2

Differentiation

I turn away with fright and horror from the lamentable evil of functions which do not have derivatives.

Charles Hermite,
in a letter to Thomas Jan Stieltjes

This chapter extends the principles of differential calculus for functions of one variable to functions of several variables. We begin in Section 2.1 with the geometry of real-valued functions and study the graphs of these functions as an aid in visualizing them. Section 2.2 gives some basic definitions relating to limits and continuity. This subject is treated briefly, because it requires time and mathematical maturity to develop fully and is therefore best left to a more advanced course. Fortunately, a complete understanding of all the subtleties of the limit concept is not necessary for our purposes; the student who has difficulty with this section should bear this in mind. However, we hasten to add that the notion of a limit is central to the definition of the derivative, but not to the computation of most derivatives in specific problems, as we already know from one-variable calculus. Sections 2.3 and 2.5 deal with the definition of the derivative, and establish some basic rules of calculus: namely, how to differentiate a sum, product, quotient, or composition. In Section 2.6, we study directional derivatives and tangent planes, relating these ideas to those in Section 2.1. Finally, the Internet supplement gives some of the technical proofs.

In generalizing calculus from one dimension to several, it is often convenient to use the language of matrix algebra. What we shall need has been summarized in Section 1.5.

2.1 The Geometry of Real-Valued Functions

We launch our investigation of real-valued functions by developing methods for visualizing them. In particular, we introduce the notions of a graph, a level curve, and a level surface of such functions.

Functions and Mappings

Let f be a function whose domain is a subset A of $\mathbb{R}^n$ and with a range contained in $\mathbb{R}^m$. By this we mean that to each $\mathbf{x} = (x_1, \dots, x_n) \in A$, f assigns a value $f(\mathbf{x})$, an m-tuple in $\mathbb{R}^m$. Such functions f are called ***vector-valued functions***[1] if $m > 1$, and ***scalar-valued functions*** if $m = 1$. For example, the scalar-valued function $f(x, y, z) = (x^2 + y^2 + z^2)^{-3/2}$ maps the set A of $(x, y, z) \neq (0, 0, 0)$ in $\mathbb{R}^3$ ($n = 3$ in this case) to $\mathbb{R}$ ($m = 1$). To denote f we sometimes write

$$f\colon (x, y, z) \mapsto (x^2 + y^2 + z^2)^{-3/2}.$$

Note that in $\mathbb{R}^3$ we often use the notation (x, y, z) instead of (x_1, x_2, x_3). In general, the notation $\mathbf{x} \mapsto f(\mathbf{x})$ is useful for indicating the value to which a point $\mathbf{x} \in \mathbb{R}^n$ is sent. We write $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ to signify that A is the domain of f (a subset of $\mathbb{R}^n$) and the range is contained in $\mathbb{R}^m$. We also use the expression f maps A *into* $\mathbb{R}^m$. Such functions f are called ***functions of several variables*** if $A \subset \mathbb{R}^n$, $n > 1$.

As another example we can take the vector-valued function $g\colon \mathbb{R}^6 \to \mathbb{R}^2$ defined by the rule

$$g(\mathbf{x}) = g(x_1, x_2, x_3, x_4, x_5, x_6) = \left(x_1x_2x_3x_4x_5x_6, \sqrt{x_1^2 + x_6^2}\right).$$

The first coordinate of the value of g at $\mathbf{x}$ is the product of the coordinates of $\mathbf{x}$.

Functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ are not just mathematical abstractions, they arise naturally in problems studied in all the sciences. For example, to specify the temperature T in a region A of space requires a function $T\colon A \subset \mathbb{R}^3 \to \mathbb{R}$ ($n = 3$, $m = 1$); thus, $T(x, y, z)$ is the temperature at the point (x, y, z). To specify the velocity of a fluid moving in space requires a map $\mathbf{V}\colon \mathbb{R}^4 \to \mathbb{R}^3$, where $\mathbf{V}(x, y, z, t)$ is the velocity vector of the fluid at the point (x, y, z) in space at time t (see Figure 2.1.1). To

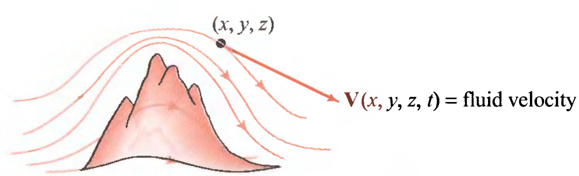


Figure 2.1.1 A fluid in motion defines a vector field $\mathbf{V}$ by specifying the velocity of the fluid particles at each point in space and time.

[1] Some mathematicians would write such an f in boldface, using the notation $\mathbf{f}(\mathbf{x})$, because the function is vector-valued. We did not do so, as a matter of personal taste. We use boldface primarily for mappings that are vector fields, introduced later. The notion of function was developed over many centuries, with the definition extended to cover more cases as they arose. For example, in 1667 James Gregory defined a function as "a quantity obtained from other quantities by a succession of algebraic operations or by any other operation imaginable." In 1755 Euler gave the following definition: "If some quantities depend on others in such a way as to undergo variation when the latter are varied then the former are called functions of the latter."

specify the reaction rate of a solution consisting of six reacting chemicals A, B, C, D, E, F in proportions x, y, z, w, u, v requires a map $\sigma: U \subset \mathbb{R}^6 \to \mathbb{R}$, where $\sigma(x, y, z, w, u, v)$ gives the rate when the chemicals are in the indicated proportions. To specify the cardiac vector (the vector giving the magnitude and direction of electric current flow in the heart) at time t requires a map $\mathbf{c}: \mathbb{R} \to \mathbb{R}^3$, $t \mapsto \mathbf{c}(t)$.

When $f: U \subset \mathbb{R}^n \to \mathbb{R}$, we say that f is a ***real-valued function of n variables with domain*** U. The reason we say "n variables" is simply that we regard the coordinates of a point $\mathbf{x} = (x_1, \ldots, x_n) \in U$ as n variables, and $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ depends on these variables. We say "real-valued" because $f(x_1, \ldots, x_n)$ is a real number. A good deal of our work will be with real-valued functions, so we give them special attention.

Graphs of Functions

For $f: U \subset \mathbb{R} \to \mathbb{R}$ ($n = 1$), the ***graph*** of f is the subset of $\mathbb{R}^2$ consisting of all points $(x, f(x))$ in the plane, for x in U. This subset can be thought of as a curve in $\mathbb{R}^2$. In symbols, we write this as

$$\operatorname{graph} f = \{(x, f(x)) \in \mathbb{R}^2 \mid x \in U\},$$

where the curly braces mean "the set of all" and the vertical bar is read "such that." Drawing the graph of a function of one variable is a useful device to help visualize how the function actually behaves. (See Figure 2.1.2.) It will be helpful to generalize

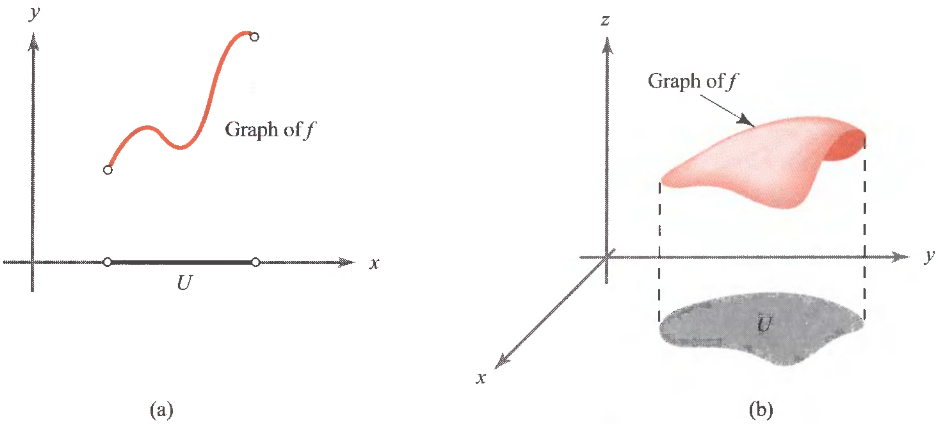


Figure 2.1.2 The graphs of (a) a function of one variable, and (b) a function of two variables.

the idea of a graph to functions of several variables. This leads to the following definition:

DEFINITION: Graph of a Function Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$. Define the ***graph*** of f to be the subset of $\mathbb{R}^{n+1}$ consisting of all the points

$$(x_1, \dots, x_n, f(x_1, \dots, x_n))$$

in $\mathbb{R}^{n+1}$ for $(x_1, \dots, x_n)$ in U. In symbols,

$$\text{graph } f = \{(x_1, \dots, x_n, f(x_1, \dots, x_n)) \in \mathbb{R}^{n+1} \mid (x_1, \dots, x_n) \in U\}.$$

For the case $n = 1$ the graph is a curve in $\mathbb{R}^2$, while for $n = 2$ it is a surface in $\mathbb{R}^3$ (see Figure 2.1.2). For $n = 3$ it is difficult to visualize the graph, because, since we are humans living in a three-dimensional world, it is hard for us to envisage sets in $\mathbb{R}^4$. To help overcome this handicap, we introduce the idea of a level set.

Level Sets, Curves, and Surfaces

Suppose $f(x, y, z) = x^2 + y^2 + z^2$. A ***level set*** is a subset of $\mathbb{R}^3$ on which f is constant; for instance, the set where $x^2 + y^2 + z^2 = 1$ is a level set for f. This we can visualize: It is just a sphere of radius 1 in $\mathbb{R}^3$. Formally, a level set is the set of (x, y, z) such that $f(x, y, z) = c$, where c is a constant. The behavior or structure of a function is determined in part by the shape of its level sets; consequently, understanding these sets aids us in understanding the function in question. Level sets are also useful for understanding functions of two variables $f(x, y)$, in which case we speak of ***level curves*** or ***level contours***.

The idea is similar to that used to prepare contour maps, where one draws lines to represent constant altitudes; walking along such a line would mean walking on a level path. In the case of a hill rising from the xy plane, a graph of all the level curves gives us a good idea of the function $h(x, y)$, which represents the height of the hill at point (x, y) (see Figure 2.1.3).

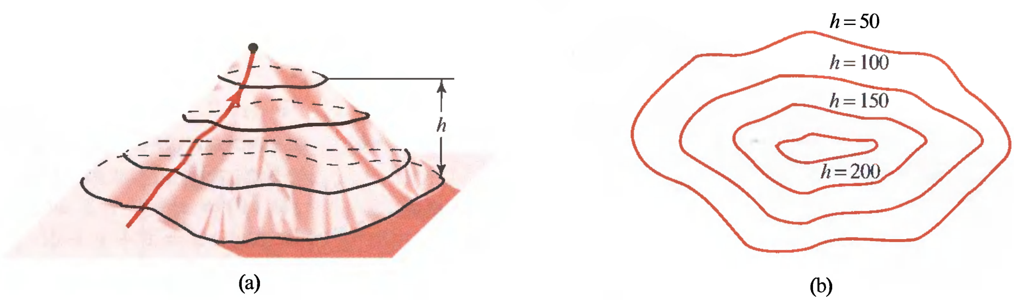


Figure 2.1.3 Level contours of a function are defined in the same manner as contour lines for a topographical map.

EXAMPLE 1 The constant function $f: \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto 2$, that is, the function $f(x, y) = 2$, has as its graph the horizontal plane $z = 2$ in $\mathbb{R}^3$. The level curve of value c is empty if $c \neq 2$, and is the whole xy plane if $c = 2$. ▲

EXAMPLE 2 The function $f: \mathbb{R}^2 \to \mathbb{R}$, defined by $f(x, y) = x + y + 2$ has as its graph the inclined plane $z = x + y + 2$. This plane intersects the xy plane ($z = 0$) in the line $y = -x - 2$ and the z axis at the point $(0, 0, 2)$. For any value $c \in \mathbb{R}$, the level curve of value c is the straight line $y = -x + (c - 2)$; or in symbols, the set

$$L_c = \{(x, y) \mid y = -x + (c - 2)\} \subset \mathbb{R}^2.$$

We indicate a few of the level curves of the function in Figure 2.1.4. This is a contour map of the function f.

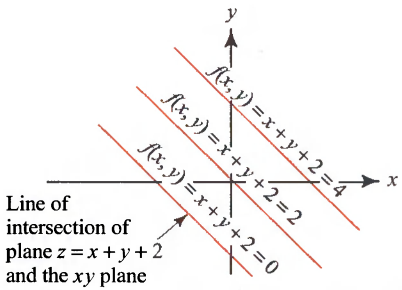


Figure 2.1.4 The level curves of $f(x, y) = x + y + 2$ show the sets on which f takes a given value.

From level curves labeled with the value or "height" of the function, the shape of the graph may be inferred by mentally elevating each level curve to the appropriate height, without stretching, tilting, or sliding it. If this procedure is visualized for all level curves, L_c, that is, for all values $c \in \mathbb{R}$, they will assemble to give the entire graph of f, as indicated by the shaded plane in Figure 2.1.5. If the graph is visualized

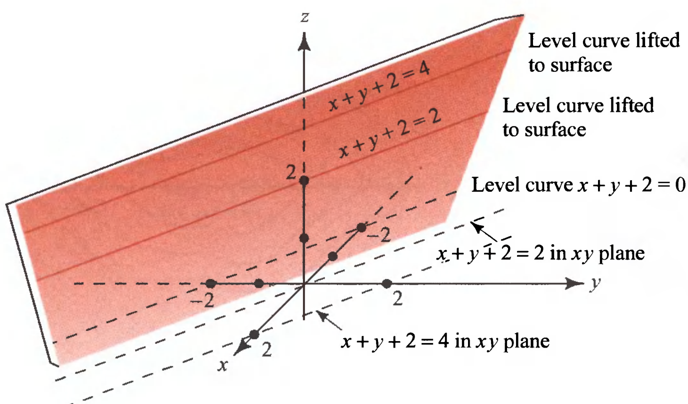


Figure 2.1.5 The relationship of level curves of Figure 2.1.4 to the graph of the function $f(x, y) = x + y + 2$, which is the plane $z = x + y + 2$.

using a finite number of level curves, a contour model is produced. If f is a smooth function, its graph will be a smooth surface, and so the contour model, mentally smoothed over, gives a good impression of the graph. ▲

DEFINITION: Level Curves and Surfaces Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ and let $c \in \mathbb{R}$. Then the ***level set of value*** c is defined to be the set of those points $\mathbf{x} \in U$ at which $f(\mathbf{x}) = c$. If $n = 2$, we speak of a ***level curve*** (of value c); and if $n = 3$, we speak of a ***level surface***. In symbols, the level set of value c is written

$$\{\mathbf{x} \in U \mid f(\mathbf{x}) = c\} \subset \mathbb{R}^n.$$

Note that the level set is always in the domain space.

EXAMPLE 3 Describe the graph of the quadratic function

$$f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2.$$

SOLUTION The graph is the ***paraboloid of revolution*** $z = x^2 + y^2$, oriented upward from the origin, around the z axis. The level curve of value c is empty for $c < 0$; for $c > 0$ the level curve of value c is the set $\{(x, y) \mid x^2 + y^2 = c\}$, a circle of radius $\sqrt{c}$ centered at the origin. Thus, raised to height c above the xy plane, the level set is a circle of radius $\sqrt{c}$, indicating a parabolic shape (see Figures 2.1.6 and 2.1.7). ▲

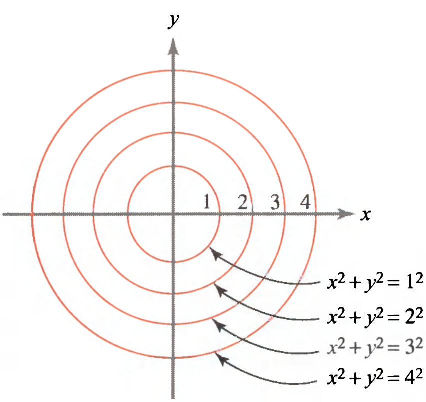


Figure 2.1.6 Some level curves for the function $f(x, y) = x^2 + y^2$.

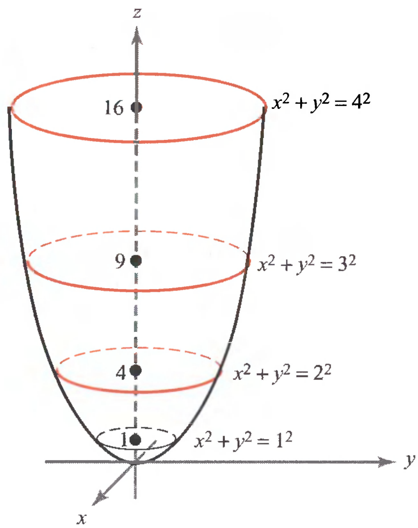


Figure 2.1.7 Level curves in Figure 2.1.6 raised to the graph.

The Method of Sections

By a ***section*** of the graph of f we mean the intersection of the graph and a (vertical) plane. For example, if P_1 is the xz plane in $\mathbb{R}^3$, defined by $y = 0$, then the section of f in Example 3 is the set

$$P_1 \cap \text{graph } f = \{(x, y, z) \mid y = 0, z = x^2\},$$

which is a parabola in the xz plane. Similarly, if P_2 denotes the yz plane, defined by $x = 0$, then the section

$$P_2 \cap \text{graph } f = \{(x, y, z) \mid x = 0, z = y^2\}$$

is a parabola in the yz plane (see Figure 2.1.8). It is usually helpful to compute at least one section to complement the information given by the level sets.

EXAMPLE 4 The graph of the quadratic function

$$f: \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 - y^2$$

is called a ***hyperbolic paraboloid***, or ***saddle***, centered at the origin. Sketch the graph.

SOLUTION To visualize this surface, we first draw the level curves. To determine the level curves, we solve the equation $x^2 - y^2 = c$. Consider the values $c = 0, \pm 1, \pm 4$. For $c = 0$, we have $y^2 = x^2$, or $y = \pm x$, so that this level set consists of two straight lines through the origin. For $c = 1$, the level curve is $x^2 - y^2 = 1$, or $y = \pm\sqrt{x^2 - 1}$, which is a hyperbola that passes vertically through the x axis at the points $(\pm 1, 0)$ (see Figure 2.1.9). Similarly, for $c = 4$, the level curve is defined by $y = \pm\sqrt{x^2 - 4}$, the hyperbola passing vertically through the x axis at $(\pm 2, 0)$. For

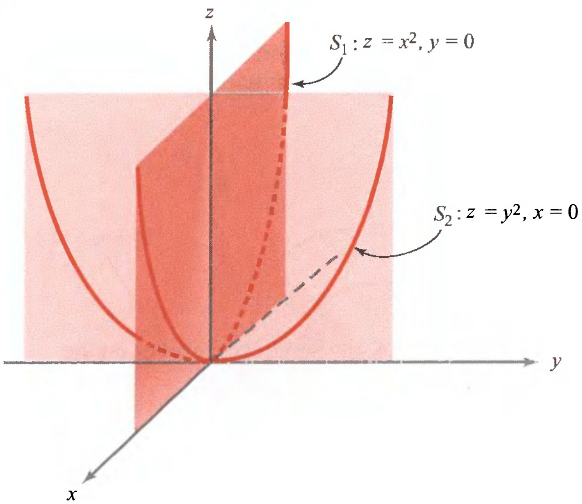


Figure 2.1.8 Two sections of the graph of $f(x, y) = x^2 + y^2$.

$c = -1$, we obtain the curve $x^2 - y^2 = -1$, that is, $x = \pm\sqrt{y^2 - 1}$, the hyperbola passing horizontally through the y axis at $(0, \pm 1)$. And for $c = -4$, the hyperbola through $(0, \pm 2)$ is obtained. These level curves are shown in Figure 2.1.9. Because it is not easy to visualize the graph of f from these data alone, we shall compute two sections, as in the previous example. For the section in the xz plane, we have

$$P_1 \cap \text{graph of } f = \{(x, y, z) \mid y = 0, z = x^2\},$$

which is a parabola opening upward; and for the yz plane,

$$P_2 \cap \text{graph } f = \{(x, y, z) \mid x = 0, z = -y^2\},$$

which is a parabola opening downward. The graph may now be visualized by lifting the level curves to the appropriate heights and smoothing out the resulting surface. Their

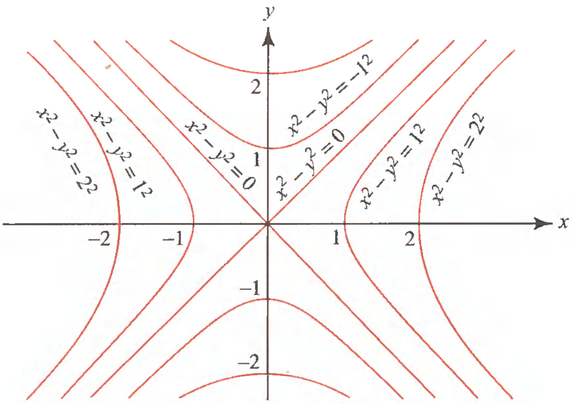


Figure 2.1.9 Level curves for the function $f(x, y) = x^2 - y^2$.

placement is aided by computing the parabolic sections. This procedure generates the hyperbolic saddle indicated in Figure 2.1.10. Compare this with the computer-generated graphs in Figure 2.1.11 (note that the orientation of the axes has been changed). ▲

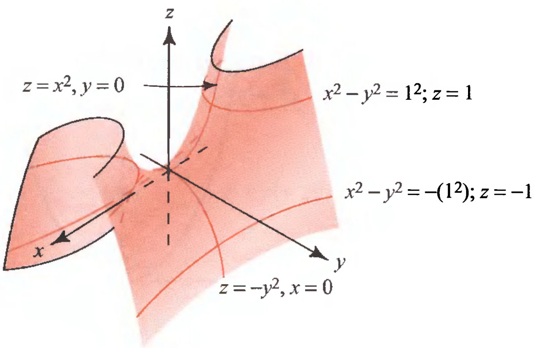


Figure 2.1.10 Some level curves on the graph of $f(x, y) = x^2 - y^2$.

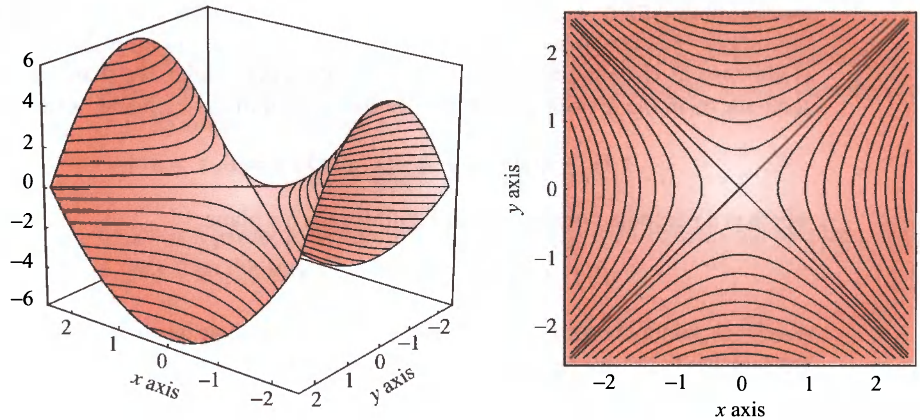


Figure 2.1.11 The graph of $z = x^2 - y^2$ and its level curves.

EXAMPLE 5 Describe the level sets of the function

$$f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto x^2 + y^2 + z^2.$$

SOLUTION This is the three-dimensional analogue of Example 3. In this context, level sets are surfaces in the three-dimensional domain $\mathbb{R}^3$. The graph, in $\mathbb{R}^4$, cannot be visualized directly, but sections can nevertheless be computed.

The level set with value c is the set

$$L_c = \{(x, y, z) \mid x^2 + y^2 + z^2 = c\},$$

which is the sphere centered at the origin with radius $\sqrt{c}$ for $c > 0$, is a single point at the origin for $c = 0$, and is empty for $c < 0$. The level sets for $c = 0, 1, 4$, and 9 are indicated in Figure 2.1.12. ▲

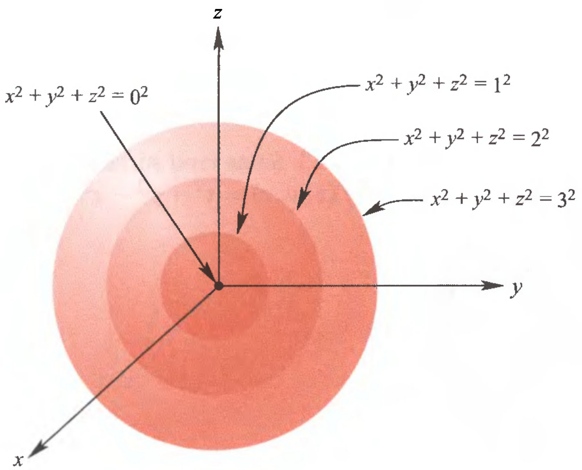


Figure 2.1.12 Some level surfaces for $f(x, y, z) = x^2 + y^2 + z^2$.

EXAMPLE 6 Describe the graph of the function $f\colon \mathbb{R}^3 \to \mathbb{R}$ defined by $f(x, y, z) = x^2 + y^2 - z^2$, which is the three-dimensional analogue of Example 4, and is also called a ***saddle***.

SOLUTION Formally, the graph of f is a subset of four-dimensional space. If we denote points in this space by (x, y, z, t), then the graph is given by

$$\{(x, y, z, t) \mid t = x^2 + y^2 - z^2\}.$$

The level surfaces of f are defined by

$$L_c = \{(x, y, z) \mid x^2 + y^2 - z^2 = c\}.$$

For $c = 0$, this is the cone $z = \pm\sqrt{x^2 + y^2}$ centered on the z axis. For c negative, say, $c = -a^2$, we obtain $z = \pm\sqrt{x^2 + y^2 + a^2}$, which is a hyperboloid of two sheets around the z axis, passing through the z axis at the points $(0, 0, \pm a)$. For c positive, say, $c = b^2$, the level surface is the ***single-sheeted hyperboloid of revolution*** around the z axis defined by $z = \pm\sqrt{x^2 + y^2 - b^2}$, which intersects the xy plane in the circle of radius $|b|$. These level surfaces are sketched in Figure 2.1.13.

Another view of the graph may be obtained from a section. For example, the subspace $S_{y=0} = \{(x, y, z, t) \mid y = 0\}$ intersects the graph in the section

$$S_{y=0} \cap \text{graph } f = \{(x, y, z, t) \mid y = 0, t = x^2 - z^2\},$$

that is, the set of points of the form $(x, 0, z, x^2 - z^2)$, which may be considered to be a surface in xzt space (see Figure 2.1.14). ▲

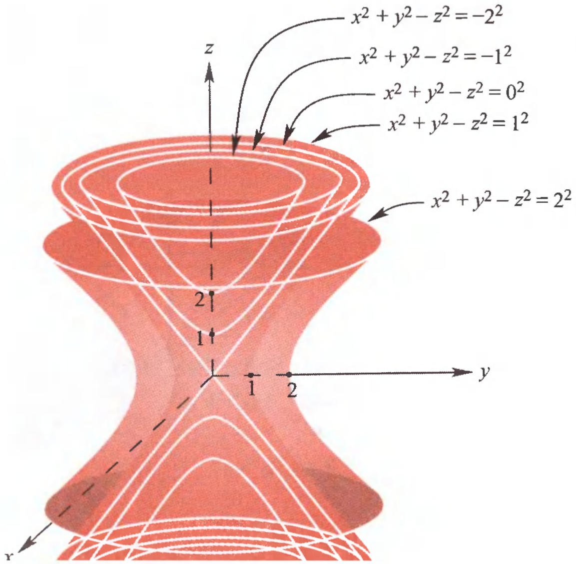
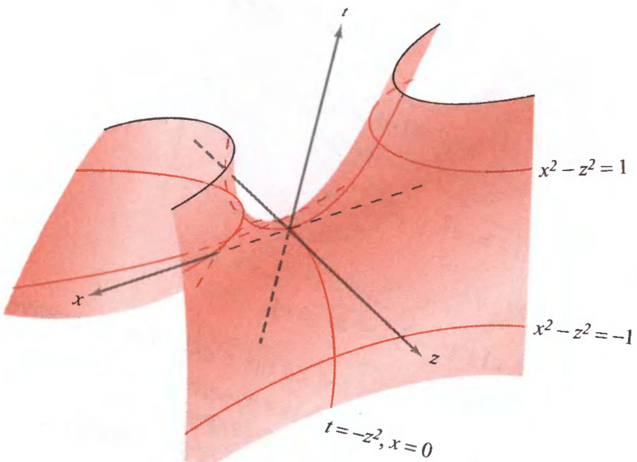


Figure 2.1.13 Some level surfaces of the function $f(x, y, z) = x^2 + y^2 - z^2$.



gure 2.1.14 The $y = 0$ section the graph of $f(x, y, z) =$ $+ y^2 - z^2$.

