

## 8.3: The Chain Rule

We all know that the position of a point in space can be specified with two coordinates,  $x$  and  $y$ , called the cartesian coordinates. We also know that we can choose instead to specify the position of the point using the distance from the origin ( $r$ ) and the angle that the vector makes with the  $x$  axis ( $\theta$ ). The latter are what we call plane polar coordinates, which we will cover in much more detail in [Chapter 10](#).

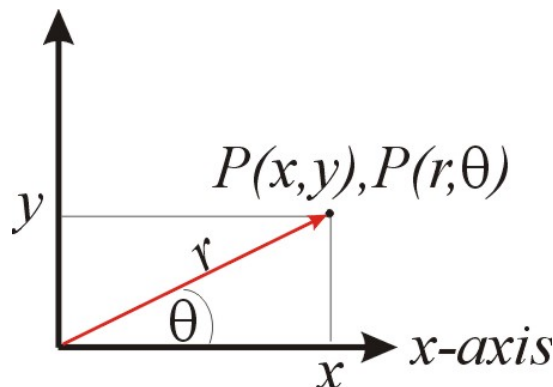


Figure 8.3.1: Cartesian and polar coordinates. (CC BY-NC-SA; [Marcia Levitus](#))

The two coordinate systems are related by:

$$x = r \cos \theta; \quad y = r \sin \theta \quad (8.3.1)$$

$$r = \sqrt{x^2 + y^2}; \quad \theta = \tan^{-1}(y/x) \quad (8.3.2)$$

Let's assume that we are given a function in polar coordinates, for example  $f(r, \theta) = e^{-3r} \cos \theta$ , and we are asked to find the partial derivatives in cartesian coordinates,  $(\partial f / \partial x)_y$  and  $(\partial f / \partial y)_x$ . We can of course re-write the function in terms of  $x$  and  $y$  and find the derivatives we need, but wouldn't it be wonderful if we had a universal formula that converts the derivatives in polar coordinates ( $(\partial f / \partial r)_\theta$  and  $(\partial f / \partial \theta)_r$ ) to the derivatives in cartesian coordinates? This would allow us to take the derivatives in the system the equation is expressed in (which is easy), and then translate the derivatives to the other system without thinking too much. The chain rule will allow us to create these 'universal' relationships between the derivatives of different coordinate systems.

Before using the chain rule, let's obtain  $(\partial f / \partial x)_y$  and  $(\partial f / \partial y)_x$  by re-writing the function in terms of  $x$  and  $y$ . I want to show you how much work this would involve, so you can appreciate how useful using the chain rule is. Using Equations 8.3.1 and 8.3.2, we can rewrite  $f(r, \theta) = e^{-3r} \cos \theta$  as

$$f(x, y) = \frac{e^{-3(x^2+y^2)^{1/2}} x}{(x^2 + y^2)^{1/2}}$$

We can easily obtain  $(\partial f / \partial x)_y$  and  $(\partial f / \partial y)_x$ , but it is certainly quite a bit of work. What if I told you that  $(\partial f / \partial x)_y$  is simply

$$\left( \frac{\partial f}{\partial x} \right)_y = \cos \theta \left( \frac{\partial f}{\partial r} \right)_\theta - \frac{\sin \theta}{r} \left( \frac{\partial f}{\partial \theta} \right)_r \quad (8.3.3)$$

independently of the function  $f$ ? We will derive this result shortly, but for now let me just mention that the procedure involves using the chain rule. You are probably sighing in relief, because the derivatives  $(\partial f / \partial r)_\theta$  and  $(\partial f / \partial \theta)_r$  are much easier to obtain:

$$\begin{aligned} \left( \frac{\partial f}{\partial r} \right)_\theta &= -3e^{-3r} \cos \theta \\ \left( \frac{\partial f}{\partial \theta} \right)_r &= -e^{-3r} \sin \theta \end{aligned}$$

and using Equation 8.3.3, we can obtain the derivative we are looking for:

$$\left( \frac{\partial f}{\partial x} \right)_y = -\cos \theta \times 3e^{-3r} \cos \theta + \frac{\sin \theta}{r} e^{-3r} \sin \theta$$

$$\left(\frac{\partial f}{\partial x}\right)_y = -\cos^2 \theta \times 3e^{-3r} + \frac{\sin^2 \theta}{r} e^{-3r} = e^{-3r} \left(\frac{\sin^2 \theta}{r} - 3\cos^2 \theta\right)$$

$$\left(\frac{\partial f}{\partial x}\right)_y = e^{-3(x^2+y^2)^{1/2}} \left(\frac{y^2}{(x^2+y^2)^{3/2}} - 3\frac{x^2}{(x^2+y^2)^{3/2}}\right)$$

Hopefully this wasn't too painful, or at least, less tedious than it would have been had we used the chain rule. What about  $(\partial f / \partial y)_x$ ? We can create an expression similar to Equation 8.3.3 and use it to relate  $(\partial f / \partial y)_x$  with  $(\partial f / \partial r)_\theta$  and  $(\partial f / \partial \theta)_r$ .

