Recall from our work in single variable Calculus:

**Theorem: (Chain Rule)**
If $h, k : \mathbb{R} \mapsto \mathbb{R}$ are differentiable functions, then

$$(k \circ h)' = \underline{\quad (k' \circ h) \cdot h' \quad}$$

1. Using what we know about derivatives of different types of functions and compositions of functions we will now conjecture the Chain Rule Theorem for computing the derivative of the composition of $f \circ \mathbf{g}$ where $f : \mathbb{R}^2 \mapsto \mathbb{R}$ and $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ are differentiable functions.

   To do this, we will create analogies between the Chain Rule in the case we know from Calculus 1 (i.e. $h, k : \mathbb{R} \mapsto \mathbb{R}$ are differentiable functions) and this new case (i.e. $f : \mathbb{R}^2 \mapsto \mathbb{R}$ and $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ are differentiable functions). Note that $k$ is analogous to $f$, and $h$ is analogous to $\mathbf{g}$.

   (a) We have the function $\boxed{f : \mathbb{R}^2 \mapsto \mathbb{R}^1}$. The derivative of $f$ is $\boxed{\underline{\nabla f} : \mathbb{R}^2 \mapsto \mathbb{R}^2}$.

   (b) The function $f$ is composed with the function $\boxed{\mathbf{g} : \mathbb{R}^1 \mapsto \mathbb{R}^2}$. The derivative of $\mathbf{g}$ is $\boxed{\underline{\mathbf{g'}} : \mathbb{R}^1 \mapsto \mathbb{R}^2}$.

   (c) $k' \circ h$ is analogous to $\underline{\nabla f \circ \mathbf{g}}$. This composition maps $\boxed{\mathbb{R}^1}$ to $\boxed{\mathbb{R}^2}$.

   (d) $h'$ is analogous to $\underline{\mathbf{g'}}$. This derivative function maps $\boxed{\mathbb{R}^1}$ to $\boxed{\mathbb{R}^2}$.

   (e) The operation of multiplication of the composition $k' \circ h$ and $h'$ is analogous to the operation of $\boxed{\underline{\text{the dot product}}}$ of the composition $\underline{\nabla f \circ \mathbf{g}}$ and $\underline{\mathbf{g'}}$.

Therefore, we can now conjecture:

**Theorem: (Chain Rule)**
$f : \mathbb{R}^2 \mapsto \mathbb{R}$ and $\mathbf{g} : \mathbb{R} \mapsto \mathbb{R}^2$ are differentiable functions, then

$$(f \circ \mathbf{g})' = \underline{\quad (\nabla f \circ g) \cdot \mathbf{g'} \quad}$$

Note that $\boxed{(f \circ \mathbf{g})' : \mathbb{R}^1 \mapsto \mathbb{R}^1}$.

2. Using the Chain Rule, compute $(f \circ \mathbf{g})'$ where $f(x, y) = 3xy^2$ and $\mathbf{g}(t) = \langle 2, 0\rangle + \langle 4, 5\rangle t$.

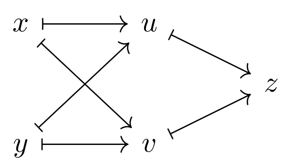
   **Solution:** $\nabla f = \langle 3y^2, 6xy\rangle$ and $\mathbf{g'} = \langle 4, 5\rangle$. Then following our chain rule, we have

$$\begin{aligned}(\nabla f \circ \mathbf{g}) \cdot \mathbf{g'} &= \nabla f(2 + 4t, 5t) \cdot \langle 4, 5\rangle \\ &= \langle 3(5t)^2, 6(2 + 4t)(5t)\rangle \cdot \langle 4, 5\rangle \\ &= 12(5t)^2 + 30(2 + 4t)(5t)\end{aligned}$$

We will apply similar reasoning to guess the Chain Rule for other composition mappings with different domains (input) and codomain (output).

3. Consider the composition function $h = (g \circ \mathbf{f})$ for the differentiable functions $\mathbf{f}(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}^2$ and $g(u, v) : \mathbb{R}^2 \mapsto \mathbb{R}$.