We have seen how the methods of sections and level sets can be used to understand the behavior of a function and its graph; these techniques can be quite useful to people who desire comprehensive visualization of complicated data. There are many computer programs available to do this, and we show the results of one such program in Figure 2.1.15.

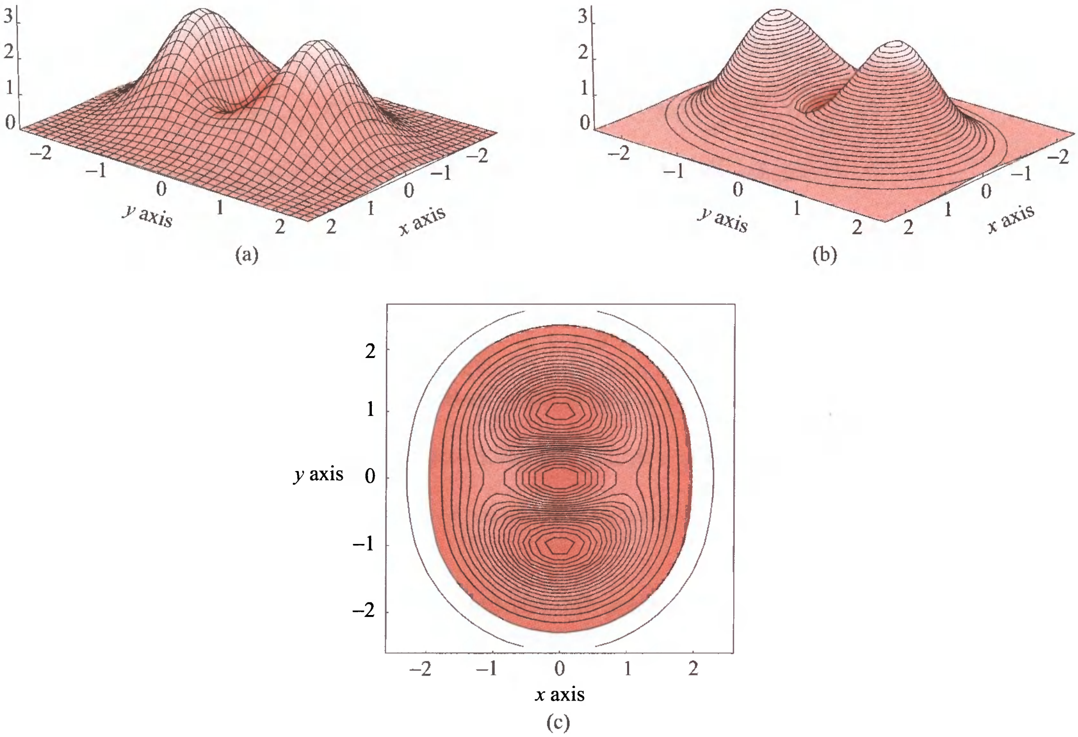


Figure 2.1.15 Computer-generated graph of $z = (x^2 + 3y^2)\exp(1 - x^2 - y^2)$ represented in three ways: (a) by sections, (b) by level curves on a graph, and (c) by level curves in the xy plane.

EXERCISES

1. Sketch the level curves and graphs of the following functions:

(a) $f\colon \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x - y + 2$

(b) $f\colon \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto x^2 + 4y^2$

(c) $f\colon \mathbb{R}^2 \to \mathbb{R},\ (x, y) \mapsto -xy$

2. Describe the behavior, as c varies, of the level curve $f(x, y) = c$ for each of these functions:

(a) $f(x, y) = x^2 + y^2 + 1$ (b) $f(x, y) = 1 - x^2 - y^2$ (c) $f(x, y) = x^3 - x$

3. For the functions in Examples 2, 3, and 4, compute the section of the graph defined by the plane

$$S_\theta = \{(x, y, z) \mid y = x \tan\theta\}$$

for a given *constant* θ. Do this by expressing z as a function of r, where $x = r\cos\theta$, $y = r\sin\theta$. Determine which of these functions f have the property that the shape of the section $S_\theta \cap \text{graph } f$ is independent of θ. (The solution for Example 3 only is in the Study Guide.)

In Exercises 4 to 10, draw the level curves (in the xy plane) for the given function f and specified values of c. Sketch the graph of $z = f(x, y)$.

4. $f(x, y) = 4 - 3x + 2y$, $c = 0, 1, 2, 3, -1, -2, -3$

5. $f(x, y) = (100 - x^2 - y^2)^{1/2}$, $c = 0, 2, 4, 6, 8, 10$

6. $f(x, y) = (x^2 + y^2)^{1/2}$, $c = 0, 1, 2, 3, 4, 5$

7. $f(x, y) = x^2 + y^2$, $c = 0, 1, 2, 3, 4, 5$

8. $f(x, y) = 3x - 7y$, $c = 0, 1, 2, 3, -1, -2, -3$

9. $f(x, y) = x^2 + xy$, $c = 0, 1, 2, 3, -1, -2, -3$

10. $f(x, y) = x/y$, $c = 0, 1, 2, 3, -1, -2, -3$

In Exercises 11 to 13, sketch or describe the level surfaces and a section of the graph of each function.

11. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto -x^2 - y^2 - z^2$

12. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto 4x^2 + y^2 + 9z^2$

13. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto x^2 + y^2$

In Exercises 14 to 18, describe the graph of each function by computing some level sets and sections.

14. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto xy$

15. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto xy + yz$

16. $f\colon \mathbb{R}^3 \to \mathbb{R}, (x, y, z) \mapsto xy + z^2$

17. $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto |y|$

18. $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto \max(|x|, |y|)$

Sketch or describe the surfaces in $\mathbb{R}^3$ of the equations presented in Exercises 19 to 31.

19. $4x^2 + y^2 = 16$

20. $x + 2z = 4$

21. $z^2 = y^2 + 4$

22. $x^2 + y^2 - 2x = 0$

23. $\dfrac{x}{4} = \dfrac{y^2}{4} + \dfrac{z^2}{9}$

24. $\dfrac{y^2}{9} + \dfrac{z^2}{4} = 1 + \dfrac{x^2}{16}$

25. $z = x^2$

26. $y^2 + z^2 = 4$

27. $z = \frac{y^2}{4} - \frac{x^2}{9}$

28. $y^2 = x^2 + z^2$

29. $4x^2 - 3y^2 + 2z^2 = 0$

30. $\frac{x^2}{9} + \frac{y^2}{12} + \frac{z^2}{9} = 1$

31. $x^2 + y^2 + z^2 + 4x - by + 9z - b = 0$, where b is a constant

32. Using polar coordinates, describe the level curves of the function defined by

$$f(x, y) = 2xy/(x^2 + y^2) \text{ if } (x, y) \neq (0, 0) \text{ and } f(0, 0) = 0.$$

33. Let $f\colon \mathbb{R}^2\backslash\{\mathbf{0}\} \to \mathbb{R}$ be given in polar coordinates by $f(r, \theta) = (\cos 2\theta)/r^2$. Sketch a few level curves in the xy plane. Here, $\mathbb{R}^2\backslash\{\mathbf{0}\} = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} \neq \mathbf{0}\}$.

34. Show that in Figure 2.1.15, the level "curve" $z = 3$ consists of two points.

2.2 Limits and Continuity

This section develops the concepts of open sets, limits, and continuity; open sets are needed to understand limits, and limits are in turn needed to understand continuity and differentiability.

As in elementary calculus, it is not necessary to completely master the limit concept in order to work problems in differentiation. For this reason, instructors may treat the following material with varying degrees of rigor. The student should consult with the instructor about the depth of understanding required.

Open Sets

We begin formulating the concept of an open set by defining an open disk. Let $\mathbf{x}_0 \in \mathbb{R}^n$ and let r be a positive real number. The ***open disk*** (or ***open ball***) of radius r and center $\mathbf{x}_0$ is defined to be the set of all points $\mathbf{x}$ such that $\|\mathbf{x} - \mathbf{x}_0\| < r$. This set is denoted $D_r(\mathbf{x}_0)$, and is the set of points $\mathbf{x}$ in $\mathbb{R}^n$ whose distance from $\mathbf{x}_0$ is less than r. Notice that we include only those $\mathbf{x}$ for which *strict* inequality holds. The disk $D_r(\mathbf{x}_0)$ is illustrated in Figure 2.2.1 for $n = 1, 2, 3$. For the case $n = 1$ and $x_0 \in \mathbb{R}$, the open disk $D_r(x_0)$ is the open interval $(x_0 - r, x_0 + r)$, which consists of all numbers $x \in \mathbb{R}$ *strictly* between $x_0 - r$ and $x_0 + r$. For the case $n = 2$, $\mathbf{x}_0 \in \mathbb{R}^2$, $D_r(\mathbf{x}_0)$ is the "inside" of the disk of radius r centered at $\mathbf{x}_0$. For the case $n = 3$, $\mathbf{x}_0 \in \mathbb{R}^3$, $D_r(\mathbf{x}_0)$ is the part strictly "inside" of the ball of radius r centered at $\mathbf{x}_0$.

DEFINITION: Open Sets Let $U \subset \mathbb{R}^n$ (that is, let U be a subset of $\mathbb{R}^n$). We call U an ***open set*** when for every point $\mathbf{x}_0$ in U there exists some $r > 0$ such that $D_r(\mathbf{x}_0)$ is contained within U; symbolically, we write $D_r(\mathbf{x}_0) \subset U$ (see Figure 2.2.2).

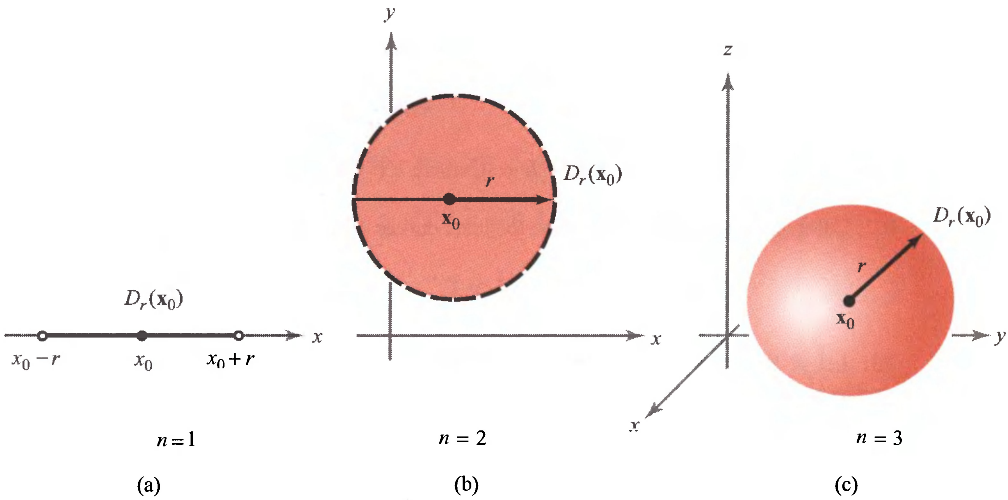


Figure 2.2.1 What disks $D_r(\mathbf{x}_0)$ look like in (a) one, (b) two, and (c) three dimensions.

The number $r > 0$ can depend on the point $\mathbf{x}_0$, and generally r will shrink as $\mathbf{x}_0$ gets closer to the "edge" of U. Intuitively speaking, a set U is open when the "boundary" points of U do not lie in U. In Figure 2.2.2, the dashed line is *not* included in U.

We establish the convention that the empty set $\varnothing$ (the set consisting of no elements) is open.

We have defined an open disk and an open set. From our choice of terms it would seem that an open disk should also be an open set. A little thought shows that this fact requires some proof. The following theorem does this.

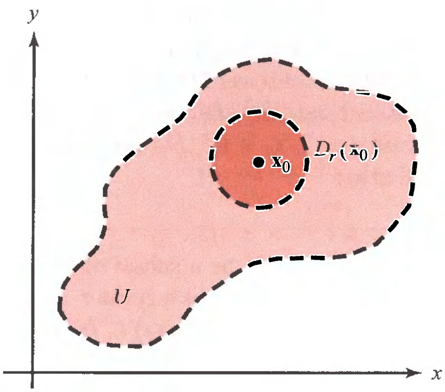


Figure 2.2.2 An open set U is one that completely encloses some disk $D_r(\mathbf{x}_0)$ about each of its points $\mathbf{x}_0$.

THEOREM 1 For each $\mathbf{x}_0 \in \mathbb{R}^n$ and $r > 0$, $D_r(\mathbf{x}_0)$ is an open set.

PROOF Let $\mathbf{x} \in D_r(\mathbf{x}_0)$, that is, let $\|\mathbf{x} - \mathbf{x}_0\| < r$. According to the definition of an open set, we must find an $s > 0$ such that $D_s(\mathbf{x}) \subset D_r(\mathbf{x}_0)$. Referring to Figure 2.2.3, we see that $s = r - \|\mathbf{x} - \mathbf{x}_0\|$ is a reasonable choice; note that $s > 0$, but that s becomes smaller if $\mathbf{x}$ is nearer the edge of $D_r(\mathbf{x}_0)$.

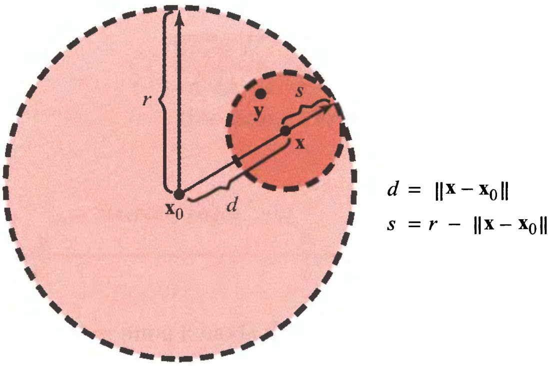


Figure 2.2.3 The geometry of the proof that an open disk is an open set.

To prove that $D_s(\mathbf{x}) \subset D_r(\mathbf{x}_0)$, let $\mathbf{y} \in D_s(\mathbf{x})$; that is, let $\|\mathbf{y} - \mathbf{x}\| < s$. We want to prove that $\mathbf{y} \in D_r(\mathbf{x}_0)$ as well. Proving this, in view of the definition of an r-disk, entails showing that $\|\mathbf{y} - \mathbf{x}_0\| < r$. This is done by using the triangle inequality for vectors in $\mathbb{R}^n$:

$$\|\mathbf{y} - \mathbf{x}_0\| = \|(\mathbf{y} - \mathbf{x}) + (\mathbf{x} - \mathbf{x}_0)\| \leq \|\mathbf{y} - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_0\| < s + \|\mathbf{x} - \mathbf{x}_0\| = r.$$

Hence, $\|\mathbf{y} - \mathbf{x}_0\| < r$. ■

The following example illustrates some techniques that are useful in establishing the openness of sets.

EXAMPLE 1 Prove that $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ is an open set.

SOLUTION The set is pictured in Figure 2.2.4.

Intuitively, this set is open, because no points on the "boundary," $x = 0$, are contained in the set. Such an argument will often suffice after one becomes accustomed to the concept of openness. At first, however, we should give details. To prove that A is open, we show that for every point $(x, y) \in A$ there exists an $r > 0$ such that $D_r(x, y) \subset A$. If $(x, y) \in A$, then $x > 0$. Choose $r = x$. If $(x_1, y_1) \in D_r(x, y)$, we have

$$|x_1 - x| = \sqrt{(x_1 - x)^2} \leq \sqrt{(x_1 - x)^2 + (y_1 - y)^2} < r = x,$$

and so $x_1 - x < x$ and $x - x_1 < x$. The latter inequality implies $x_1 > 0$, that is, $(x_1, y_1) \in A$. Hence $D_r(x, y) \subset A$, and therefore A is open (see Figure 2.2.5). ▲

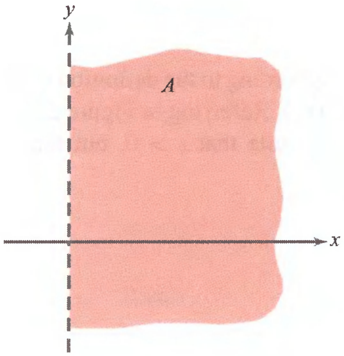


Figure 2.2.4 Show that A is an open set.

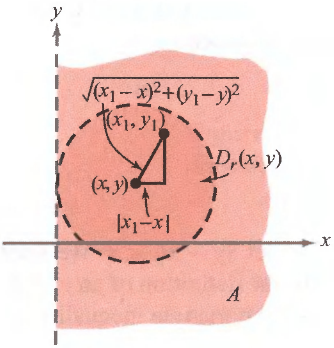


Figure 2.2.5 The construction of a disk about a point in A that is completely enclosed in A.

It is useful to have a special name for an open set containing a given point $\mathbf{x}$, because this idea arises often in the study of limits and continuity. Thus, by a ***neighborhood*** of $\mathbf{x} \in \mathbb{R}^n$ we merely mean an open set U containing the point $\mathbf{x}$. For example, $D_r(\mathbf{x}_0)$ is a neighborhood of $\mathbf{x}_0$ for any $r > 0$. The set A in Example 1 is a neighborhood of the point $\mathbf{x}_0 = (3, -10)$.

Boundary

Let us formally introduce the concept of a boundary point, which we alluded to in Example 1.

DEFINITION: Boundary Points Let $A \subset \mathbb{R}^n$. A point $\mathbf{x} \in \mathbb{R}^n$ is called a ***boundary point*** of A if every neighborhood of $\mathbf{x}$ contains at least one point in A and at least one point not in A.

In this definition, $\mathbf{x}$ itself may or may not be in A; if $\mathbf{x} \in A$, then $\mathbf{x}$ is a boundary point if every neighborhood of $\mathbf{x}$ contains at least one point *not* in A (it already contains a point of A, namely, $\mathbf{x}$). Similarly, if $\mathbf{x}$ is not in A, it is a boundary point if every neighborhood of $\mathbf{x}$ contains at least one point of A.

We shall be particularly interested in boundary points of open sets. By the definition of an open set, no point of an open set A can be a boundary point of A. Thus, *a point* $\mathbf{x}$ *is a boundary point of an open set* A *if and only if* $\mathbf{x}$ *is not in* A *and every neighborhood of* $\mathbf{x}$ *has a nonempty intersection with* A.

This expresses in precise terms the intuitive idea that a boundary point of A is a point just on the "edge" of A. In many examples it is perfectly clear what the boundary points are.

EXAMPLE 2 (a) Let $A = (a, b)$ in $\mathbb{R}$. Then the boundary points of A consist of the points a and b. A consideration of Figure 2.2.6 and the definition will make this clear. [The reader will be asked to prove this in Exercise 20(c).]

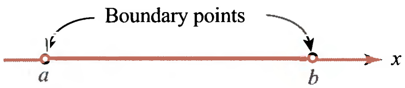


Figure 2.2.6 The boundary points of the interval (a, b).

(b) Let $A = D_r(x_0, y_0)$ be an r-disk about (x_0, y_0) in the plane. The boundary consists of points (x, y) with $(x - x_0)^2 + (y - y_0)^2 = r^2$ (Figure 2.2.7).

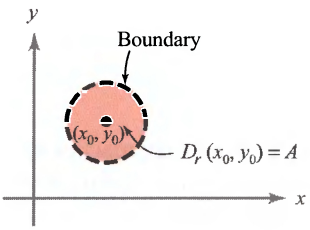


Figure 2.2.7 The boundary of A consists of points on the edge of A.

(c) Let $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$. Then the boundary of A consists of all points on the y axis (the student should draw a figure).

(d) Let A be $D_r(\mathbf{x}_0)$ minus the point $\mathbf{x}_0$ (a "punctured" disk about $\mathbf{x}_0$). Then $\mathbf{x}_0$ is a boundary point of A. ▲

Limits

We now turn our attention to the concept of a limit. *Throughout the following discussions the domain of definition of the function* f *will be an open set* A. We are interested in finding the limit of f as $\mathbf{x} \in A$ approaches either a point of A or a boundary point of A.

The reader should appreciate the fact that the limit concept is a basic and useful tool for the analysis of functions; it enables us to study derivatives, and hence maxima and minima, asymptotes, improper integrals, and other important features of functions, as well as being useful for infinite series and sequences. We will present a theory of limits for functions of several variables that includes the theory for functions of one variable as a special case.

In one-variable calculus, the student has encountered the notion of $\operatorname*{limit}_{x\to x_0} f(x) = l$ for a function $f\colon A \subset \mathbb{R} \to \mathbb{R}$ from a subset A of the real numbers to the real numbers. Intuitively, this means that as x gets closer and closer to x_0, the values $f(x)$ get closer and closer to (the limiting value) l. To put this intuitive idea on a firm, mathematical foundation, either the "epsilon (ε) and delta (δ) method" or the "neighborhood method" is usually introduced. The same is true for functions of several variables. In what follows we develop the neighborhood approach to limits. The epsilon-delta approach is left for optional study at the end of this section.

DEFINITION: Limit Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$, where A is an open set. Let $\mathbf{x}_0$ be in A or be a boundary point of A, and let N be a neighborhood of $\mathbf{b} \in \mathbb{R}^m$. We say f is ***eventually in*** N ***as*** $\mathbf{x}$ ***approaches*** $\mathbf{x}_0$ if there exists a neighborhood U of $\mathbf{x}_0$ such that $\mathbf{x} \neq \mathbf{x}_0$, $\mathbf{x} \in U$, and $\mathbf{x} \in A$ imply $f(\mathbf{x}) \in N$. [The geometric meaning of this assertion is illustrated in Figure 2.2.8; note that $\mathbf{x}_0$ need not be in the set A, so that $f(\mathbf{x}_0)$ is not necessarily defined.] We say $f(\mathbf{x})$ ***approaches*** $\mathbf{b}$ as $\mathbf{x}$ approaches $\mathbf{x}_0$, or, in symbols,

$$\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} \qquad \text{or} \qquad f(\mathbf{x}) \to \mathbf{b} \quad \text{as} \quad \mathbf{x} \to \mathbf{x}_0,$$

when, given *any* neighborhood N of $\mathbf{b}$, f is eventually in N as $\mathbf{x}$ approaches $\mathbf{x}_0$ [that is, "$f(\mathbf{x})$ is close to $\mathbf{b}$ if $\mathbf{x}$ is close to $\mathbf{x}_0$"]. It may be that as $\mathbf{x}$ approaches $\mathbf{x}_0$, the values $f(\mathbf{x})$ do not get close to any particular number. In this case, we say that $\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ ***does not exist***.

Henceforth, whenever we consider the notion $\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$, we shall always assume that $\mathbf{x}_0$ either belongs to some open set on which f is defined or is on the boundary of such a set.

One reason we insist on $\mathbf{x} \neq \mathbf{x}_0$ in the definition of limit will become clear if we remember from one-variable calculus that we want to be able to define the derivative $f'(x_0)$ of a function f at a point x_0 by

$$f'(x_0) = \operatorname*{limit}_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

and this expression is not defined at $x = x_0$.

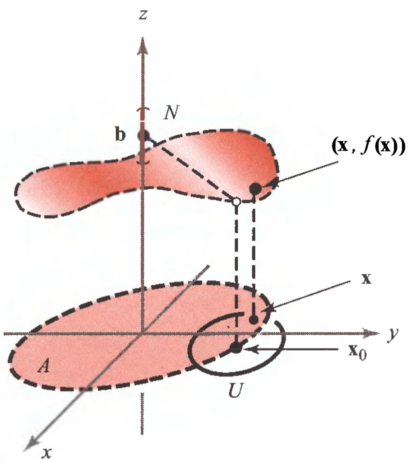


Figure 2.2.8 Limits in terms of neighborhoods; if $\mathbf{x}$ is in U, then $f(\mathbf{x})$ will be in N. (The little open circle denotes that the point does not lie on the graph.) In the figure, $f\colon A = \{(x, y) \mid x^2 + y^2 < 1\} \to \mathbb{R}$. (The dashed line is not in the graph of f.)

EXAMPLE 3 (a) This example illustrates a limit that does not exist. Consider the function $f\colon \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x \le 0. \end{cases}$$

The $\lim_{x\to 0} f(x)$ does not exist, since there are points x_1 arbitrarily close to 0 with $f(x_1) = 1$ and also points x_2 arbitrarily close to 0 with $f(x_2) = -1$; that is, there is no single number that f is close to when x is close to 0 (see Figure 2.2.9). If f is restricted to the domain (0, 1) or (−1, 0), then the limit does exist. Can you say why?

(b) This example illustrates a function whose limit does exist, but whose limiting value does not equal its value at the limiting point. Define $f\colon \mathbb{R} \to \mathbb{R}$ by

$$f(x) = \begin{cases} 0 & \text{if } x \ne 0 \\ 1 & \text{if } x = 0. \end{cases}$$

It is true that $\lim_{x\to 0} f(x) = 0$, since for any neighborhood U of 0, $x \in U$ and $x \neq 0$ implies that $f(x) = 0$. One sees from the graph in Figure 2.2.10 that f approaches 0 as $x \to 0$; we do not care that f happens to take on some other value at 0. ▲

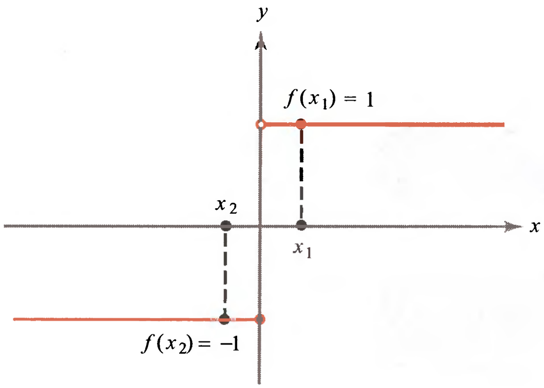


Figure 2.2.9 The limit of this function as $x \to 0$ does not exist.

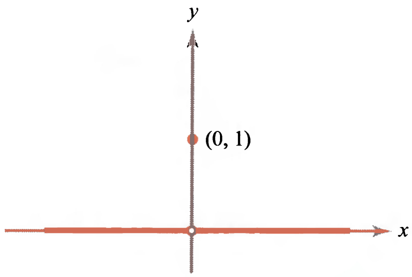


Figure 2.2.10 The limit of this function as $x \to 0$ is zero.

EXAMPLE 4 Use the definition to verify that the "obvious" $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{x} = \mathbf{x}_0$ holds, where $\mathbf{x}$ and $\mathbf{x}_0 \in \mathbb{R}^n$.

SOLUTION Let f be the function defined by $f(\mathbf{x}) = \mathbf{x}$, and let N be any neighborhood of $\mathbf{x}_0$. We must show that $f(\mathbf{x})$ is eventually in N as $\mathbf{x} \to \mathbf{x}_0$. According to the definition, we must find a neighborhood U of $\mathbf{x}_0$ with the property that if $\mathbf{x} \neq \mathbf{x}_0$ and $\mathbf{x} \in U$, then $f(\mathbf{x}) \in N$. Pick $U = N$. If $\mathbf{x} \in U$, then $\mathbf{x} \in N$; because $\mathbf{x} = f(\mathbf{x})$, it follows that $f(\mathbf{x}) \in N$. Thus, we have shown that $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{x} = \mathbf{x}_0$. In a similar way, we have

$$\lim_{(x,y)\to(x_0,y_0)} x = x_0, \qquad \text{etc.} \quad \blacktriangle$$

In what follows, the student may assume, without proof, the validity of limits from one-variable calculus. For example, $\lim_{x\to 1} \sqrt{x} = \sqrt{1} = 1$ and $\lim_{\theta\to 0} \sin\theta = \sin 0 = 0$ may be used.

EXAMPLE 5 (This example demonstrates another case in which the limit cannot simply be "read off" from the function.) Find $\lim_{x\to 1} g(x)$ where

$$g\colon x \mapsto \frac{x-1}{\sqrt{x}-1}.$$

SOLUTION This function is graphed in Figure 2.2.11(a).