At this point you may be thinking that this all worked well because the function we had was easier to derive in polar coordinates than in cartesian coordinates. True, but this is the whole point. Many physical systems are described in polar coordinates more naturally than in cartesian coordinates (especially in three dimensions). This has to do with the symmetry of the system. For an atom, for example, it is much more natural to use spherical coordinates than cartesian coordinates. We could use cartesian, but the expressions would be much more complex and hard to work with. If we have equations that are more easily expressed in polar coordinates, getting the derivatives in polar coordinates will always be easier. But why would we want the derivatives in cartesian coordinates then? A great example is the Schrödinger equation, which is at the core of quantum mechanics. We will talk more about this when we discuss operators, but for now, the Schrödinger equation is a partial differential equation (unless the particle moves in one dimension) that can be written as:

$$E\psi(\vec{r}) = -\frac{\hbar}{2m}\nabla^2\psi(\vec{r}) + V(\vec{r})\psi(\vec{r})$$

Because of the symmetry of the system, for atoms and molecules it is simpler to express the position of the particle ( $\vec{r}$ ) in spherical coordinates. However, the operator  $\nabla^2$  (known as the Laplacian) is defined in cartesian coordinates as:

$$\nabla^2 f(x, y, z) = \left(\frac{\partial^2 f}{\partial x^2}\right)_{y,z} + \left(\frac{\partial^2 f}{\partial y^2}\right)_{x,z} + \left(\frac{\partial^2 f}{\partial z^2}\right)_{x,y}$$

In other words, the Laplacian instructs you to take the second derivatives of the function with respect to  $x$ , with respect to  $y$  and with respect to  $z$ , and add the three together. We could express the functions  $V(\vec{r})$  and  $\psi(\vec{r})$  in cartesian coordinates, but again, this would lead to a terribly complex differential equation. Instead, we can express the Laplacian in spherical coordinates, and this is in fact the best approach. To do this, we would need to relate the derivatives in spherical coordinates to the derivatives in cartesian coordinates, and this is done using the chain rule.

Hopefully all this convinced you of the uses of the chain rule in the physical sciences, so now we just need to see how to use it for our purposes. In two dimensions, the chain rule states that if we have a function in one coordinate system  $u(x, y)$ , and these coordinates are functions of two other variables (e.g.  $x = x(\theta, r)$  and  $y = y(\theta, r)$ ) then:

$$\left(\frac{\partial u}{\partial r}\right)_\theta = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial r}\right)_\theta + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial r}\right)_\theta$$

$$\left(\frac{\partial u}{\partial \theta}\right)_r = \left(\frac{\partial u}{\partial x}\right)_y \left(\frac{\partial x}{\partial \theta}\right)_r + \left(\frac{\partial u}{\partial y}\right)_x \left(\frac{\partial y}{\partial \theta}\right)_r$$

Some students find the following 'tree' constructions useful:

