14.4 The Chain Rule

The Chain Rule for functions of a single variable studied in Section 3.5 said that when $w = f(x)$ was a differentiable function of x and $x = g(t)$ was a differentiable function of t, w became a differentiable function of t and dw/dt could be calculated with the formula

$$\frac{dw}{dt} = \frac{dw}{dx}\frac{dx}{dt}.$$

For functions of two or more variables the Chain Rule has several forms. The form depends on how many variables are involved but works like the Chain Rule in Section 3.5 once we account for the presence of additional variables.

Functions of Two Variables

The Chain Rule formula for a function $w = f(x, y)$ when $x = x(t)$ and $y = y(t)$ are both differentiable functions of t is given in the following theorem.

> **THEOREM 5 Chain Rule for Functions of Two Independent Variables**
> If $w = f(x, y)$ has continuous partial derivatives f_x and f_y and if $x = x(t), y = y(t)$ are differentiable functions of t, then the composite $w = f(x(t), y(t))$ is a differentiable function of t and
>
> $$\frac{df}{dt} = f_x(x(t), y(t)) \cdot x'(t) + f_y(x(t), y(t)) \cdot y'(t),$$
>
> or
>
> $$\frac{dw}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Proof The proof consists of showing that if x and y are differentiable at $t = t_0$, then w is differentiable at t_0 and

$$\left(\frac{dw}{dt}\right)_{t_0} = \left(\frac{\partial w}{\partial x}\right)_{P_0}\left(\frac{dx}{dt}\right)_{t_0} + \left(\frac{\partial w}{\partial y}\right)_{P_0}\left(\frac{dy}{dt}\right)_{t_0},$$

where $P_0 = (x(t_0), y(t_0))$. The subscripts indicate where each of the derivatives are to be evaluated.

Let Δx, Δy, and Δw be the increments that result from changing t from t_0 to $t_0 + \Delta t$. Since f is differentiable (see the definition in Section 14.3),

$$\Delta w = \left(\frac{\partial w}{\partial x}\right)_{P_0}\Delta x + \left(\frac{\partial w}{\partial y}\right)_{P_0}\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y,$$

where $\epsilon_1, \epsilon_2 \to 0$ as $\Delta x, \Delta y \to 0$. To find dw/dt, we divide this equation through by Δt and let Δt approach zero. The division gives

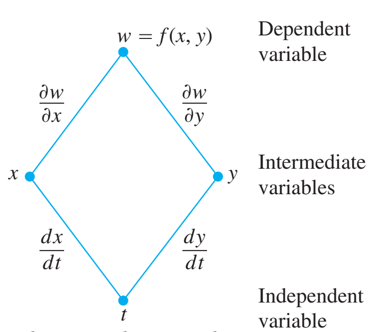
$$\frac{\Delta w}{\Delta t} = \left(\frac{\partial w}{\partial x}\right)_{P_0}\frac{\Delta x}{\Delta t} + \left(\frac{\partial w}{\partial y}\right)_{P_0}\frac{\Delta y}{\Delta t} + \epsilon_1\frac{\Delta x}{\Delta t} + \epsilon_2\frac{\Delta y}{\Delta t}.$$

Letting Δt approach zero gives

$$\begin{aligned}\left(\frac{dw}{dt}\right)_{t_0} &= \lim_{\Delta t \to 0}\frac{\Delta w}{\Delta t}\\ &= \left(\frac{\partial w}{\partial x}\right)_{P_0}\left(\frac{dx}{dt}\right)_{t_0} + \left(\frac{\partial w}{\partial y}\right)_{P_0}\left(\frac{dy}{dt}\right)_{t_0} + 0\cdot\left(\frac{dx}{dt}\right)_{t_0} + 0\cdot\left(\frac{dy}{dt}\right)_{t_0}. \qquad \blacksquare\end{aligned}$$

To remember the Chain Rule picture the diagram below. To find dw/dt, start at w and read down each route to t, multiplying derivatives along the way. Then add the products.

Chain Rule



$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt}$$

The **tree diagram** in the margin provides a convenient way to remember the Chain Rule. From the diagram, you see that when $t = t_0$, the derivatives dx/dt and dy/dt are evaluated at t_0. The value of t_0 then determines the value x_0 for the differentiable function x and the value y_0 for the differentiable function y. The partial derivatives $\partial w/\partial x$ and $\partial w/\partial y$ (which are themselves functions of x and y) are evaluated at the point $P_0(x_0, y_0)$ corresponding to t_0. The "true" independent variable is t, whereas x and y are *intermediate variables* (controlled by t) and w is the dependent variable.