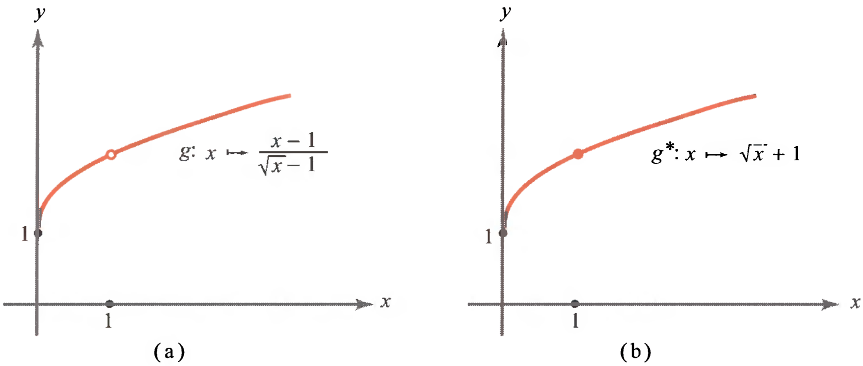


Figure 2.2.11 These graphs are the same except that in part (a), g is undefined at $x = 1$, whereas in part (b), g^* is defined for all $x \geq 0$.

We see that $g(1)$ is not defined, because division by zero is not defined. However, if we multiply the numerator and denominator of $g(x)$ by $\sqrt{x} + 1$, we find that for all x in the domain of g we have

$$g(x) = \frac{x-1}{\sqrt{x}-1} = \sqrt{x} + 1, \qquad x \neq 1.$$

The expression $g^*(x) = \sqrt{x} + 1$ is defined and takes the value 2 at $x = 1$; from one-variable calculus, $g^*(x) \to 2$ as $x \to 1$. But because $g^*(x) = g(x)$ for all $x \geq 0$, $x \neq 1$, we must have as well that $g(x) \to 2$ as $x \to 1$. ▲

We will consider other examples in two variables shortly.

Properties of Limits

To properly speak of *the* limit, we should establish that f can have *at most one* limit as $\mathbf{x} \to \mathbf{x}_0$. This is intuitively clear and we now state it formally. (See the Internet supplement for the proof.)

THEOREM 2: Uniqueness of Limits

$$\text{If} \quad \operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_1 \quad \text{and} \quad \operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_2, \quad \text{then} \quad \mathbf{b}_1 = \mathbf{b}_2.$$

To carry out practical computations with limits, we require some rules for limits, for example, that the limit of a sum is the sum of the limits. These rules are summarized in the following theorem (see the Internet supplement for Chapter 2 for the proof).

THEOREM 3: Properties of Limits Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$, $g\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x}_0$ be in A or be a boundary point of A, $\mathbf{b} \in \mathbb{R}^m$, and $c \in \mathbb{R}$; then

(i) If $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}$, then $\lim_{\mathbf{x}\to\mathbf{x}_0} cf(\mathbf{x}) = c\mathbf{b}$, where $cf\colon A \to \mathbb{R}^m$ is defined by $\mathbf{x} \mapsto c(f(\mathbf{x}))$.

(ii) If $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}_1$ and $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = \mathbf{b}_2$, then $\lim_{\mathbf{x}\to\mathbf{x}_0} (f+g)(\mathbf{x}) = \mathbf{b}_1 + \mathbf{b}_2$, where $(f+g)\colon A \to \mathbb{R}^m$ is defined by $\mathbf{x} \mapsto f(\mathbf{x}) + g(\mathbf{x})$.

(iii) If $m = 1$, $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = b_1$, and $\lim_{\mathbf{x}\to\mathbf{x}_0} g(\mathbf{x}) = b_2$, then $\lim_{\mathbf{x}\to\mathbf{x}_0} (fg)(\mathbf{x}) = b_1 b_2$, where $(fg)\colon A \to \mathbb{R}$ is defined by $\mathbf{x} \mapsto f(\mathbf{x})g(\mathbf{x})$.

(iv) If $m = 1$, $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = b \neq 0$, and $f(\mathbf{x}) \neq 0$ for all $\mathbf{x} \in A$, then $\lim_{\mathbf{x}\to\mathbf{x}_0} 1/f(\mathbf{x}) = 1/b$, where $1/f\colon A \to \mathbb{R}$ is defined by $\mathbf{x} \mapsto 1/f(\mathbf{x})$.

(v) If $f(\mathbf{x}) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ where $f_i\colon A \to \mathbb{R}, i = 1, \ldots, m$, are the component functions of f, then $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b} = (b_1, \ldots, b_m)$ if and only if $\lim_{\mathbf{x}\to\mathbf{x}_0} f_i(\mathbf{x}) = b_i$ for each $i = 1, \ldots, m$.

These results ought to be intuitively clear. For instance, rule (ii) says that if $f(\mathbf{x})$ is close to $\mathbf{b}_1$ and $g(\mathbf{x})$ is close to $\mathbf{b}_2$ when $\mathbf{x}$ is close to $\mathbf{x}_0$, then $f(\mathbf{x}) + g(\mathbf{x})$ is close to $\mathbf{b}_1 + \mathbf{b}_2$ when $\mathbf{x}$ is close to $\mathbf{x}_0$. The following example illustrates how this works.

EXAMPLE 6 Let $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 + y^2 + 2$. Compute the limit

$$\lim_{(x,y)\to(0,1)} f(x, y).$$

SOLUTION Here f is the sum of the three functions $(x, y) \mapsto x^2$, $(x, y) \mapsto y^2$, and $(x, y) \mapsto 2$. The limit of a sum is the sum of the limits, and the limit of a product is the product of the limits (Theorem 3). Hence, using the fact that $\lim_{(x,y)\to(x_0,y_0)} x = x_0$ (Example 4), we obtain

$$\lim_{(x,y)\to(x_0,y_0)} x^2 = \left(\lim_{(x,y)\to(x_0,y_0)} x\right)\left(\lim_{(x,y)\to(x_0,y_0)} x\right) = x_0^2$$

and, using the same reasoning, $\lim_{(x,y)\to(x_0,y_0)} y^2 = y_0^2$. Consequently,

$$\lim_{(x,y)\to(0,1)} f(x, y) = 0^2 + 1^2 + 2 = 3. \quad \blacktriangle$$

Continuous Functions

In single-variable calculus we learned that the idea of a continuous function is based on the intuitive notion of a function whose graph is an unbroken curve, that is, a curve that has no *jumps*, or the kind of curve that would be traced by a particle in motion or by a moving pencil point that is not lifted from the paper.

To perform a detailed analysis of functions, we need concepts more precise than this rather vague notion. An example may clarify these ideas. Consider the specific function $f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = -1$ if $x \leq 0$ and $f(x) = 1$ if $x > 0$. The graph of f is shown in Figure 2.2.12(a). [The little open circle denotes the fact that the point (0, 1) does *not* lie on the graph of f]. Clearly, the graph of f is broken at $x = 0$. Consider also the function $g: x \mapsto x^2$. This function is pictured in Figure 2.2.12(b). The graph of g is not broken at any point.

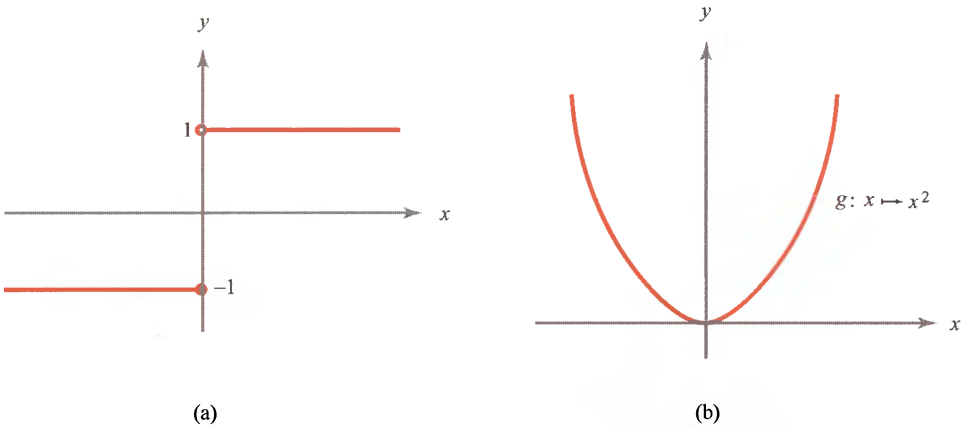


Figure 2.2.12 The function f in part (a) is not continuous, because its value jumps as x crosses 0, whereas the function g in part (b) is continuous.

If one examines examples of functions like f, whose graphs are broken at some point x_0, and functions like g, whose graphs are not broken, one sees that the principal difference between them is that for a function like g, the values of $g(x)$ get closer to $g(x_0)$ as x gets closer and closer to x_0. The same idea works for functions of several variables. But the notion of closer and closer does not suffice as a mathematical definition; thus, we shall formulate these concepts precisely in terms of limits.

Because the condition $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$ means that $f(\mathbf{x})$ is close to $f(\mathbf{x}_0)$ when $\mathbf{x}$ is close to $\mathbf{x}_0$, we see that this limit condition does indeed correspond to the requirement that the graph of f be unbroken (see Figure 2.2.13, where we illustrate the case $f: \mathbb{R} \to \mathbb{R}$). The case of several variables is easiest to visualize if we deal with real-valued functions, say $f: \mathbb{R}^2 \to \mathbb{R}$. In this case, we can visualize f by drawing its graph, which consists of all points (x, y, z) in $\mathbb{R}^3$ with $z = f(x, y)$. The continuity of f thus means that its graph has no "breaks" in it (see Figure 2.2.14).

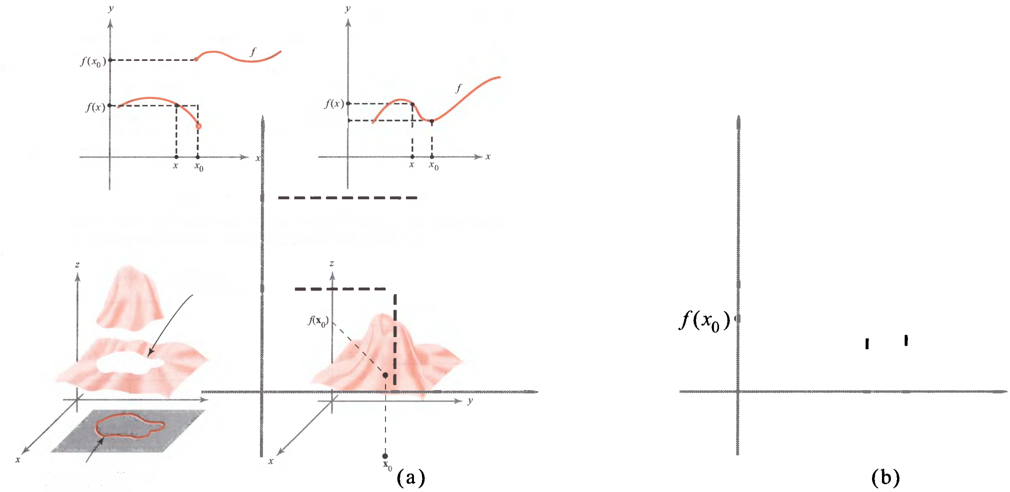


Figure 2.2.13 (a) Discontinuous function for which $\text{limit}_{x \to x_0} f(x)$ does not exist. (b) Continuous function for which this limit exists and equals $f(x_0)$.

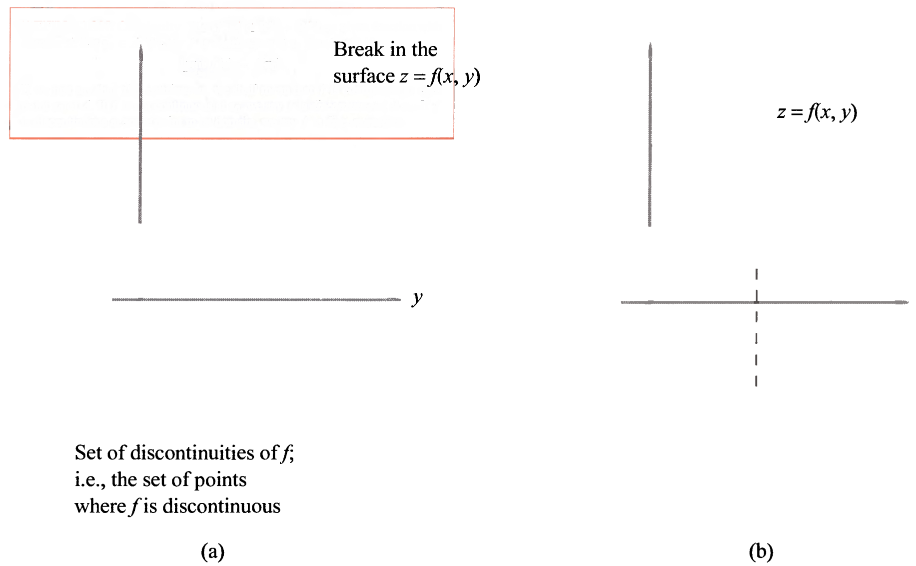


Figure 2.2.14 (a) A discontinuous function of two variables. (b) A continuous function.

DEFINITION: Continuity Let $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ be a given function with domain A. Let $\mathbf{x}_0 \in A$. We say f is ***continuous*** at $\mathbf{x}_0$ if and only if

$$\underset{\mathbf{x} \to \mathbf{x}_0}{\text{limit}} f(\mathbf{x}) = f(\mathbf{x}_0).$$

If we just say that f is ***continuous***, we shall mean that f is continuous at each point $\mathbf{x}_0$ of A. If f is not continuous at $\mathbf{x}_0$, we say f is ***discontinuous*** at $\mathbf{x}_0$. If f is discontinuous at some point in its domain, we say f is ***discontinuous***.

EXAMPLE 7 Any polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is continuous from $\mathbb{R}$ to $\mathbb{R}$. Indeed, from Theorem 3 and Example 4,

$$\begin{aligned}\operatorname*{limit}_{x \to x_0} (a_0 + a_1 x + \cdots + a_n x^n) &= \operatorname*{limit}_{x \to x_0} a_0 + \operatorname*{limit}_{x \to x_0} a_1 x + \cdots + \operatorname*{limit}_{x \to x_0} a_n x^n \\ &= a_0 + a_1 x_0 + \cdots + a_n x_0^n,\end{aligned}$$

because the limit of a product is the product of the limits, which gives

$$\operatorname*{limit}_{x \to x_0} x^n = \left(\operatorname*{limit}_{x \to x_0} x\right)^n = x_0^n. \quad \blacktriangle$$

EXAMPLE 8 Let $f\colon \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = xy$. Then f is continuous, because, by the limit theorems and Example 4,

$$\operatorname*{limit}_{(x,y)\to(x_0,y_0)} xy = \left(\operatorname*{limit}_{(x,y)\to(x_0,y_0)} x\right)\left(\operatorname*{limit}_{(x,y)\to(x_0,y_0)} y\right) = x_0 y_0. \quad \blacktriangle$$

One can see by the same method that any polynomial $p(x, y)$ [for example, $p(x, y) = 3x^2 - 6xy^2 + y^3$] in x and y is continuous.

EXAMPLE 9 The function $f\colon \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} 1 & \text{if } x \leq 0 \text{ or } y \leq 0 \\ 0 & \text{otherwise} \end{cases}$$

is not continuous at (0, 0) or at any point on the positive x axis or positive y axis. Indeed, if $(x_0, y_0) = \mathbf{u}$ is such a point (i.e., $x_0 = 0$ and $y_0 \geq 0$, or $y_0 = 0$ and $x_0 \geq 0$) and $\delta > 0$, there are points $(x, y) \in D_\delta(\mathbf{u})$, a neighborhood of $\mathbf{u}$, with $f(x, y) = 1$ and other points $(x, y) \in D_\delta(\mathbf{u})$ with $f(x, y) = 0$. Thus, it is *not* true that $f(x, y) \to f(x_0, y_0) = 1$ as $(x, y) \to (x_0, y_0)$. $\blacktriangle$

To prove that specific functions are continuous, we can avail ourselves of the limit theorems (see Theorem 3 and Example 7). If we transcribe those results in terms of continuity, we are led to the following:

THEOREM 4: Properties of Continuous Functions Suppose that $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$, $g\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$, and let c be a real number.

(i) If f is continuous at $\mathbf{x}_0$, so is cf, where $(cf)(\mathbf{x}) = c[f(\mathbf{x})]$.

(ii) If f and g are continuous at $\mathbf{x}_0$, so is $f + g$, where the sum of f and g is defined by $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$.

(iii) If f and g are continuous at $\mathbf{x}_0$ and $m = 1$, then the product function fg defined by $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ is continuous at $\mathbf{x}_0$.

(iv) If $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$ is continuous at $\mathbf{x}_0$ and nowhere zero on A, then the quotient $1/f$ is continuous at $\mathbf{x}_0$, where $(1/f)(\mathbf{x}) = 1/f(\mathbf{x})$.

(v) If $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ and $f(\mathbf{x}) = (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))$, then f is continuous at $\mathbf{x}_0$ if and only if each of the real-valued functions $f_1, \dots, f_m$ is continuous at $\mathbf{x}_0$.

A variant of (iv) is often used: If $f(\mathbf{x}_0) \neq 0$ and f is continuous, then $f(\mathbf{x}) \neq 0$ in a neighborhood of $\mathbf{x}_0$ and so $1/f$ is defined in that neighborhood, and $1/f$ is continuous at $\mathbf{x}_0$.

EXAMPLE 10 Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^2y, (y + x^3)/(1 + x^2))$. Show that f is continuous.

SOLUTION To see this, it is sufficient, by property (v) of Theorem 4, to show that each component is continuous. As we have mentioned, any polynomial in two variables is continuous; thus, the map $(x, y) \mapsto x^2y$ is continuous. Because $1 + x^2$ is continuous and nonzero, by property (iv), we know that $1/(1 + x^2)$ is continuous; hence, $(y + x^3)/(1 + x^2)$ is a product of continuous functions, and by (iii) is continuous. ▲

Similar reasoning applies to examples like the function $\mathbf{c}\colon \mathbb{R} \to \mathbb{R}^3$ given by $\mathbf{c}(t) = (t^2, 1, t^3/(1 + t^2))$ to show they are continuous as well.

Composition

Next we discuss *composition*, another basic operation that can be performed on functions. If g maps A to B and f maps B to C, the ***composition of*** g ***with*** f, or of f on g, denoted by $f \circ g$, maps A to C by sending $\mathbf{x} \mapsto f(g(\mathbf{x}))$ (see Figure 2.2.15). For example, $\sin(x^2)$ is the composition of $x \mapsto x^2$ with $y \mapsto \sin y$.

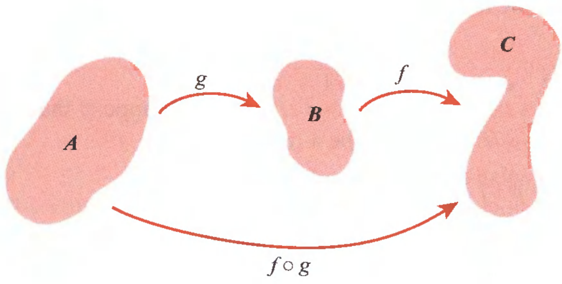


Figure 2.2.15 The composition of f on g.

THEOREM 5: Continuity of Compositions Let $g\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ and let $f\colon B \subset \mathbb{R}^m \to \mathbb{R}^p$. Suppose $g(A) \subset B$, so that $f \circ g$ is defined on A. If g is continuous at $\mathbf{x}_0 \in A$ and f is continuous at $\mathbf{y}_0 = g(\mathbf{x}_0)$, then $f \circ g$ is continuous at $\mathbf{x}_0$.

The intuition behind this is easy; the formal proof in the Internet supplement follows a similar pattern. Intuitively, we must show that as $\mathbf{x}$ gets close to $\mathbf{x}_0$, $f(g(\mathbf{x}))$ gets close to $f(g(\mathbf{x}_0))$. But as $\mathbf{x}$ gets close to $\mathbf{x}_0$, $g(\mathbf{x})$ gets close to $g(\mathbf{x}_0)$ (by continuity of g at $\mathbf{x}_0$); and as $g(\mathbf{x})$ gets close to $g(\mathbf{x}_0)$, $f(g(\mathbf{x}))$ gets close to $f(g(\mathbf{x}_0))$ [by continuity of f at $g(\mathbf{x}_0)$].

EXAMPLE 11 Let $f(x, y, z) = (x^2 + y^2 + z^2)^{30} + \sin z^3$. Show that f is continuous.

SOLUTION Here we can write f as a sum of the two functions $(x^2 + y^2 + z^2)^{30}$ and $\sin z^3$, so it suffices to show that each is continuous. The first is the composite of $(x, y, z) \mapsto (x^2 + y^2 + z^2)$ with $u \mapsto u^{30}$, and the second is the composite of $(x, y, z) \mapsto z^3$ with $u \mapsto \sin u$, and so we have continuity by Theorem 5. ▲

Limits in Terms of ε's and δ's

We now state a theorem (proved in the Internet supplement for Chapter 2) giving a useful formulation of the notion of limit in terms of epsilons and deltas that is often taken as the *definition* of limit. This is, in fact, another way of making precise the intuitive statement that "$f(\mathbf{x})$ is close to $\mathbf{b}$ when $\mathbf{x}$ is close to $\mathbf{x}_0$." To help understand this formulation, the reader should consider it with respect to each of the examples already presented.

THEOREM 6 Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ and let $\mathbf{x}_0$ be in A or be a boundary point of A. Then $\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \mathbf{b}$ if and only if for every number $\varepsilon > 0$ there is a $\delta > 0$ such that for any $\mathbf{x} \in A$ satisfying $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$, we have $\|f(\mathbf{x}) - \mathbf{b}\| < \varepsilon$ (see Figure 2.2.16).

To illustrate the methodology of the epsilon-delta technique in Theorem 6, we consider the following examples.

EXAMPLE 12 Show that $\operatorname*{limit}_{(x,y)\to(0,0)} x = 0$ using the ε-δ method.

SOLUTION Note that if $\delta > 0$, $\|(x, y) - (0, 0)\| = \sqrt{x^2 + y^2} < \delta$ implies $|x - 0| = |x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2} < \delta$. Thus, if $\|(x, y) - (0, 0)\| < \delta$, then $|x - 0|$ is also less than δ. Given $\varepsilon > 0$, we are required to find a $\delta > 0$ (generally depending on ε) with the property that $0 < \|(x, y) - (0, 0)\| < \delta$ implies $|x - 0| < \varepsilon$. What are we to pick as our δ? From the preceding calculation, we see that if we choose $\delta = \varepsilon$,

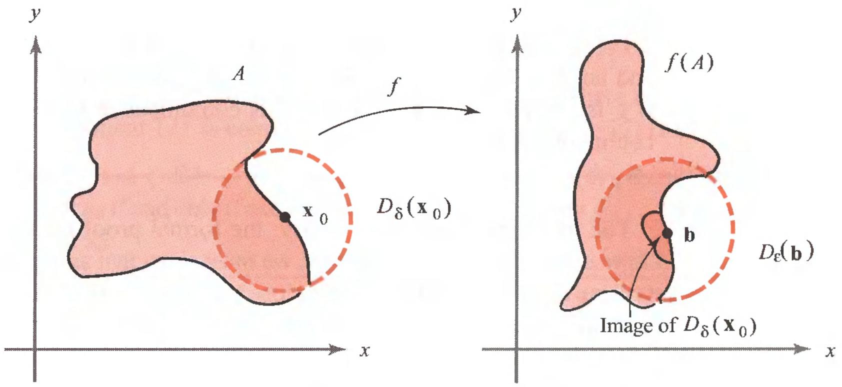


Figure 2.2.16 The geometry of the ε-δ definition of limit.

then $\|(x, y) - (0, 0)\| < \delta$ implies $|x - 0| < \varepsilon$. This shows that $\operatorname*{limit}_{(x,y)\to(0,0)} x = 0$. Given $\varepsilon > 0$, we could have also chosen $\delta = \varepsilon/2$ or $\varepsilon/3$, but it suffices to find just one δ satisfying the requirements of the definition of a limit. ▲

EXAMPLE 13 Consider the function

$$f(x, y) = \frac{\sin(x^2 + y^2)}{x^2 + y^2}.$$

Even though f is not defined at (0, 0), determine whether $f(x, y)$ approaches some number as (x, y) approaches (0, 0).

SOLUTION From one-variable calculus or L'Hôpital's rule we know that

$$\operatorname*{limit}_{\alpha\to 0} \frac{\sin \alpha}{\alpha} = 1.$$

Thus, it is reasonable to guess that

$$\operatorname*{limit}_{\mathbf{v}\to(0,0)} f(\mathbf{v}) = \operatorname*{limit}_{\mathbf{v}\to(0,0)} \frac{\sin \|\mathbf{v}\|^2}{\|\mathbf{v}\|^2} = 1.$$

Indeed, because $\operatorname*{limit}_{\alpha\to 0} (\sin \alpha)/\alpha = 1$, given $\varepsilon > 0$ we are able to find a $\delta > 0$, with $0 < \delta < 1$, such that $0 < |\alpha| < \delta$ implies that $|(\sin \alpha)/\alpha - 1| < \varepsilon$. If $0 < \|\mathbf{v}\| < \delta$, then $0 < \|\mathbf{v}\|^2 < \delta^2 < \delta$, and therefore

$$|f(\mathbf{v}) - 1| = \left| \frac{\sin \|\mathbf{v}\|^2}{\|\mathbf{v}\|^2} - 1 \right| < \varepsilon.$$

Thus, $\operatorname*{limit}_{\mathbf{v}\to(0,0)} f(\mathbf{v}) = 1$. If we plot $[\sin(x^2 + y^2)]/(x^2 + y^2)$ on a computer, we get a graph that is indeed well behaved near (0, 0) (Figure 2.2.17). ▲

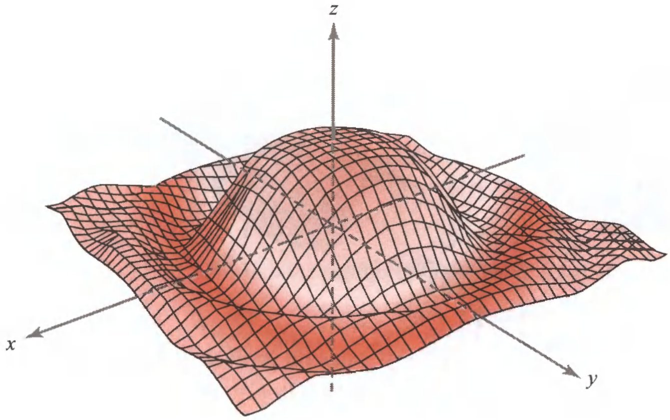


Figure 2.2.17 Graph of the function $f(x, y) = [\sin(x^2 + y^2)]/(x^2 + y^2)$.

EXAMPLE 14 Show that

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}} = 0.$$

SOLUTION We must show that $x^2/\sqrt{x^2+y^2}$ is small when (x, y) is close to the origin. To do this, we use the following inequality:

$$\begin{aligned} 0 \le \frac{x^2}{\sqrt{x^2+y^2}} &\le \frac{x^2+y^2}{\sqrt{x^2+y^2}} \qquad \text{(because } y^2 \ge 0\text{)} \\ &= \sqrt{x^2+y^2}. \end{aligned}$$

Given $\varepsilon > 0$, choose $\delta = \varepsilon$. Then $\|(x, y) - (0, 0)\| = \|(x, y)\| = \sqrt{x^2+y^2}$, and so $\|(x, y) - (0, 0)\| < \delta$ implies that

$$\left| \frac{x^2}{\sqrt{x^2+y^2}} - 0 \right| = \frac{x^2}{\sqrt{x^2+y^2}} \le \sqrt{x^2+y^2} = \|(x, y) - (0, 0)\| < \delta = \varepsilon.$$

Thus, the conditions of Theorem 6 have been fulfilled and the limit is verified. ▲

EXAMPLE 15 (a) Does

$$\lim_{(x,y)\to(0,0)} x^2/(x^2+y^2)$$

exist? [See Figure 2.2.18(a).]