   Observe $h$ maps from $\mathbb{R}^2$ to $\mathbb{R}$. To help us visualize the domain and codomain for this composition of functions, we can diagram the mapping of the variables for the composition function $h = (g \circ \mathbf{f})$:




$$\mathbb{R}^2 \overset{\mathbf{f}}{\longmapsto} \mathbb{R}^2 \overset{g}{\longmapsto} \mathbb{R}$$

   We will use what we know about the Chain Rule and derivatives to conjecture the derivative of the composition function $h = (g \circ \mathbf{f})$.

   (a) We have the composition function $h : \mathbb{R}^2 \mapsto \mathbb{R}$, therefore we expect the derivative of $h$ to be $\boxed{\underline{\quad \nabla h \quad} : \mathbb{R}^2 \mapsto \mathbb{R}^2}$.

   (b) The derivative of $g$ is $\boxed{\underline{\quad \nabla g \quad} : \mathbb{R}^2 \mapsto \mathbb{R}^2}$.

   (c) The derivative of $\mathbf{f}$ is $\boxed{\underline{\quad D\mathbf{f} \quad} : \mathbb{R}^2 \mapsto \mathbb{R}^4}$, a $\underline{\quad 2 \quad}$ × $\underline{\quad 2 \quad}$ matrix.

   We conjecture the Chain Rule for this case to be:

   **Theorem: (Chain Rule)**
   $g : \mathbb{R}^2 \mapsto \mathbb{R}$ and $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ are differentiable functions, then

$$(g \circ \mathbf{f})' = \underline{\qquad (\nabla g \circ \mathbf{f}) D\mathbf{f} \qquad}$$

   Note that $\boxed{(g \circ \mathbf{f})' : \mathbb{R}^2 \mapsto \mathbb{R}^2}$.

   (d) Following the idea of the Chain Rule, we want to "multiply" $\nabla g \circ \mathbf{f}$ by $D\mathbf{f}$. To do this, we must interpret $\nabla g \circ \mathbf{f}$ as a $\underline{\quad 1 \quad}$ × $\underline{\quad 2 \quad}$ row matrix.

   (e) Does your expectation for the derivative of $h$ from part (a) match the result of the Chain Rule differentiation of $h$?

   **Solution:** We have the composition of differentiable functions $\mathbf{f}(x, y) : \mathbb{R}^2 \mapsto \mathbb{R}^2$ and $g(u, v) : \mathbb{R}^2 \mapsto \mathbb{R}$. The derivative of $\mathbf{f}$ is $D\mathbf{f}$, a $2 \times 2$ matrix. The derivative of $g$ is $\nabla g$, the gradient vector, which we can interpret as a $1 \times 2$ matrix. Then following the chain rule we would have

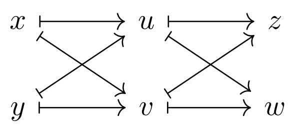
$$h' = (\nabla g \circ \mathbf{f}) \cdot D\mathbf{f}$$

   where the "multiplication" operation is matrix multiplication. The result of the chain rule to gives us a $1 \times 2$ row matrix. This row matrix can be thought of as a vector in $\mathbb{R}^2$, matching our expectations for the derivative $h$ from part (a).

Let's continue our work examining how we can apply the Chain Rule in other cases of composition mappings.

4. Consider the composition of differentiable functions $\mathbf{f}(x,y) : \mathbb{R}^2 \mapsto \mathbb{R}^2$ and $\mathbf{g}(u,v) : \mathbb{R}^2 \mapsto \mathbb{R}^2$.

   (a) Create a similar mapping diagram to the one shown in Problem 3 to represent the mapping of the variables for the composition function $\mathbf{h} = (\mathbf{g} \circ \mathbf{f})$.




$$\mathbb{R}^2 \xmapsto{\mathbf{f}} \mathbb{R}^2 \xmapsto{\mathbf{g}} \mathbb{R}^2$$

   (b) Based off of your work, conjecture the Chain Rule in the case if $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ and $\mathbf{g} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ are both differentiable functions, and $\mathbf{h} = \mathbf{g} \circ \mathbf{f}$.