A more precise notation for the Chain Rule shows how the various derivatives in Theorem 5 are evaluated:

$$\frac{dw}{dt}(t_0) = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{dy}{dt}(t_0).$$

EXAMPLE 1 Applying the Chain Rule

Use the Chain Rule to find the derivative of

$$w = xy$$

with respect to t along the path $x = \cos t, y = \sin t$. What is the derivative's value at $t = \pi/2$?

Solution We apply the Chain Rule to find dw/dt as follows:

$$\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} \\
&= \frac{\partial(xy)}{\partial x} \cdot \frac{d}{dt}(\cos t) + \frac{\partial(xy)}{\partial y} \cdot \frac{d}{dt}(\sin t) \\
&= (y)(-\sin t) + (x)(\cos t) \\
&= (\sin t)(-\sin t) + (\cos t)(\cos t) \\
&= -\sin^2 t + \cos^2 t \\
&= \cos 2t.
\end{aligned}$$

In this example, we can check the result with a more direct calculation. As a function of t,

$$w = xy = \cos t \sin t = \frac{1}{2}\sin 2t,$$

so

$$\frac{dw}{dt} = \frac{d}{dt}\left(\frac{1}{2}\sin 2t\right) = \frac{1}{2} \cdot 2\cos 2t = \cos 2t.$$

In either case, at the given value of t,

$$\left(\frac{dw}{dt}\right)_{t=\pi/2} = \cos\left(2 \cdot \frac{\pi}{2}\right) = \cos \pi = -1.$$

Functions of Three Variables

You can probably predict the Chain Rule for functions of three variables, as it only involves adding the expected third term to the two-variable formula.

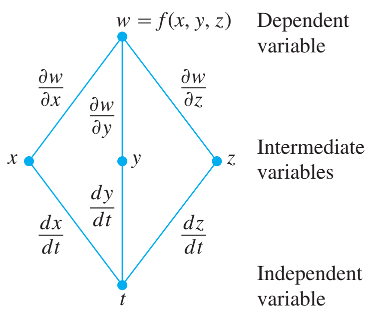
THEOREM 6 Chain Rule for Functions of Three Independent Variables

If $w = f(x, y, z)$ is differentiable and x, y, and z are differentiable functions of t, then w is a differentiable function of t and

$$\frac{dw}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}.$$

Here we have three routes from w to t instead of two, but finding dw/dt is still the same. Read down each route, multiplying derivatives along the way; then add.

Chain Rule



$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$$

The proof is identical with the proof of Theorem 5 except that there are now three intermediate variables instead of two. The diagram we use for remembering the new equation is similar as well, with three routes from w to t.

EXAMPLE 2 Changes in a Function's Values Along a Helix

Find dw/dt if

$$w = xy + z, \qquad x = \cos t, \qquad y = \sin t, \qquad z = t.$$

In this example the values of w are changing along the path of a helix (Section 13.1). What is the derivative's value at $t = 0$?

Solution

$$\begin{aligned}
\frac{dw}{dt} &= \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt} \\
&= (y)(-\sin t) + (x)(\cos t) + (1)(1) \\
&= (\sin t)(-\sin t) + (\cos t)(\cos t) + 1 \qquad \text{Substitute for the intermediate variables.} \\
&= -\sin^2 t + \cos^2 t + 1 = 1 + \cos 2t.
\end{aligned}$$

$$\left(\frac{dw}{dt}\right)_{t=0} = 1 + \cos(0) = 2.$$

■

Here is a physical interpretation of change along a curve. If $w = T(x, y, z)$ is the temperature at each point (x, y, z) along a curve C with parametric equations $x = x(t)$, $y = y(t)$, and $z = z(t)$, then the composite function $w = T(x(t), y(t), z(t))$ represents the temperature relative to t along the curve. The derivative dw/dt is then the instantaneous rate of change of temperature along the curve, as calculated in Theorem 6.

Functions Defined on Surfaces

If we are interested in the temperature $w = f(x, y, z)$ at points (x, y, z) on a globe in space, we might prefer to think of x, y, and z as functions of the variables r and s that give the points' longitudes and latitudes. If $x = g(r, s)$, $y = h(r, s)$, and $z = k(r, s)$, we could then express the temperature as a function of r and s with the composite function

$$w = f(g(r, s), h(r, s), k(r, s)).$$

Under the right conditions, w would have partial derivatives with respect to both r and s that could be calculated in the following way.

THEOREM 7 Chain Rule for Two Independent Variables and Three Intermediate Variables

Suppose that $w = f(x, y, z)$, $x = g(r, s)$, $y = h(r, s)$, and $z = k(r, s)$. If all four functions are differentiable, then w has partial derivatives with respect to r and s, given by the formulas

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r}$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}.$$

The first of these equations can be derived from the Chain Rule in Theorem 6 by holding s fixed and treating r as t. The second can be derived in the same way, holding r fixed and treating s as t. The tree diagrams for both equations are shown in Figure 14.19.

