Partial Derivatives

Functions of Several Variables

Note. We now consider functions whose domains are sets of ordered pairs, ordered triples, or in general ordered n-tuples of real numbers, and whose ranges are subsets of the real numbers. For example, a function of two variables might be of the form $F(x, y) = x^2 + y^2$. You have seen this when defining what it means for a function $y = f(x)$ to be *implicit* to an equation $F(x, y) = 0$. We now define this more clearly.

Definition. Suppose D is a set of n-tuples of real numbers $\{(x_1, x_2, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ for } 1 \leq i \leq n\}$. A *real-valued function* f on D is a rule that assigns a unique (single) real number $w = f(x_1, x_2, \ldots, x_n)$ to each element in D. The set D is the function's *domain*. The set of w-values taken on by f is the function's *range*. The symbol w is the *dependent variable* of f, and f is said to be a function of the n *independent variables* x_1 to x_n.

Note. We can think of f as a mapping from the "space" $\mathbb{R}^n$ of ordered n-tuples of real numbers to the set $\mathbb{R}$ of real numbers.

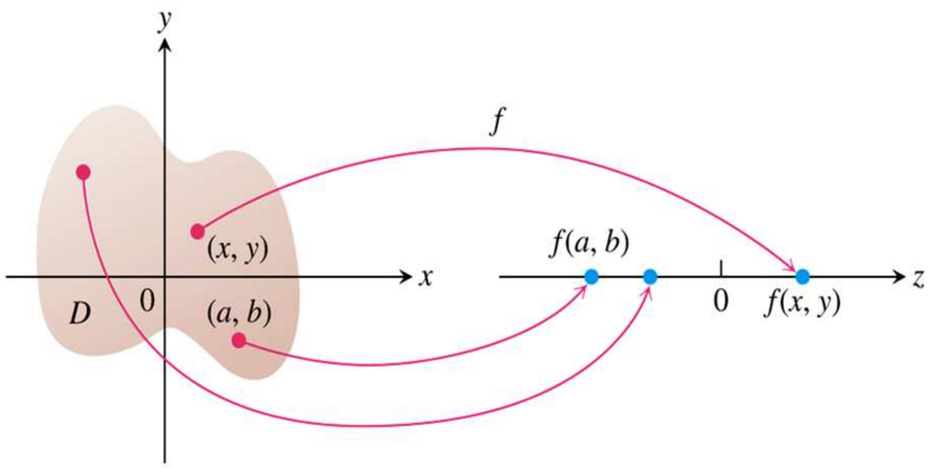


Figure 1

Definition. A point (x_0, y_0) in a region R in the xy-plane is an *interior point* of R if it is the center of a disk of positive radius that lies entirely in R. A point (x_0, y_0) is a *boundary point* of R if every disk centered at (x_0, y_0) contains points that lie outside of R as well as points that lie in R. A point (x_0, y_0) is a *limit point* of R if every disk centered at (x_0, y_0) contains a point that lies in R other than (x_0, y_0) itself. (Boundary points and limit points may or may not be in R). The interior points of a region make up the *interior* of the region. The region's boundary point make up its *boundary*. A region is *open* if it consists entirely of interior points. A region is *closed* if it contains all of its boundary points. The *closure* of a region consists of all the points in the set and all limit points of the set.

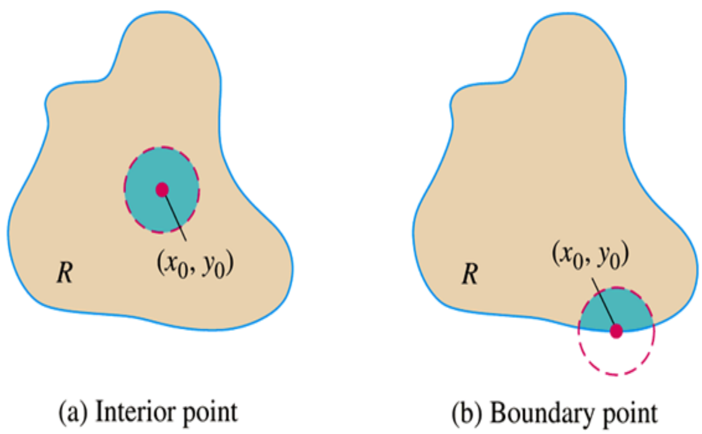


Figure 2

Definition. A region in the plane is *bounded* if it lies inside a disk of fixed radius. A region is *unbounded* if it is not bounded.

Note. There are two ways to picture the values of a function $f(x, y)$. One is to draw and label curves in the domain on which f has a constant value (we concentrate on this technique). The other is to sketch the surface $z = f(x, y)$ in space.

Definition. The set of points in the plane where a function $f(x, y)$ has a constant value $f(x, y) = c$ is called a *level curve* of f. The set of all points $(x, y, f(x, y))$ in space, for (x, y) in the domain of f, is called the *graph* of f. The graph of f is also called the *surface* $z = f(x, y)$.

A relief map of Johnson City and Buffalo Mountain (from Google Maps).

Definition. The set of points (x, y, z) in space where a function of three independent variables has a constant value $f(x, y, z) = c$ is called a *level surface* of f.

Example. Describe the level surfaces of the function $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$.

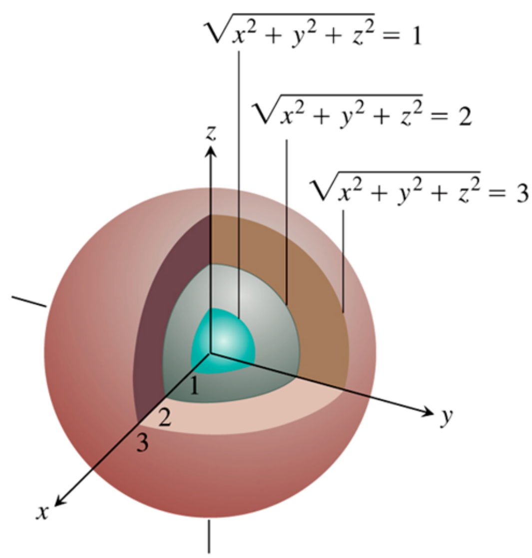


Figure 3

Definition. A point (x_0, y_0, z_0) in a region R in (3-D) space is an *interior point* of R if it is the center of a solid ball (by "solid ball" we mean the set of points lying within a distance $r > 0$ of a given point) that lies entirely in R. A point (x_0, y_0, z_0) is a *boundary point* of R if every solid ball centered at (x_0, y_0, z_0) contains points that lie outside of R as well as points that lie inside of R. A point (x_0, y_0, z_0) is a *limit point* of R if every solid ball centered at (x_0, y_0) contains a point that lies in R other than (x_0, y_0, z_0) itself. (Boundary points and limit points may or may not be in R). The *interior* of R is the set of interior points of R. The *boundary* of R is the set of boundary points of R. A region is *open* if it consists entirely of interior points. A region is *closed* if it contains all of its boundary points. The *closure* of a region consists of all the points in the set and all limit points of the set.

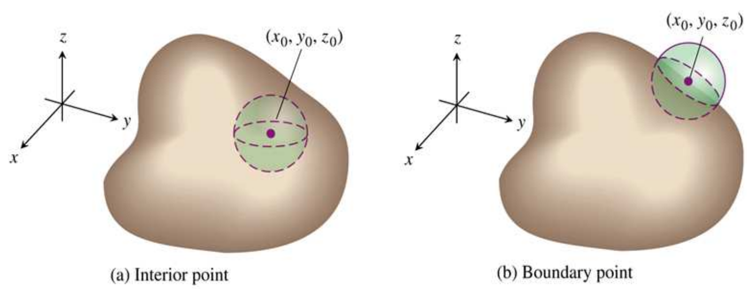


Figure 4

Partial Derivatives

Limits and Continuity in Higher Dimensions

Note. Analogous to the behavior of a function of a single variable, we wish to cleanly define the concept of *limit* for a function of "several" variables (in this section "several" means two, but the ideas are easily extended to more than two variables). If the values of $f(x,y)$ lie *arbitrarily close* to a fixed real number L for all points (x,y) *sufficiently close* to a point (x_0,y_0), we say that f approaches the limit L as (x,y) approaches (x_0,y_0). As in Calculus 1, we just need to clearly define the "arbitrarily/sufficiently" stuff. However, the textbook somewhat deviates from the definition of limit from Calculus 1 and this has some weird consequences!

