We will prove the Chain Rule, including the proof that the composition of two differentiable functions is differentiable. We will do it for compositions of functions of two variables. The idea is the same for other combinations of finite numbers of variables.

Theorem 1. (Chain Rule) *Denote $w = w(u, v)$; $u = u(x, y)$; and $v = v(x, y)$, where w, u, and v are assumed to be differentiable functions, with the composition $w(u(x, y), v(x, y))$ assumed to be well–defined. Then the* composite function $w(u(x, y), v(x, y))$ *is a* differentiable *function of x and y, and the partial derivatives are given as follows:*

$$w_x = w_u u_x + w_v v_x,$$
$$w_y = w_u u_y + w_v v_y.$$

Proof. We will begin by proving that the *composite* function $w(u(x, y), v(x, y))$ is differentiable. From this the formulas for the partial derivatives will follow quickly. We will proceed by applying the hypothesis of differentiability for each of the functions w, u, and v, prior to the composition.

We will take increments Δx and Δy in the independent variables x and y, which results in increments Δu and Δv in u and v. As $(\Delta x, \Delta y) \to (0, 0)$, it will follow that $(\Delta u, \Delta v) \to 0$ as well, because differentiable functions must be continuous. All ϵ–terms will come from the definition of differentiability, and they must approach zero as the increments in their independent variables approach zero.

$$\begin{aligned}
\Delta w &= (w_u + \epsilon_1)\Delta u + (w_v + \epsilon_2)\Delta v \\
&= (w_u + \epsilon_1)[(u_x + \epsilon_3)\Delta x + (u_y + \epsilon_4)\Delta y] \\
&\quad + (w_v + \epsilon_2)[(v_x + \epsilon_5)\Delta x + (v_y + \epsilon_6)\Delta y] \\
&= (w_u u_x + w_v v_x)\Delta x + (w_u u_y + w_v v_y)\Delta y + \mathcal{E}_1 \Delta x + \mathcal{E}_2 \Delta y, \qquad (1)
\end{aligned}$$

where we must show that $\mathcal{E}_i \to 0$ for $i = 1, 2$, as $(\Delta x, \Delta y) \to (0, 0)$ in order to establish differentiability. We have the following information from Equation (1) as $(\Delta x, \Delta y) \to 0$:

$$\mathcal{E}_1 = (w_u + \epsilon_1)\epsilon_3 + (\epsilon_1 u_x + \epsilon_2 v_x) + (w_v + \epsilon_2)\epsilon_5 \to 0,$$
$$\mathcal{E}_2 = (w_u + \epsilon_1)\epsilon_4 + (\epsilon_1 u_y + \epsilon_2 v_y) + (w_v + \epsilon_2)\epsilon_6 \to 0.$$

Hence we will know that the composite function is differentiable, once we have established the partial differentiation formulas. Then we will know that Equation (1) states that

$$\Delta w = dw + \mathcal{E}_1 \Delta x + \mathcal{E}_2 \Delta y.$$

To find the partial derivative of w with respect to x, set $\Delta y = 0$ in Equation (1), divide both side by Δx, and obtain

$$\frac{\partial w}{\partial x} = \lim_{\Delta x \to 0} (w_u u_x + w_v v_x + \mathcal{E}_1) = w_u u_x + w_v v_x.$$

The other formula is obtained similarly. $\square$