(b) Prove that [see Figure 2.2.18(b)]

$$\lim_{(x,y)\to(0,0)} \frac{2x^2y}{x^2+y^2} = 0.$$

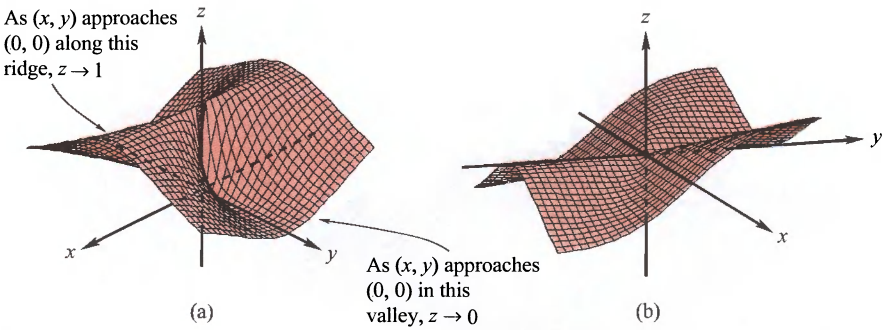


Figure 2.2.18 (a) The function $z = x^2/(x^2 + y^2)$ has no limit at (0, 0). (b) The function $z = (2x^2 y)/(x^2 + y^2)$ has limit 0 at (0, 0).

SOLUTION (a) If the limit exists, $x^2/(x^2 + y^2)$ should approach a definite value, say a, as (x, y) gets near (0, 0). In particular, if (x, y) approaches zero along any given path, then $x^2/(x^2 + y^2)$ should approach the limiting value a. If (x, y) approaches (0, 0) along the line $y = 0$, the limiting value is clearly 1 (just set $y = 0$ in the preceding expression to get $x^2/x^2 = 1$). If (x, y) approaches (0, 0) along the line $x = 0$, the limiting value is

$$\lim_{y\to 0} \frac{0^2}{0^2 + y^2} = 0 \neq 1.$$

Hence, $\lim\limits_{(x,y)\to(0,0)} x^2/(x^2 + y^2)$ does not exist.

(b) Note that

$$\left|\frac{2x^2 y}{x^2 + y^2}\right| \leq \left|\frac{2x^2 y}{x^2}\right| = 2|y|.$$

Thus, given $\varepsilon > 0$, choose $\delta = \varepsilon/2$; then $0 < \|(x, y) - (0, 0)\| = \sqrt{x^2 + y^2} < \delta$ implies $|y| < \delta$, and thus

$$\left|\frac{2x^2 y}{x^2 + y^2} - 0\right| < 2\delta = \varepsilon. \quad \blacktriangle$$

Using the ε-δ notation, we are led to the following reformulation of the definition of continuity.

THEOREM 7 Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ be given. Then f is continuous at $\mathbf{x}_0 \in A$ if and only if for every number $\varepsilon > 0$ there is a number $\delta > 0$ such that

$$\mathbf{x} \in A \qquad \text{and} \qquad \|\mathbf{x} - \mathbf{x}_0\| < \delta \qquad \text{implies} \qquad \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon.$$

The proof is almost immediate. Notice that in Theorem 6 we insisted that $0 < \|\mathbf{x} - \mathbf{x}_0\|$, that is, $\mathbf{x} \neq \mathbf{x}_0$. That is *not* imposed here; indeed, the conclusion of Theorem 7 is certainly valid when $\mathbf{x} = \mathbf{x}_0$, and so there is no need to exclude this case. Here we do care about the value of f at $\mathbf{x}_0$; we want f at nearby points to be close to *this* value.

EXERCISES

In the following exercises the reader may assume that the exponential, sine, and cosine functions are continuous and may freely use techniques from one-variable calculus, such as L'Hôpital's rule.

Show that the subsets of the plane in Exercises 1–4 are open:

1. $A = \{(x, y) \mid -1 < x < 1, -1 < y < 1\}$

2. $B = \{(x, y) \mid y > 0\}$

3. $C = \{(x, y) \mid 2 < x^2 + y^2 < 4\}$

4. $D = \{(x, y) \mid x \neq 0 \text{ and } y \neq 0\}$

5. Compute the limits:

(a) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,1)} x^3 y$

(b) $\displaystyle \operatorname*{limit}_{x\to 0} \frac{\cos x - 1}{x^2}$

(c) $\displaystyle \operatorname*{limit}_{h\to 0} \frac{e^h - 1}{h}$.

6. Compute the following limits:

(a) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,1)} e^x y$

(b) $\displaystyle \operatorname*{limit}_{x\to 0} \frac{\sin^2 x}{x}$

(c) $\displaystyle \operatorname*{limit}_{x\to 0} \frac{\sin^2 x}{x^2}$

7. Compute the following limits:

(a) $\displaystyle \operatorname*{limit}_{x\to 3}(x^2 - 3x + 5)$

(b) $\displaystyle \operatorname*{limit}_{x\to 0} \sin x$

(c) $\displaystyle \operatorname*{limit}_{h\to 0} \frac{(x+h)^2 - x^2}{h}$

8. Compute the following limits if they exist:

(a) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{(x+y)^2 - (x-y)^2}{xy}$

(b) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{\sin xy}{y}$

(c) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{x^3 - y^3}{x^2 + y^2}$

9. Compute the following limits if they exist:

(a) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{e^{xy} - 1}{y}$

(b) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{\cos(xy) - 1}{x^2 y^2}$

(c) $\displaystyle \operatorname*{limit}_{(x,y)\to(0,0)} \frac{xy}{x^2 + y^2 + 2}$

10. Compute the following limits, if they exist:

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{e^{xy}}{x+1}$

(b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\cos x - 1 - (x^2/2)}{x^4+y^4}$

(c) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{(x-y)^2}{x^2+y^2}$

11. Compute the following limits if they exist:

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin xy}{xy}$

(b) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{\sin(xyz)}{xyz}$

(c) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} f(x,y,z)$, where $f(x,y,z) = (x^2+3y^2)/(x+1)$.

12. Compute the following limits if they exist:

(a) $\displaystyle\lim_{x\to 0} \frac{\sin 2x - 2x}{x^3}$

(b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin 2x - 2x + y}{x^3+y}$

(c) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{2x^2 y \cos z}{x^2+y^2}$

13. Compute $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$, if it exists, for the following cases:

(a) $f\colon \mathbb{R}\to\mathbb{R}, x\mapsto |x|, x_0 = 1$

(b) $f\colon \mathbb{R}^n\to\mathbb{R}, \mathbf{x}\mapsto \|\mathbf{x}\|$, arbitrary $\mathbf{x}_0$

(c) $f\colon \mathbb{R}\to\mathbb{R}^2, x\mapsto (x^2, e^x), x_0 = 1$

(d) $f\colon \mathbb{R}^2\backslash\{(0,0)\}\to\mathbb{R}^2, (x,y)\mapsto (\sin(x-y), e^{x(y+1)} - x - 1)/\|(x,y)\|, \mathbf{x}_0 = (0,0)$.

14. Let $A\subset\mathbb{R}^2$ be the open unit disk $D_1(0,0)$ with the point $\mathbf{x}_0 = (1,0)$ added, and let $f\colon A\to\mathbb{R}, \mathbf{x}\mapsto f(\mathbf{x})$ be the constant function $f(\mathbf{x}) = 1$. Show that $\displaystyle\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = 1$.

15. If $f\colon\mathbb{R}^n\to\mathbb{R}$ and $g\colon\mathbb{R}^n\to\mathbb{R}$ are continuous, show that the functions

$$f^2 g\colon \mathbb{R}^n\to\mathbb{R}, \mathbf{x}\mapsto [f(\mathbf{x})]^2 g(\mathbf{x})$$

and

$$f^2+g\colon \mathbb{R}^n\to\mathbb{R}, x\mapsto [f(\mathbf{x})]^2 + g(\mathbf{x})$$

are continuous.

16. (a) Show that $f\colon\mathbb{R}\to\mathbb{R}, x\mapsto (1-x)^8 + \cos(1+x^3)$ is continuous.

(b) Show that the map $f\colon\mathbb{R}\to\mathbb{R}, x\mapsto x^2 e^x/(2-\sin x)$ is continuous.

17. (a) Can $[\sin(x+y)]/(x+y)$ be made continuous by suitably defining it at $(0,0)$?

(b) Can $xy/(x^2+y^2)$ be made continuous by suitably defining it at $(0,0)$?

(c) Prove that $f\colon\mathbb{R}^2\to\mathbb{R}, (x,y)\mapsto ye^x + \sin x + (xy)^4$ is continuous.

18. Using either ε's and δ's or spherical coordinates, show that

$$\lim_{(x,y,z)\to(0,0,0)} \frac{xyz}{x^2+y^2+z^2} = 0.$$

19. Use the ε-δ formulation of limits to prove that $x^2 \to 4$ as $x \to 2$. Give another proof using Theorem 3.

20. (a) Prove that for $\mathbf{x} \in \mathbb{R}^n$ and $s < t$, $D_s(\mathbf{x}) \subset D_t(\mathbf{x})$.
(b) Prove that if U and V are neighborhoods of $\mathbf{x} \in \mathbb{R}^n$, then so are $U \cap V$ and $U \cup V$.
(c) Prove that the boundary points of an open interval $(a, b) \subset \mathbb{R}$ are the points a and b.

21. Suppose $\mathbf{x}$ and $\mathbf{y}$ are in $\mathbb{R}^n$ and $\mathbf{x} \neq \mathbf{y}$. Show that there is a continuous function $f\colon \mathbb{R}^n \to \mathbb{R}$ with $f(\mathbf{x}) = 1$, $f(\mathbf{y}) = 0$, and $0 \le f(\mathbf{z}) \le 1$ for every $\mathbf{z}$ in $\mathbb{R}^n$.

22. Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}$ be given and let $\mathbf{x}_0$ be a boundary point of A. We say that $\lim_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = \infty$ if for every $N > 0$ there is a $\delta > 0$ such that $0 < \|\mathbf{x} - \mathbf{x}_0\| < \delta$ and $\mathbf{x} \in A$ implies $f(\mathbf{x}) > N$.

(a) Prove that $\lim_{x\to 1} (x-1)^{-2} = \infty$.
(b) Prove that $\lim_{x\to 0} 1/|x| = \infty$. Is it true that $\lim_{x\to 0} 1/x = \infty$?
(c) Prove that $\lim_{(x,y)\to(0,0)} 1/(x^2 + y^2) = \infty$.

23. Let $b \in \mathbb{R}$ and $f\colon \mathbb{R}\backslash[b] \to \mathbb{R}$ be a function. We write $\lim_{x\to b-} f(x) = L$ and say that L is the ***left-hand limit*** of f at b, if for every $\varepsilon > 0$, there is a $\delta > 0$ such that $x < b$ and $0 < |x - b| < \delta$ implies $|f(x) - L| < \varepsilon$.

(a) Formulate a definition of ***right-hand limit***, or $\lim_{x\to b+} f(x)$.
(b) Find $\lim_{x\to 0-} 1/(1 + e^{1/x})$ and $\lim_{x\to 0+} 1/(1 + e^{1/x})$.
(c) Sketch the graph of $1/(1 + e^{1/x})$.

24. Show that f is continuous at $\mathbf{x}_0$ if and only if

$$\lim_{\mathbf{x}\to\mathbf{x}_0} \|f(\mathbf{x}) - f(\mathbf{x}_0)\| = 0.$$

25. Let $f\colon A \subset \mathbb{R}^n \to \mathbb{R}^m$ satisfy $\|f(\mathbf{x}) - f(\mathbf{y})\| \le K\|\mathbf{x} - \mathbf{y}\|^\alpha$ for all $\mathbf{x}$ and $\mathbf{y}$ in A for positive constants K and α. Show that f is continuous. (Such functions are called ***Hölder-continuous*** or, if $\alpha = 1$, ***Lipschitz-continuous***.)

26. Show that $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is continuous at all points if and only if the inverse image of every open set is open.

27. (a) Find a specific number $\delta > 0$ such that if $|a| < \delta$, then $|a^3 + 3a^2 + a| < 1/100$.
(b) Find a specific number $\delta > 0$ such that if $x^2 + y^2 < \delta^2$, then

$$|x^2 + y^2 + 3xy + 180xy^5| < 1/10{,}000.$$

2.3 Differentiation

In Section 2.1 we considered a few methods for graphing functions. By these methods alone it may be impossible to compute enough information to grasp even the general features of a complicated function. From elementary calculus we know that the idea of the derivative can greatly aid us in this task; for example, it enables us to locate maxima and minima and to compute rates of change. The derivative also has many applications beyond this, as the student surely has discovered in elementary calculus.

Intuitively, we know from our work in Section 2.2 that a continuous function is one that has no "breaks" in its graph. A differentiable function from $\mathbb{R}^2$ to $\mathbb{R}$ ought to be such that not only are there no breaks in its graph, but there is a well-defined plane tangent to the graph at each point. Thus, there must not be any sharp folds, corners, or peaks in the graph (see Figure 2.3.1). In other words, the graph must be *smooth*.

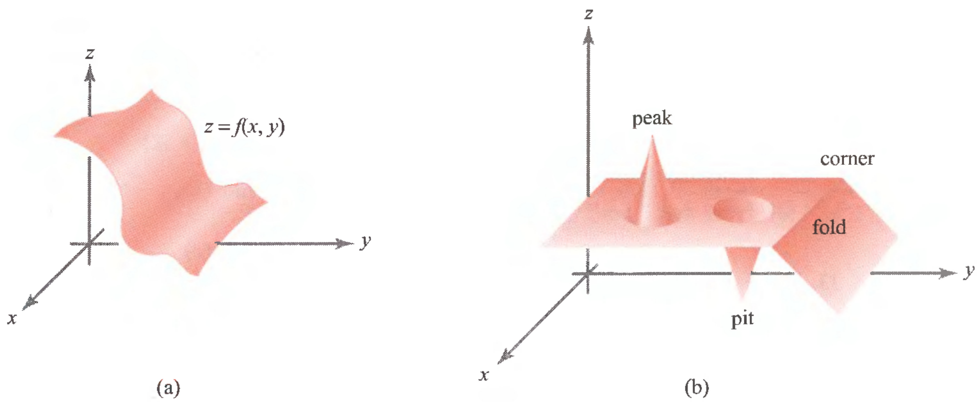


Figure 2.3.1 (a) A smooth graph and (b) a nonsmooth one.

Partial Derivatives

To make these ideas precise, we need a sound definition of what we mean by the phrase "$f(x_1, \ldots, x_n)$ is differentiable at $\mathbf{x} = (x_1, \ldots, x_n)$." Actually, this definition is not quite as simple as one might think. Toward this end, however, let us introduce the notion of the *partial derivative*. This notion relies only on our knowledge of one-variable calculus. (A quick review of the definition of the derivative in a one-variable calculus text might be advisable at this point.)

DEFINITION: Partial Derivatives Let $U \subset \mathbb{R}^n$ be an open set and suppose $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is a real-valued function. Then $\partial f/\partial x_1, \ldots, \partial f/\partial x_n$, the ***partial derivatives*** of f with respect to the first, second, ..., nth variable, are the real-valued functions of n variables, which, at the point $(x_1, \ldots, x_n) = \mathbf{x}$, are defined by

$$\frac{\partial f}{\partial x_j}(x_1, \ldots, x_n) = \lim_{h\to 0} \frac{f(x_1, x_2, \ldots, x_j + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

$$= \lim_{h\to 0} \frac{f(\mathbf{x} + h\mathbf{e}_j) - f(\mathbf{x})}{h}$$

if the limits exist, where $1 \le j \le n$ and $\mathbf{e}_j$ is the jth standard basis vector defined by $\mathbf{e}_j = (0, \dots, 1, \dots, 0)$, with 1 in the jth slot (see Section 1.5). The domain of the function $\partial f/\partial x_j$ is the set of $\mathbf{x} \in \mathbb{R}^n$ for which the limit exists.

In other words, $\partial f/\partial x_j$ is just the derivative of f with respect to the variable x_j, with the other variables held fixed. If $f\colon \mathbb{R}^3 \to \mathbb{R}$, we shall often use the notation $\partial f/\partial x$, $\partial f/\partial y$, $\partial f/\partial z$ in place of $\partial f/\partial x_1$, $\partial f/\partial x_2$, $\partial f/\partial x_3$. If $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$, then we can write

$$f(x_1, \dots, x_n) = (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)),$$

so that we can speak of the partial derivatives of each component; for example, $\partial f_m/\partial x_n$ is the partial derivative of the mth *component* with respect to x_n, the nth variable.

EXAMPLE 1 If $f(x, y) = x^2y + y^3$, find $\partial f/\partial x$ and $\partial f/\partial y$.

SOLUTION To find $\partial f/\partial x$ we hold y constant (think of it as some number, say 1) and differentiate only with respect to x; this yields

$$\frac{\partial f}{\partial x} = \frac{d(x^2y + y^3)}{dx} = 2xy.$$

Similarly, to find $\partial f/\partial y$ we hold x constant and differentiate only with respect to y:

$$\frac{\partial f}{\partial y} = \frac{d(x^2y + y^3)}{dy} = x^2 + 3y^2. \quad \blacktriangle$$

To indicate that a partial derivative is to be evaluated at a particular point, for example, at (x_0, y_0), we write

$$\frac{\partial f}{\partial x}(x_0, y_0) \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{x=x_0, y=y_0} \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{(x_0, y_0)}.$$

When we write $z = f(x, y)$ for the dependent variable, we sometimes write $\partial z/\partial x$ for $\partial f/\partial x$. Strictly speaking, this is an abuse of notation, but it is common practice to use these two notations interchangeably.

EXAMPLE 2 If $z = \cos xy + x \cos y = f(x, y)$, find the two partial derivatives $(\partial z/\partial x)(x_0, y_0)$ and $(\partial z/\partial y)(x_0, y_0)$.

SOLUTION First we fix y_0 and differentiate with respect to x, giving

$$\begin{aligned}\frac{\partial z}{\partial x}(x_0, y_0) &= \left.\frac{d(\cos xy_0 + x \cos y_0)}{dx}\right|_{x=x_0} \\ &= (-y_0 \sin xy_0 + \cos y_0)|_{x=x_0} \\ &= -y_0 \sin x_0y_0 + \cos y_0.\end{aligned}$$

Similarly, we fix x_0 and differentiate with respect to y to obtain

$$\begin{aligned}\frac{\partial z}{\partial y}(x_0, y_0) &= \left.\frac{d(\cos x_0 y + x_0 \cos y)}{dy}\right|_{y=y_0} \\ &= (-x_0 \sin x_0 y - x_0 \sin y)|_{y=y_0} \\ &= -x_0 \sin x_0 y_0 - x_0 \sin y_0. \quad \blacktriangle\end{aligned}$$

EXAMPLE 3 Find $\partial f/\partial x$ if $f(x, y) = xy/\sqrt{x^2 + y^2}$.

SOLUTION By the quotient rule,

$$\frac{\partial f}{\partial x} = \frac{y\sqrt{x^2+y^2} - xy(x/\sqrt{x^2+y^2})}{x^2+y^2} = \frac{y(x^2+y^2) - x^2 y}{(x^2+y^2)^{3/2}} = \frac{y^3}{(x^2+y^2)^{3/2}}. \quad \blacktriangle$$

A definition of differentiability that requires only the existence of partial derivatives turns out to be insufficient. Many standard results, such as the chain rule for functions of several variables would not follow, as Example 4 shows. Below, we shall see how to rectify this situation.

EXAMPLE 4 Let $f(x, y) = x^{1/3}y^{1/3}$. By definition,

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0,$$

and, similarly, $(\partial f/\partial y)(0, 0) = 0$ (these are not indeterminate forms!). It is necessary to use the original definition of partial derivatives, because the functions $x^{1/3}$ and $y^{1/3}$ are not themselves differentiable at 0. Suppose we restrict f to the line $y = x$ to get $f(x, x) = x^{2/3}$ (see Figure 2.3.2). We can view the substitution $y = x$ as the composition $f \circ g$ of the function g: $\mathbb{R} \to \mathbb{R}^2$, defined by $g(x) = (x, x)$, and f: $\mathbb{R}^2 \to \mathbb{R}$, defined by $f(x, y) = x^{1/3}y^{1/3}$.

Thus, the composite $f \circ g$ is given by $(f \circ g)(x) = x^{2/3}$. Each component of g is differentiable in x, and f has partial derivatives at $(0, 0)$, but *$f \circ g$ is not differentiable* at $x = 0$, in the sense of one-variable calculus. In other words, *the composition of f with g is not differentiable* in contrast to the calculus of functions of one variable, where the composition of differentiable functions *is* differentiable. Below, we shall give a definition of differentiability that has the pleasant consequence that the composition of differentiable functions *is* differentiable.

There is another reason for being dissatisfied with the mere existence of partial derivatives of $f(x, y) = x^{1/3}y^{1/3}$: There is no plane tangent, in any reasonable sense, to the graph at $(0, 0)$. The xy plane is tangent to the graph along the x and y axes because f has slope zero at $(0, 0)$ along these axes; that is, $\partial f/\partial x = 0$ and $\partial f/\partial y = 0$

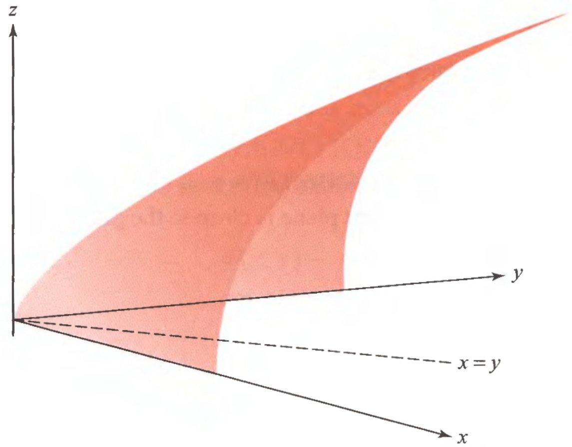


Figure 2.3.2 The portion of the graph of $x^{1/3}y^{1/3}$ in the first quadrant.

at (0, 0). Thus, if there is a tangent plane, it must be the xy plane. However, as is evident from Figure 2.3.2, the xy plane is not tangent to the graph in other directions, because the graph has a severe crinkle, and so the xy plane cannot be said to be tangent to the graph of f. ▲

The Linear Approximation

To "motivate" our definition of differentiability, let us compute what the equation of the plane tangent to the graph of $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto f(x, y)$ at (x_0, y_0) ought to be if f is smooth enough. In $\mathbb{R}^3$, a nonvertical plane has an equation of the form

$$z = ax + by + c.$$

If it is to be the plane tangent to the graph of f, the slopes along the x and y axes must be equal to $\partial f/\partial x$ and $\partial f/\partial y$, the rates of change of f with respect to x and y. Thus, $a = \partial f/\partial x$, $b = \partial f/\partial y$ [evaluated at (x_0, y_0)]. Finally, we may determine the constant c from the fact that $z = f(x_0, y_0)$ when $x = x_0$, $y = y_0$. Thus, we get the ***linear approximation***:

$$z = f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0), \tag{1}$$

which should be the equation of the plane tangent to the graph of f at (x_0, y_0), if f is "smooth enough" (see Figure 2.3.3).

Our definition of differentiability will mean in effect that the plane defined by the linear approximation (1) is a "good" approximation of f near (x_0, y_0). To get an idea of what one might mean by a good approximation, let us return for a moment to

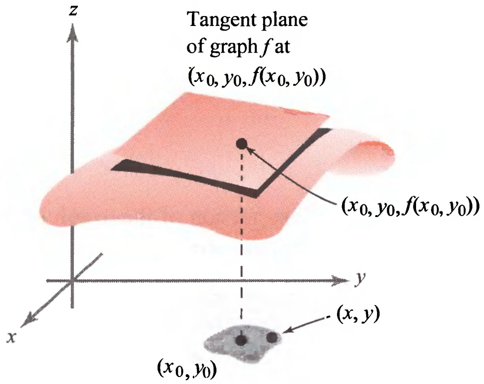


Figure 2.3.3 For points (x, y) near (x_0, y_0), the graph of the tangent plane is close to the graph of f.

one-variable calculus. If f is differentiable at a point x_0, then we know that

$$\operatorname*{limit}_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} = f'(x_0).$$

Let $x = x_0 + \Delta x$ and rewrite this as

$$\operatorname*{limit}_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Using the trivial limit $\operatorname*{limit}_{x \to x_0} f'(x_0) = f'(x_0)$, we can rewrite the preceding equation as

$$\operatorname*{limit}_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \operatorname*{limit}_{x \to x_0} f'(x_0);$$

that is,

$$\operatorname*{limit}_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right] = 0;$$

that is,

$$\operatorname*{limit}_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0.$$

Thus, the tangent line l through $(x_0, f(x_0))$ with slope $f'(x_0)$ is close to f in the sense that the difference between $f(x)$ and $l(x) = f(x_0) + f'(x_0)(x - x_0)$, the equation of the tangent line goes to zero *even* when divided by $x - x_0$ as x goes to x_0. This is the notion of a "good approximation" that we will adapt to functions of several variables, with the tangent line replaced by the tangent plane [see equation (1), given earlier].

Differentiability for Functions of Two Variables

Using the linear approximation, we are ready to define the notion of differentiability.

DEFINITION: Differentiable: Two Variables Let $f\colon \mathbb{R}^2 \to \mathbb{R}$. We say f is ***differentiable*** at (x_0, y_0), if $\partial f/\partial x$ and $\partial f/\partial y$ exist at (x_0, y_0) and if

$$\frac{f(x, y) - f(x_0, y_0) - \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) - \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0)}{\|(x, y) - (x_0, y_0)\|} \to 0 \tag{2}$$

as $(x, y) \to (x_0, y_0)$. This equation expresses what we mean by saying that

$$f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0)$$

is a ***good approximation*** to the function f.

It is not always easy to use this definition to see whether f is differentiable, but it will be easy to use another criterion, given shortly in Theorem 9.

Tangent Plane

We have used the informal notion of the plane tangent to the graph of a function to motivate our definition of differentiability. Now we are ready to adopt a formal definition of the tangent plane.

DEFINITION: Tangent Plane Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be differentiable at $\mathbf{x}_0 = (x_0, y_0)$. The plane in $\mathbb{R}^3$ defined by the equation

$$z = f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0),$$

is called the ***tangent plane*** of the graph of f at the point (x_0, y_0).

EXAMPLE 5 Compute the plane tangent to the graph of $z = x^2 + y^4 + e^{xy}$ at the point $(1, 0, 2)$.

SOLUTION Use formula (1), with $x_0 = 1$, $y_0 = 0$, and $z_0 = f(x_0, y_0) = 2$. The partial derivatives are

$$\frac{\partial z}{\partial x} = 2x + ye^{xy} \qquad \text{and} \qquad \frac{\partial z}{\partial y} = 4y^3 + xe^{xy}.$$

At (1, 0, 2), these partial derivatives are 2 and 1, respectively. Thus, by formula (1), the tangent plane is

$$z = 2(x-1) + 1(y-0) + 2, \qquad \text{that is,} \qquad z = 2x + y. \quad \blacktriangle$$

Let us write $\mathbf{D}f(x_0, y_0)$ for the row matrix

$$\left[\frac{\partial f}{\partial x}(x_0, y_0) \quad \frac{\partial f}{\partial y}(x_0, y_0)\right],$$

so that the definition of differentiability asserts that

$$\begin{aligned} f(x_0, y_0) &+ \mathbf{D}f(x_0, y_0)\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0) \end{aligned} \tag{3}$$

is our good approximation to f near (x_0, y_0). As earlier, "good" is taken in the sense that expression (3) differs from $f(x, y)$ by something small times $\sqrt{(x-x_0)^2 + (y-y_0)^2}$. We say that expression (3) is the ***best linear approximation*** to f near (x_0, y_0).

Differentiability: The General Case

Now we are ready to give a definition of differentiability for maps f of $\mathbb{R}^n$ to $\mathbb{R}^m$, using the preceding discussion as motivation. The derivative $\mathbf{D}f(\mathbf{x}_0)$ of $f = (f_1, \ldots, f_m)$ at a point $\mathbf{x}_0$ is a matrix $\mathbf{T}$ whose elements are $t_{ij} = \partial f_i/\partial x_j$ evaluated at $\mathbf{x}_0$.[2]

DEFINITION: Differentiable, *n* Variables, *m* Functions Let U be an open set in $\mathbb{R}^n$ and let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a given function. We say that f is ***differentiable*** at $\mathbf{x}_0 \in U$ if the partial derivatives of f exist at $\mathbf{x}_0$ and if

$$\underset{\mathbf{x}\to\mathbf{x}_0}{\text{limit}} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{T}(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0, \tag{4}$$

where $\mathbf{T} = \mathbf{D}f(\mathbf{x}_0)$ is the $m \times n$ matrix with matrix elements $\partial f_i/\partial x_j$ evaluated at $\mathbf{x}_0$ and $\mathbf{T}(\mathbf{x} - \mathbf{x}_0)$ means the product of $\mathbf{T}$ with $\mathbf{x} - \mathbf{x}_0$ (regarded as a column matrix). We call $\mathbf{T}$ the ***derivative*** of f at $\mathbf{x}_0$.