Definition. We say that a function $f(x,y)$ approaches the

limit L as (x,y) approaches (x_0,y_0), denoted $\displaystyle\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L$, if for every number $\epsilon > 0$, there exists a corresponding number $\delta > 0$ such that for all (x,y) in the domain of f,

$$|f(x,y) - L| < \epsilon \text{ whenever } 0 < \sqrt{(x-x_0)^2 + (y-y_0)^2} < \delta.$$

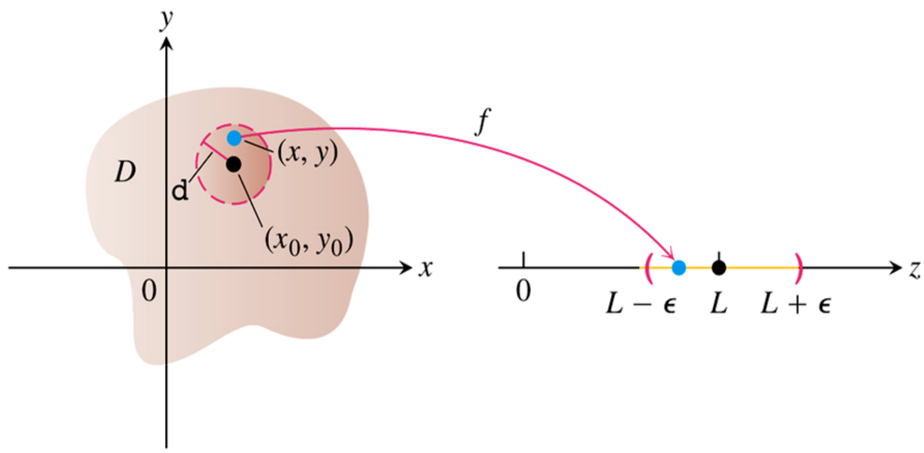


Figure 6 (the "d" here should be a δ—this is a typo in this figure, though your text has this correctly labeled).

Note. Notice the restriction of consideration to points (x, y) **in the domain of** f!!! This is different from the definition of $\lim_{x \to x_0} f(x) = L$ where it is required that the function "$f(x)$ be defined on an open interval containing x_0 except possibly at x_0 itself."

—

Theorem 1. Properties of Limits of Functions of Two Variables.

The following rules hold if L, M, and k are real numbers and

$$\lim_{(x,y)\to(x_0,y_0)} f(x,y) = L \text{ and } \lim_{(x,y)\to(x_0,y_0)} g(x,y) = M.$$

1. *Sum Rule*: $\lim_{(x,y)\to(x_0,y_0)} (f(x,y) + g(x,y)) = L + M$

2. *Difference Rule*: $\lim_{(x,y)\to(x_0,y_0)} (f(x,y) - g(x,y)) = L - M$

3. *Constant Multiple Rule*: $\lim_{(x,y)\to(x_0,y_0)} kf(x,y) = kL$ (any number k)

4. *Product Rule*: $\lim_{(x,y)\to(x_0,y_0)} (f(x,y)g(x,y)) = LM$

5. *Quotient Rule*: $\lim_{(x,y)\to(x_0,y_0)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{L}{M}$, $M \neq 0$

6. *Power Rule*: $\lim_{(x,y)\to(x_0,y_0)} (f(x,y))^n = L^n$, n a positive integer

7. *Root Rule*: $\lim_{(x,y)\to(x_0,y_0)} \sqrt[n]{f(x,y)} = \sqrt[n]{L} = L^{1/n}$, n a positive integer and if n is even, we assume $L \geq 0$.

Note. The textbook makes a bit of an error here. In the Root Rule, the book state that it requires $L > 0$ when n is even. However, with the book's definition of limit (as well as with our Alternate Definition 2) we can also allow $L = 0$. Were we to take Alternate Definition 1, then we would need the strict inequality $L > 0$. All of this is the result of whether or not we consider only values of the independent variable(s) which are in the domain or not and the issue of square roots of negatives (an issue which potentially arises when n is even and $L = 0$).

Definition. A function $f(x, y)$ is *continuous at the point* (x_0, y_0) if

1. f is defined at (x_0, y_0),

2. $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$ exists, and

3. $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$.

A function is *continuous* if it is continuous at every point of its domain.

Note. To actually evaluate limits, we can use Theorem 1, along with the standard "factor, cancel, substitute" ("FCS") method. However, it can be difficult to establish that a particular limit *does not* exist. In Calculus 1, you could test left-hand and right-hand limits to see if the "regular" two-sided limit exists. However, if a function consists of two (or more) variables, then there are an infinite number of directions from which we can approach a point (x_0, y_0). We probably cannot test *all* of these directions to see if they are the same, but we can cleverly check two of them to see if they are *different*. That's the idea behind the following.

Theorem. Two-Path Test for Nonexistence of a Limit.

If a function $f(x, y)$ has different limits along two different paths in the domain of f as (x, y) approaches (x_0, y_0), then $\lim\limits_{(x,y)\to(x_0,y_0)} f(x, y)$ does not exist. (NOTE: You'll be relieved to hear that this holds regardless of which of the many possible definitions we take of limit!)

Theorem. Continuity of Composites.

If f is continuous at (x_0, y_0) and g is a single-variable function continuous at $f(x_0, y_0)$, then the composite function $h(x, y) = g(f(x, y)) = g \circ f$ is continuous at (x_0, y_0).

Partial Derivatives

Note. When we hold all but one of the independent variables of a function constant and differentiate with respect to that one variable, we get a "partial" derivative.

Definition. The *partial derivative of* $f(x, y)$ *with respect to* x at the point (x_0, y_0) is

$$\left.\frac{\partial f}{\partial x}\right|_{(x_0,y_0)} = \left.\frac{d}{dx}[f(x, y_0)]\right|_{x=x_0} = \lim_{h\to 0}\frac{f(x_0+h, y_0) - f(x_0, y_0)}{h},$$

provided the limit exists.

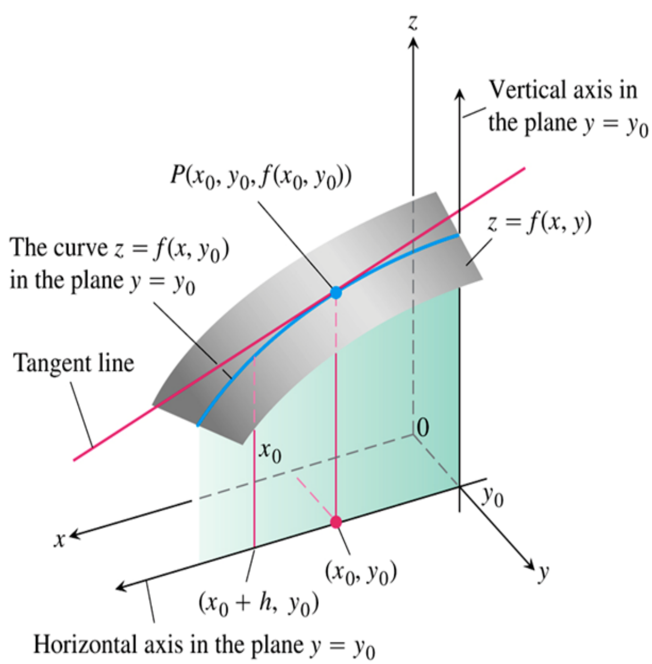


Figure 7

Note. There are several standard notations for the partial derivative of $z = f$ with respect to x:

$$\frac{\partial f}{\partial x}(x_0, y_0) \text{ or } f_x(x_0, y_0), \text{ and } f_x, \frac{\partial f}{\partial x}, z_x, \text{ or } \frac{\partial z}{\partial x}.$$

Definition. The *partial derivative of $f(x, y)$ with respect to y* at the point (x_0, y_0) is

$$\left.\frac{\partial f}{\partial y}\right|_{(x_0,y_0)} = \left.\frac{d}{dy}[f(x_0, y)]\right|_{y=y_0} = \lim_{h\to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h},$$

provided the limit exists.