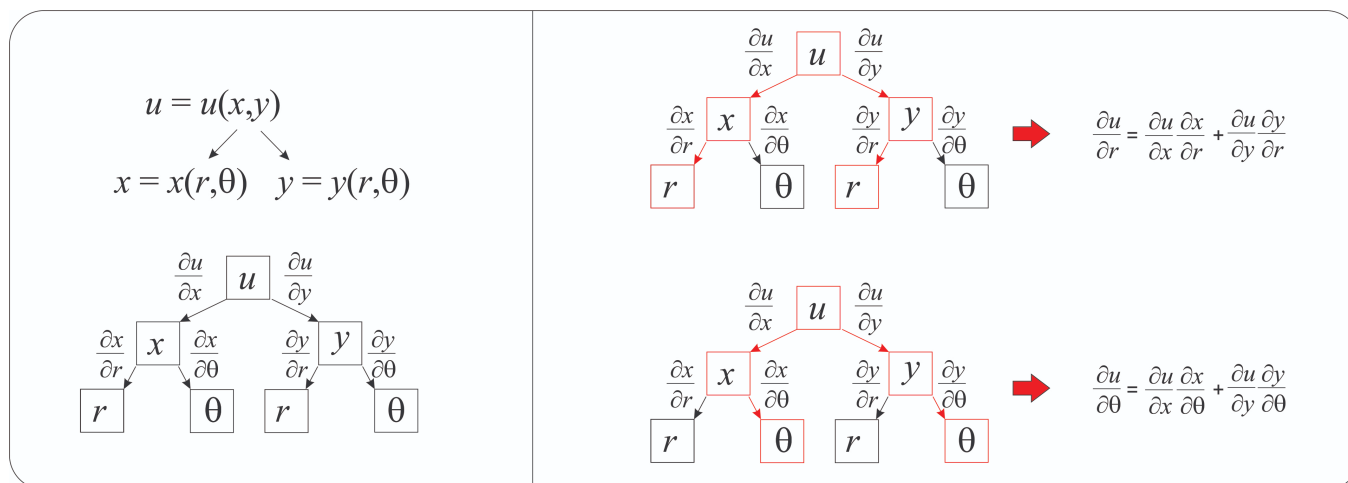


Figure 8.3.2: The chain rule (CC BY-NC-SA; Marcia Levitus)

We can also consider  $u = u(r, \theta)$ , and  $\theta = \theta(x, y)$  and  $r = r(x, y)$ , which gives:

$$\left(\frac{\partial u}{\partial x}\right)_y = \left(\frac{\partial u}{\partial r}\right)_\theta \left(\frac{\partial r}{\partial x}\right)_y + \left(\frac{\partial u}{\partial \theta}\right)_r \left(\frac{\partial \theta}{\partial x}\right)_y$$

$$\left(\frac{\partial u}{\partial y}\right)_x = \left(\frac{\partial u}{\partial r}\right)_\theta \left(\frac{\partial r}{\partial y}\right)_x + \left(\frac{\partial u}{\partial \theta}\right)_r \left(\frac{\partial \theta}{\partial y}\right)_x$$

### ✓ Example 8.3.1

Derive Equation 8.3.3.

#### Solution

We need to prove  $\left(\frac{\partial f}{\partial x}\right)_y = \cos \theta \left(\frac{\partial f}{\partial r}\right)_\theta - \frac{\sin \theta}{r} \left(\frac{\partial f}{\partial \theta}\right)_r$ . Using the chain rule:

$$\left(\frac{\partial f}{\partial x}\right)_y = \left(\frac{\partial f}{\partial r}\right)_\theta \left(\frac{\partial r}{\partial x}\right)_y + \left(\frac{\partial f}{\partial \theta}\right)_r \left(\frac{\partial \theta}{\partial x}\right)_y$$

From Equation 8.3.1 and 8.3.2

$$\left(\frac{\partial r}{\partial x}\right)_y = \frac{1}{2}(x^2 + y^2)^{-1/2}(2x) = \frac{1}{2}(r^2)^{-1/2}(2r \cos \theta) = \cos \theta$$

$$\left(\frac{\partial \theta}{\partial x}\right)_y = \frac{1}{1 + (y/x)^2} \frac{(-y)}{x^2} = -\frac{1}{1 + (y/x)^2} \frac{y}{x} \frac{1}{x} = -\frac{1}{1 + \tan^2 \theta} \tan \theta \frac{1}{r \cos \theta} = -\frac{1}{1 + \frac{\sin^2 \theta}{\cos^2 \theta}} \frac{\sin \theta}{\cos \theta} \frac{1}{r \cos \theta} =$$

$$-\frac{\sin \theta}{r}$$

Therefore,

$$\left(\frac{\partial f}{\partial x}\right)_y = \cos \theta \left(\frac{\partial f}{\partial r}\right)_\theta - \frac{\sin \theta}{r} \left(\frac{\partial f}{\partial \theta}\right)_r$$

Need help? The videos below contain examples of how to use the chain rule for partial derivatives:

- Example 1: <http://www.youtube.com/watch?v=HOYA0-pOHsg>
- Example 2: <http://www.youtube.com/watch?v=kCr13iTRN7E> (tree diagrams)

This page titled [8.3: The Chain Rule](#) is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by [Marcia Levitus](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.