

A.7**The Mixed Derivative Theorem and the Increment Theorem**

This appendix derives the Mixed Derivative Theorem (Theorem 2, Section 14.3) and the Increment Theorem for Functions of Two Variables (Theorem 3, Section 14.3). Euler first published the Mixed Derivative Theorem in 1734, in a series of papers he wrote on hydrodynamics.

THEOREM 2 The Mixed Derivative Theorem

If $f(x, y)$ and its partial derivatives f_x, f_y, f_{xy} , and f_{yx} are defined throughout an open region containing a point (a, b) and are all continuous at (a, b) , then $f_{xy}(a, b) = f_{yx}(a, b)$.

Proof The equality of $f_{xy}(a, b)$ and $f_{yx}(a, b)$ can be established by four applications of the Mean Value Theorem (Theorem 4, Section 4.2). By hypothesis, the point (a, b) lies in the interior of a rectangle R in the xy -plane on which f, f_x, f_y, f_{xy} , and f_{yx} are all defined. We let h and k be the numbers such that the point $(a + h, b + k)$ also lies in R , and we consider the difference

$$\Delta = F(a + h) - F(a), \quad (1)$$

where

$$F(x) = f(x, b + k) - f(x, b). \quad (2)$$

We apply the Mean Value Theorem to F , which is continuous because it is differentiable. Then Equation (1) becomes

$$\Delta = hF'(c_1), \quad (3)$$

where c_1 lies between a and $a + h$. From Equation (2),

$$F'(x) = f_x(x, b + k) - f_x(x, b),$$

so Equation (3) becomes

$$\Delta = h[f_x(c_1, b + k) - f_x(c_1, b)]. \quad (4)$$

Now we apply the Mean Value Theorem to the function $g(y) = f_x(c_1, y)$ and have

$$g(b + k) - g(b) = kg'(d_1),$$

or

$$f_x(c_1, b + k) - f_x(c_1, b) = kf_{xy}(c_1, d_1)$$

for some d_1 between b and $b + k$. By substituting this into Equation (4), we get

$$\Delta = hkf_{xy}(c_1, d_1) \quad (5)$$

for some point (c_1, d_1) in the rectangle R' whose vertices are the four points (a, b) , $(a + h, b)$, $(a + h, b + k)$, and $(a, b + k)$. (See Figure A.12.)

By substituting from Equation (2) into Equation (1), we may also write

$$\begin{aligned} \Delta &= f(a + h, b + k) - f(a + h, b) - f(a, b + k) + f(a, b) \\ &= [f(a + h, b + k) - f(a, b + k)] - [f(a + h, b) - f(a, b)] \\ &= \phi(b + k) - \phi(b), \end{aligned} \quad (6)$$

where

$$\phi(y) = f(a + h, y) - f(a, y). \quad (7)$$

The Mean Value Theorem applied to Equation (6) now gives

$$\Delta = k\phi'(d_2) \quad (8)$$

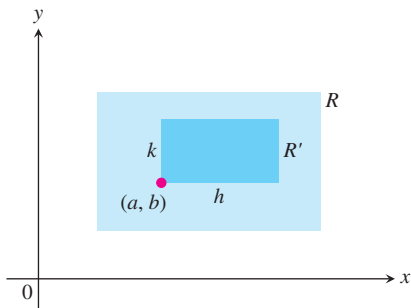


FIGURE A.12 The key to proving $f_{xy}(a, b) = f_{yx}(a, b)$ is that no matter how small R' is, f_{xy} and f_{yx} take on equal values somewhere inside R' (although not necessarily at the same point).

for some d_2 between b and $b + k$. By Equation (7),

$$\phi'(y) = f_y(a + h, y) - f_y(a, y). \quad (9)$$

Substituting from Equation (9) into Equation (8) gives

$$\Delta = k[f_y(a + h, d_2) - f_y(a, d_2)].$$

Finally, we apply the Mean Value Theorem to the expression in brackets and get

$$\Delta = khf_{yx}(c_2, d_2) \quad (10)$$

for some c_2 between a and $a + h$.

Together, Equations (5) and (10) show that

$$f_{xy}(c_1, d_1) = f_{yx}(c_2, d_2), \quad (11)$$

where (c_1, d_1) and (c_2, d_2) both lie in the rectangle R' (Figure A.12). Equation (11) is not quite the result we want, since it says only that f_{xy} has the same value at (c_1, d_1) that f_{yx} has at (c_2, d_2) . The numbers h and k in our discussion, however, may be made as small as we wish. The hypothesis that f_{xy} and f_{yx} are both continuous at (a, b) means that $f_{xy}(c_1, d_1) = f_{xy}(a, b) + \epsilon_1$ and $f_{yx}(c_2, d_2) = f_{yx}(a, b) + \epsilon_2$, where each of $\epsilon_1, \epsilon_2 \rightarrow 0$ as both $h, k \rightarrow 0$. Hence, if we let h and $k \rightarrow 0$, we have $f_{xy}(a, b) = f_{yx}(a, b)$. ■

The equality of $f_{xy}(a, b)$ and $f_{yx}(a, b)$ can be proved with hypotheses weaker than the ones we assumed. For example, it is enough for f, f_x , and f_y to exist in R and for f_{xy} to be continuous at (a, b) . Then f_{yx} will exist at (a, b) and equal f_{xy} at that point.