[2]It turns out that we need to postulate the existence of only *some* matrix giving the best linear approximation near $\mathbf{x}_0 \in \mathbb{R}^n$, because in fact this matrix is *necessarily* the matrix whose ijth entry is $\partial f_i/\partial x_j$ (see the Internet supplement for Chapter 2).

We shall always denote the derivative $\mathbf{T}$ of f at $\mathbf{x}_0$ by $\mathbf{D}f(\mathbf{x}_0)$, although in some books it is denoted $df(\mathbf{x}_0)$ and referred to as the ***differential*** of f. In the case where $m = 1$, the matrix $\mathbf{T}$ is just the row matrix

$$\left[\frac{\partial f}{\partial x_1}(\mathbf{x}_0) \quad \cdots \quad \frac{\partial f}{\partial x_n}(\mathbf{x}_0)\right].$$

(Sometimes, when there is danger of confusion, we separate the entries by commas.) Furthermore, setting $n = 2$ and putting the result back into equation (4), we see that conditions (2) and (4) do agree. Thus, if we let $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$, a real-valued function f of n variables is differentiable at a point $\mathbf{x}_0$ if

$$\underset{\mathbf{h}\to\mathbf{0}}{\text{limit}} \frac{1}{\|\mathbf{h}\|}\left|f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}_0)h_j\right| = 0,$$

because

$$\mathbf{Th} = \sum_{j=1}^{n} h_j \frac{\partial f}{\partial x_j}(\mathbf{x}_0).$$

For the general case of f mapping a subset of $\mathbb{R}^n$ to $\mathbb{R}^m$, the derivative is the $m \times n$ matrix given by

$$\mathbf{D}f(\mathbf{x}_0) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix},$$

where $\partial f_i/\partial x_j$ is evaluated at $\mathbf{x}_0$. The matrix $\mathbf{D}f(\mathbf{x}_0)$ is, appropriately, called the ***matrix of partial derivatives of f at $\mathbf{x}_0$***.

EXAMPLE 6 Calculate the matrices of partial derivatives for these functions.

(a) $f(x, y) = (e^{x+y} + y, y^2x)$

(b) $f(x, y) = (x^2 + \cos y, ye^x)$

(c) $f(x, y, z) = (ze^x, -ye^z)$

SOLUTION

(a) Here $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $f_1(x, y) = e^{x+y} + y$ and $f_2(x, y) = y^2x$. Hence, $\mathbf{D}f(x, y)$ is the 2×2 matrix

$$\mathbf{D}f(x, y) = \begin{bmatrix} e^{x+y} & e^{x+y} + 1 \\ y^2 & 2xy \end{bmatrix}.$$

(b) We have

$$\mathbf{D}f(x, y) = \begin{bmatrix} 2x & -\sin y \\ ye^x & e^x \end{bmatrix}.$$

(c) In this case,

$$\mathbf{D}f(x, y, z) = \begin{bmatrix} ze^x & 0 & e^x \\ 0 & -e^z & -ye^z \end{bmatrix}. \quad \blacktriangle$$

Gradients

For real-valued functions we use special terminology for the derivative.

DEFINITION: Gradient Consider the special case $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$. Here $\mathbf{D}f(\mathbf{x})$ is a $1 \times n$ matrix:

$$\mathbf{D}f(\mathbf{x}) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} & \cdots & \dfrac{\partial f}{\partial x_n} \end{bmatrix}.$$

We can form the corresponding vector $(\partial f/\partial x_1, \ldots, \partial f/\partial x_n)$, called the ***gradient*** of f and denoted by ∇f or grad f.

From the definition, we see that for $f\colon \mathbb{R}^3 \to \mathbb{R}$,

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

while for $f\colon \mathbb{R}^2 \to \mathbb{R}$,

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

The geometric significance of the gradient will be discussed in Section 2.6. In terms of inner products, we can write the derivative of f as

$$\mathbf{D}f(\mathbf{x})(\mathbf{h}) = \nabla f(\mathbf{x}) \cdot \mathbf{h}.$$

EXAMPLE 7 Let $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = xe^y$. Then

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (e^y, xe^y, 0). \quad \blacktriangle$$

EXAMPLE 8 If $f\colon \mathbb{R}^2 \to \mathbb{R}$ is given by $(x, y) \mapsto e^{xy} + \sin xy$, then

$$\begin{aligned}\nabla f(x, y) &= (ye^{xy} + y\cos xy)\mathbf{i} + (xe^{xy} + x\cos xy)\mathbf{j} \\ &= (e^{xy} + \cos xy)(y\mathbf{i} + x\mathbf{j}). \quad \blacktriangle\end{aligned}$$

In one-variable calculus it is shown that if f is differentiable, then f is continuous. We will state in Theorem 8 that this is also true for differentiable functions of several variables. As we know, there are plenty of functions of one variable that are continuous but not differentiable, such as $f(x) = |x|$. Before stating the result, let us give an example of a function of two variables whose *partial derivatives exist at a point, but which is not continuous at that point.*

EXAMPLE 9 Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(x, y) = \begin{cases} 1 & \text{if } x = 0 \text{ or if } y = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Because f is constant on the x and y axes, where it equals 1,

$$\frac{\partial f}{\partial x}(0, 0) = 0 \qquad \text{and} \qquad \frac{\partial f}{\partial y}(0, 0) = 0.$$

But f is not continuous at $(0, 0)$, because $\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist. ▲

Some Basic Theorems

The first of these basic theorems relates differentiability and continuity.

THEOREM 8 Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0 \in U$. Then f is continuous at $\mathbf{x}_0$.

This result is very reasonable, because "differentiability" means that there is enough smoothness to have a tangent plane, which is stronger than just being continuous. Consult the Internet supplement for Chapter 2 for the formal proof.

As we have seen, it is usually easy to tell when the partial derivatives of a function exist using what we know from one-variable calculus. However, the definition of differentiability looks somewhat complicated, and the required approximation condition in equation (4) may seem, and sometimes is, difficult to verify. Fortunately, there is a simple criterion, given in the following theorem, that tells us when a function is differentiable.

THEOREM 9 Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$. Suppose the partial derivatives $\partial f_i/\partial x_j$ of f all exist and are continuous in a neighborhood of a point $\mathbf{x} \in U$. Then f is differentiable at $\mathbf{x}$.

We give the proof in the Internet supplement for Chapter 2. Notice the following hierarchy:

$$\begin{tikzcd}[column sep=small]
 & \text{Theorem 9} \arrow[d] & & \substack{\text{Definition}\\ \text{of derivative}} \arrow[d] & \\
\text{Continuous partials} & \Rightarrow & \text{Differentiable} & \Rightarrow & \text{Partials exist}
\end{tikzcd}$$

Each converse statement, obtained by reversing an implication, is invalid. [For a counterexample to the converse of the first implication, use $f(x) = x^2 \sin(1/x)$, $f(0) = 0$; for the second, see Example 1 in the Internet supplement for Chapter 2 or use Example 4 in this section.]

A function whose partial derivatives exist and are continuous is said to be of ***class*** C^1. Thus, Theorem 9 says that *any C^1 function is differentiable.*

EXAMPLE 10 Let

$$f(x, y) = \frac{\cos x + e^{xy}}{x^2 + y^2}.$$

Show that f is differentiable at all points $(x, y) \neq (0, 0)$.

SOLUTION Observe that the partial derivatives

$$\frac{\partial f}{\partial x} = \frac{(x^2 + y^2)(ye^{xy} - \sin x) - 2x(\cos x + e^{xy})}{(x^2 + y^2)^2}$$

$$\frac{\partial f}{\partial y} = \frac{(x^2 + y^2)xe^{xy} - 2y(\cos x + e^{xy})}{(x^2 + y^2)^2}$$

are continuous except when $x = 0$ and $y = 0$ (by the results in Section 2.2). Thus, f is differentiable by Theorem 9. ▲

In the Internet supplement we show that $f(x, y) = xy/\sqrt{x^2 + y^2}$ [with $f(0, 0) = 0$] is continuous, has partial derivatives at (0, 0), yet is *not* differentiable there. See Figure 2.3.4. By Theorem 9, its partial derivatives cannot be continuous at (0, 0).

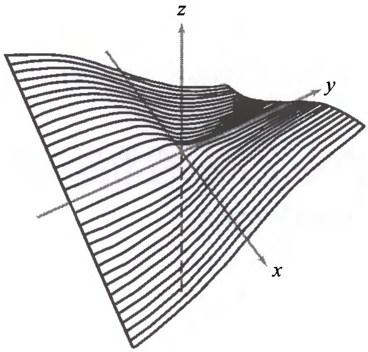


Figure 2.3.4 This function is not differentiable at (0, 0), because it is "crinkled."

EXERCISES

1. Find $\partial f/\partial x$, $\partial f/\partial y$ if

(a) $f(x, y) = xy$
(b) $f(x, y) = e^{xy}$
(c) $f(x, y) = x \cos x \cos y$
(d) $f(x, y) = (x^2 + y^2) \log (x^2 + y^2)$

2. Evaluate the partial derivatives $\partial z/\partial x$, $\partial z/\partial y$ for the given function at the indicated points.

(a) $z = \sqrt{a^2 - x^2 - y^2}$; $(0, 0)$, $(a/2, a/2)$
(b) $z = \log \sqrt{1 + xy}$; $(1, 2)$, $(0, 0)$
(c) $z = e^{ax} \cos (bx + y)$; $(2\pi/b, 0)$

3. In each case following, find the partial derivatives $\partial w/\partial x$, $\partial w/\partial y$.

(a) $w = xe^{x^2+y^2}$
(b) $w = \dfrac{x^2 + y^2}{x^2 - y^2}$
(c) $w = e^{xy} \log (x^2 + y^2)$
(d) $w = x/y$
(e) $w = \cos (ye^{xy}) \sin x$

4. Show that each of the following functions is differentiable at each point in its domain. Decide which of the functions are C^1.

(a) $f(x, y) = \dfrac{2xy}{(x^2 + y^2)^2}$
(b) $f(x, y) = \dfrac{x}{y} + \dfrac{y}{x}$
(c) $f(r, \theta) = \frac{1}{2} r \sin 2\theta$, $r > 0$
(d) $f(x, y) = \dfrac{xy}{\sqrt{x^2 + y^2}}$
(e) $f(x, y) = \dfrac{x^2 y}{x^4 + y^2}$

5. Find the equation of the plane tangent to the surface $z = x^2 + y^3$ at $(3, 1, 10)$.

6. Using the respective functions in Exercise 1, compute the plane tangent to the graphs at the indicated points.

(a) $(0, 0)$ (b) $(0, 1)$ (c) $(0, \pi)$ (d) $(0, 1)$

7. Compute the matrix of partial derivatives of the following functions:

(a) $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (x, y)$
(b) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (xe^y + \cos y, x, x + e^y)$
(c) $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, $f(x, y, z) = (x + e^z + y, yx^2)$
(d) $f\colon \mathbb{R}^2 \to \mathbb{R}^3$, $f(x, y) = (xye^{xy}, x \sin y, 5xy^2)$

8. Compute the matrix of partial derivatives of

(a) $f(x, y) = (e^x, \sin xy)$
(b) $f(x, y, z) = (x - y, y + z)$
(c) $f(x, y) = (x + y, x - y, xy)$
(d) $f(x, y, z) = (x + z, y - 5z, x - y)$

9. Where does the plane tangent to $z = e^{x-y}$ at $(1, 1, 1)$ meet the z axis?

10. Why should the graphs of $f(x, y) = x^2 + y^2$ and $g(x, y) = -x^2 - y^2 + xy^3$ be called "tangent" at $(0, 0)$?

11. Let $f(x, y) = e^{xy}$. Show that $x(\partial f/\partial x) = y(\partial f/\partial y)$.

12. Use the linear approximation to approximate a suitable function $f(x, y)$ and thereby estimate the following:

(a) $(0.99e^{0.02})^8$
(b) $(0.99)^3 + (2.01)^3 - 6(0.99)(2.01)$
(c) $\sqrt{(4.01)^2 + (3.98)^2 + (2.02)^2}$

13. Compute the gradients of the following functions:

(a) $f(x, y, z) = x \exp(-x^2 - y^2 - z^2)$ (Note that $\exp u = e^u$.)
(b) $f(x, y, z) = \dfrac{xyz}{x^2 + y^2 + z^2}$
(c) $f(x, y, z) = z^2 e^x \cos y$

14. Compute the tangent plane at $(1, 0, 1)$ for each of the functions in Exercise 13. [The solution to part (c) only is in the Study Guide.]

15. Find the equation of the tangent plane to $z = x^2 + 2y^3$ at $(1, 1, 3)$.

16. Calculate $\nabla h(1, 1, 1)$ if $h(x, y, z) = (x + z)e^{x-y}$.

17. Let $f(x, y, z) = x^2 + y^2 - z^2$. Calculate $\nabla f(0, 0, 1)$.

18. Evaluate the gradient of $f(x, y, z) = \log(x^2 + y^2 + z^2)$ at $(1, 0, 1)$.

19. Describe all Hölder-continuous functions with $\alpha > 1$ (see Exercise 25, Section 2.2). (HINT: What is the derivative of such a function?)

20. Suppose $f: \mathbb{R}^n \to \mathbb{R}^m$ is a linear map. What is the derivative of f?

2.4 Introduction to Paths and Curves

In this section, we introduce some of the basic geometry and computational methods for paths in the plane and space. This will be an important ingredient for the chain rule treated in the next section. We will return to paths with additional topics in Chapter 4.

Paths and Curves

One often thinks of a curve as a line drawn on paper, such as a straight line, a circle, or a sine curve. It is useful to think of a curve C mathematically as the set of values of a function that maps an interval of real numbers into the plane or space. We shall call such a map a ***path***. We usually denote a path by $\mathbf{c}$. The image C of the path then corresponds to the curve we see on paper (see Figure 2.4.1). Often we write t for the independent variable and imagine it to be *time*, so that $\mathbf{c}(t)$ is the position at time t of a moving particle, which ***traces out*** a curve as t varies. We also say $\mathbf{c}$ ***parametrizes*** C. Strictly speaking, we should distinguish between $\mathbf{c}(t)$ as a *point* in space and as a *vector* based at the origin.

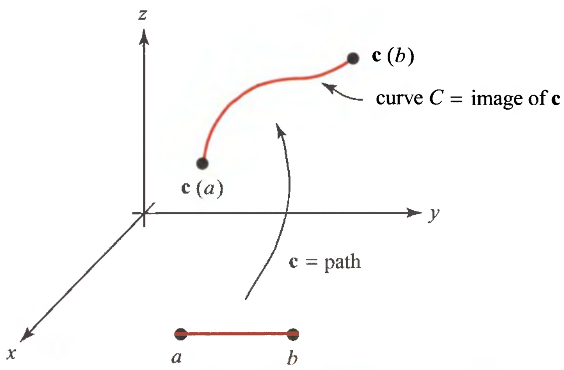


Figure 2.4.1 The map $\mathbf{c}$ is the path; its image C is the curve we "see."

EXAMPLE 1 The straight line L in $\mathbb{R}^3$ through the point (x_0, y_0, z_0) in the direction of vector $\mathbf{v}$ is the image of the path

$$\mathbf{c}(t) = (x_0, y_0, z_0) + t\mathbf{v}$$

for $t \in \mathbb{R}$ (see Figure 2.4.2). Thus, our notion of curve includes straight lines as special cases. ▲

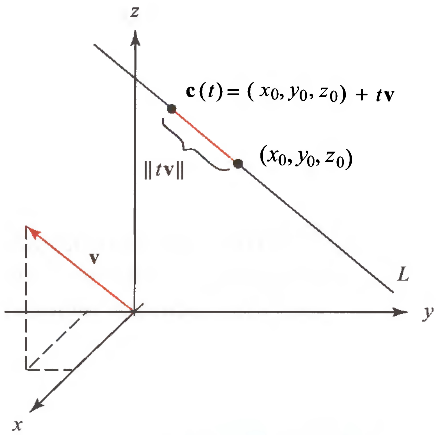


Figure 2.4.2 L is the straight line in space through (x_0, y_0, z_0) and in direction $\mathbf{v}$; its equation is $\mathbf{c}(t) = (x_0, y_0, z_0) + t\mathbf{v}$.

EXAMPLE 2 The unit circle C: $x^2 + y^2 = 1$ in the plane is the image of the path

$$\mathbf{c}: \mathbb{R} \to \mathbb{R}^2, \qquad \mathbf{c}(t) = (\cos t, \sin t), \qquad 0 \le t \le 2\pi,$$

(see Figure 2.4.3). The unit circle is also the image of the path $\tilde{\mathbf{c}}(t) = (\cos 2t, \sin 2t)$, $0 \le t \le \pi$. Thus, different paths may parametrize the same curve. ▲

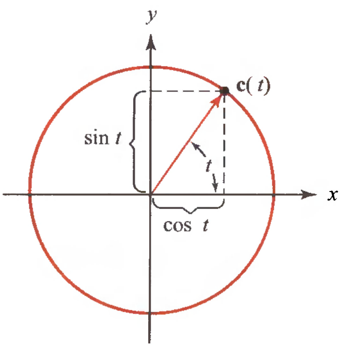


Figure 2.4.3 $\mathbf{c}(t) = (\cos t, \sin t)$ is a path whose image C is the unit circle.

Paths and Curves A ***path*** in $\mathbb{R}^n$ is a map $\mathbf{c}: [a, b] \to \mathbb{R}^n$; it is a ***path in the plane*** if $n = 2$ and a ***path in space*** if $n = 3$. The collection C of points $\mathbf{c}(t)$ as t varies in $[a, b]$ is called a ***curve***, and $\mathbf{c}(a)$ and $\mathbf{c}(b)$ are its ***endpoints***. The path $\mathbf{c}$ is said to ***parametrize*** the curve C. We also say $\mathbf{c}(t)$ ***traces out*** C as t varies.

If $\mathbf{c}$ is a path in $\mathbb{R}^3$, we can write $\mathbf{c}(t) = (x(t), y(t), z(t))$ and we call $x(t)$, $y(t)$, and $z(t)$ the ***component functions*** of $\mathbf{c}$. We form component functions similarly in $\mathbb{R}^2$ or, generally, in $\mathbb{R}^n$.

EXAMPLE 3 The path $\mathbf{c}(t) = (t, t^2)$ traces out a parabolic arc. This curve coincides with the graph $f(x) = x^2$ (see Figure 2.4.4). ▲

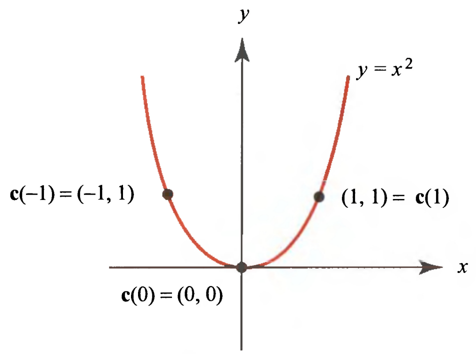


Figure 2.4.4 The image of $\mathbf{c}(t) = (t, t^2)$ is the parabola $y = x^2$.

EXAMPLE 4 A wheel of radius R rolls to the right along a straight line at speed v. Use vector methods to find the path $\mathbf{c}(t)$ of the point on the wheel that initially lies at a distance r below the center.

SOLUTION We place the wheel in the xy plane with its center initially at $(0, R)$, so that the position of the center at time t is given by the path $\mathbf{C}(t) = (vt, R)$. (Refer to Figure 2.4.5.)

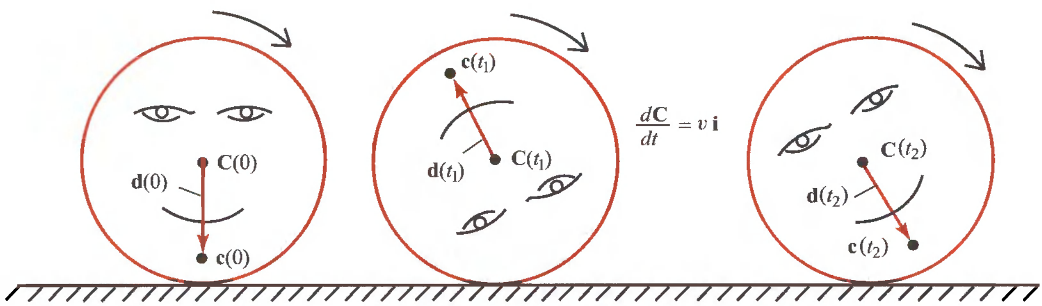


Figure 2.4.5 The vector $\mathbf{d}(t)$ points from the wheel's center, $\mathbf{C}(t)$, to the position $\mathbf{c}(t)$ of a point on the wheel and rotates in the clockwise direction while the wheel moves to the right.

The position of the point $\mathbf{c}(t)$ *relative to the center* is given by the vector $\mathbf{d}(t) = \mathbf{c}(t) - \mathbf{C}(t)$ that has the initial value $-r\mathbf{j}$ and rotates in the *clockwise* direction. The rate of rotation is such that the wheel makes a full rotation after the center has moved a distance $2\pi R$ (equal to the circumference of the wheel). This takes a time $2\pi R/v$, so the angular velocity $d\theta/dt$ of the wheel is v/R. Because the rotation is clockwise, the vector function $\mathbf{d}(t)$ is of the form

$$\mathbf{d}(t) = r\left(\cos\left[-\frac{v}{R}t + \theta\right]\mathbf{i} + \sin\left[-\frac{v}{R}t + \theta\right]\mathbf{j}\right)$$

for some initial angle θ. Because $\mathbf{d}(0) = -r\mathbf{j}$, we have $\cos\theta = 0$ and $\sin\theta = -1$, so $\theta = -\pi/2$, and hence

$$\mathbf{d}(t) = r\left(\cos\left[-\frac{v}{R}t - \frac{\pi}{2}\right]\mathbf{i} + \sin\left[-\frac{v}{R}t - \frac{\pi}{2}\right]\mathbf{j}\right).$$

Using $\cos(\varphi - \pi/2) = \sin\varphi$ and $\sin(\varphi - \pi/2) = -\cos\varphi$, along with $\cos(-\varphi) = \cos\varphi$ and $\sin(-\varphi) = -\sin\varphi$, we get

$$\mathbf{d}(t) = r\left(-\sin\frac{vt}{R}\mathbf{i} - \cos\frac{vt}{R}\mathbf{j}\right).$$

Finally, the path $\mathbf{c}(t)$ is given by adding the components of the vector function $\mathbf{d}(t)$ to the coordinates of the path $\mathbf{C}(t)$; the result is

$$\mathbf{c}(t) = \left(vt - r\sin\frac{vt}{R},\ R - r\cos\frac{vt}{R}\right).$$

In the special case $v = R = r = 1$, we get $\mathbf{c}(t) = (t - \sin t, 1 - \cos t)$. The image curve C of this path $\mathbf{c}$ is shown in Figure 2.4.6; it is called a ***cycloid***. ▲

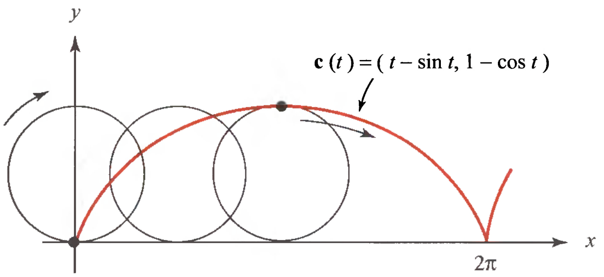


Figure 2.4.6 The curve traced by a point moving on the rim of a rolling circle is called a cycloid.

The preceding example considered the path of a point not necessarily on the rim of a wheel rolling along a straight line. When the wheel rolls on a circle, the resulting curve is called an ***epicycle***. These are the epicycles discussed in the Ptolemaic theory in the introduction. If the wheel is outside the circle and the point is on the rim, the curve is called an ***epicycloid***, and when the wheel is inside the circle it is a ***hypocycloid***. An example of the latter is shown in Figure 2.4.7.

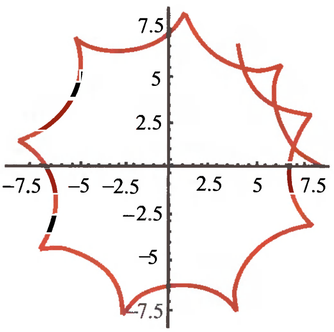


Figure 2.4.7 An example of a hypocycloid.

—Historical Note—

The French mathematician Blaise Pascal studied the cycloid in 1649 as a way of distracting himself at a time when he was suffering from a painful toothache. When the pain disappeared, he took it as a sign that God was not displeased with his thoughts. Pascal's results stimulated other mathematicians to investigate this curve, and subsequently numerous remarkable properties were found. One of these was discovered by the Dutchman Christian Huygens, who used it in the construction of a "perfect" pendulum clock.

Velocity and Tangents to Paths

If we think of $\mathbf{c}(t)$ as the curve traced out by a particle and t as time, it is reasonable to define the velocity vector as follows.

DEFINITION: Velocity Vector If $\mathbf{c}$ is a path and it is differentiable, we say $\mathbf{c}$ is a ***differentiable path***. The ***velocity*** of $\mathbf{c}$ at time t is defined by[3]

$$\mathbf{c}'(t) = \lim_{h \to 0} \frac{\mathbf{c}(t+h) - \mathbf{c}(t)}{h}.$$

We normally draw the vector $\mathbf{c}'(t)$ with its tail at the point $\mathbf{c}(t)$. The ***speed*** of the path $\mathbf{c}(t)$ is $s = \|\mathbf{c}'(t)\|$, the length of the velocity vector. If $\mathbf{c}(t) = (x(t), y(t))$ in $\mathbb{R}^2$, then

$$\mathbf{c}'(t) = (x'(t), y'(t)) = x'(t)\mathbf{i} + y'(t)\mathbf{j}$$

and if $\mathbf{c}(t) = (x(t), y(t), z(t))$ in $\mathbb{R}^3$, then

$$\mathbf{c}'(t) = (x'(t), y'(t), z'(t)) = x'(t)\mathbf{i} + y'(t)\mathbf{j} + z'(t)\mathbf{k}.$$

Here, $x'(t)$ is the one-variable derivative dx/dt. If we accept limits of vectors interpreted componentwise, the formulas for the velocity vector follow from the definition of the derivative. However, the limit can be interpreted in the sense of vectors as well. In Figure 2.4.8, we see that $[\mathbf{c}(t+h) - \mathbf{c}(t)]/h$ approaches the tangent to the path as $h \to 0$.

[3] If t lies at the endpoint of an interval, one should, as in one-variable calculus, take right- or left-handed limits.

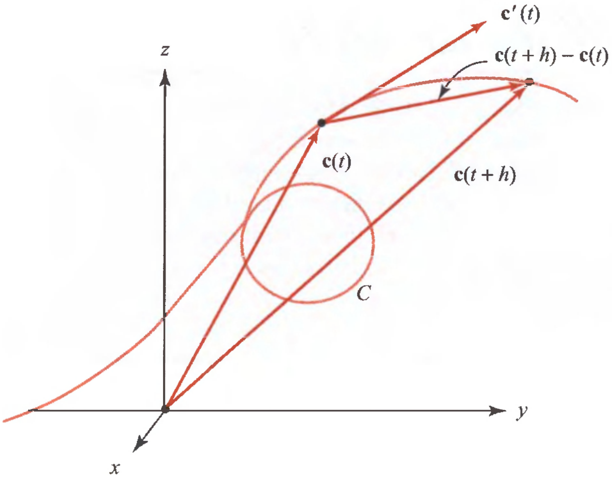


Figure 2.4.8 The vector $\mathbf{c}'(t)$ is tangent to the path $\mathbf{c}(t)$.

> **Tangent Vector** The velocity $\mathbf{c}'(t)$ is a vector ***tangent*** to the path $\mathbf{c}(t)$ at time t. If C is a curve traced out by $\mathbf{c}$ and if $\mathbf{c}'(t)$ is not equal to $\mathbf{0}$, then $\mathbf{c}'(t)$ is a vector tangent to the curve C at the point $\mathbf{c}(t)$.