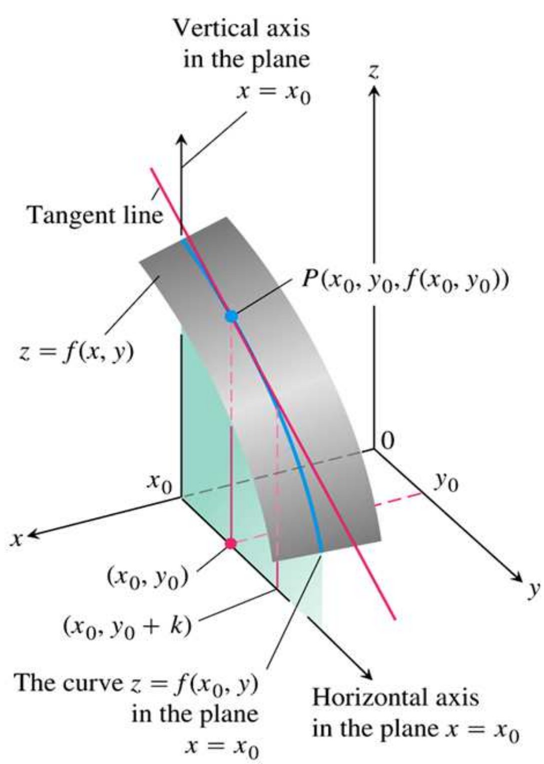


Figure 8

Note. There are several standard notations for the partial derivative of $z = f$ with respect to y:

$$\frac{\partial f}{\partial y}(x_0, y_0) \text{ or } f_y(x_0, y_0), \text{ and } f_y, \frac{\partial f}{\partial y}, z_y, \text{ or } \frac{\partial z}{\partial y}.$$

Note. Notice that we can use the two partial derivatives, $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$, to find lines tangent to the surface $z = f(x, y)$ at the point $P(x_0, y_0, f(x_0, y_0))$. We will use this later to find the equation of a tangent plane to a surface (in Section 14.6).

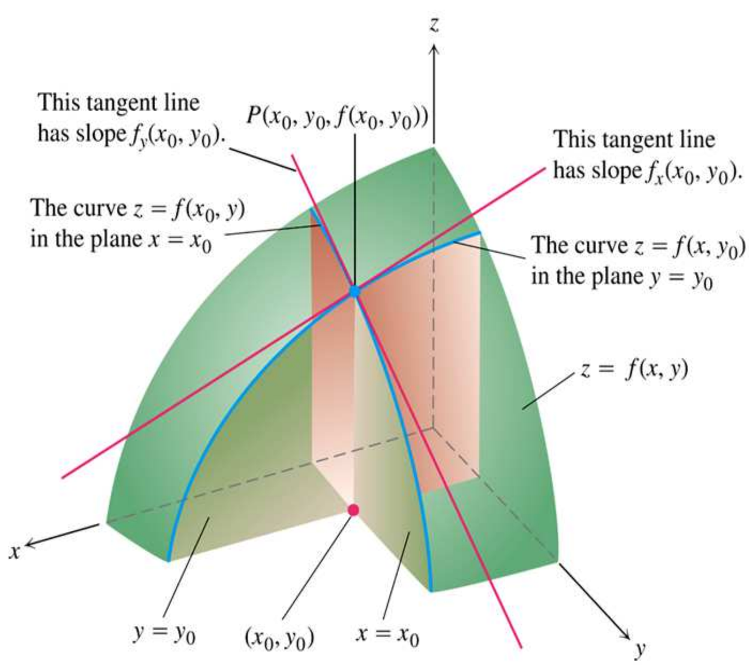


Figure 9

Note. Functions of three variables are partially differentiated in a similar way.

Note. As with ordinary derivatives, we can take higher order partial derivatives. There are four possible second order derivatives of $f(x,y)$:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx}, \frac{\partial^2 f}{\partial y^2} = f_{yy}, \frac{\partial^2 f}{\partial x \partial y} = f_{yx}, \text{ and } \frac{\partial^2 f}{\partial y \partial x} = f_{xy}.$$

Notice the particular order in the last two second order partial derivatives (which are called *mixed* partial derivatives). For example,

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left[\frac{\partial f}{\partial y}\right] = f_{yx} = (f_y)_x.$$

Note. The mixed partials f_{xy} and f_{yx} may not be equal. However, they often are as given in the following theorem.

Theorem 2. The Mixed Derivative Theorem (Clairaut's 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 $f_{xy}(a,b) = f_{yx}(a,b)$.

Note. Of course, we can take higher order partial derivatives as well. We just need to (maybe) be careful about the order of differentiation. When using the partial symbol ∂ in the fractional notation, derivatives are calculated by reading the variables from right to left, whereas when we use the subscript notation, the order of differentiation is read from left to right. For example:

$$\frac{\partial^3 f}{\partial x \partial y^2} = f_{yyx} \text{ and } \frac{\partial^4 f}{\partial x^2 \partial y^2} = f_{yyxx}.$$

Note. Recall from Calculus 1 that if a function is differentiable at a point, then it is continuous at that point. We want a similar result for functions of several variables. For Example that the function

$$f(x,y) = \begin{cases} 0, & xy \neq 0 \\ 1, & xy = 0 \end{cases}$$

is not continuous at $(0,0)$, yet both of the partial derivatives exist at $(0,0)$. So to get continuity at a point, we need a condition stronger than the existence of the partial derivatives.

Theorem 3. If the partial derivatives f_x and f_y of a function $f(x, y)$ are continuous throughout an open region R, then f is differentiable at every point of R.

Theorem 4. Differentiability Implies Continuity.

If a function $f(x, y)$ is differentiable at (x_0, y_0), then f is continuous at (x_0, y_0). (Remember that "differentiable" here means what is stated in the definition above.)

Partial Derivatives

The Chain Rule

Note. We now wish to find derivatives of functions of several variables when the variables themselves are functions of additional variables. That is, we want to deal with compositions of functions of several variables. This requires Chain Rules.

Theorem 5. Chain Rule for Functions of Two Independent Variables.

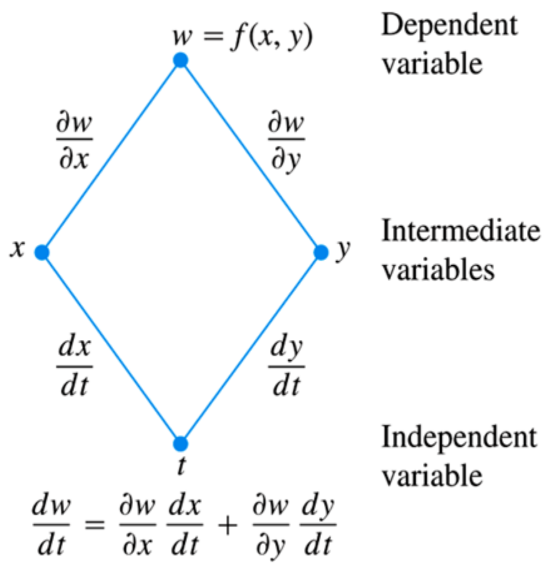
If $w = f(x, y)$ is differentiable and if $x = x(t)$, $y = y(t)$ are differentiable functions of t, then the composite $w = f(x(t), y(t))$ is a differentiable function of t and

$$\frac{dw}{dt} = f_x(x(t), y(t))[x'(t)] + f_y(x(t), y(t))[y'(t)],$$

or

$$\frac{dw}{dt} = \frac{\partial f}{\partial x}\left[\frac{dx}{dt}\right] + \frac{\partial f}{\partial y}\left[\frac{dy}{dt}\right] = \frac{\partial w}{\partial x}\left[\frac{dx}{dt}\right] + \frac{\partial w}{\partial y}\left[\frac{dy}{dt}\right].$$

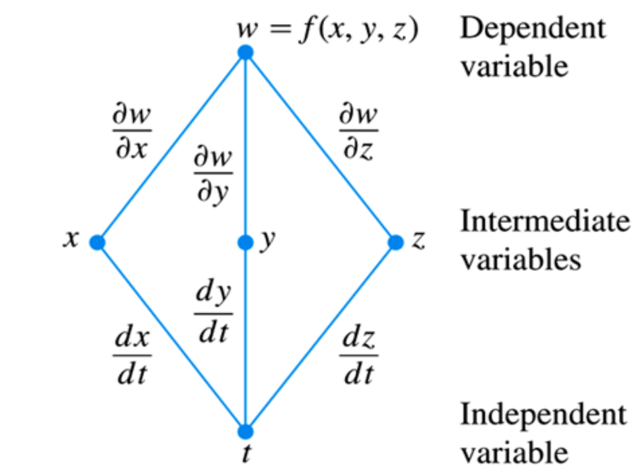
Note. We will remember the various versions of the Chain Rule which we address in this section using "branch diagrams" which reflect the relationships between each of the variables. For example, Theorem 5 can be illustrated as:



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

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\left[\frac{dx}{dt}\right] + \frac{\partial w}{\partial y}\left[\frac{dy}{dt}\right] + \frac{\partial w}{\partial z}\left[\frac{dz}{dt}\right].$$



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

Note. To motivate other function compositions, the textbook describes the following. Suppose we are interested in the temperature $w = f(x, y, z)$ at points (x, y, z) on the earth's surface, 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 $w = f(g(r, s), h(r, s), k(r, s))$. Therefore w has partial derivatives with respect to r and s, as given in the following theorem.