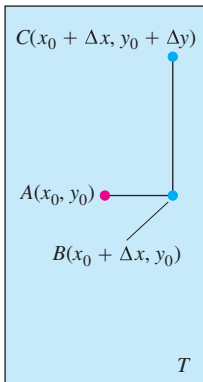


FIGURE A.13 The rectangular region T in the proof of the Increment Theorem. The figure is drawn for Δx and Δy positive, but either increment might be zero or negative.

THEOREM 3 The Increment Theorem for Functions of Two Variables

Suppose that the first partial derivatives of $z = f(x, y)$ are defined throughout an open region R containing the point (x_0, y_0) and that f_x and f_y are continuous at (x_0, y_0) . Then the change $\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0)$ in the value of f that results from moving from (x_0, y_0) to another point $(x_0 + \Delta x, y_0 + \Delta y)$ in R satisfies an equation of the form

$$\Delta z = f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y,$$

in which each of $\epsilon_1, \epsilon_2 \rightarrow 0$ as both $\Delta x, \Delta y \rightarrow 0$.

Proof We work within a rectangle T centered at $A(x_0, y_0)$ and lying within R , and we assume that Δx and Δy are already so small that the line segment joining A to $B(x_0 + \Delta x, y_0)$ and the line segment joining B to $C(x_0 + \Delta x, y_0 + \Delta y)$ lie in the interior of T (Figure A.13).

We may think of Δz as the sum $\Delta z = \Delta z_1 + \Delta z_2$ of two increments, where

$$\Delta z_1 = f(x_0 + \Delta x, y_0) - f(x_0, y_0)$$

is the change in the value of f from A to B and

$$\Delta z_2 = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)$$

is the change in the value of f from B to C (Figure A.14).

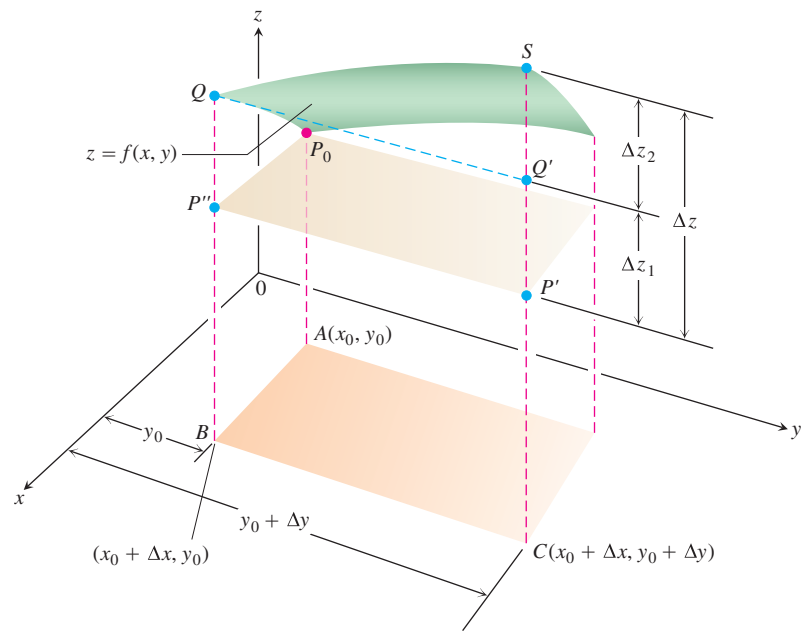


FIGURE A.14 Part of the surface $z = f(x, y)$ near $P_0(x_0, y_0, f(x_0, y_0))$. The points P_0 , P' , and P'' have the same height $z_0 = f(x_0, y_0)$ above the xy -plane. The change in z is $\Delta z = P'S$. The change

$$\Delta z_1 = f(x_0 + \Delta x, y_0) - f(x_0, y_0),$$

shown as $P''Q = P'Q'$, is caused by changing x from x_0 to $x_0 + \Delta x$ while holding y equal to y_0 . Then, with x held equal to $x_0 + \Delta x$,

$$\Delta z_2 = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0)$$

is the change in z caused by changing y_0 from y_0 to $y_0 + \Delta y$, which is represented by $Q'S$? The total change in z is the sum of Δz_1 and Δz_2 .

