Suppose you are interested in a composition
\[
\begin{array}{ccc}
\mathbb{R}^p & \xrightarrow{G} & \mathbb{R}^n \\
\mathbb{R}^n & \xrightarrow{F} & \mathbb{R}^m
\end{array}
\]
and want the total derivative \( H'(\hat{x}) : \mathbb{R}^p \rightarrow \mathbb{R}^m \) at \( \hat{x} = \hat{a} \). If it exists, this is a linear transformation with the property that \( H(\hat{a} + \hat{u}) - H(\hat{a}) = H'(\hat{a}) \hat{u} + \| \hat{u} \| \ E_H(\hat{a}; \hat{u}) \) where \( E_H(\hat{a}; \hat{u}) \rightarrow 0 \) with \( \hat{u} \). (This is the definition.) As we saw last time, when \( H'(\hat{a}) \) exists it is represented by the "Jacobian matrix" \( J_H(\hat{a}) = \left[ \frac{\partial h_i}{\partial x_j}(\hat{a}) \right] \).

**Theorem:** If \( F \) is differentiable at \( \hat{b} = G(\hat{a}) \) and \( G \) is differentiable at \( \hat{a} \), then \( H \) is differentiable at \( \hat{a} \) and
\[
\begin{align*}
H'(\hat{a}) &= F'(\hat{b}) \cdot G'(\hat{a}) \quad \text{(composition of L.T.s)} \\
J_H(\hat{a}) &= J_F(\hat{b}) \cdot J_G(\hat{a}) \quad \text{(matrix multiplication)}
\end{align*}
\]
Proof: \( H(\tilde{a}+\tilde{v}) - H(\tilde{a}) = \mathcal{F}(\mathcal{G}(\tilde{a}+\tilde{v})) - \mathcal{F}(\mathcal{G}(\tilde{a})) \)

\[ \text{Since } \mathcal{G}(\tilde{a}) = \mathcal{G}(\tilde{a}) + \mathcal{G}(\tilde{a}) \tilde{v} + \frac{1}{2} \|	ilde{v}\|^2 E_\mathcal{G}(\tilde{a}; \tilde{v}) \]

\[ = \mathcal{F}(\tilde{b}+\tilde{v}) - \mathcal{F}(\tilde{b}) \]

\[ = \mathcal{F}'(\tilde{b}) \tilde{v} + \frac{1}{2} \|	ilde{v}\|^2 E_f(\tilde{b}; \tilde{v}) \]

\[ = \mathcal{F}'(\tilde{b}) \mathcal{G}(\tilde{a}) \tilde{v} + \frac{1}{2} \|	ilde{v}\|^2 \left( \frac{\partial}{\partial \tilde{v}} E_\mathcal{G}(\tilde{a}; \tilde{v}) \right) \]

So in matrix form we get

\[ \left[ \frac{\partial h_i(\tilde{a})}{\partial x_j} \right] = \left[ \frac{\partial \mathcal{F}(\mathcal{G}(\tilde{a}))}{\partial y_k} \right] \left[ \frac{\partial \mathcal{G}(\tilde{a})}{\partial x_j} \right] \]

i.e., for each fixed \( i, j \)

\[ \frac{\partial h_i}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial \mathcal{F}}{\partial y_k} \frac{\partial \mathcal{G}_{y_k}}{\partial x_j} \]

Ex. Consider the composition \( IR^2 \xrightarrow{G} IR^2 \xrightarrow{f} IR \)

where \( G(\tau, \theta) = (\cos \theta, \sin \theta) \).

Then \( f' \) means \( \nabla f = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) \), \( f' \) means \( \nabla f' = \left( \frac{\partial f'}{\partial x}, \frac{\partial f'}{\partial \theta} \right) \),

and \( G' = \left( \frac{\partial G_1}{\partial \tau}, \frac{\partial G_2}{\partial \theta} \right) = (\sin \theta, \cos \theta) \). The chain rule says

\[ \frac{\partial f'}{\partial \theta} = \nabla f' \cdot G' = \left( \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta, -\frac{\partial f}{\partial x} \sin \theta + \frac{\partial f}{\partial y} \cos \theta \right) \]

A more efficient way to work this out/memorise device is given by the picture:
which you are supposed to read as
\[
\frac{\partial^2 z}{\partial r^2} = \frac{\partial}{\partial x} \left( \frac{\partial z}{\partial r} \right) + \frac{\partial}{\partial y} \left( \frac{\partial z}{\partial r} \right)
\]

Sum over paths from $r$ to $\theta$
the products of labels along the path.