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:

$$\frac{\partial w}{\partial r} = \frac{\partial w}{\partial x}\left[\frac{\partial x}{\partial r}\right] + \frac{\partial w}{\partial y}\left[\frac{\partial y}{\partial r}\right] + \frac{\partial w}{\partial z}\left[\frac{\partial z}{\partial r}\right]$$

$$\frac{\partial w}{\partial s} = \frac{\partial w}{\partial x}\left[\frac{\partial x}{\partial s}\right] + \frac{\partial w}{\partial y}\left[\frac{\partial y}{\partial s}\right] + \frac{\partial w}{\partial z}\left[\frac{\partial z}{\partial s}\right].$$

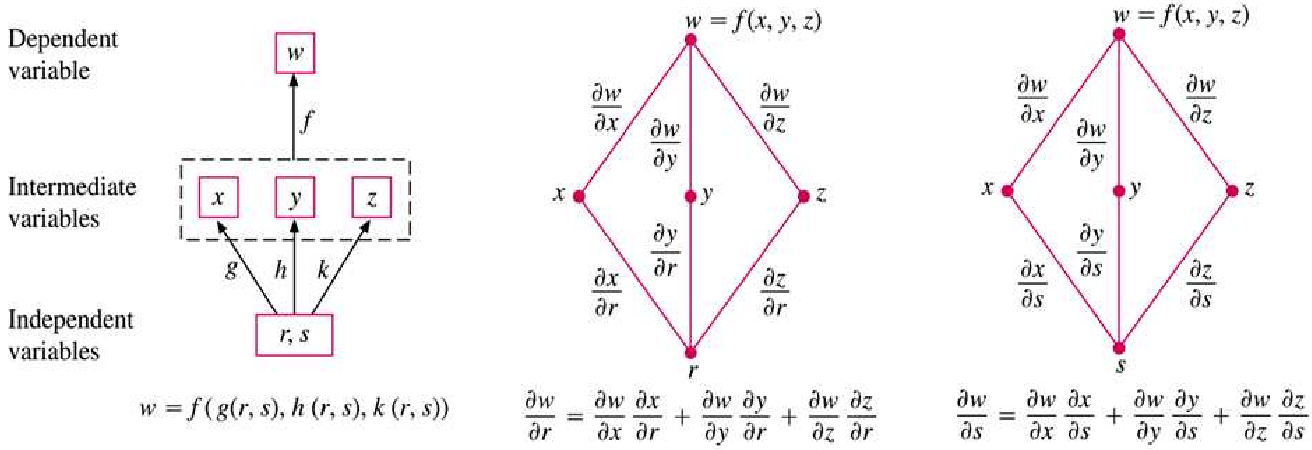


Figure 12

Note. Implicit Differentiation Revisited.

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

1.T he function $F(x, y)$ is differentiable and

2.T he equation $F(x, y) = 0$ defines y implicitly as a differentiable function of x, say $y = h(x)$.

Since $w = F(x, y) = 0$, the derivative dw/dx must be zero. Computing the derivative from the Chain Rule (see Figure 13 below), we find

$$\begin{aligned} 0 &= \frac{dw}{dx} = F_x \frac{dx}{dx} + F_y \frac{dy}{dx} \\ &= F_x + F_y \frac{dy}{dx}. \end{aligned}$$

If $F_y = \partial w / \partial y \neq 0$, we can solve this equation for dy/dx to get $\dfrac{dy}{dx} = -\dfrac{F_x}{F_y}$. In summary, we have the following 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. That at any point where $F_y \neq 0$, $\dfrac{dy}{dx} = -\dfrac{F_x}{F_y}$.

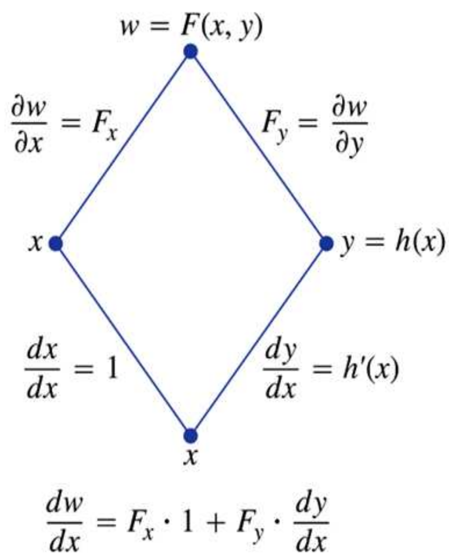


Figure 13

Note. Theorem 8 can be extended to three variables. Suppose that the equation $F(x, y, z) = 0$ defines the variable z implicitly as a function $z = f(x, y)$. Then partial derivatives of z with respect to z and y are (when $F_z \neq 0$) given by:

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} \text{ and } \frac{\partial z}{\partial y} = -\frac{F_y}{F_z}.$$

Note. We have claimed that we would address functions of *several* variables, but have really only concentrated on functions of two or three variables. Suppose that $w = f(x, y, \ldots, v)$ is a differentiable function of the variables $x, y, \ldots, v$ (a finite number of variables) and that $x, y, \ldots, v$ are differentiable functions of $p, q, \ldots, t$ (a finite number of variables). Then w is a differentiable function of the variables $p, q, \ldots, t$ and the partial derivatives of w with respect to these variables are given by equations of the form:

$$\frac{\partial w}{\partial p} = \frac{\partial w}{\partial x} \overset{\curvearrowright}{\left[\frac{\partial x}{\partial p}\right]} + \frac{\partial w}{\partial y} \overset{\curvearrowright}{\left[\frac{\partial y}{\partial p}\right]} + \cdots + \frac{\partial w}{\partial v} \overset{\curvearrowright}{\left[\frac{\partial v}{\partial p}\right]}.$$

Definition. The *derivative* of $f(x, y)$ at point $P_0(x_0, y_0)$ in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$ is the number

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = \lim_{s\to 0} \frac{f(x_0 + su_1, y_0 + su_2) - f(x_0, y_0)}{s} = (D_{\mathbf{u}}f)P_0,$$

provided the limit exists.

Note. df/ds as given above is the rate of change of f at P_0 in the direction $\mathbf{u}$.

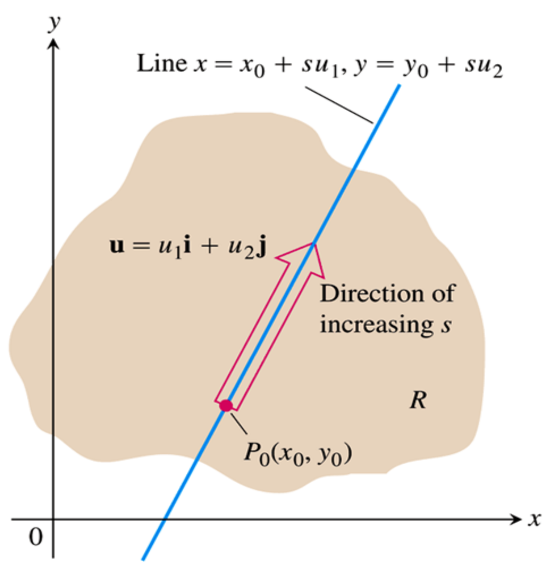


Figure 15

Note. The equation $x = f(x, y)$ represents a surface S in space. If $z_0 = f(x_0, y_0)$, then the point $P(x_0, y_0, z_0)$ lies on S. The vertical plane that passes through P and $P_0(x_0, y_0, 0)$ parallel to $\mathbf{u}$ intersects S in a curve C. The rate of change of f in the direction $\mathbf{u}$ is the slope of the tangent to C at P. When $\mathbf{u} = \mathbf{i}$, the directional derivative at P_0 is $\partial f/\partial x$ evaluated at (x_0, y_0). Then $\mathbf{u} = \mathbf{j}$, the directional derivative at P_0 is $\partial f/\partial y$ evaluated at (x_0, y_0).