On the closed interval of x -values joining x_0 to $x_0 + \Delta x$, the function $F(x) = f(x, y_0)$ is a differentiable (and hence continuous) function of x , with derivative

$$F'(x) = f_x(x, y_0).$$

By the Mean Value Theorem (Theorem 4, Section 4.2), there is an x -value c between x_0 and $x_0 + \Delta x$ at which

$$F(x_0 + \Delta x) - F(x_0) = F'(c)\Delta x$$

or

$$f(x_0 + \Delta x, y_0) - f(x_0, y_0) = f_x(c, y_0)\Delta x$$

or

$$\Delta z_1 = f_x(c, y_0)\Delta x. \tag{12}$$

Similarly, $G(y) = f(x_0 + \Delta x, y)$ is a differentiable (and hence continuous) function of y on the closed y -interval joining y_0 and $y_0 + \Delta y$, with derivative

$$G'(y) = f_y(x_0 + \Delta x, y).$$

Hence, there is a y -value d between y_0 and $y_0 + \Delta y$ at which

$$G(y_0 + \Delta y) - G(y_0) = G'(d)\Delta y$$

or

$$f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y) = f_y(x_0 + \Delta x, d)\Delta y$$

or

$$\Delta z_2 = f_y(x_0 + \Delta x, d)\Delta y. \tag{13}$$

Now, as both Δx and $\Delta y \rightarrow 0$, we know that $c \rightarrow x_0$ and $d \rightarrow y_0$. Therefore, since f_x and f_y are continuous at (x_0, y_0) , the quantities

$$\begin{aligned} \epsilon_1 &= f_x(c, y_0) - f_x(x_0, y_0), \\ \epsilon_2 &= f_y(x_0 + \Delta x, d) - f_y(x_0, y_0) \end{aligned} \tag{14}$$

both approach zero as both Δx and $\Delta y \rightarrow 0$.

Finally,

$$\begin{aligned} \Delta z &= \Delta z_1 + \Delta z_2 \\ &= f_x(c, y_0)\Delta x + f_y(x_0 + \Delta x, d)\Delta y && \text{From Equations (12) and (13)} \\ &= [f_x(x_0, y_0) + \epsilon_1]\Delta x + [f_y(x_0, y_0) + \epsilon_2]\Delta y && \text{From Equation (14)} \\ &= f_x(x_0, y_0)\Delta x + f_y(x_0, y_0)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y, \end{aligned}$$

where both ϵ_1 and $\epsilon_2 \rightarrow 0$ as both Δx and $\Delta y \rightarrow 0$, which is what we set out to prove. ■

Analogous results hold for functions of any finite number of independent variables. Suppose that the first partial derivatives of $w = f(x, y, z)$ are defined throughout an open region containing the point (x_0, y_0, z_0) and that f_x, f_y , and f_z are continuous at (x_0, y_0, z_0) . Then

$$\begin{aligned} \Delta w &= f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) - f(x_0, y_0, z_0) \\ &= f_x\Delta x + f_y\Delta y + f_z\Delta z + \epsilon_1\Delta x + \epsilon_2\Delta y + \epsilon_3\Delta z, \end{aligned} \tag{15}$$

where $\epsilon_1, \epsilon_2, \epsilon_3 \rightarrow 0$ as $\Delta x, \Delta y$, and $\Delta z \rightarrow 0$.

The partial derivatives f_x, f_y, f_z in Equation (15) are to be evaluated at the point (x_0, y_0, z_0) .

Equation (15) can be proved by treating Δw as the sum of three increments,

$$\Delta w_1 = f(x_0 + \Delta x, y_0, z_0) - f(x_0, y_0, z_0) \tag{16}$$

$$\Delta w_2 = f(x_0 + \Delta x, y_0 + \Delta y, z_0) - f(x_0 + \Delta x, y_0, z_0) \tag{17}$$

$$\Delta w_3 = f(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z) - f(x_0 + \Delta x, y_0 + \Delta y, z_0), \tag{18}$$

and applying the Mean Value Theorem to each of these separately. Two coordinates remain constant and only one varies in each of these partial increments $\Delta w_1, \Delta w_2, \Delta w_3$. In Equation (17), for example, only y varies, since x is held equal to $x_0 + \Delta x$ and z is held equal to z_0 . Since $f(x_0 + \Delta x, y, z_0)$ is a continuous function of y with a derivative f_y , it is subject to the Mean Value Theorem, and we have

$$\Delta w_2 = f_y(x_0 + \Delta x, y_1, z_0)\Delta y$$

for some y_1 between y_0 and $y_0 + \Delta y$.