If we think of the derivative $\mathbf{Dc}(t)$ as a matrix, it will be a column vector with the entries $x'(t)$, $y'(t)$, and $z'(t)$. Thus, the derivative here is consistent with our earlier notion.

EXAMPLE 5 Compute the tangent vector to the path $\mathbf{c}(t) = (t, t^2, e^t)$ at $t = 0$.

SOLUTION Here $\mathbf{c}'(t) = (1, 2t, e^t)$, and so at $t = 0$ we obtain the tangent vector $(1, 0, 1)$. ▲

EXAMPLE 6 Describe the path $\mathbf{c}(t) = (\cos t, \sin t, t)$. Find the velocity vector at the point on the image curve where $t = \pi/2$.

SOLUTION For a given t, the point $(\cos t, \sin t, 0)$ lies on the circle $x^2 + y^2 = 1$ in the xy plane. Therefore, the point $(\cos t, \sin t, t)$ lies t units above the point $(\cos t, \sin t, 0)$ if t is positive and $-t$ units below $(\cos t, \sin t, 0)$ if t is negative. As t increases, $(\cos t, \sin t, t)$ wraps around the cylinder $x^2 + y^2 = 1$ with the z coordinate increasing. The curve this traces out is called a ***helix***, which is depicted in Figure 2.4.9. At $t = \pi/2$, $\mathbf{c}'(\pi/2) = (-\sin\pi/2, \cos\pi/2, 1) = (-1, 0, 1) = -\mathbf{i} + \mathbf{k}$. ▲

EXAMPLE 7 The cycloidal path of a particle on the edge of a wheel of radius R with velocity v is given by $\mathbf{c}(t) = (vt - R\sin(vt/R), R - R\cos(vt/R))$. (See Example 4.) Find the velocity $\mathbf{c}'(t)$ of the particle as a function of t. When is the velocity zero? Is the velocity vector ever vertical?

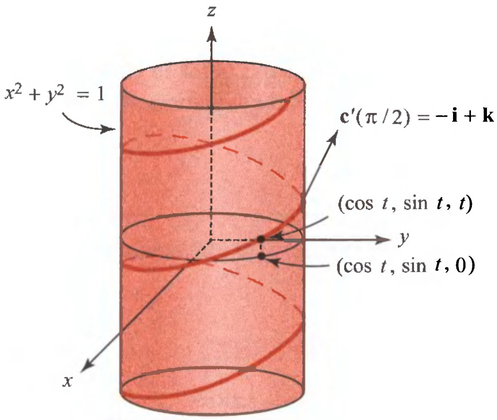


Figure 2.4.9 The helix $\mathbf{c}(t) = (\cos t, \sin t, t)$ wraps around the cylinder $x^2 + y^2 = 1$.

SOLUTION To find the velocity, we differentiate:

$$\mathbf{c}'(t) = \left(\frac{d}{dt}\left(vt - R\sin\frac{vt}{R}\right), \frac{d}{dt}\left(R - R\cos\frac{vt}{R}\right)\right)$$
$$= \left(v - v\cos\frac{vt}{R}, v\sin\frac{vt}{R}\right).$$

In vector notation, $\mathbf{c}'(t) = (v - v\cos(vt/R))\mathbf{i} + (v\sin(vt/R))\mathbf{j}$. The component in the direction of $\mathbf{i}$ is $v(1 - \cos(vt/R))$, which is zero whenever vt/R is an integer multiple of 2π. For such values of t, $\sin(vt/R)$ is zero as well, so the only times at which the velocity is zero are when $t = 2\pi nR/v$ for some integer n. At such times, $\mathbf{c}(t) = (2\pi nR, 0)$, so the moving point is touching the ground. These moments occur at time intervals of $2\pi R/v$ (more frequently for small wheels, as well as for rapidly rolling ones).

The velocity vector is never vertical, because the horizontal component vanishes only when the vertical one does as well. ▲

Figure 2.4.10 shows some velocity vectors superimposed on the cycloidal path of Figure 2.4.6.

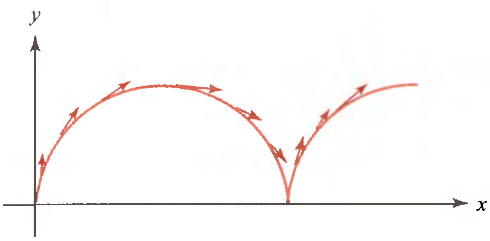


Figure 2.4.10 Velocity vectors for the curve traced out by a point on the rim of a rolling wheel.

Tangent Line

The tangent line to a path at a point is the line through the point in the direction of the tangent vector. Using the point-direction form of the equation of a line, we obtain the parametric equation for the tangent line.

Tangent Line to a Path If $\mathbf{c}(t)$ is a path, and if $\mathbf{c}'(t_0) \neq \mathbf{0}$, the equation of its ***tangent line*** at the point $\mathbf{c}(t_0)$ is

$$\boldsymbol{l}(t) = \mathbf{c}(t_0) + (t - t_0)\mathbf{c}'(t_0).$$

If C is the curve traced out by $\mathbf{c}$, then the line traced out by $\boldsymbol{l}$ is the tangent line to the curve C at $\mathbf{c}(t_0)$.

Notice that we have written the equation in such a way that $\boldsymbol{l}$ goes through the point $\mathbf{c}(t_0)$ at $t = t_0$ (rather than $t = 0$). See Figure 2.4.11.

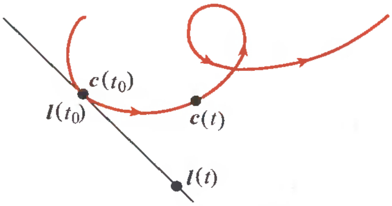


Figure 2.4.11 The tangent line to a path.

EXAMPLE 8 A path in $\mathbb{R}^3$ goes through the point $(3, 6, 5)$ at $t = 0$ with tangent vector $\mathbf{i} - \mathbf{j}$. Find the equation of the tangent line.

SOLUTION The equation of the tangent line is

$$\boldsymbol{l}(t) = (3, 6, 5) + t(\mathbf{i} - \mathbf{j}) = (3, 6, 5) + t(1, -1, 0) = (3 + t, 6 - t, 5).$$

In (x, y, z) coordinates, the tangent line is $x = 3 + t, y = 6 - t, z = 5$. ▲

Physically, we can interpret motion along the tangent line as the path that a particle on a curve would follow if it were set free at a certain moment.

EXAMPLE 9 Suppose that a particle follows the path $\mathbf{c}(t) = (e^t, e^{-t}, \cos t)$ until it flies off on a tangent at $t = 1$. Where is it at $t = 3$?

SOLUTION The velocity vector is $(e^t, -e^{-t}, -\sin t)$, which at $t = 1$ is the vector $(e, -1/e, -\sin 1)$. The particle is at $(e, 1/e, \cos 1)$ at $t = 1$. The equation of the tangent line is $\boldsymbol{l}(t) = (e, 1/e, \cos 1) + (t - 1)(e, -1/e, -\sin 1)$. At $t = 3$, the position on this line is

$$\begin{aligned}\boldsymbol{l}(3) &= \left(e, \frac{1}{e}, \cos 1\right) + 2\left(e, -\frac{1}{e}, -\sin 1\right) = \left(3e, -\frac{1}{e}, \cos 1 - 2\sin 1\right)\\ &\cong (8.155, -0.368, -1.143). \quad \blacktriangle\end{aligned}$$

EXERCISES

Sketch the curves that are the images of the paths in Exercises 1 to 4.

1. $x = \sin t$, $y = 4\cos t$, where $0 \le t \le 2\pi$

2. $x = 2\sin t$, $y = 4\cos t$, where $0 \le t \le 2\pi$

3. $\mathbf{c}(t) = (2t - 1, t + 2, t)$

4. $\mathbf{c}(t) = (-t, 2t, 1/t)$, where $1 \le t \le 3$

In Exercises 5 to 8, determine the velocity vector of the given path.

5. $\mathbf{c}(t) = 6t\mathbf{i} + 3t^2\mathbf{j} + t^3\mathbf{k}$

6. $\mathbf{c}(t) = (\sin 3t)\mathbf{i} + (\cos 3t)\mathbf{j} + 2t^{3/2}\mathbf{k}$

7. $\mathbf{r}(t) = (\cos^2 t, 3t - t^3, t)$

8. $\mathbf{r}(t) = (4e^t, 6t^4, \cos t)$

In Exercises 9 to 12, compute the tangent vector to the given path.

9. $\mathbf{c}(t) = (e^t, \cos t)$

10. $\mathbf{c}(t) = (3t^2, t^3)$

11. $\mathbf{c}(t) = (t\sin t, 4t)$

12. $\mathbf{c}(t) = (t^2, e^2)$

13. When is the velocity vector of a point on the rim of a rolling wheel *horizontal*? What is the speed at this point?

14. If the position of a particle in space is $(6t, 3t^2, t^3)$ at time t, what is its velocity vector at $t = 0$?

In Exercises 15 and 16, determine the equation of the tangent line to the given path at the specified value of t.

15. $(\sin 3t, \cos 3t, 2t^{5/2}); t = 1$

16. $(\cos^2 t, 3t - t^3, t); t = 0$

In Exercises 17 to 20, suppose that a particle following the given path $\mathbf{c}(t)$ flies off on a tangent at $t = t_0$. Compute the position of the particle at the given time t_1.

17. $\mathbf{c}(t) = (t^2, t^3 - 4t, 0)$, where $t_0 = 2, t_1 = 3$

18. $\mathbf{c}(t) = (e^t, e^{-t}, \cos t)$, where $t_0 = 1, t_1 = 2$

19. $\mathbf{c}(t) = (4e^t, 6t^4, \cos t)$, where $t_0 = 0, t_1 = 1$

20. $\mathbf{c}(t) = (\sin e^t, t, 4 - t^3)$, where $t_0 = 1, t_1 = 2$

2.5 Properties of the Derivative

In elementary calculus, we learn how to differentiate sums, products, quotients, and composite functions. We now generalize these ideas to functions of several variables, paying particular attention to the differentiation of composite functions. The rule for differentiating composites, called the *chain rule*, takes on a more profound form for functions of several variables than for those of one variable.

If f is a real-valued function of one variable, written as $z = f(y)$, and y is a function of x, written $y = g(x)$, then z becomes a function of x through substitution, namely, $z = f(g(x))$, and we have the familiar chain rule:

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = f'(g(x))g'(x).$$

If f is a real-valued function of three variables u, v, and w, written in the form $z = f(u, v, w)$, and the variables u, v, w are each functions of x, $u = g(x)$, $v = h(x)$, and $w = k(x)$, then by substituting $g(x)$, $h(x)$, and $k(x)$ for u, v, and w, we obtain z as a function of x: $z = f(g(x), h(x), k(x))$. The chain rule in this case reads:

$$\frac{dz}{dx} = \frac{\partial z}{\partial u}\frac{du}{dx} + \frac{\partial z}{\partial v}\frac{dv}{dx} + \frac{\partial z}{\partial w}\frac{dw}{dx}.$$

One of the goals of this section is to explain such formulas in detail.

Sums, Products, Quotients

These rules work just as they do in one-variable calculus.

THEOREM 10: Sums, Products, Quotients

(i) **Constant Multiple Rule.** Let $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0$ and let c be a real number. Then $h(\mathbf{x}) = cf(\mathbf{x})$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0) \qquad \text{(equality of matrices)}.$$

(ii) **Sum Rule.** Let $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $g: U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0$. Then $h(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = \mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0) \qquad \text{(sum of matrices)}.$$

(iii) **Product Rule.** Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$ and $g: U \subset \mathbb{R}^n \to \mathbb{R}$ be differentiable at $\mathbf{x}_0$ and let $h(\mathbf{x}) = g(\mathbf{x})f(\mathbf{x})$. Then $h: U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) + f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0).$$

(Note that each side of this equation is a $1 \times n$ matrix; a more general product rule is presented in Exercise 29 at the end of this section.)

(iv) **Quotient Rule.** With the same hypotheses as in rule (iii), let $h(\mathbf{x}) = f(\mathbf{x})/g(\mathbf{x})$ and suppose g is never zero on U. Then h is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}h(\mathbf{x}_0) = \frac{g(\mathbf{x}_0)\mathbf{D}f(\mathbf{x}_0) - f(\mathbf{x}_0)\mathbf{D}g(\mathbf{x}_0)}{[g(\mathbf{x}_0)]^2}.$$

PROOF The proofs of rules (i) through (iv) proceed almost exactly as in the one-variable case with a slight difference in notation. We shall prove rules (i) and (ii), leaving the proofs of rules (iii) and (iv) as Exercise 25.

(i) To show that $\mathbf{D}h(\mathbf{x}_0) = c\mathbf{D}f(\mathbf{x}_0)$, we must show that

$$\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} \frac{\|h(\mathbf{x}) - h(\mathbf{x}_0) - c\mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

that is, that

$$\operatorname*{limit}_{\mathbf{x}\to\mathbf{x}_0} \frac{\|cf(\mathbf{x}) - cf(\mathbf{x}_0) - c\mathbf{D}f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0,$$

[see equation (4) of Section 2.3]. This is certainly true, since f is differentiable and the constant c can be factored out [see Theorem 3(i), Section 2.2].

(ii) By the triangle inequality, we may write

$$\begin{aligned}
&\frac{\|h(\mathbf{x}) - h(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0) + \mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\
&= \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0) + g(\mathbf{x}) - g(\mathbf{x}_0) - [\mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} \\
&\leq \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - [\mathbf{D}f(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} + \frac{\|g(\mathbf{x}) - g(\mathbf{x}_0) - [\mathbf{D}g(\mathbf{x}_0)](\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|},
\end{aligned}$$

and each term approaches 0 as $\mathbf{x} \to \mathbf{x}_0$. Hence, rule (ii) holds. ■

EXAMPLE 1 Verify the formula for $\mathbf{D}h$ in rule (iv) of Theorem 10 with

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

SOLUTION Here

$$h(x, y, z) = \frac{x^2 + y^2 + z^2}{x^2 + 1},$$

so that by direct differentiation

$$\begin{aligned}
\mathbf{D}h(x, y, z) &= \left[\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z}\right] = \left[\frac{(x^2+1)2x - (x^2 + y^2 + z^2)2x}{(x^2+1)^2}, \frac{2y}{x^2+1}, \frac{2z}{x^2+1}\right] \\
&= \left[\frac{2x(1 - y^2 - z^2)}{(x^2+1)^2}, \frac{2y}{x^2+1}, \frac{2z}{x^2+1}\right].
\end{aligned}$$

By rule (iv), we get

$$\mathbf{D}h = \frac{g\mathbf{D}f - f\mathbf{D}g}{g^2} = \frac{(x^2+1)[2x, 2y, 2z] - (x^2 + y^2 + z^2)[2x, 0, 0]}{(x^2+1)^2},$$

which is the same as what we obtained directly. ▲

Chain Rule

As we mentioned earlier, it is in the differentiation of composite functions that we meet apparently substantial alterations of the formula from one-variable calculus. However, if we use the **D** notation, that is, matrix notation for derivatives, the chain rule for functions of several variables looks similar to the one-variable rule.

THEOREM 11: Chain Rule Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets. Let $g: U \subset \mathbb{R}^n \to \mathbb{R}^m$ and $f: V \subset \mathbb{R}^m \to \mathbb{R}^p$ be given functions such that g maps U into V, so that $f \circ g$ is defined. Suppose g is differentiable at $\mathbf{x}_0$ and f is differentiable at $\mathbf{y}_0 = g(\mathbf{x}_0)$. Then $f \circ g$ is differentiable at $\mathbf{x}_0$ and

$$\mathbf{D}(f \circ g)(\mathbf{x}_0) = \mathbf{D}f(\mathbf{y}_0)\mathbf{D}g(\mathbf{x}_0). \tag{1}$$

The right-hand side is the matrix product of $\mathbf{D}f(\mathbf{y}_0)$ with $\mathbf{D}g(\mathbf{x}_0)$.

We shall now give a proof of the chain rule *under the additional assumption that the partial derivatives of f are continuous*, building up to the general case by developing two special cases that are themselves important. (The complete proof of Theorem 11 without the additional assumption of continuity is given in the Internet supplement for Chapter 2.)

First Special Case of the Chain Rule

Suppose $\mathbf{c}: \mathbb{R} \to \mathbb{R}^3$ is a differentiable path and $f: \mathbb{R}^3 \to \mathbb{R}$. Let $h(t) = f(\mathbf{c}(t)) = f(x(t), y(t), z(t))$, where $\mathbf{c}(t) = (x(t), y(t), z(t))$. Then

$$\frac{dh}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}. \tag{2}$$

That is,

$$\frac{dh}{dt} = \nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t),$$

where $\mathbf{c}'(t) = (x'(t), y'(t), z'(t))$.

This is the special case of Theorem 11 in which we take $\mathbf{c} = g$ and f to be real-valued, and $m = 3$. Notice that

$$\nabla f(\mathbf{c}(t)) \cdot \mathbf{c}'(t) = \mathbf{D}f(\mathbf{c}(t))\mathbf{D}\mathbf{c}(t),$$

where the product on the left-hand side is the dot product of vectors, while the product on the right-hand side is matrix multiplication, and where we regard $\mathbf{D}f(\mathbf{c}(t))$ as a *row* matrix and $\mathbf{D}\mathbf{c}(t)$ as *a column* matrix. The vectors $\nabla f(\mathbf{c}(t))$ and $\mathbf{c}'(t)$ have the same components as their matrix equivalents; the notational change indicates the switch from matrices to vectors.

PROOF OF EQUATION (2). By definition,

$$\frac{dh}{dt}(t_0) = \underset{t \to t_0}{\text{limit}} \frac{h(t) - h(t_0)}{t - t_0}.$$

Adding and subtracting two terms, we write

$$\begin{aligned}\frac{h(t)-h(t_0)}{t-t_0} &= \frac{f(x(t),y(t),z(t))-f(x(t_0),y(t_0),z(t_0))}{t-t_0}\\ &= \frac{f(x(t),y(t),z(t))-f(x(t_0),y(t),z(t))}{t-t_0}\\ &\quad+\frac{f(x(t_0),y(t),z(t))-f(x(t_0),y(t_0),z(t))}{t-t_0}\\ &\quad+\frac{f(x(t_0),y(t_0),z(t))-f(x(t_0),y(t_0),z(t_0))}{t-t_0}.\end{aligned}$$

Now we invoke the *mean-value theorem* from one-variable calculus, which states: *If $g\colon [a,b]\to\mathbb{R}$ is continuous and is differentiable on the open interval (a,b), then there is a point c in (a,b) such that $g(b)-g(a)=g'(c)(b-a)$.* Applying this to f as a function of x, we can assert that for some c between x and x_0,

$$f(x,y,z)-f(x_0,y,z)=\left[\frac{\partial f}{\partial x}(c,y,z)\right](x-x_0).$$

In this way, we find that

$$\begin{aligned}\frac{h(t)-h(t_0)}{t-t_0} &= \left[\frac{\partial f}{\partial x}(c,y(t),z(t))\right]\frac{x(t)-x(t_0)}{t-t_0}+\left[\frac{\partial f}{\partial y}(x(t_0),d,z(t))\right]\frac{y(t)-y(t_0)}{t-t_0}\\ &\quad+\left[\frac{\partial f}{\partial z}(x(t_0),y(t_0),e)\right]\frac{z(t)-z(t_0)}{t-t_0},\end{aligned}$$

where c, d, and e lie between $x(t)$ and $x(t_0)$, between $y(t)$ and $y(t_0)$, and between $z(t)$ and $z(t_0)$, respectively. Taking the limit $t\to t_0$, using the continuity of the partials $\partial f/\partial x$, $\partial f/\partial y$, $\partial f/\partial z$, and the fact that c, d, and e converge to $x(t_0)$, $y(t_0)$, and $z(t_0)$, respectively, we obtain formula (2). ■

Second Special Case of the Chain Rule

Let $f\colon\mathbb{R}^3\to\mathbb{R}$ and let $g\colon\mathbb{R}^3\to\mathbb{R}^3$. Write

$$g(x,y,z)=(u(x,y,z),v(x,y,z),w(x,y,z))$$

and define $h\colon\mathbb{R}^3\to\mathbb{R}$ by setting

$$h(x,y,z)=f(u(x,y,z),v(x,y,z),w(x,y,z)).$$

In this case, the chain rule states that

$$\begin{bmatrix}\dfrac{\partial h}{\partial x} & \dfrac{\partial h}{\partial y} & \dfrac{\partial h}{\partial z}\end{bmatrix}=\begin{bmatrix}\dfrac{\partial f}{\partial u} & \dfrac{\partial f}{\partial v} & \dfrac{\partial f}{\partial w}\end{bmatrix}\begin{bmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z}\\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial z}\\ \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} & \dfrac{\partial w}{\partial z}\end{bmatrix}. \qquad (3)$$

In this special case, we have taken $n = m = 3$ and $p = 1$ for concreteness, and $U = \mathbb{R}^3$ and $V = \mathbb{R}^3$ for simplicity, and have written out the matrix product $[\mathbf{D}f(\mathbf{y}_0)][\mathbf{D}g(\mathbf{x}_0)]$ explicitly (with the arguments $\mathbf{x}_0$ and $\mathbf{y}_0$ suppressed in the matrices).

PROOF OF THE SECOND SPECIAL CASE OF THE CHAIN RULE. By definition, $\partial h/\partial x$ is obtained by differentiating h with respect to x, holding y and z fixed. But then $(u(x, y, z), v(x, y, z), w(x, y, z))$ may be regarded as a vector function of the single variable x. The first special case applies to this situation and, after the variables are renamed, gives

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x}. \tag{3$'$}$$

Similarly,

$$\frac{\partial h}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial y} \tag{3$''$}$$

and

$$\frac{\partial h}{\partial z} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial z} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial z} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial z}. \tag{3$'''$}$$

These equations are exactly what would be obtained by multiplying out the matrices in equation (3). ■

PROOF OF THEOREM 11. The general case in equation (1) may be proved in two steps. First, equation (2) is generalized to m variables; that is, for $f(x_1, \dots, x_m)$ and $\mathbf{c}(t) = (x_1(t), \dots, x_m(t))$, one has

$$\frac{dh}{dt} = \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}\frac{dx_i}{dt},$$

where $h(t) = f(x_1(t), \dots, x_m(t))$. Second, the result obtained in the first step is used to obtain the formula

$$\frac{\partial h_j}{\partial x_i} = \sum_{k=1}^{m} \frac{\partial f_j}{\partial y_k}\frac{\partial y_k}{\partial x_i},$$

where $f = (f_1, \dots, f_p)$ is a vector function of arguments $y_1, \dots, y_m$; $g(x_1, \dots, x_n) = (y_1(x_1, \dots, x_n), \dots, y_m(x_1, \dots, x_n))$; and $h_j(x_1, \dots, x_n) = f_j(y_1(x_1, \dots, x_n), \dots, y_m(x_1, \dots, x_n))$. (Using the letter y for both functions and arguments is an abuse of notation, but it can help one remember the formula.) This formula is equivalent to formula (1) after the matrices are multiplied out. ■

The pattern of the chain rule will become clear once the student has worked some additional examples. For instance,

$$\frac{\partial}{\partial x} f(u(x,y), v(x,y), w(x,y), z(x,y)) = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial x},$$

with a similar formula for $\partial f/\partial y$.

The chain rule can help us understand the relationship between the geometry of a mapping $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ and the geometry of curves in $\mathbb{R}^2$. (Similar statements may be made about $\mathbb{R}^3$ or, generally, $\mathbb{R}^n$.) If $\mathbf{c}(t)$ is a path in the plane, then as we saw in Section 2.4, $\mathbf{c}'(t)$ represents the tangent (or velocity) vector of the path $\mathbf{c}(t)$, and this tangent (or velocity) vector is thought of as beginning at $\mathbf{c}(t)$. Now let $\mathbf{p}(t) = f(\mathbf{c}(t))$, where $f\colon \mathbb{R}^2 \to \mathbb{R}^2$. The path $\mathbf{p}$ represents the image of the path $\mathbf{c}(t)$ under the mapping f. The tangent vector to $\mathbf{p}$ is given by the chain rule:

matrix multiplication

$$\mathbf{p}'(t) = \underbrace{\mathbf{D}f(\mathbf{c}(t))}_{\text{matrix}}\ \underbrace{\mathbf{c}'(t)}_{\text{column vector}}.$$

In other words, *the derivative matrix of f maps the tangent* (*or velocity*) *vector of a path* $\mathbf{c}$ *to the tangent* (*or velocity*) *vector of the corresponding image path* $\mathbf{p}$ (see Figure 2.5.1). Thus, points are mapped by f, while tangent vectors to curves are mapped by the derivative of f, evaluated at the base point of the tangent vector in the domain.

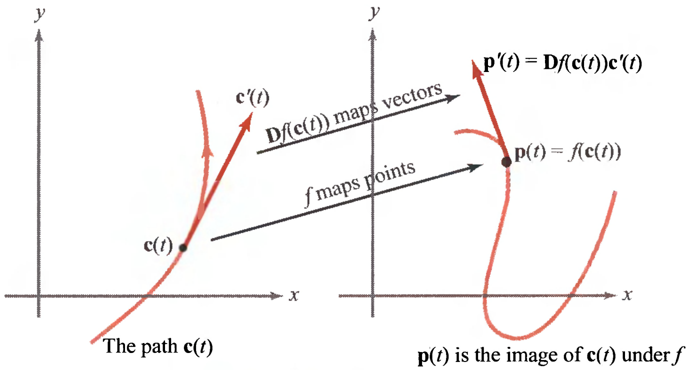


Figure 2.5.1 Tangent vectors are mapped by the derivative matrix.

EXAMPLE 2 Verify the chain rule in the form of formula (3′) for

$$f(u, v, w) = u^2 + v^2 - w,$$

where

$$u(x, y, z) = x^2 y, \qquad v(x, y, z) = y^2, \qquad w(x, y, z) = e^{-xz}.$$

SOLUTION Here

$$\begin{aligned} h(x, y, z) &= f(u(x, y, z), v(x, y, z), w(x, y, z)) \\ &= (x^2y)^2 + y^4 - e^{-xz} = x^4y^2 + y^4 - e^{-xz}. \end{aligned}$$

Thus, differentiating directly,

$$\frac{\partial h}{\partial x} = 4x^3y^2 + ze^{-xz}.$$

On the other hand, using the chain rule,

$$\begin{aligned} \frac{\partial h}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x} = 2u(2xy) + 2v \cdot 0 + (-1)(-ze^{-xz}) \\ &= (2x^2y)(2xy) + ze^{-xz}, \end{aligned}$$

which is the same as the preceding equation. ▲

EXAMPLE 3 Given $g(x, y) = (x^2 + 1, y^2)$ and $f(u, v) = (u + v, u, v^2)$, compute the derivative of $f \circ g$ at the point $(x, y) = (1, 1)$ using the chain rule.

SOLUTION The matrices of partial derivatives are

$$\mathbf{D}f(u, v) = \begin{bmatrix} \dfrac{\partial f_1}{\partial u} & \dfrac{\partial f_1}{\partial v} \\ \dfrac{\partial f_2}{\partial u} & \dfrac{\partial f_2}{\partial v} \\ \dfrac{\partial f_3}{\partial u} & \dfrac{\partial f_3}{\partial v} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2v \end{bmatrix} \quad \text{and} \quad \mathbf{D}g(x, y) = \begin{bmatrix} 2x & 0 \\ 0 & 2y \end{bmatrix}.$$

When $(x, y) = (1, 1)$, note that $g(x, y) = (u, v) = (2, 1)$. Hence,

$$\mathbf{D}(f \circ g)(1, 1) = \mathbf{D}f(2, 1)\mathbf{D}g(1, 1) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 0 \\ 0 & 4 \end{bmatrix}$$

is the required derivative. ▲

EXAMPLE 4 Let $f(x, y)$ be given and make the substitution $x = r\cos\theta$, $y = r\sin\theta$ (polar coordinates). Write a formula for $\partial f/\partial\theta$.

SOLUTION By the chain rule,

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta}$$

that is,

$$\frac{\partial f}{\partial \theta} = -r \sin\theta \frac{\partial f}{\partial x} + r \cos\theta \frac{\partial f}{\partial y}. \quad \blacktriangle$$

EXAMPLE 5 Let $f(x, y) = (\cos y + x^2, e^{x+y})$ and $g(u, v) = (e^{u^2}, u - \sin v)$. (a) Write a formula for $f \circ g$. (b) Calculate $\mathbf{D}(f \circ g)(0, 0)$ using the chain rule.