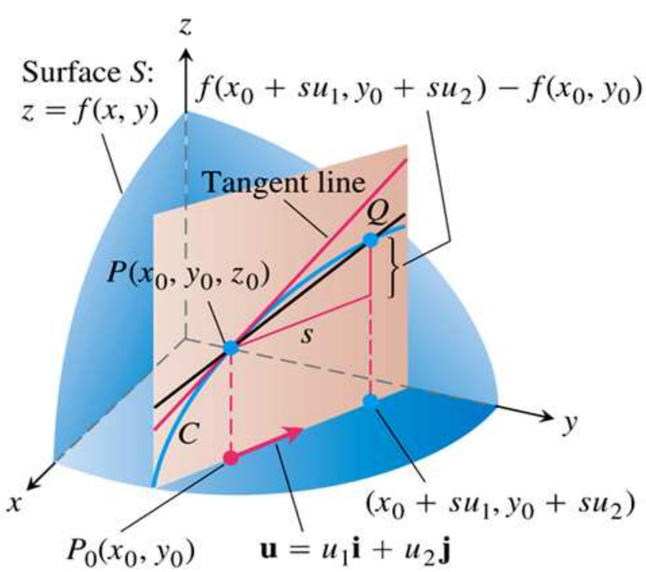


Figure 16

Note. We now need an easy way to calculate directional derivatives without using the limit definition. Consider the parametric line $x = x_0 + su_1$, $y = y_0 + su_2$ through point $P_0(x_0, y_0)$, parametrized with respect to arc length parameter s increasing in the direction of the unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$. Then by the Chain Rule we find:

$$\begin{aligned}
\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} &= \left(\frac{\partial f}{\partial x}\right)_{P_0}^{\curvearrowright}\left[\frac{dx}{ds}\right] + \left(\frac{\partial f}{\partial y}\right)_{P_0}^{\curvearrowright}\left[\frac{dy}{ds}\right] \\
&= \left(\frac{\partial f}{\partial x}\right)_{P_0}[u_1] + \left(\frac{\partial f}{\partial y}\right)_{P_0}[u_2] \\
&= \left[\left(\frac{\partial f}{\partial x}\right)_{P_0}\mathbf{i} + \left(\frac{\partial f}{\partial y}\right)_{P_0}\mathbf{j}\right]\cdot[u_1\mathbf{i} + u_2\mathbf{j}].
\end{aligned}$$

Definition. The *gradient vector* (*gradient*) of $f(x, y)$ at a point $P_0(x_0, y_0)$ is the vector

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}$$

obtained be evaluating the partial derivatives of f at P_0.

Theorem 9. The Directional Derivative Is a Dot Product.

If $f(x, y)$ is differentiable in an open region containing $P_0(x_0, y_0)$, then

$$\left(\frac{df}{ds}\right)_{\mathbf{u},P_0} = (\nabla f)_{P_0} \cdot \mathbf{u},$$

the dot product of the gradient ∇f at P_0 and unit vector $\mathbf{u}$.

Theorem. Properties of the Directional Derivative $D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = |\nabla f| \cos\theta$

1. The function f increases most rapidly when $\cos\theta = 1$ or when $\theta = 0$ and $\mathbf{u}$ is the direction of ∇f. That is, at each point P of its domain, f increases most rapidly in the direction of the gradient vector ∇f at P. The derivative in this direction is $D_{\mathbf{u}}f = |\nabla f| \cos(0) = |\nabla f|$.

2. Similarly, f decreases most rapidly in the direction $-\nabla f$. The derivative in this direction is $D_{\mathbf{u}}f = |\nabla f| \cos(\pi) = -|\nabla f|$.

3. Any direction $\mathbf{u}$ orthogonal to a gradient $\nabla f \neq 0$ is a direction of zero change in f because θ then equals $\pi/2$ and $D_{\mathbf{u}}f = |\nabla f| \cos(\pi/2) = |\nabla f| 0 = 0$.

Theorem. At every point (x_0, y_0) in the domain of a differentiable function $f(x, y)$, the gradient of f is normal to the level curve through (x_0, y_0).

Proof. If a differentiable function $f(x, y)$ has a constant value c along a smooth curve $\mathbf{r} = g(t)\mathbf{i} + h(t)\mathbf{j}$, then $f(g(t), h(t)) = c$. Differentiating both sides of this equation with respect to t leads to the equations

$$\begin{aligned} \frac{d}{dt}[f(g(t), h(t)) &= \frac{d}{dt}[c] \\ \frac{\partial f}{\partial x}\frac{dg}{dt} + \frac{\partial f}{\partial y}\frac{d}{dt} &= 0 \\ \left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j}\right) \cdot \left(\frac{dg}{dt}\mathbf{i} + \frac{dh}{dt}\mathbf{j}\right) &= 0 \\ \nabla f \cdot \frac{d\mathbf{r}}{dt} &= 0. \end{aligned}$$

Since $d\mathbf{r}/dt$ is tangent to the level curve $\mathbf{r} = g(t)\mathbf{i} + h(t)\mathbf{j}$ and $\nabla f \cdot \dfrac{d\mathbf{r}}{dt} = 0$, then the gradient of f is normal to the level curve.

Note. The previous theorem allows us to find equations for tangent lines to level curves. They are lines normal to the gradients. The line through a point $P_0(x_0, y_0)$ normal to a vector $\mathbf{N} = A\mathbf{i} + B\mathbf{j}$ has the equation

$$A(x - x_0) + B(y - y_0) = 0.$$

If $\mathbf{N}$ is the gradient $(\nabla f)_{x_0,y_0)} = f_x(x_0, y_0)\mathbf{i} + f_y(x_0, y_0)\mathbf{j}$, the equation for the tangent line is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$$

Theorem. Algebraic Rules for Gradients

1. *Sum Rule*: $\nabla(f + g) = \nabla f + \nabla g$

2. *Difference Rule*: $\nabla(f - g) = \nabla f - \nabla g$

3. *Constant Multiple Rule*: $\nabla(kf) = k\nabla f$ (for any number k)

4. *Product Rule*: $\nabla(fg) = f\nabla g + g\nabla f$

5. *Quotient Rule*: $\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}$

Note. For a differentiable function $f(x, y, z)$ and a unit vector $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$ in space, we have

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}$$

and

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = \frac{\partial f}{\partial x}u_1 + \frac{\partial f}{\partial y}u_2 + \frac{\partial f}{\partial z}u_3.$$

Partial Derivatives

Tangent Planes and Differentials

Note. If $\mathbf{r} = g(t)\mathbf{i} + h(t)\mathbf{j} + k(t)\mathbf{k}$ is a smooth curve on the level surface $f(x, y, z) = c$ of a differentiable function f, then $f(g(t), h(t), k(t)) = c$. Differentiating both sides of this equation with respect to t gives

$$\begin{aligned}
\frac{d}{dt}[f(g(t), h(t), k(t)) &= \frac{d}{dt}[c] \\
\frac{\partial f}{\partial x}\left[\frac{dg}{dt}\right] + \frac{\partial f}{\partial y}\left[\frac{dh}{dt}\right] + \frac{\partial f}{\partial z}\left[\frac{dk}{dt}\right] &= 0 \\
\left(\frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}\right) \cdot \left(\frac{dg}{dt}\mathbf{i} + \frac{dh}{dt}\mathbf{j} + \frac{dk}{dt}\mathbf{k}\right) &= 0 \\
(\nabla f) \cdot \left(\frac{d\mathbf{r}}{dt}\right) &= 0.
\end{aligned}$$

At every point along the curve, ∇f is orthogonal to the curve's velocity vector. In the figure below, we see that all the velocity vectors at point P_0 are orthogonal to ∇f at P_0, so the curves' tangent lines all lie in the plane through P_0 normal to ∇f. Therefore, the gradient of f at P_0 will act as a normal vector to the tangent plane to the surface at P_0.

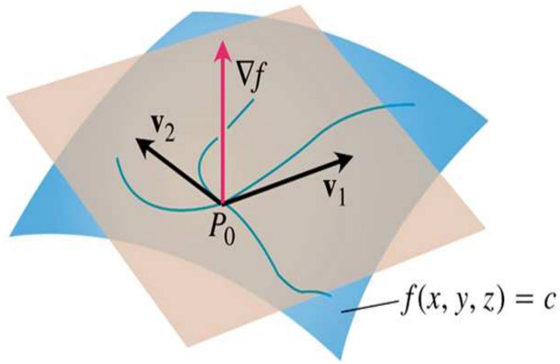


Figure 17

Definition. The *tangent plane* at the point $P_0(x_0, y_0, z_0)$ on the level surface $f(x, y, z) = c$ of a differentiable function f is the plane through P_0 normal to $\nabla f|_{P_0}$. The *normal line* of the surface at P_0 is the line through P_0 parallel to $\nabla f|_{P_0}$.

Note. The equation of the tangent plane to $f(x, y, z) = c$ at $P_0(x_0, y_0, z_0)$ is

$$f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0) = 0.$$

The equation of the normal line to $f(x, y, z) = c$ at $P_0(x_0, y_0, z_0)$ is

$$x = x_0 + f_x(P_0)t, \quad y = y_0 + f_y(P_0)t, \quad z = z_0 + f_z(P_0)t.$$

Note. If we consider the function $z = f(x, y)$, then the tangent plane to this surface at the point $(x_0, y_0, f(x_0, y_0))$ is

$$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) - (z - z_0) = 0.$$

Note. We now use differentials to estimate changes in functions, similar to what was done for functions of a single variable. To

estimate the change in the value of a differentiable function f when we move a small distance ds from a point P_0 in a particular direction $\mathbf{u}$, we use the differential

$$df = \left(\nabla f|_{P_0} \cdot \mathbf{u} \right) ds.$$

Notice that df is the directional derivative of f times the distance increment ds.