Clairaut's Theorem

Let $f(x,y)$ be a function whose second partials exist in a neighborhood of $(0,0)$. Is there a difference between
\[
\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = f_{yx} \quad \text{and} \quad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = f_{xy}?
\]

The definition says that $f_x(0,y) = \lim_{\Delta x \to 0} \frac{f(x,y) - f(0,y)}{\Delta x}$, so
\[
f_x(0,y) = \lim_{\Delta x \to 0} \left( \frac{f(x,y) - f(0,y)}{\Delta x} \right)
\]

\[
\frac{f_{xy}}{\partial y} = \frac{\lim_{\Delta y \to 0} \left( \lim_{\Delta x \to 0} \frac{f(x,\Delta y) - f(0,\Delta y)}{\Delta x} \right)}{\partial y}
\]

This is symmetric in $x$ and $y$ except for the order of the limits. The question is whether we can switch those.

Ex.
\[
f(x,y) = \begin{cases} 
\frac{xy(x^2-y^2)}{x^2+y^2} & (x,y) \neq (0,0) \\
0 & (x,y) = (0,0)
\end{cases}
\]
\[ f_x(0, y) = \frac{x^4y + 4x^3y^3 - y^5}{(x^2+y^2)^2} \bigg|_{x=0} = -\frac{y^5}{y^4} = -y \]

\[ \Rightarrow f_x(0, 0) = -1 \]

\[ f_y(x, 0) = \frac{x^5 - 4x^3y^2 - xy^4}{(x^2+y^2)^2} \bigg|_{y=0} = \frac{x^5}{x^2} = x \]

\[ \Rightarrow f_y x(0, 0) = 1. \quad \text{Note also that } f_x, f_y \text{ are continuous about } (0, 0) \text{ (use the polar form + squeeze thm.)} \]

hence \( f \) is actually \textbf{differentiable there!} What is going on?

Well, the full \( f_{xy}(x, y) = \frac{x^6 + 9x^4y^2 - 9x^2y^4 + 15y^6}{(x^2+y^2)^3} \)

is evidently \textbf{not continuous at } (0, 0) \text{ (consider the limits along the vertical & horizontal axes...)} This suggests a way at:

\textbf{Theorem:} Assume \( f_{xy} \) & \( f_{yx} \) are continuous in a neighborhood of \( (0, 0) \). Then \( f_{xy}(0, 0) = f_{yx}(0, 0) \).

\textbf{Proof:} Write \( \Delta(h) := f(h, h) - f(h, 0) - f(0, h) + f(0, 0) \).

Setting \( g(x) := f(x, h) - f(x, 0) \), for some \( a \in [0, h] \)

\[ \frac{\Delta(h)}{h} = \frac{g(h) - g(0)}{h} \quad \text{MVT} \]

\[ = g'(a) = f_x(a, h) - f_x(a, 0). \]

Setting \( G(y) := f_x(a, y) \), for some \( b \in [0, h] \)

\[ \frac{\Delta(h)}{h^2} = \frac{G(h) - G(0)}{h} \quad \text{MVT} \]

\[ = G'(b) = f_{xy}(a, b) \].
and \( \lim_{h \to 0} \frac{\Delta(h)}{h^2} = \lim_{(x,y) \to (0,0)} f_{xy}(x,y) = f_{xy}(0,0) \) by continuity of \( f_{xy} \).

An exactly symmetric argument (in \( x \delta y \)) gives

\[ \lim_{h \to 0} \frac{\Delta(h)}{h^2} = f_{yx}(0,0), \] proving the Theorem. \( \square \)