SOLUTION (a) We have

$$\begin{aligned}(f \circ g)(u, v) &= f(e^{u^2}, u - \sin v)\\ &= \left(\cos(u - \sin v) + e^{2u^2}, e^{e^{u^2} + u - \sin v}\right).\end{aligned}$$

(b) By the chain rule,

$$\mathbf{D}(f \circ g)(0, 0) = [\mathbf{D}f(g(0, 0))][\mathbf{D}g(0, 0)] = [\mathbf{D}f(1, 0)][\mathbf{D}g(0, 0)].$$

Now

$$\mathbf{D}g(0, 0) = \begin{bmatrix} 2ue^{u^2} & 0 \\ 1 & -\cos v \end{bmatrix}_{(u,v)=(0,0)} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$$

and

$$\mathbf{D}f(1, 0) = \begin{bmatrix} 2x & -\sin y \\ e^{x+y} & e^{x+y} \end{bmatrix}_{(x,y)=(1,0)} = \begin{bmatrix} 2 & 0 \\ e & e \end{bmatrix}.$$

[Remember that $\mathbf{D}f$ is evaluated at $g(0, 0)$, not at $(0, 0)$!] Thus,

$$\mathbf{D}(f \circ g)(0, 0) = \begin{bmatrix} 2 & 0 \\ e & e \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ e & -e \end{bmatrix}. \quad \blacktriangle$$

EXAMPLE 6 Let $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ be differentiable, with $f = (f_1, \ldots, f_m)$, and let $g(\mathbf{x}) = \sin[f(\mathbf{x}) \cdot f(\mathbf{x})]$. Compute $\mathbf{D}g(\mathbf{x})$.

SOLUTION By the chain rule, $\mathbf{D}g(\mathbf{x}) = \cos[f(\mathbf{x}) \cdot f(\mathbf{x})]\mathbf{D}h(\mathbf{x})$, where $h(\mathbf{x}) = [f(\mathbf{x}) \cdot f(\mathbf{x})] = f_1^2(\mathbf{x}) + \cdots + f_m^2(\mathbf{x})$. Then

$$\begin{aligned}\mathbf{D}h(x) &= \begin{bmatrix} \dfrac{\partial h}{\partial x_1} & \cdots & \dfrac{\partial h}{\partial x_n} \end{bmatrix}\\ &= \begin{bmatrix} 2f_1 \dfrac{\partial f_1}{\partial x_1} + \cdots + 2f_m \dfrac{\partial f_m}{\partial x_1} & \cdots & 2f_1 \dfrac{\partial f_1}{\partial x_n} + \cdots + 2f_m \dfrac{\partial f_m}{\partial x_n} \end{bmatrix},\end{aligned}$$

which can be written $2f(\mathbf{x})\mathbf{D}f(\mathbf{x})$, where we regard f as a row matrix,

$$f = [f_1 \quad \cdots \quad f_m] \qquad \text{and} \qquad \mathbf{D}f = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

Thus, $\mathbf{D}g(\mathbf{x}) = 2[\cos(f(\mathbf{x}) \cdot f(\mathbf{x}))]f(\mathbf{x})\mathbf{D}f(\mathbf{x})$. ▲

EXERCISES

1. If $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable, prove that $\mathbf{x} \mapsto f^2(\mathbf{x}) + 2f(\mathbf{x})$ is differentiable as well, and compute its derivative in terms of $\mathbf{D}f(\mathbf{x})$.

2. Prove that the following functions are differentiable, and find their derivatives at an arbitrary point:

(a) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto 2$
(b) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x + y$
(c) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto 2 + x + y$
(d) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + y^2$
(e) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto e^{xy}$
(f) $f\colon U \to \mathbb{R}, (x, y) \mapsto \sqrt{1 - x^2 - y^2}$, where $U = \{(x, y) \mid x^2 + y^2 < 1\}$
(g) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^4 - y^4$

3. Write out the chain rule for each of the following functions and justify your answer in each case using Theorem 11.

(a) $\partial h/\partial x$ where $h(x, y) = f(x, u(x, y))$
(b) dh/dx where $h(x) = f(x, u(x), v(x))$
(c) $\partial h/\partial x$ where $h(x, y, z) = f(u(x, y, z), v(x, y), w(x))$

4. Verify the chain rule for $\partial h/\partial x$, where $h(x, y) = f(u(x, y), v(x, y))$ and

$$f(u, v) = \frac{u^2 + v^2}{u^2 - v^2}, \qquad u(x, y) = e^{-x-y}, \qquad v(x, y) = e^{xy}.$$

5. Verify the first special case of the chain rule for the composition $f \circ \mathbf{c}$ in each of the cases:

(a) $f(x, y) = xy$, $\mathbf{c}(t) = (e^t, \cos t)$
(b) $f(x, y) = e^{xy}$, $\mathbf{c}(t) = (3t^2, t^3)$
(c) $f(x, y) = (x^2 + y^2) \log \sqrt{x^2 + y^2}$, $\mathbf{c}(t) = (e^t, e^{-t})$
(d) $f(x, y) = x \exp(x^2 + y^2)$, $\mathbf{c}(t) = (t, -t)$

6. What is the velocity vector for each path $\mathbf{c}(t)$ in Exercise 5? [The solution to part (b) only is in the Study Guide to this text.]

7. Let $f: \mathbb{R}^3 \to \mathbb{R}$ and $g: \mathbb{R}^3 \to \mathbb{R}$ be differentiable. Prove that

$$\nabla(fg) = f\nabla g + g\nabla f.$$

8. Let $f: \mathbb{R}^3 \to \mathbb{R}$ be differentiable. Making the substitution

$$x = \rho\cos\theta\sin\phi, \qquad y = \rho\sin\theta\sin\phi, \qquad z = \rho\cos\phi$$

(spherical coordinates) into $f(x, y, z)$, compute $\partial f/\partial\rho$, $\partial f/\partial\theta$, and $\partial f/\partial\phi$ in terms of $\partial f/\partial x$, $\partial f/\partial y$, and $\partial f/\partial z$.

9. Let $f(u, v) = (\tan(u-1) - e^v, u^2 - v^2)$ and $g(x, y) = (e^{x-y}, x - y)$. Calculate $f \circ g$ and $\mathbf{D}(f \circ g)(1, 1)$.

10. Let $f(u, v, w) = (e^{u-w}, \cos(v + u) + \sin(u + v + w))$ and $g(x, y) = (e^x, \cos(y - x), e^{-y})$. Calculate $f \circ g$ and $\mathbf{D}(f \circ g)(0, 0)$.

11. Find $(\partial/\partial s)(f \circ T)(1, 0)$, where $f(u, v) = \cos u \sin v$ and $T: \mathbb{R}^2 \to \mathbb{R}^2$ is defined by $T(s, t) = (\cos(t^2 s), \log\sqrt{1 + s^2})$.

12. Suppose that the temperature at the point (x, y, z) in space is $T(x, y, z) = x^2 + y^2 + z^2$. Let a particle follow the right-circular helix $\sigma(t) = (\cos t, \sin t, t)$ and let $T(t)$ be its temperature at time t.

(a) What is $T'(t)$?
(b) Find an approximate value for the temperature at $t = (\pi/2) + 0.01$.

13. Suppose that a duck is swimming in the circle $x = \cos t$, $y = \sin t$ and that the water temperature is given by the formula $T = x^2 e^y - xy^3$. Find dT/dt, the rate of change in temperature the duck might feel: (a) by the chain rule; (b) by expressing T in terms of t and differentiating.

14. Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be a linear mapping so that (by Exercise 20, Section 2.3) $\mathbf{D}f(\mathbf{x})$ is the matrix of f. Check the validity of the chain rule directly for linear mappings.

15. Let $f: \mathbb{R}^2 \to \mathbb{R}^2; (x, y) \mapsto (e^{x+y}, e^{x-y})$. Let $\mathbf{c}(t)$ be a path with $\mathbf{c}(0) = (0, 0)$ and $\mathbf{c}'(0) = (1, 1)$. What is the tangent vector to the image of $\mathbf{c}(t)$ under f at $t = 0$?

16. Let $f(x, y) = 1/\sqrt{x^2 + y^2}$. Compute $\nabla f(x, y)$.

17. (a) Let $y(x)$ be defined implicitly by $G(x, y(x)) = 0$, where G is a given function of two variables. Prove that if $y(x)$ and G are differentiable, then

$$\frac{dy}{dx} = -\frac{\partial G/\partial x}{\partial G/\partial y} \quad \text{if} \quad \frac{\partial G}{\partial y} \neq 0.$$

(b) Obtain a formula analogous to that in part (a) if y_1, y_2 are defined implicitly by

$$G_1(x, y_1(x), y_2(x)) = 0,$$
$$G_2(x, y_1(x), y_2(x)) = 0.$$

(c) Let y be defined implicitly by

$$x^2 + y^3 + e^y = 0.$$

Compute dy/dx in terms of x and y.

18. Thermodynamics texts[4] use the relationship

$$\left(\frac{\partial y}{\partial x}\right)\left(\frac{\partial z}{\partial y}\right)\left(\frac{\partial x}{\partial z}\right) = -1.$$

Explain the meaning of this equation and prove that it is true. [HINT: Start with a relationship $F(x, y, z) = 0$ that defines $x = f(y, z)$, $y = g(x, z)$, and $z = h(x, y)$ and differentiate implicitly.]

19. Dieterici's equation of state for a gas is

$$P(V - b)e^{a/RVT} = RT,$$

where a, b, and R are constants. Regard volume V as a function of temperature T and pressure P and prove that

$$\frac{\partial V}{\partial T} = \left(R + \frac{a}{TV}\right)\Big/\left(\frac{RT}{V - b} - \frac{a}{V^2}\right).$$

20. This exercise gives another example of the fact that the chain rule is not applicable if f is not differentiable. Consider the function

$$f(x, y) = \begin{cases} \dfrac{xy^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

Show that

(a) $\partial f/\partial x$ and $\partial f/\partial y$ exist at $(0, 0)$.

(b) If $\mathbf{g}(t) = (at, bt)$ for constants a and b, then $f \circ \mathbf{g}$ is differentiable and $(f \circ \mathbf{g})'(0) = ab^2/(a^2 + b^2)$, but $\nabla f(0, 0) \cdot \mathbf{g}'(0) = 0$.

21. Prove that if $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at $\mathbf{x}_0 \in U$, there is a neighborhood V of $\mathbf{0} \in \mathbb{R}^n$ and a function $R_1\colon V \to \mathbb{R}$ such that for all $\mathbf{h} \in V$, we have $\mathbf{x}_0 + \mathbf{h} \in U$,

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + [\mathbf{D}f(\mathbf{x}_0)]\mathbf{h} + R_1(\mathbf{h})$$

---

[4]See S. M. Binder, "Mathematical Methods in Elementary Thermodynamics," *J. Chem. Educ.* 43 (1966): 85–92. A proper understanding of partial differentiation can be of significant use in applications; for example, see M. Feinberg, "Constitutive Equation for Ideal Gas Mixtures and Ideal Solutions as Consequences of Simple Postulates," *Chem. Eng. Sci.* 32 (1977): 75–78.

and

$$\frac{R_1(\mathbf{h})}{\|\mathbf{h}\|} \to 0 \quad \text{as} \quad \mathbf{h} \to \mathbf{0}.$$

22. Suppose $\mathbf{x}_0 \in \mathbb{R}^n$ and $0 \leq r_1 < r_2$. Show that there is a C^1 function $f: \mathbb{R}^n \to \mathbb{R}$ such that $f(\mathbf{x}) = 0$ for $\|\mathbf{x} - \mathbf{x}_0\| \geq r_2$; $0 < f(\mathbf{x}) < 1$ for $r_1 < \|\mathbf{x} - \mathbf{x}_0\| < r_2$; and $f(\mathbf{x}) = 1$ for $\|\mathbf{x} - \mathbf{x}_0\| \leq r_1$. [HINT: Apply a cubic polynomial with $g(r_1^2) = 1$ and $g(r_2^2) = g'(r_2^2) = g'(r_1^2) = 0$ to $\|\mathbf{x} - \mathbf{x}_0\|^2$ when $r_1 < \|\mathbf{x} - \mathbf{x}_0\| < r_2$.]

23. Find a C^1 mapping $f: \mathbb{R}^3 \to \mathbb{R}^3$ that takes the vector $\mathbf{i} + \mathbf{j} + \mathbf{k}$ emanating from the origin to $\mathbf{i} - \mathbf{j}$ emanating from $(1, 1, 0)$ and takes $\mathbf{k}$ emanating from $(1, 1, 0)$ to $\mathbf{k} - \mathbf{i}$ emanating from the origin.

24. What is wrong with the following argument? Suppose $w = f(x, y, z)$ and $z = g(x, y)$. By the chain rule,

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

Hence, $0 = (\partial w/\partial z)(\partial z/\partial x)$, and so $\partial w/\partial z = 0$ or $\partial z/\partial x = 0$, which is, in general, absurd.

25. Prove rules (iii) and (iv) of Theorem 10. (HINT: Use the same addition and subtraction tricks as in the one-variable case and Theorem 8.)

26. Show that $h: \mathbb{R}^n \to \mathbb{R}^m$ is differentiable if and only if each of the m components $h_i: \mathbb{R}^n \to \mathbb{R}$ is differentiable. (HINT: Use the coordinate projection function and the chain rule for one implication and consider

$$\left[\frac{\|h(\mathbf{x}) - h(\mathbf{x}_0) - \mathbf{D}h(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|}\right]^2 = \frac{\sum_{i=1}^m [h_i(\mathbf{x}) - h_i(\mathbf{x}_0)\mathbf{D}h_i(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)]^2}{\|\mathbf{x} - \mathbf{x}_0\|^2}$$

to obtain the other.)

27. Use the chain rule and differentiation under the integral sign, namely,

$$\frac{d}{dx}\int_a^b f(x, y)\, dy = \int_a^b \frac{\partial f}{\partial x}(x, y)\, dy,$$

to show that

$$\frac{d}{dx}\int_0^x f(x, y)\, dy = f(x, x) + \int_0^x \frac{\partial f}{\partial x}(x, y)\, dy.$$

28. For what integers $p > 0$ is

$$f(x) = \begin{cases} x^p \sin(1/x) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

differentiable? For what p is the derivative continuous?

29. Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ and $g\colon \mathbb{R}^n \to \mathbb{R}^m$ are differentiable. Show that the product function $h(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ is differentiable and that if $\mathbf{x}_0$ and $\mathbf{y}$ are in $\mathbb{R}^n$, then $[\mathbf{D}h(\mathbf{x}_0)]\mathbf{y} = f(\mathbf{x}_0)\{[\mathbf{D}g(\mathbf{x}_0)]\mathbf{y}\} + \{[\mathbf{D}f(\mathbf{x}_0)]\mathbf{y}\}g(\mathbf{x}_0)$.

2.6 Gradients and Directional Derivatives

In Section 2.1 we studied the graphs of real-valued functions. Now we take up this study again, using the methods of calculus. Specifically, gradients will be used to obtain a formula for the plane tangent to a level surface.

Gradients in $\mathbb{R}^3$

Let us recall the definition.

DEFINITION: The Gradient If $f\colon U \subset \mathbb{R}^3 \to \mathbb{R}$ is differentiable, the ***gradient*** of f at (x, y, z) is the vector in space given by

$$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right).$$

This vector is also denoted $\nabla f(x, y, z)$. Thus, ∇f is just the matrix of the derivative $\mathbf{D}f$, written as a vector.

EXAMPLE 1 Let $f(x, y, z) = \sqrt{x^2 + y^2 + z^2} = r$, the distance from $\mathbf{0}$ to (x, y, z). Then

$$\begin{aligned}\nabla f(x, y, z) &= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right)\\ &= \left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = \frac{\mathbf{r}}{r},\end{aligned}$$

where $\mathbf{r}$ is the point (x, y, z). Thus, ∇f is the unit vector in the direction of (x, y, z). ▲

EXAMPLE 2 If $f(x, y, z) = xy + z$, then

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = (y, x, 1). \quad ▲$$

Suppose $f\colon \mathbb{R}^3 \to \mathbb{R}$ is a real-valued function. Let $\mathbf{v}$ and $\mathbf{x} \in \mathbb{R}^3$ be fixed vectors and consider the function from $\mathbb{R}$ to $\mathbb{R}$ defined by $t \mapsto f(\mathbf{x} + t\mathbf{v})$. The set of points of the form $\mathbf{x} + t\mathbf{v}$, $t \in \mathbb{R}$, is the line L through the point $\mathbf{x}$ parallel to the vector $\mathbf{v}$ (see Figure 2.6.1).

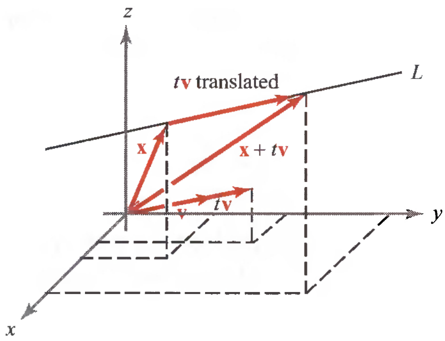


Figure 2.6.1 The equation of L is $\mathbf{l}(t) = \mathbf{x} + t\mathbf{v}$.

Directional Derivatives

The function $t \mapsto f(\mathbf{x} + t\mathbf{v})$ represents the function f restricted to the line L. For example, if a bird flies along this line with velocity $\mathbf{v}$ so that $\mathbf{x} + t\mathbf{v}$ is its position at time t, and if f represents the temperature as a function of position, then $f(\mathbf{x} + t\mathbf{v})$ is the temperature at time t. We may ask: How fast are the values of f changing along the line L at the point $\mathbf{x}$? Because the rate of change of a function is given by a derivative, we could say that the answer to this question is the value of the derivative of this function of t at $t = 0$ (when $t = 0$, $\mathbf{x} + t\mathbf{v}$ reduces to $\mathbf{x}$). This would be the derivative of f at the point $\mathbf{x}$ in the direction of L, that is, of $\mathbf{v}$. We can formalize this concept as follows.

DEFINITION: Directional Derivatives If $f: \mathbb{R}^3 \to \mathbb{R}$, the ***directional derivative*** of f at $\mathbf{x}$ along the vector $\mathbf{v}$ is given by

$$\left.\frac{d}{dt} f(\mathbf{x} + t\mathbf{v})\right|_{t=0}$$

if this exists.

In the definition of a directional derivative, we normally choose $\mathbf{v}$ to be a *unit* vector. In this case we are moving in the direction $\mathbf{v}$ with unit speed and we refer to $\frac{d}{dt} f(\mathbf{x} + t\mathbf{v})|_{t=0}$ as the ***directional derivative of f in the direction*** $\mathbf{v}$.

We now elaborate on why a *unit vector* is chosen in the definition of the directional derivative. Suppose that f measures the temperature in degrees and that we are interested in how fast the temperature changes as we move in a particular direction. If we are measuring distance in meters, then the rate of change of temperature will be measured in degrees per meter. Suppose, for simplicity, that the temperature is changing at a constant rate—say, two degrees per meter—as we move in a given direction $\mathbf{v}$ starting at $\mathbf{x}$. Thus, when we go one meter ahead, the temperature changes by two degrees. That is,

$$f(\mathbf{x}+\mathbf{v}) - f(\mathbf{x}) = 2$$

Such a relation is going to hold only when $\mathbf{v}$ is a unit vector, reflecting the fact that we are going ahead by *one* meter. More generally, the definition of the directional derivative is going to truly measure only the rate of change of f with respect to distance along a line in a given direction if $\mathbf{v}$ is a unit vector.

From the definition, we can see that the directional derivative can also be defined by the formula

$$\operatorname*{limit}_{h\to 0} \frac{f(\mathbf{x}+h\mathbf{v}) - f(\mathbf{x})}{h}.$$

THEOREM 12 If $f\colon \mathbb{R}^3 \to \mathbb{R}$ is differentiable, then all directional derivatives exist. The directional derivative at $\mathbf{x}$ in the direction $\mathbf{v}$ is given by

$$\mathbf{D}f(\mathbf{x})\mathbf{v} = \operatorname{grad} f(\mathbf{x})\cdot\mathbf{v} = \nabla f(\mathbf{x})\cdot\mathbf{v} = \left[\frac{\partial f}{\partial x}(\mathbf{x})\right]v_1 + \left[\frac{\partial f}{\partial y}(\mathbf{x})\right]v_2 + \left[\frac{\partial f}{\partial z}(\mathbf{x})\right]v_3,$$

where $\mathbf{v} = (v_1, v_2, v_3)$.

PROOF Let $\mathbf{c}(t) = \mathbf{x} + t\mathbf{v}$, so that $f(\mathbf{x}+t\mathbf{v}) = f(\mathbf{c}(t))$. By the first special case of the chain rule, $(d/dt)f(\mathbf{c}(t)) = \nabla f(\mathbf{c}(t))\cdot\mathbf{c}'(t)$. However, $\mathbf{c}(0) = \mathbf{x}$ and $\mathbf{c}'(0) = \mathbf{v}$, and so

$$\left.\frac{d}{dt} f(\mathbf{x}+t\mathbf{v})\right|_{t=0} = \nabla f(\mathbf{x})\cdot\mathbf{v},$$

as we were required to prove. ■

Notice that one does not have to use straight lines when computing the rate of change of f in a specific direction $\mathbf{v}$. Indeed, for a general path $\mathbf{c}(t)$ with $\mathbf{c}(0) = \mathbf{x}$ and $\mathbf{c}'(0) = \mathbf{v}$, we have from the chain rule,

$$\left.\frac{d}{dt} f(\mathbf{c}(t))\right|_{t=0} = \left.\nabla f(\mathbf{c}(t))\cdot\mathbf{c}'(t)\right|_{t=0} = \nabla f(\mathbf{x})\cdot\mathbf{v}.$$

EXAMPLE 3 Let $f(x, y, z) = x^2 e^{-yz}$. Compute the rate of change of f in the direction of the unit vector

$$\mathbf{v} = \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) \quad \text{at the point} \quad (1, 0, 0).$$

SOLUTION The required rate of change is, using Theorem 12,

$$\nabla f \cdot \mathbf{v} = (2xe^{-yz}, -x^2ze^{-yz}, -x^2ye^{-yz}) \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right),$$

which, at the point (1, 0, 0), becomes

$$(2, 0, 0) \cdot \left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right) = \frac{2}{\sqrt{3}}. \quad \blacktriangle$$

Directions of Fastest Increase

From Theorem 12 we can also obtain the geometrical significance of the gradient:

THEOREM 13 Assume $\nabla f(\mathbf{x}) \neq \mathbf{0}$. Then $\nabla f(\mathbf{x})$ points in the direction along which f is increasing the fastest.

PROOF If $\mathbf{n}$ is a unit vector, the rate of change of f in direction $\mathbf{n}$ is given by $\nabla f(\mathbf{x}) \cdot \mathbf{n} = \|\nabla f(\mathbf{x})\| \cos\theta$, where θ is the angle between $\mathbf{n}$ and $\nabla f(\mathbf{x})$. This is maximum when $\theta = 0$; that is, when $\mathbf{n}$ and ∇f are parallel. [If $\nabla f(\mathbf{x}) = \mathbf{0}$ this rate of change is 0 for any $\mathbf{n}$.] ■

In other words, if one wishes to move in a direction in which f will increase most quickly, one should proceed in the direction $\nabla f(\mathbf{x})$. Analogously, if one wishes to move in a direction in which f *decreases* the fastest, one should proceed in the direction $-\nabla f(\mathbf{x})$.

EXAMPLE 4 In what direction from (0, 1) does $f(x, y) = x^2 - y^2$ increase the fastest?

SOLUTION The gradient is

$$\nabla f = 2x\mathbf{i} - 2y\mathbf{j},$$

and so at (0, 1) this is

$$\nabla f|_{(0,1)} = -2\mathbf{j}.$$

By Theorem 13, f increases fastest in the direction $-\mathbf{j}$. (Can you see why this answer is consistent with Figure 2.1.9?) ▲

Gradients and Tangent Planes to Level Sets

Now we find the relationship between the gradient of a function f and its level surfaces. The gradient points in the direction in which the values of f change most rapidly, whereas a level surface lies in the directions in which they do not change at all. If f is reasonably well behaved, the gradient and the level surface will be perpendicular.

THEOREM 14: The Gradient is Normal to Level Surfaces Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ be a C^1 map and let (x_0, y_0, z_0) lie on the level surface S defined by $f(x, y, z) = k$, for k a constant. Then $\nabla f(x_0, y_0, z_0)$ is normal to the level surface in the following sense: If $\mathbf{v}$ is the tangent vector at $t = 0$ of a path $\mathbf{c}(t)$ in S with $\mathbf{c}(0) = (x_0, y_0, z_0)$, then $\nabla f(x_0, y_0, z_0) \cdot \mathbf{v} = 0$ (see Figure 2.6.2).

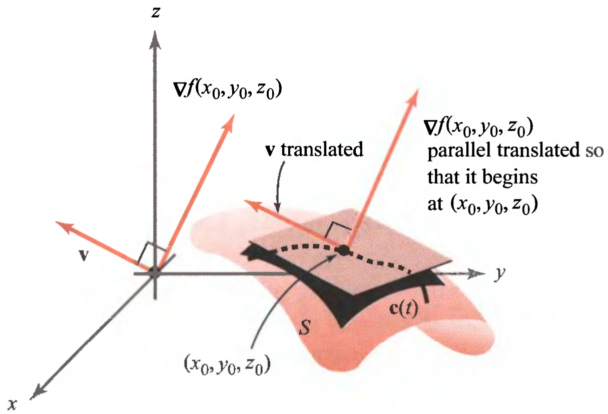


Figure 2.6.2 Geometric significance of the gradient: ∇f is orthogonal to the surface S on which f is constant.

PROOF Let $\mathbf{c}(t)$ lie in S; then $f(\mathbf{c}(t)) = k$. Let $\mathbf{v}$ be as in the hypothesis; then $\mathbf{v} = \mathbf{c}'(0)$. Hence, the fact that $f(\mathbf{c}(t))$ is constant in t, and the chain rule give

$$0 = \frac{d}{dt} f(\mathbf{c}(t))\bigg|_{t=0} = \nabla f(\mathbf{c}(0)) \cdot \mathbf{v}. \quad \blacksquare$$

If we study the conclusion of Theorem 14, we see that it is reasonable to *define* the plane tangent to S as the orthogonal plane to the gradient.

DEFINITION: Tangent Planes to Level Surfaces Let S be the surface consisting of those (x, y, z) such that $f(x, y, z) = k$, for k a constant. The ***tangent plane*** of S at a point (x_0, y_0, z_0) of S is defined by the equation

$$\nabla f(x_0, y_0, z_0) \cdot (x - x_0, y - y_0, z - z_0) = 0 \tag{1}$$

if $\nabla f(x_0, y_0, z_0) \neq \mathbf{0}$. That is, the tangent plane is the set of points (x, y, z) that satisfy equation (1).

This extends the definition we gave earlier for the tangent plane of the graph of a function (see Exercise 11 at the end of this section).

EXAMPLE 5 Compute the equation of the plane tangent to the surface defined by $3xy + z^2 = 4$ at $(1, 1, 1)$.

SOLUTION Here $f(x, y, z) = 3xy + z^2$ and $\nabla f = (3y, 3x, 2z)$, which at $(1, 1, 1)$ is the vector $(3, 3, 2)$. Thus, the tangent plane is

$$(3, 3, 2) \cdot (x - 1, y - 1, z - 1) = 0;$$

that is,

$$3x + 3y + 2z = 8. \quad \blacktriangle$$

In Theorem 14 and the definition following it, we could just as well have worked in two dimensions as in three. Thus, if we have $f\colon \mathbb{R}^2 \to \mathbb{R}$ and consider a level *curve*

$$C = \{(x, y) \mid f(x, y) = k\},$$

then $\nabla f(x_0, y_0)$ is perpendicular to C for any point (x_0, y_0) on C. Likewise, the tangent line to C at (x_0, y_0) has the equation

$$\nabla f(x_0, y_0) \cdot (x - x_0, y - y_0) = 0 \tag{2}$$

if $\nabla f(x_0, y_0) \neq \mathbf{0}$; that is, the tangent line is the set of points (x, y) that satisfy equation (2) (see Figure 2.6.3).