Definition. The *linearization* of a function $f(x, y)$ at a point (x_0, y_0) where f is differentiable is the function

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

The approximation $f(x, y) \approx L(x, y)$ is the *standard linear approximation* of f at (x_0, y_0).

Note. In fact, the plane $z = L(x, y)$ is tangent to the surface $z = f(x, y)$ at the point (x_0, y_0) (just as the line $y = L(x)$ was the tangent line to $y = f(x)$ at the point of approximation in section 3.11). Thus, the linearization of a function of two variables is a tangent-plane approximation. As long as (x, y) is "close to" (x_0, y_0) (that is, if Δx and Δy are small), then $L(x, y)$ will take on approximately the same values as $f(x, y)$.

Note. If f has continuous first and second partial derivatives throughout an open set containing a rectangle R centered at (x_0, y_0) and if M is any upper bound for the values of $|f_{xx}|$, $|f_{yy}|$, and $|f_{xy}|$ on R, then the error $E(x, y)$ incurred in replacing $f(x, y)$ on R by it linearization

$$L(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

satisfies the inequality

$$|E(x,y)| \leq \frac{1}{2}M(|x-x_0|+|y-y_0|)^2.$$

Notice that the error is small when M, Δx, and/or Δy are small (especially Δx and Δy).

Definition. The *differentials* dx and dy are independent variables (so they can take on any values). Often we take $dx = \Delta x = x - x_0$ and $dy = \Delta y = y - y_0$. If we move from (x_0, y_0) to a point $(x_0 + dx, y_0 + dy)$ nearby, the resulting change

$$df = f_x(x_0,y_0)dx + f_y(x_0,y_0)dy$$

in the linearization of f is called the *total differential* of f.

Note. We can extend the ideas of this section to functions of more than two variables. For functions of three variables, we have the following.

1. The *linearization* of $f(x, y, z)$ at a point $P_0(x_0, y_0, z_0)$ is

$$L(x, y, z) = f(P_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + f_z(P_0)(z - z_0).$$

2. Suppose that R is a closed rectangular solid centered at P_0 and lying in an open region on which the second partial derivatives of f are continuous. Suppose also that $|f_{xx}|$, $|f_{yy}|$, $|f_{zz}|$, $|f_{xy}|$, $|f_{xz}|$, and $|f_{yz}|$ are all less than or equal to M throughout R. Then the error $E(x, y, z) = f(x, y, z) - L(x, y, z)$ in the approximation of f by L is bounded throughout R by

$$|E| \leq \frac{1}{2} M(|x - x_0| + |y - y_0| + |z - z_0|)^2.$$

3. If the second partial derivatives of f are continuous and if x, y, and z change from x_0, y_0, and z_0 by "small" amounts dx, dy, and dz, the total differential

$$df = f_x(P_0)dx + f_y(P_0)dy + f_z(P_0)dz$$

gives a "good" approximation of the resulting change in f.

Partial Derivatives

Extreme Values and Saddle Points

Definition. Let $f(x, y)$ be defined on a region R containing the point (a, b). Then

1. $f(a, b)$ is a *local maximum* value of f if $f(a, b) \geq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

2. $f(a, b)$ is a *local minimum* value of f if $f(a, b) \leq f(x, y)$ for all domain points (x, y) in an open disk centered at (a, b).

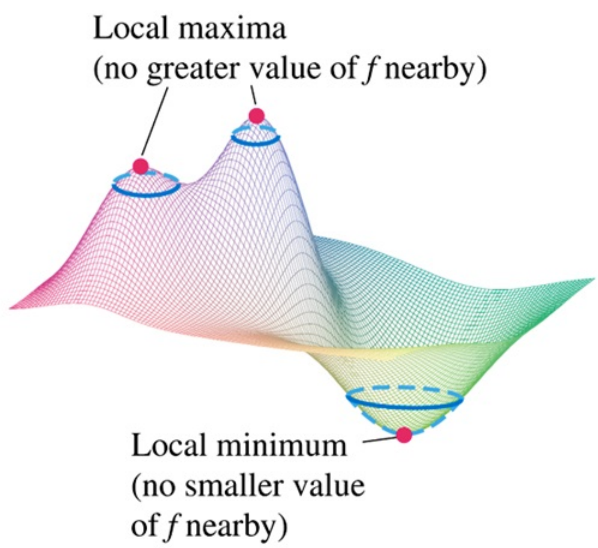


Figure 18

Theorem 10. First Derivative Test for Local Extreme Values
If $f(x, y)$ has a local maximum or minimum value at an interior point (a, b) of its domain and if the first partial derivatives exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.

Proof. If f has a local extremum at (a, b), then the function $g(x) = f(x, b)$ has a local extremum at $x = a$. Therefore $g'(a) = 0$. Now $g'(a) = f_x(a, b)$, so $f_x(a, b) = 0$. A similar argument with the function $h(y) = f(a, y)$ shows that $f_y(a, b) = 0$. *Q.E.D.*

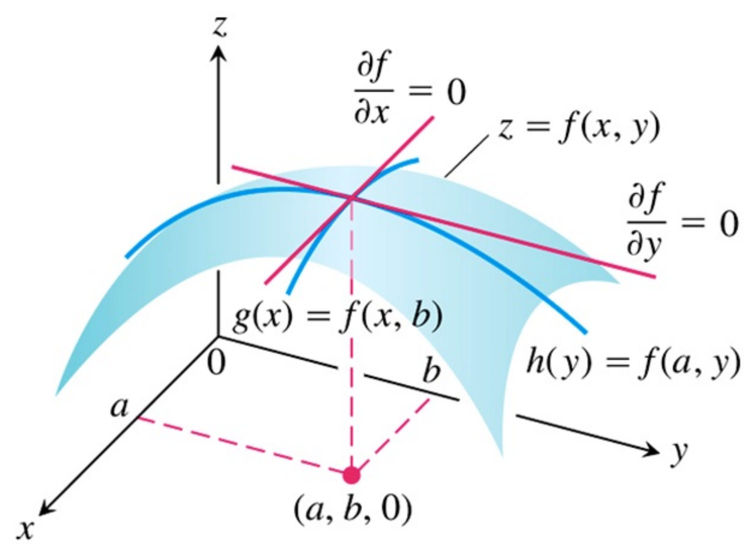
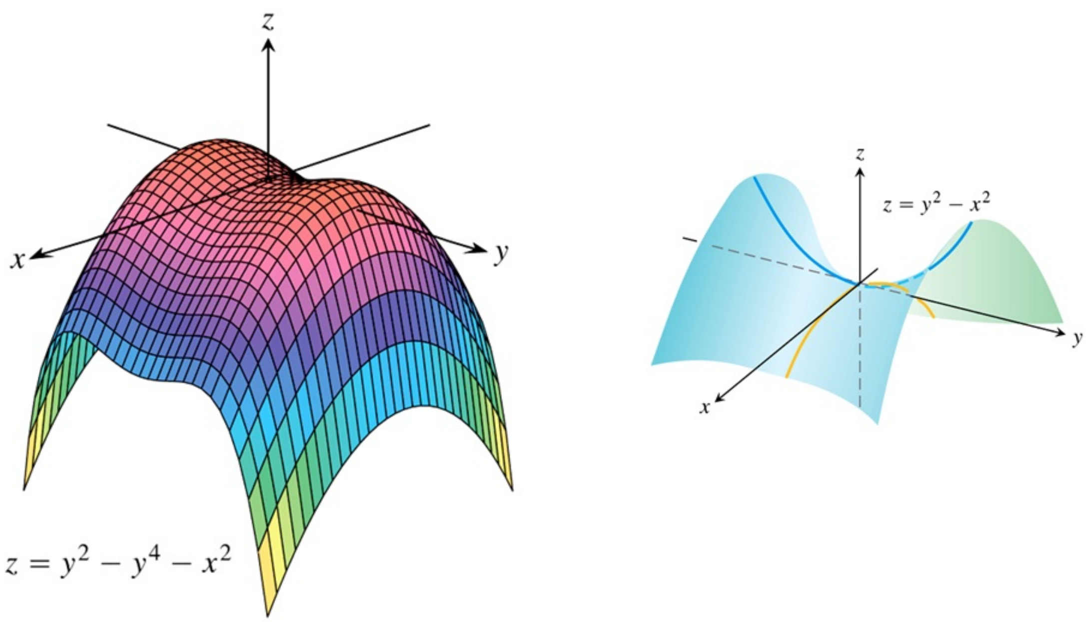


Figure 19

Definition. An interior point of the domain of a function $f(x, y)$ where both f_x and f_y are zero or where one or both of f_x and f_y do not exist is a *critical point* of f. A differentiable function $f(x, y)$ has a *saddle point* at a critical point (a, b) if in every open disk centered at (a, b) there are domain points (x, y) where $f(x, y) > f(a, b)$ and domain points (s, y) where $f(x, y) < f(a, b)$. The corresponding point $(a, b, f(a, b))$ on the surface $z = f(x, y)$ is called a *saddle point* of the surface.



