) \circ Dg(x) \) whereas as matrices we usually write \( Df(g(x))Dg(x) \). The product of the matrices induces the composition of the linear mappings. Because of this people often denote the composition of linear mappings without writing \( \circ \).
Consider the function \( f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \).

Show that

a) \( \frac{\partial f}{\partial x} \) and \( \frac{\partial f}{\partial y} \) exist at \((0, 0)\).

b) If \( g(t) = \langle at, bt \rangle \) for all \( t \in \mathbb{R} \), and for constants \( a, b \in \mathbb{R} \), \( ab \neq 0 \),

then \( f \circ g \) is differentiable and \( D(f \circ g)(0) = ab^2/(a^2 + b^2) \) but

\( Df(g(0))Dg(0) = 0 \). Hence the conclusion of the chain rule fails to hold, so \( f \) must not be differentiable at \((0, 0)\).

**Solution:** (This was done in class.)

a) \( \frac{\partial f}{\partial x}(0, 0) = \lim_{h \to 0} \frac{f(0+h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0 - 0}{h} = 0 \). So it exists.

Similarly, \( \frac{\partial f}{\partial y}(0, 0) = \lim_{h \to 0} \frac{f(0, 0+h) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0 - 0}{h} = 0 \). So it exists.

b) \( (f \circ g)(t) = f\langle at, bt \rangle = \begin{cases} \frac{(at)(bt)^2}{(at)^2 + (bt)^2} & t \neq 0 \\ 0 & t = 0 \end{cases} = ab^2t/(a^2 + b^2) \). This function is clearly differentiable and \( D(f \circ g)(t) = ab^2/(a^2 + b^2) \) for all \( t \), including \( t = 0 \). \( Dg(t) = \langle a, b \rangle \) for all \( t \), including \( t = 0 \). \( Df(g(0)) = Df(0, 0) = (\frac{\partial f}{\partial x}(0, 0), \frac{\partial f}{\partial y}(0, 0)) = (0, 0) \) from part a). Performing the matrix product: \( Df(g(0))Dg(0) = (0, 0)\langle a, b \rangle = 0 \) as desired. So to summarize:

\( D(f \circ g)(0) = ab^2/(a^2 + b^2) \neq 0 = Df(g(0))Dg(0) \),

violating the conclusion of the chain rule.