Chain Rule—Type II

For this situation, consider a function $f(x, y)$ of two variables and suppose that the variables $x$ and $y$ are functions of other variables.

For example, consider $x$ and $y$ as a function of the polar coordinates $r$ and $\theta$. That is,

$$x = r \cos \theta \quad \text{and} \quad y = r \sin \theta.$$
Directional derivatives

Partial derivatives only measure rates of change along paths parallel to the axes. Directional derivatives measure the rate of change of a function in any direction.

**Example.** On pages 788 and 790 of your textbook, there is a temperature map for parts of California and Nevada at 3:00 P.M. on a day in October. Let’s estimate the rate that the temperature increases if we leave Reno traveling southeast (toward Las Vegas).
Example. Consider the function \( f(x, y) = 2x^2 + y^2 \). Here is its graph and level sets.

Let’s calculate its derivative at the point \((2, 1)\) in the “northeast” direction.
On the web site there are links to two animations that illustrate the concept of a directional derivative.

**Definition of a Directional Derivative.** We start with the two-variable case. Define the “directional derivative of $f(x, y)$ at the point $(a, b)$ in the $u$-direction” by parametrizing the line through $(a, b)$ using the direction vector $u$. In other words, if $u = u_1i + u_2j$, then the line is written in vector form as $L(h) = (a_1i + b_1j) + hu$ or in parametric form as $(x, y) = (a + u_1h, b + u_2h)$.

Then we compute

$$Du_f(a, b) = \lim_{h \to 0} \frac{f(x, y) - f(a, b)}{h}.$$  

Using vector notation with $P = a_i + b_j$, the same limit is written as

$$Du_f(P) = \lim_{h \to 0} \frac{f(P + hu) - f(P)}{h}.$$  

This vector notation generalizes nicely to functions of three variables or, in fact, to any number of variables.
Computing directional derivatives. A directional derivative for \( f(x, y) \) at the point \((a, b)\) can be computed by applying the Chain Rule to the composition \( f(L(h)) \). Note that the vector \( \mathbf{u} = u_1 \mathbf{i} + u_2 \mathbf{j} \) is used only to indicate the direction, and consequently, it is always a unit vector. In other words, \( u_1^2 + u_2^2 = 1 \).

**Theorem.** \( D_{\mathbf{u}}f(a, b) = [\nabla f(a, b)] \cdot \mathbf{u} \).

**Example.** Calculate the directional derivative of \( f(x, y) = e^x \sin y \) at the point \((\ln 2, \pi/6)\) in the direction of \(2i + j\).
This theorem tells us how a function changes in any given direction, and in particular, it indicates directions of most rapid increase or decrease for the function. Since \( \mathbf{u} \) is a unit vector,

\[
D_{\mathbf{u}} f(a, b) = (\nabla f(a, b)) \cdot \mathbf{u} \\
= |\nabla f(a, b)| |\mathbf{u}| \cos \theta \\
= |\nabla f(a, b)| \cos \theta,
\]

where \( \theta \) is the angle between \( \mathbf{u} \) and the gradient vector \( \nabla f(a, b) \).

For what values of \( \theta \) is this number largest? smallest? zero?

**Theorem.** The function \( f(x, y) \) increases most rapidly in the direction of the gradient. The function is “constant” in directions perpendicular to the gradient.