Figures 20

Theorem 11. Second Derivative Test for Local Extreme Values

Suppose that $f(x, y)$ and its first and second partial derivatives are continuous throughout a disk centered at (a, b) and that $f_x(a, b) = f_y(a, b) = 0$. Then

(i) f has a *local maximum* at (a, b) if $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(ii) f has a *local minimum* at (a, b) if $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(iii) f has a *saddle point* at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a, b).

(iv) The test is inconclusive at (a, b) if $f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b). In this case, we must find some other way to determine the behavior of f at (a, b).

Note. The expression $f_{xx}f_{yy} - f_{xy}^2$ is called the *discriminant* or *Hessian* of f. It is sometimes easier to remember it in determinant form:

$$f_{xx}f_{yy} - f_{xy}^2 = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{vmatrix}.$$

Theorem 11 makes the most since if we explore the topic of *curvature*! The "Hessian" is really (related to) the curvature of the surface. When curvature is positive at point (a, b) (as in cases (i) and (ii) of Theorem 11), then the surface lies entirely one one side of a tangent plane to the surface at point (a, b) (in some neighborhood of (a, b)). Since at the critical point we have $f_x(a, b) = f_y(a, b) = 0$, then the tangent plane is a horizontal plane so the critical point corresponds to a local maximum if the surface lies above the tangent plane, and the critical point corresponds to a local minimum if the surface lies below the tangent plane. We can determine which is the case be considering the sign of f_{xx} to determine the "concavity" of the surface. A surface is of negative curvature at point (a, b) (as in case (iii) of Theorem 11) if part of the surface lies on one side of the tangent plane to the surface at (a, b) and another part of the surface lies on the other side of the tangent plane (in all open neighborhoods of (a, b)). This is why there is no local extremum in case (iii). In case (iv) of Theorem 11, the surface has zero curvature and we cannot determine whether the surface has a local maximum, local minimum, or saddle point at (a, b).

Note. Absolute Maxima and Minima of Closed Bounded Regions

We organize the search for the absolute extrema of a continuous function $f(x, y)$ on a closed and bounded region R into three steps:

1. *List the interior points of* R where f may have local maxima and minima and evaluate f at these points. These are the critical points of f.

2. *List the boundary points of* R where f has local maxima and minima and evaluate f at these points. (Details to follow in the next example.)

3. *Look through the lists* for the maximum and minimum values of f. These will be the absolute maximum and minimum values of f on

R. Since absolute maxima and minima are also local maxima and minima, the absolute maximum and minimum values of f appear somewhere in the lists made in Steps 1 and 2.

Note. Solving the extreme value problems with algebraic constraints on the variables usually requires the method of Lagrange multipliers introduced in the next section. But sometimes we can solve such problems directly.

Note. Summary of Max-Min Tests

The extreme values of $f(x, y)$ can occur only at

(i) boundary points of the domain of f, and

(ii) critical points (interior points where $f_x = f_y = 0$ or points where f_x or f_y fails to exist).

If the first- and second-order partial derivatives of f are continuous throughout a disk centered at a point (a, b) and $f_x(a, b) = f_y(a, b) = 0$, the nature of $f(a, b)$ can be tested with the Second Derivative Test:

(i) $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b) \Rightarrow$ local maximum.

(ii) $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at $(a, b) \Rightarrow$ local minimum.

(iii) $f_{xx}f_{yy} - f_{xy}^2 < 0$ at $(a, b) \Rightarrow$ saddle point.

(iv) $f_{xx}f_{yy} - f_{xy}^2 = 0$ at $(a, b) \Rightarrow$ test is inconclusive.

Partial Derivatives

Lagrange Multipliers

Note. In this section, we consider extrema of functions of several (well, two) variables where the is an added constraint (i.e., relationship) between the two variables. We start with an example.

Example. Find the points on the hyperbolic cylinder $x^2 - z^2 - 1 = 0$ that are closest to the origin.

Solution. These are the points whose coordinates minimize the value of the functions $f(x, y, z) = x^2 + y^2 + z^2$ subject to the constraint that $x^2 - z^2 - 1 = 0$. Notice that the function $f(x, y, z)$ has as its level curves spheres of various radii centered at the origin. So we geometrically consider a small sphere centered at the origin which expands. At the instant when the sphere contacts the hyperbolic cylinder, both surfaces will have the same tangent plane and normal line at the points of contact.

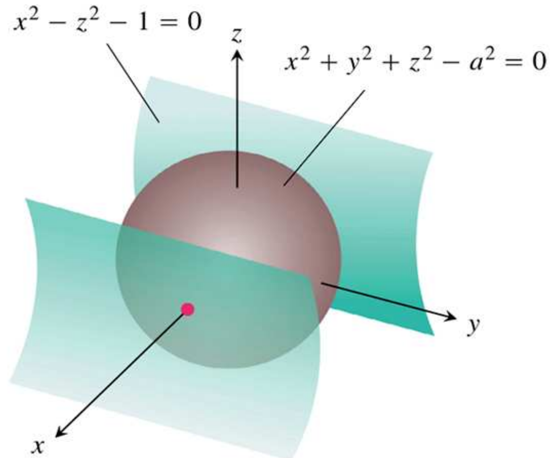


Figure 22

Therefore, if the sphere and cylinder are represented as the level surfaces obtained by setting

$$f(x, y, z) = x^2 + y^2 + z^2 - a^2 \text{ and } g(x, y, z) = x^2 - z^2 - 1$$

equal to 0, then the gradients ∇f and ∇g will be parallel where the surfaces touch. At any point of contact, we should be able to find a scalar λ such that $\nabla f = \lambda \nabla g$, or in this case, $2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k} = \lambda(2x\mathbf{i} - 2z\mathbf{k})$. Solving, we find (see the text for details) that the desired points are $(-1, 0, 0)$ and $(1, 0, 0)$.

Note. The method used above is the method of *Lagrange multipliers*. It implies that the extreme values of function $f(x, y, z)$ whose variables are subject to a constraint $g(x, y, z) = 0$ are to be found on the surface $g = 0$ among the points where $\nabla f = \lambda \nabla g$ for some scalar λ (called the *Lagrange multiplier*).

Theorem 12. The Orthogonal Gradient Theorem. Suppose that $f(x, y, z)$ is differentiable in a region whose interior contains a smooth curve

$$C : \mathbf{r}(t) = g(t)\mathbf{i} + h(t)\mathbf{j} + k(t)\mathbf{k}.$$

If P_0 is a point on C where f has a local maximum or minimum relative to its values on C, then ∇f is orthogonal to C at P_0.

Proof. The values of f on C are given by the composite $f(g(t), h(t), k(t))$, whose derivative with respect to t is

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dg}{dt} + \frac{\partial f}{\partial y}\frac{dh}{dt} + \frac{\partial f}{\partial z}\frac{dk}{dt} = \nabla f \cdot \mathbf{v}.$$

At any point P_0 where f has a local extrema relative to its values on the curve, we have $df/dt = 0$ and so $\nabla f \cdot \mathbf{v} = 0$.

Note. Theorem 12 is the key to the method of Lagrange multipliers. Suppose that $f(x, y, z)$ and $g(x, y, z)$ are differentiable and that point P_0 is a point on the surface $g(x, y, z) = 0$ where f has a local maximum or minimum value relative to its other values on the surface. We assume also that $\nabla g \neq \mathbf{0}$ at the points on the surface $g(x, y, z) = 0$. Then f takes on a local maximum or minimum at P_0 relative to its values on every differentiable curve through P_0 on the surface $g(x, y, z) = 0$. Therefore,

∇f is orthogonal to the velocity vector of every such differentiable curve through P_0. So is ∇g, since ∇g is orthogonal to the level surface $g = 0$. Therefore, at P_0, ∇f is some scalar multiple λ of ∇g. In summary:

The Method of Lagrange Multipliers. Suppose that $f(x, y, z)$ and $g(x, y, z)$ are differentiable and $\nabla g \neq 0$ when $g(x, y, z) = 0$. to find the local maximum and minimum values of f subject to the constraint $g(x, y, z) = 0$ (if these exist), find the values of x, y, z, and λ that simultaneously satisfy the equations $\nabla f = \lambda \nabla g$ and $g(x, y, z) = 0$. For functions of two independent variables, the condition is similar, without the variable z.