**Solution:** Since $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ its derivative is a $2 \times 2$

$$D\mathbf{f} = \begin{bmatrix} u_x(x,y) & u_y(x,y) \\ v_x(x,y) & v_y(x,y) \end{bmatrix}$$

Since $\mathbf{f} : \mathbb{R}^2 \mapsto \mathbb{R}^2$ its derivative is a $2 \times 2$

$$D\mathbf{g} = \begin{bmatrix} z_u(u,v) & z_v(u,v) \\ w_u(u,v) & w_v(u,v) \end{bmatrix}$$

Following the chain rule, we have

$$D\mathbf{h} = (D\mathbf{g} \circ \mathbf{f})D\mathbf{f}$$

which is a $2 \times 2$ matrix multiplied by a $2 \times 2$ matrix. The result of this matrix multiplication is a $2 \times 2$ matrix. Therefore, $D\mathbf{h}$ is a $2 \times 2$ matrix.

Apply the Chain Rule to compute the derivative of the following composition functions:

5. $f(x,y) = 2x^2 - y^2$
   $\mathbf{g}(s,t) = \langle 2s+5t, 3st\rangle$
   With the composition function $f \circ \mathbf{g}$.

   **Solution:**
   $\nabla f = \langle 4x, -2y\rangle$
   $D\mathbf{g} = \begin{bmatrix} 2 & 5 \\ 3t & 3s \end{bmatrix}$
   $\nabla f \circ \mathbf{g} = \langle 4(2s+5t), -2(3st)\rangle$
   $(\nabla f \circ \mathbf{g})D\mathbf{g} = \begin{bmatrix} 8s+20t & -6st \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 3t & 3s \end{bmatrix} = \begin{bmatrix} 16s+40t-18st^2 & 40s+100t-18s^2t \end{bmatrix}$

6. $\mathbf{f}(t) = \langle t^2, \tan(t)\rangle$
   $\mathbf{g}(u,v) = \langle u-v, u^2v\rangle$
   With the composition function $\mathbf{g} \circ \mathbf{f}$.

   **Solution:**
   $\mathbf{f}' = \langle 2t, \sec^2(t)\rangle$
   $D\mathbf{g} = \begin{bmatrix} 1 & -1 \\ 2uv & u^2 \end{bmatrix}$
   $D\mathbf{g} \circ \mathbf{f} = \begin{bmatrix} 1 & -1 \\ 2t^2\tan(t) & t^4 \end{bmatrix}$
   $(D\mathbf{g} \circ \mathbf{f})\mathbf{f}' = \begin{bmatrix} 1 & -1 \\ 2t^2\tan(t) & t^4 \end{bmatrix}\begin{bmatrix} 2t \\ \sec^2(t) \end{bmatrix} = \begin{bmatrix} 2t-\sec^2(t) \\ 4t^3\tan(t)+t^4\sec^2(t) \end{bmatrix}$
   Note that to carry out our multiplication, we must interpret $\mathbf{f}'$ as $2 \times 1$ matrix.

7. $\mathbf{w}(x,y) = \langle x^2, y\rangle$
   $\mathbf{g}(r,s,t) = \langle r^2s, t^r + 3s^2\rangle$
   With the composition function $\mathbf{w} \circ \mathbf{g}$.

   **Solution:**
   $D\mathbf{w} = \begin{bmatrix} 2x & 0 \\ 0 & 1 \end{bmatrix}$
   $D\mathbf{g} = \begin{bmatrix} 2rs & r^2 & 0 \\ \ln|t|t^r & 6s & rt^{r-1} \end{bmatrix}$
   $D\mathbf{w} \circ \mathbf{g} = \begin{bmatrix} 2r^2s & 0 \\ 0 & 1 \end{bmatrix}$
   $(D\mathbf{w} \circ \mathbf{g})D\mathbf{g} = \begin{bmatrix} 2r^2s & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2rs & r^2 & 0 \\ \ln|t|t^r & 6s & rt^{r-1} \end{bmatrix} = \begin{bmatrix} 4r^3s^2 & 2r^4s & 0 \\ \ln|t|t^r & 6s & rt^{r-1} \end{bmatrix}$