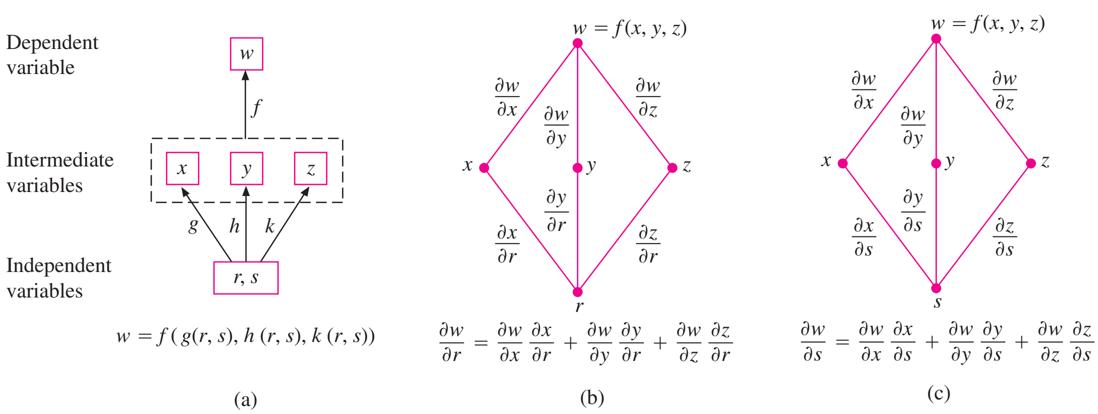


FIGURE 14.19 Composite function and tree diagrams for Theorem 7.

EXAMPLE 3 Partial Derivatives Using Theorem 7

Express $\partial w/\partial r$ and $\partial w/\partial s$ in terms of r and s if

$$w = x + 2y + z^2, \qquad x = \frac{r}{s}, \qquad y = r^2 + \ln s, \qquad z = 2r.$$

Solution

$$\begin{aligned}\frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial r}\\ &= (1)\left(\frac{1}{s}\right) + (2)(2r) + (2z)(2)\\ &= \frac{1}{s} + 4r + (4r)(2) = \frac{1}{s} + 12r\end{aligned}$$

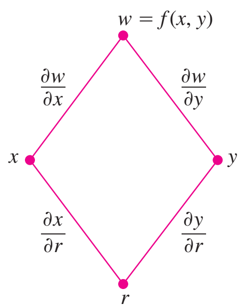
Substitute for intermediate variable z.

$$\begin{aligned}\frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} + \frac{\partial w}{\partial z}\frac{\partial z}{\partial s}\\ &= (1)\left(-\frac{r}{s^2}\right) + (2)\left(\frac{1}{s}\right) + (2z)(0) = \frac{2}{s} - \frac{r}{s^2}\end{aligned}$$

■

If f is a function of two variables instead of three, each equation in Theorem 7 becomes correspondingly one term shorter.

Chain Rule



$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r}$$

FIGURE 14.20 Tree diagram for the equation

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r}.$$

> If $w = f(x, y)$, $x = g(r, s)$, and $y = h(r, s)$, then
>
> $$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} \quad \text{and} \quad \frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s}.$$

Figure 14.20 shows the tree diagram for the first of these equations. The diagram for the second equation is similar; just replace r with s.

EXAMPLE 4 More Partial Derivatives

You Try It

Express $\partial w/\partial r$ and $\partial w/\partial s$ in terms of r and s if

$$w = x^2 + y^2, \qquad x = r - s, \qquad y = r + s.$$

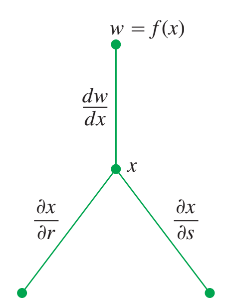
Solution

$$\begin{aligned} \frac{\partial w}{\partial r} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial r} \\ &= (2x)(1) + (2y)(1) \\ &= 2(r - s) + 2(r + s) \\ &= 4r \end{aligned} \qquad \begin{aligned} \frac{\partial w}{\partial s} &= \frac{\partial w}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial s} \\ &= (2x)(-1) + (2y)(1) \\ &= -2(r - s) + 2(r + s) \\ &= 4s \end{aligned}$$

Substitute for the intermediate variables.

If f is a function of x alone, our equations become even simpler.

Chain Rule



$$\frac{\partial w}{\partial r} = \frac{dw}{dx}\frac{\partial x}{\partial r}$$

$$\frac{\partial w}{\partial s} = \frac{dw}{dx}\frac{\partial x}{\partial s}$$

FIGURE 14.21 Tree diagram for differentiating f as a composite function of r and s with one intermediate variable.