Note. Many problems require us to find the extreme values of a differentiable function $f(x, y, z)$ whose variables are subject to two constraints. If the constraints are

$$g_1(x, y, z) = 0 \text{ and } g_2(x, y, z) = 0$$

and g_1 and g_2 are differentiable, with ∇g_1 not parallel to ∇g_2, we find the constrained local maxima and minima of f by introducing two Lagrange multipliers λ and μ. That is, we locate the points $P(x, y, z)$ where f takes on its constrained extreme values by finding the values x, y, z, λ, and μ that simultaneously satisfy the equations

$$\nabla f = \lambda \nabla g_1 + \mu \nabla g_2, \quad g_1(x, y, z) = 0, \quad g_2(x, y, z) = 0.$$

Partial Derivatives

Taylor's Formula for Two Variables

Note. Let $f(x,y)$ have continuous partial derivatives in an open region R containing a point $P(a,b)$ where $f_x = f_y = 0$. Let h and k be increments small enough to put the point $S(a+h, b+k)$ and the line segment joining it to P inside R. We parametrize the segment PS as

$$x = a + th, \quad y = b + tk, \quad t \in [0,1].$$

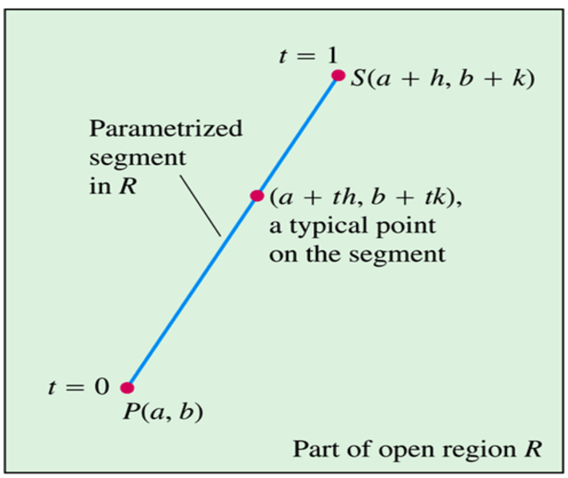


Figure 23

Define $F(t) = f(a + th, b + tk)$. The Chain Rule gives

$$F'(t) = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} = hf_x + kf_y.$$

Since f_x and f_y are differentiable (by assumption), F' is a differentiable function of t and

$$\begin{aligned} F'' = \frac{\partial F'}{\partial x}\frac{dx}{dt} + \frac{\partial F'}{\partial y}\frac{dy}{dt} &= \frac{\partial}{\partial x}[hf_x + kf_y] \cdot h + \frac{\partial}{\partial y}[hf_x + kf_y] \cdot k \\ &= h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}. \end{aligned}$$

Since F and F' are continuous on $[0, 1]$ and F' is differentiable on $(0, 1)$, we can apply Taylor's Theorem with $n = 2$ and $a = 0$ to obtain

$$F(1) = F(0) + F'(0)(1-0) + F''(c)\frac{(1-0)^2}{2} = F(0) + F'(0) + \frac{1}{2}F''(0)$$

for some c between 0 and 1. Rewriting in terms of f gives

$$\begin{aligned} f(a+h, b+k) &= f(a,b) + hf_x(a,b) + kf_y(a,b) \\ &\quad + \frac{1}{2}\left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)\Big|_{(a+ch,b+ck)}. \end{aligned} \qquad (2)$$

Since $f_x(a,b) = f_y(a,b) = 0$, this reduces to

$$f(a+h, b+k) - f(a,b) = \frac{1}{2}\left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)\Big|_{(a+ch,b+ck)}.$$

The presence of an extremum of f at (a, b) is determined by the sign of $f(a + h, b + k) - f(a, b)$. From the previous equation, we see that this is the same as the sign of

$$Q(c) \equiv \left(h^2 f_{xx} + 2hk f_{xy} + k^2 f_{yy}\right)\Big|_{(a+ch,b+ck)}.$$

If $Q(0) \neq 0$, the sign of $Q(c)$ will be the same as the sign of $Q(0)$ for sufficiently small values of h and k. We can predict the sign of

$$Q(0) = h^2 f_{xx}(a,b) + 2hk f_{xy}(a,b) + k^2 f_{yy}(a,b)$$

from the signs of f_{xx} and $f_{xx}f_{yy} - f_{xy}^2$ at (a,b). Multiply both sides of the previous equation by f_{xx} and rearrange the right-hand side to get

$$f_{xx}Q(0) = (hf_{xx} + kf_{xy})^2 + (f_{xx}f_{yy} - f_{xy}^2)k^2.$$

We see that

1. If $f_{xx} < 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a,b) then $Q(0) < 0$ for all sufficiently small nonzero values of h and k, and f has a *local maximum* value at (a,b).

2. If $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ at (a,b) then $Q(0) > 0$ for all sufficiently small nonzero values of h and k, and f has a *local minimum* value at (a,b).

3. If $f_{xx}f_{yy} - f_{xy}^2 < 0$ at (a,b) there are combinations of arbitrarily small nonzero values of h and k for which $Q(0) > 0$, and other values for which $Q(0) < 0$. Arbitrarily close to the point $P_0(a, b, f(a, b))$ on the surface $z = f(x, y)$ there are points above P_0 and points below P_0, so f has a *saddle point* at (a, b).

4. If $f_{xx}f_{yy} - f_{xy}^2 = 0$ at (a, b), another test is needed.

Note. We now justify the error formula for linearizations. Assume the function f has continuous second partial derivatives

throughout an open set containing a closed rectangular region R centered at (x_0, y_0). Let the number M be an upper bound for $|f_{xx}|$, $|f_{yy}|$, and $|f_{xy}|$ on R. The inequality we want comes from equation (2) above. We substitute x_0 and y_0 for a and b, and $x - x_0$ and $y - y_0$ for h and k (resp.), and rearrange the result as

$$f(x, y) = \underbrace{f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)}_{\text{linearization } L(x,y)}$$

$$+ \underbrace{\frac{1}{2}\left((x - x_0)^2 + f_{xx} + 2(x - x_0)(y - y - 0)f_{xy} + (y - y_0)^2 f_{yy}\right)\Big|_{(x_0 + c(x - x_0), y_0 + c(y - y_0))}}_{\text{error } E(x,y)}.$$

This equation reveals that

$$|E| \leq \frac{1}{2}\left(|x - x_0|^2|f_{xx}| + 2|x - x_0|\,|y - y_0|\,|f_{xy}| + |y - y_0||f_{yy}|\right).$$

Here, if M is an upper bound for the values of $|f_{xx}|$, $|f_{xy}|$, and $|f_{yy}|$ on R,

$$|E| \leq \frac{1}{2}\left(|x-x_0|^2 M + 2|x-x_0|\,|y-y_0| M + |y-y_0|^2 M\right)$$

$$= \frac{1}{2} M(|x-x_0| + y - y_0|)^2.$$

Note. This justifies the standard linear approximation of $f(x,y)$ at (x_0, y_0) and the error of this approximation.

Note. We can get a higher degree of approximation of $f(x,y)$ as follows.

Theorem. Taylor's Formula for $f(x,y)$ at the Origin. Suppose $f(x,y)$ and its partial derivatives through order $n+1$ are continuous throughout an open rectangular region R centered at the point $(0,0)$. then throughout R,

$$f(x,y) = f(0,0) + x f_x + y f_y + \frac{1}{2!}(x^2 f_x x + 2xy f_{xy} + y^2 f_{yy})$$

$$+\frac{1}{3!}(x^3 f_{xxx} + 3x^2 y f_{xxy} + 3xy^2 f_{xyy} + y^3 f_{yyy}) + \cdots + \frac{1}{n!}\left(x^n \frac{\partial^n f}{\partial x^n} + n x^{n-1} y \frac{\partial^n f}{\partial x^{n-1}\partial y} + \cdots + y^n \frac{\partial^n f}{\partial y^n}\right)$$

$$+\frac{1}{(n+1)!}\left(x^{n+1}\frac{\partial^{n+1} f}{\partial x^{n+1}} + (n+1)x^n y \frac{\partial^{n+1} f}{\partial x^n \partial y} + \cdots + y^{n+1}\frac{\partial^{n+1} f}{\partial y^{n+1}}\right)\Bigg|_{(cx,cy)}.$$