> If $w = f(x)$ and $x = g(r, s)$, then
>
> $$\frac{\partial w}{\partial r} = \frac{dw}{dx}\frac{\partial x}{\partial r} \quad \text{and} \quad \frac{\partial w}{\partial s} = \frac{dw}{dx}\frac{\partial x}{\partial s}.$$

In this case, we can use the ordinary (single-variable) derivative, dw/dx. The tree diagram is shown in Figure 14.21.

Implicit Differentiation Revisited

The two-variable Chain Rule in Theorem 5 leads to a formula that takes most of the work out of implicit differentiation. Suppose that

1. The function $F(x, y)$ is differentiable and
2. The equation $F(x, y) = 0$ defines y implicitly as a differentiable function of x, say $y = h(x)$.

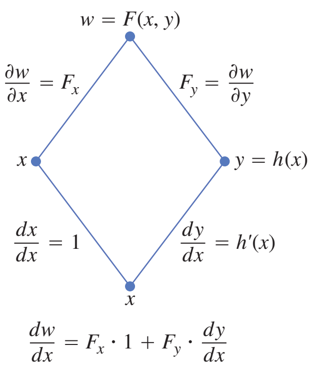


FIGURE 14.22 Tree diagram for differentiating $w = F(x, y)$ with respect to x. Setting $dw/dx = 0$ leads to a simple computational formula for implicit differentiation (Theorem 8).

Since $w = F(x, y) = 0$, the derivative dw/dx must be zero. Computing the derivative from the Chain Rule (tree diagram in Figure 14.22), we find

$$\begin{aligned} 0 = \frac{dw}{dx} &= F_x \frac{dx}{dx} + F_y \frac{dy}{dx} \qquad \text{Theorem 5 with } t = x \text{ and } f = F \\ &= F_x \cdot 1 + F_y \cdot \frac{dy}{dx}. \end{aligned}$$

If $F_y = \partial w/\partial y \neq 0$, we can solve this equation for dy/dx to get

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

This relationship gives a surprisingly simple shortcut to finding derivatives of implicitly defined functions, which we state here as a theorem.

THEOREM 8 A Formula for Implicit Differentiation

Suppose that $F(x, y)$ is differentiable and that the equation $F(x, y) = 0$ defines y as a differentiable function of x. Then at any point where $F_y \neq 0$,

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

EXAMPLE 5 Implicit Differentiation

You Try It

Use Theorem 8 to find dy/dx if $y^2 - x^2 - \sin xy = 0$.

Solution Take $F(x, y) = y^2 - x^2 - \sin xy$. Then

$$\begin{aligned} \frac{dy}{dx} &= -\frac{F_x}{F_y} = -\frac{-2x - y\cos xy}{2y - x\cos xy} \\ &= \frac{2x + y\cos xy}{2y - x\cos xy}. \end{aligned}$$

This calculation is significantly shorter than the single-variable calculation with which we found dy/dx in Section 3.6, Example 3. ■

Functions of Many Variables

We have seen several different forms of the Chain Rule in this section, but you do not have to memorize them all if you can see them as special cases of the same general formula. When solving particular problems, it may help to draw the appropriate tree diagram by placing the dependent variable on top, the intermediate variables in the middle, and the selected independent variable at the bottom. To find the derivative of the dependent variable with respect to the selected independent variable, start at the dependent variable and read down each route of the tree to the independent variable, calculating and multiplying the derivatives along each route. Then add the products you found for the different routes.

In general, suppose that $w = f(x, y, \ldots, v)$ is a differentiable function of the variables $x, y, \ldots, v$ (a finite set) and the $x, y, \ldots, v$ are differentiable functions of $p, q, \ldots, t$ (another finite set). Then w is a differentiable function of the variables p through t and the partial derivatives of w with respect to these variables are given by equations of the form

$$\frac{\partial w}{\partial p} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial p} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial p} + \cdots + \frac{\partial w}{\partial v}\frac{\partial v}{\partial p}.$$

The other equations are obtained by replacing p by $q, \ldots, t$, one at a time.

One way to remember this equation is to think of the right-hand side as the dot product of two vectors with components

$$\underbrace{\left(\frac{\partial w}{\partial x}, \frac{\partial w}{\partial y}, \ldots, \frac{\partial w}{\partial v}\right)}_{\substack{\text{Derivatives of } w \text{ with} \\ \text{respect to the} \\ \text{intermediate variables}}} \quad \text{and} \quad \underbrace{\left(\frac{\partial x}{\partial p}, \frac{\partial y}{\partial p}, \ldots, \frac{\partial v}{\partial p}\right)}_{\substack{\text{Derivatives of the intermediate} \\ \text{variables with respect to the} \\ \text{selected independent variable}}}.$$