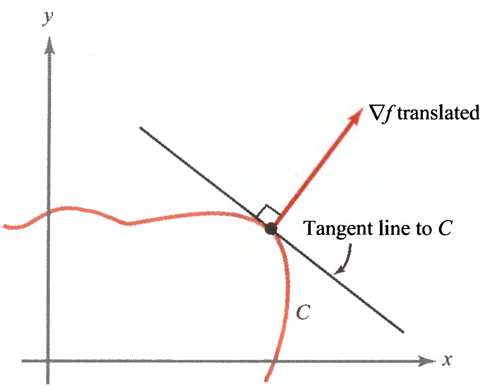


Figure 2.6.3 In the plane, the gradient ∇f is orthogonal to the curve $f = \text{constant}$.

The Gradient Vector Field

We often speak of ∇f as a ***gradient vector field***. The word "field" means that ∇f assigns a vector to each point in the domain of f. In Figure 2.6.4 we describe the gradient ∇f not by drawing its graph, which, if $f\colon \mathbb{R}^3 \to \mathbb{R}$, would be a subset of

$\mathbb{R}^6$, that is, the set of tuples $(\mathbf{x}, \nabla f(\mathbf{x}))$, but by representing $\nabla f(\mathrm{P})$, for each point P, as a vector emanating from the point P rather than from the origin. Like a graph, this pictorial method of depicting ∇f contains the point P and the value $\nabla f(\mathrm{P})$ in the same picture.

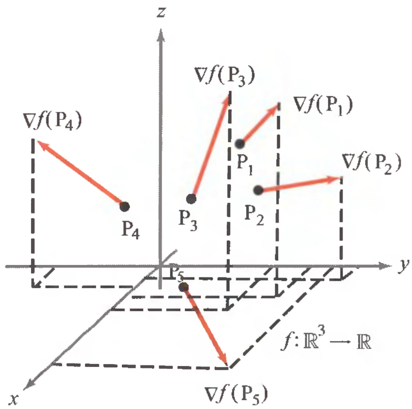


Figure 2.6.4 The gradient ∇f of a function $f: \mathbb{R}^3 \to \mathbb{R}$ is a vector field on $\mathbb{R}^3$; at each point P_i, $\nabla f(\mathrm{P}_i)$ is a vector emanating from P_i.

The gradient vector field has important geometric significance. It shows the direction in which f is increasing the fastest and the direction that is orthogonal to the level surfaces (or curves in the plane) of f. That it does both of these at once is quite plausible. To see this, imagine a hill as shown in Figure 2.6.5(a). Let h be the height function, a function of two variables. If we draw level curves of h, these are just level contours of the hill. We could imagine them as level paths on the hill [see Figure 2.6.5(b)]. One thing should be obvious to anyone who has gone for a hike: To get to the top of the hill the fastest, we should walk perpendicular to level contours.[5] This is consistent with Theorems 13 and 14, which state that the direction of fastest increase (the gradient) is orthogonal to the level curves.

EXAMPLE 6 The gravitational force on a unit mass m at (x, y, z) produced by a mass M at the origin in $\mathbb{R}^3$ is, according to Newton's law of gravitation, given by

$$\mathbf{F} = -\frac{GmM}{r^2}\mathbf{n},$$

where G is a constant; $r = \|\mathbf{r}\| = \sqrt{x^2 + y^2 + z^2}$, which is the distance of (x, y, z) from the origin; and $\mathbf{n} = \mathbf{r}/r$, the unit vector in the direction of $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, which is the position vector from the origin to (x, y, z).

Note that $\mathbf{F} = \nabla(GmM/r) = -\nabla V$, that is, $\mathbf{F}$ is the negative of the gradient of the gravitational potential $V = -GmM/r$. This can be verified as in Example 1.

[5]This discussion assumes that one walks at the same speed in all directions. Of course, hikers know that this is not necessarily realistic.

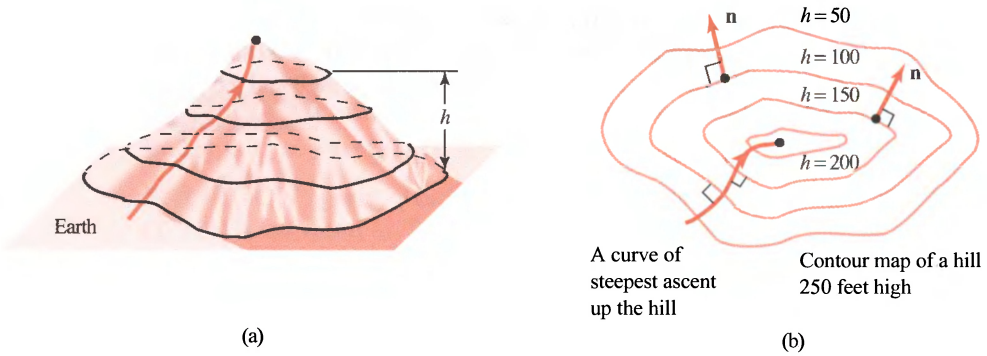


Figure 2.6.5 A physical illustration of the two facts (a) ∇f is the direction of fastest increase of f, and (b) ∇f is orthogonal to the level curves.

Notice that $\mathbf{F}$ is directed inward toward the origin. Also, the level surfaces of V are spheres. The gradient vector field $\mathbf{F}$ is normal to these spheres, which confirms the result of Theorem 14. ▲

EXAMPLE 7 Find a unit vector normal to the surface S given by $z = x^2y^2 + y + 1$ at the point (0, 0, 1).

SOLUTION Let $f(x, y, z) = x^2y^2 + y + 1 - z$, and consider the level surface defined by $f(x, y, z) = 0$. Because this is the set of points (x, y, z) with $z = x^2y^2 + y + 1$, we see that this level set coincides with the surface S. The gradient is given by

$$\nabla f(x, y, z) = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k} = 2xy^2\mathbf{i} + (2x^2y + 1)\mathbf{j} - \mathbf{k},$$

and so

$$\nabla f(0, 0, 1) = \mathbf{j} - \mathbf{k}.$$

This vector is perpendicular to S at (0, 0, 1), and so to find a unit normal $\mathbf{n}$ we divide this vector by its length to obtain

$$\mathbf{n} = \frac{\nabla f(0, 0, 1)}{\|\nabla f(0, 0, 1)\|} = \frac{1}{\sqrt{2}}(\mathbf{j} - \mathbf{k}).$$ ▲

EXAMPLE 8 Consider two conductors, one charged positively and the other negatively. Between them, an electric potential is set up. This potential is a function $\phi\colon \mathbb{R}^3 \to \mathbb{R}$ (an example of a *scalar* field). The electric field is given by $\mathbf{E} = -\nabla\phi$. From Theorem 14 we know that $\mathbf{E}$ is perpendicular to level surfaces of ϕ. These level surfaces are called ***equipotential surfaces***, because the potential is constant on them (see Figure 2.6.6). ▲

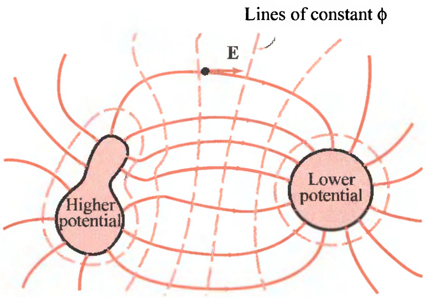


Figure 2.6.6 Equipotential surfaces (the dotted lines) are orthogonal to the electric force field **E**.

EXERCISES

1. Show that the directional derivative of $f(x, y, z) = z^2x + y^3$ at $(1, 1, 2)$ in the direction $(1/\sqrt{5})\mathbf{i} + (2/\sqrt{5})\mathbf{j}$ is $2\sqrt{5}$.

2. Compute the directional derivatives of the following functions at the indicated points in the given directions:

(a) $f(x, y) = x + 2xy - 3y^2$, $(x_0, y_0) = (1, 2)$, $\mathbf{v} = \frac{3}{5}\mathbf{i} + \frac{4}{5}\mathbf{j}$
(b) $f(x, y) = \log\sqrt{x^2 + y^2}$, $(x_0, y_0) = (1, 0)$, $\mathbf{v} = (1/\sqrt{5})(2\mathbf{i} + \mathbf{j})$
(c) $f(x, y) = e^x \cos(\pi y)$, $(x_0, y_0) = (0, -1)$, $\mathbf{v} = -(1/\sqrt{5})\mathbf{i} + (2/\sqrt{5})\mathbf{j}$
(d) $f(x, y) = xy^2 + x^3y$, $(x_0, y_0) = (4, -2)$, $\mathbf{v} = (1/\sqrt{10})\mathbf{i} + (3/\sqrt{10})\mathbf{j}$

3. Compute the directional derivatives of the following functions along unit vectors at the indicated points in directions *parallel* to the given vector:

(a) $f(x, y) = x^y$, $(x_0, y_0) = (e, e)$, $\mathbf{d} = 5\mathbf{i} + 12\mathbf{j}$
(b) $f(x, y, z) = e^x + yz$, $(x_0, y_0, z_0) = (1, 1, 1)$, $\mathbf{d} = (1, -1, 1)$
(c) $f(x, y, z) = xyz$, $(x_0, y_0, z_0) = (1, 0, 1)$, $\mathbf{d} = (1, 0, -1)$

4. Find the planes tangent to the following surfaces at the indicated points:

(a) $x^2 + 2y^2 + 3xz = 10$, at the point $(1, 2, \frac{1}{3})$
(b) $y^2 - x^2 = 3$, at the point $(1, 2, 8)$
(c) $xyz = 1$, at the point $(1, 1, 1)$

5. Find the equation for the plane tangent to each surface $z = f(x, y)$ at the indicated point:

(a) $z = x^3 + y^3 - 6xy$, at the point $(1, 2, -3)$
(b) $z = (\cos x)(\cos y)$, at the point $(0, \pi/2, 0)$
(c) $z = (\cos x)(\sin y)$, at the point $(0, \pi/2, 1)$

6. Compute the gradient ∇f for each of the following functions:

(a) $f(x, y, z) = 1/\sqrt{x^2 + y^2 + z^2}$
(b) $f(x, y, z) = xy + yz + xz$
(c) $f(x, y, z) = \dfrac{1}{x^2 + y^2 + z^2}$

7. For the functions in Exercise 6, what is the direction of fastest increase at $(1, 1, 1)$? [The solution to part (c) only is in the Study Guide to this text.]

8. Show that a unit normal to the surface $x^3y^3 + y - z + 2 = 0$ at $(0, 0, 2)$ is given by $\mathbf{n} = (1/\sqrt{2})(\mathbf{j} - \mathbf{k})$.

9. Find a unit normal to the surface $\cos(xy) = e^z - 2$ at $(1, \pi, 0)$.

10. Verify Theorems 13 and 14 for $f(x, y, z) = x^2 + y^2 + z^2$.

11. Show that the definition following Theorem 14 yields, as a special case, the formula for the plane tangent to the graph of $f(x, y)$ by regarding the graph as a level surface of $F(x, y, z) = f(x, y) - z$ (see Section 2.3).

12. Let $f(x, y) = -(1 - x^2 - y^2)^{1/2}$ for (x, y) such that $x^2 + y^2 < 1$. Show that the plane tangent to the graph of f at $(x_0, y_0, f(x_0, y_0))$ is orthogonal to the vector with components $(x_0, y_0, f(x_0, y_0))$. Interpret this geometrically.

13. For the following functions $f\colon \mathbb{R}^3 \to \mathbb{R}$ and $\mathbf{g}\colon \mathbb{R} \to \mathbb{R}^3$, find ∇f and $\mathbf{g}'$ and evaluate $(f \circ \mathbf{g})'(1)$.

(a) $f(x, y, z) = xz + yz + xy$, $\mathbf{g}(t) = (e^t, \cos t, \sin t)$
(b) $f(x, y, z) = e^{xyz}$, $\mathbf{g}(t) = (6t, 3t^2, t^3)$
(c) $f(x, y, z) = (x^2 + y^2 + z^2)\log\sqrt{x^2 + y^2 + z^2}$, $\mathbf{g}(t) = (e^t, e^{-t}, t)$

14. Compute the directional derivative of f in the given directions $\mathbf{v}$ at the given points P.

(a) $f(x, y, z) = xy^2 + y^2z^3 + z^3x$, $\mathrm{P} = (4, -2, -1)$, $\mathbf{v} = 1/\sqrt{14}(\mathbf{i} + 3\mathbf{j} + 2\mathbf{k})$
(b) $f(x, y, z) = x^{yz}$, $\mathrm{P} = (e, e, 0)$, $\mathbf{v} = \frac{12}{13}\mathbf{i} + \frac{3}{13}\mathbf{j} + \frac{4}{13}\mathbf{k}$

15. Let $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ and $r = \|\mathbf{r}\|$. Prove that

$$\nabla\left(\frac{1}{r}\right) = -\frac{\mathbf{r}}{r^3}.$$

16. Captain Ralph is in trouble near the sunny side of Mercury. The temperature of the ship's hull when he is at location (x, y, z) will be given by $T(x, y, z) = e^{-x^2-2y^2-3z^2}$, where x, y, and z are measured in meters. He is currently at $(1, 1, 1)$.

(a) In what direction should he proceed in order to decrease the temperature most rapidly?

(b) If the ship travels at e^8 meters per second, how fast will be the temperature decrease if he proceeds in that direction?

(c) Unfortunately, the metal of the hull will crack if cooled at a rate greater than $\sqrt{14}e^2$ degrees per second. Describe the set of possible directions in which he may proceed to bring the temperature down at no more than that rate.

17. A function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is said to be *independent of the second variable* if there is a function $g\colon \mathbb{R} \to \mathbb{R}$ such that $f(x, y) = g(x)$ for all x in $\mathbb{R}$. In this case, calculate ∇f in terms of g'.

18. Let f and g be functions from $\mathbb{R}^3$ to $\mathbb{R}$. Suppose f is differentiable and $\nabla f(\mathbf{x}) = g(\mathbf{x})\mathbf{x}$. Show that spheres centered at the origin are contained in the level sets for f; that is, f is constant on such spheres.

19. A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is called an *even* function if $f(\mathbf{x}) = f(-\mathbf{x})$ for every $\mathbf{x}$ in $\mathbb{R}^n$. If f is differentiable and even, find $\mathbf{D}f$ at the origin.

20. Suppose that a mountain has the shape of an elliptic paraboloid $z = c - ax^2 - by^2$, where a, b, and c are positive constants, x and y are the east-west and north-south map coordinates, and z is the altitude above sea level (x, y, z are all measured in meters). At the point $(1, 1)$, in what direction is the altitude increasing most rapidly? If a marble were released at $(1, 1)$, in what direction would it begin to roll?

21. An engineer wishes to build a railroad up the mountain of Exercise 20. Straight up the mountain is much too steep for the power of the engines. At the point $(1, 1)$, in what directions may the track be laid so that it will be climbing with a 3% grade—that is, an angle whose tangent is 0.03? (There are two possibilities.) Make a sketch of the situation indicating the two possible directions for a 3% grade at $(1, 1)$.

22. In electrostatics, the force $\mathbf{P}$ of attraction between two particles of opposite charge is given by $\mathbf{P} = k(\mathbf{r}/\|\mathbf{r}\|^3)$ (*Coulomb's law*), where k is a constant and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Show that $\mathbf{P}$ is the gradient of $f = -k/\|\mathbf{r}\|$.

23. The electrostatic potential V due to two infinite parallel filaments with linear charge densities λ and $-\lambda$ is $V = (\lambda/2\pi\varepsilon_0)\ln(r_2/r_1)$, where $r_1^2 = (x - x_0)^2 + y^2$ and $r_2^2 = (x + x_0)^2 + y^2$. We think of the filaments as being in the z direction, passing through the xy plane at $(-x_0, 0)$ and $(x_0, 0)$. Find $\nabla V(x, y)$.

24. For each of the following, find the maximum and minimum values attained by the function f along the path $\mathbf{c}(t)$:

(a) $f(x, y) = xy$; $\mathbf{c}(t) = (\cos t, \sin t)$; $0 \le t \le 2\pi$.
(b) $f(x, y) = x^2 + y^2$; $\mathbf{c}(t) = (\cos t, 2\sin t)$; $0 \le t \le 2\pi$.

25. Suppose that a particle is ejected from the surface $x^2 + y^2 - z^2 = -1$ at the point $(1, 1, \sqrt{3})$ along the normal directed toward the xy plane to the surface at time $t = 0$ with a speed of 10 units per second. When and where does it cross the xy plane?

26. Let $f\colon \mathbb{R}^3 \to \mathbb{R}$ and regard $\mathbf{D}f(x, y, z)$ as a linear map of $\mathbb{R}^3$ to $\mathbb{R}$. Show that the kernel (that is, the set of vectors mapped to zero) of $\mathbf{D}f$ is the plane in $\mathbb{R}^3$ orthogonal to ∇f.

REVIEW EXERCISES FOR CHAPTER 2

1. Describe the graphs of:

(a) $f(x, y) = 3x^2 + y^2$ (b) $f(x, y) = xy + 3x$

2. Describe some appropriate level surfaces and sections of the graphs of:

(a) $f(x, y, z) = 2x^2 + y^2 + z^2$

(b) $f(x, y, z) = x^2$
(c) $f(x, y, z) = xyz$

3. Compute the derivative $\mathbf{D}f(\mathbf{x})$ of each of the following functions:

(a) $f(x, y) = (x^2 y, e^{-xy})$
(b) $f(x) = (x, x)$
(c) $f(x, y, z) = e^x + e^y + e^z$
(d) $f(x, y, z) = (x, y, z)$

4. Suppose $f(x, y) = f(y, x)$ for all (x, y). Prove that

$$(\partial f/\partial x)(a, b) = (\partial f/\partial y)(b, a).$$

5. Let $f(x, y) = (1 - x^2 - y^2)^{1/2}$. Show that the plane tangent to the graph of f at $(x_0, y_0, f(x_0, y_0))$ is orthogonal to the vector $(x_0, y_0, f(x_0, y_0))$. Interpret geometrically.

6. Let $F(u, v)$ and $u = h(x, y, z)$, $v = k(x, y, z)$ be given (differentiable) real-valued functions and let $f(x, y, z)$ be defined by $f(x, y, z) = F(h(x, y, z), k(x, y, z))$. Write a formula for the gradient of f in terms of the partial derivatives of F, h, and k.

7. Find an equation for the tangent plane of the graph of f at the point $(x_0, y_0, f(x_0, y_0))$ for:

(a) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x - y + 2,$ $\quad (x_0, y_0) = (1, 1)$
(b) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto x^2 + 4y^2,$ $\quad (x_0, y_0) = (2, -1)$
(c) $f\colon \mathbb{R}^2 \to \mathbb{R}, (x, y) \mapsto xy,$ $\quad (x_0, y_0) = (-1, -1)$
(d) $f(x, y) = \log(x + y) + x \cos y + \arctan(x + y),$ $\quad (x_0, y_0) = (1, 0)$
(e) $f(x, y) = \sqrt{x^2 + y^2},$ $\quad (x_0, y_0) = (1, 1)$
(f) $f(x, y) = xy,$ $\quad (x_0, y_0) = (2, 1)$

8. Compute an equation for the tangent planes of the following surfaces at the indicated points.

(a) $x^2 + y^2 + z^2 = 3,$ $\quad (1, 1, 1)$
(b) $x^3 - 2y^3 + z^3 = 0,$ $\quad (1, 1, 1)$
(c) $(\cos x)(\cos y)e^z = 0,$ $\quad (\pi/2, 1, 0)$
(d) $e^{xyz} = 1,$ $\quad (1, 1, 0)$

9. Draw some level curves for the following functions:

(a) $f(x, y) = 1/xy$
(b) $f(x, y) = x^2 - xy - y^2$

10. Consider a temperature function $T(x, y) = x \sin y$. Plot a few level curves. Compute ∇T and explain its meaning.

11. Find the following limits if they exist:

(a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\cos xy - 1}{x}$ $\qquad$ (b) $\displaystyle\lim_{(x,y)\to(0,0)} \sqrt{|(x + y)/(x - y)|}, x \neq y$

12. Compute the first partial derivatives and gradients of the following functions:

(a) $f(x, y, z) = xe^z + y \cos x$
(b) $f(x, y, z) = (x + y + z)^{10}$
(c) $f(x, y, z) = (x^2 + y)/z$

13. Compute $\dfrac{\partial}{\partial x}[x \exp(1 + x^2 + y^2)]$

14. Let $y(x)$ be a differentiable function defined implicitly by $F(x, y(x)) = 0$. From Exercise 17(a), Section 2.5, we know that

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}.$$

Consider the surface $z = F(x, y)$, and suppose F is increasing as a function of x and as a function of y; that is, $\partial F/\partial x > 0$ and $\partial F/\partial y > 0$. By considering the graph and the plane $z = 0$, show that for z fixed at $z = 0$, y should *decrease* as x increases and x should *decrease* as y increases. Does this agree with the minus sign in the formula for dy/dx?

15. (a) Consider the graph of a function $f(x, y)$ [Figure 2.R.1(a)]. Let (x_0, y_0) lie on a level curve C, so $\nabla f(x_0, y_0)$ is perpendicular to this curve. Show that the tangent plane of the graph is the plane that (i) contains the line perpendicular to $\nabla f(x_0, y_0)$ and lying in the horizontal plane $z = f(x_0, y_0)$, and (ii) has slope $\|\nabla f(x_0, y_0)\|$ relative to the xy plane. (By the *slope* of a plane P relative to the xy plane we mean the tangent of the angle θ, $0 \le \theta \le \pi$, between the upward-pointing normal $\mathbf{p}$ to P and the unit vector $\mathbf{k}$.)

(b) Use this method to show that the tangent plane of the graph of $f(x, y) = (x + \cos y)x^2$ at $(1, 0, 2)$ is as sketched in Figure 2.R.1(b).

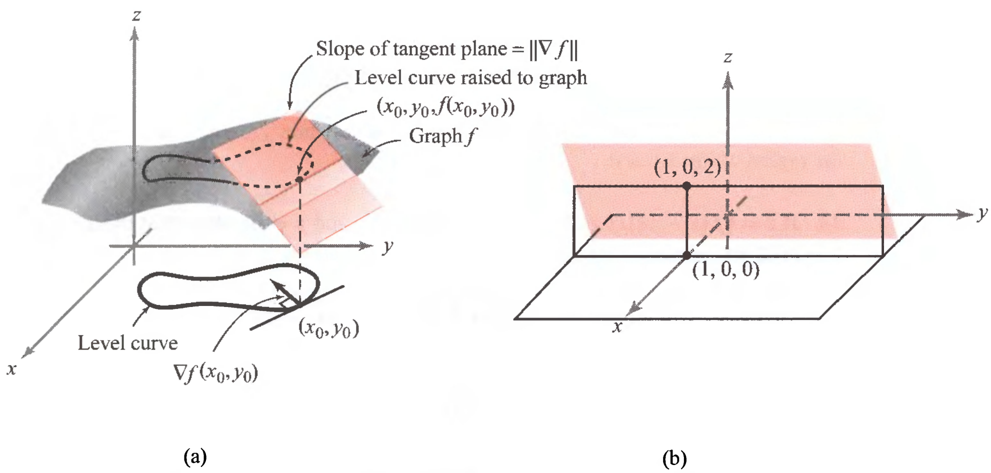


Figure 2.R.1 (a) The relationship between the gradient of a function and the tangent plane to the graph [Exercise 15(a)]. A specific instance of the plane in (b) for Exercise 15(b).

16. Find the plane tangent to the surface $z = x^2 + y^2$ at the point $(1, -2, 5)$. Explain the geometric significance, for this surface, of the gradient of $f(x, y) = x^2 + y^2$ (see Exercise 15).

17. In which direction is the directional derivative of $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$ at $(1, 1)$ equal to zero?

18. Find the directional derivative of the given function at the given point and in the direction of the given vector.

(a) $f(x, y, z) = e^x \cos(yz)$, $\quad p_0 = (0, 0, 0)$, $\quad \mathbf{v} = (2, 1, -2)$
(b) $f(x, y, z) = xy + yz + zx$, $\quad p_0 = (1, 1, 2)$, $\quad \mathbf{v} = (10, -1, 2)$

19. Find the tangent plane and normal to the hyperboloid $x^2 + y^2 - z^2 = 18$ at $(3, 5, -4)$.

20. Let $(x(t), y(t))$ be a path in the plane, $0 \le t \le 1$, and let $f(x, y)$ be a C^1 function of two variables. Assume that $(dx/dt)f_x + (dy/dt)f_y \le 0$. Show that $f(x(1), y(1)) \le f(x(0), y(0))$.

21. A bug finds itself in a toxic environment. The toxicity level is given by $T(x, y) = 2x^2 - 4y^2$. The bug is at $(-1, 2)$. In what direction should it move to lower the toxicity the fastest?

22. Find the direction in which the function $w = x^2 + xy$ increases most rapidly at the point $(-1, 1)$. What is the magnitude of ∇w at this point? Interpret this magnitude geometrically.

23. Let f be defined on an open set S in $\mathbb{R}^n$. We say that f is ***homogeneous of degree*** p over S if $f(\lambda\mathbf{x}) = \lambda^p f(\mathbf{x})$ for every real λ and for every $\mathbf{x}$ in S for which $\lambda\mathbf{x} \in S$.

(a) If such a function is differentiable at $\mathbf{x}$, show that $\mathbf{x} \cdot \nabla f(\mathbf{x}) = pf(\mathbf{x})$. This is known as ***Euler's theorem*** for homogeneous functions. [HINT: For fixed $\mathbf{x}$, define $g(\lambda) = f(\lambda\mathbf{x})$ and compute $g'(1)$.]

(b) Find p and check Euler's theorem for the function $f(x, y, z) = x - 2y - \sqrt{xz}$, on the region where $xz > 0$.

24. If $z = [f(x - y)]/y$ (where f is differentiable and $y \neq 0$), show that the identity $z + y(\partial z/\partial x) + y(\partial z/\partial y) = 0$ holds.

25. Given $z = f((x + y)/(x - y))$ for f a C^1 function, show that

$$x\frac{\partial z}{\partial x} + y\frac{\partial z}{\partial y} = 0.$$

26. Let f have partial derivatives $\partial f(\mathbf{x})/\partial x_i$, where $i = 1, 2, \ldots, n$, at each point $\mathbf{x}$ of an open set U in $\mathbb{R}^n$. If f has a local maximum or a local minimum at the point $\mathbf{x}_0$ in U, show that $\partial f(\mathbf{x}_0)/\partial x_i = 0$ for each i.

27. Consider the functions defined in $\mathbb{R}^2$ by the following formulas:

(i) $f(x, y) = xy/(x^2 + y^2)$ if $(x, y) \neq (0, 0)$, $f(0, 0) = 0$
(ii) $f(x, y) = x^2y^2/(x^2 + y^4)$ if $(x, y) \neq (0, 0)$, $f(0, 0) = 0$

(a) In each case, show that the partial derivatives $\partial f(x, y)/\partial x$ and $\partial f(x, y)/\partial y$ exist for every (x, y) in $\mathbb{R}^2$, and evaluate these derivatives explicitly in terms of x and y.

(b) Explain why the functions described in (i) and (ii) are or are not differentiable at $(0, 0)$.

28. Compute the gradient vector $\nabla f(x, y)$ at all points (x, y) in $\mathbb{R}^2$ for each of the following functions:

(a) $f(x, y) = x^2y^2 \log(x^2 + y^2)$ if $(x, y) \neq (0, 0)$, $f(0, 0) = 0$
(b) $f(x, y) = xy \sin[1/(x^2 + y^2)]$ if $(x, y) \neq (0, 0)$, $f(0, 0) = 0$

29. Find the directional derivatives of the following functions at the point $(1, 1)$ in the direction $(\mathbf{i} + \mathbf{j})/\sqrt{2}$:

(a) $f(x, y) = x \tan^{-1}(x/y)$
(b) $f(x, y) = \cos(\sqrt{x^2 + y^2})$
(c) $f(x, y) = \exp(-x^2 - y^2)$

30. (a) Let $\mathbf{u} = \mathbf{i} - 2\mathbf{j} + 2\mathbf{k}$ and $\mathbf{v} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$. Find: $\|\mathbf{u}\|$, $\mathbf{u} \cdot \mathbf{v}$, $\mathbf{u} \times \mathbf{v}$, and a vector in the same direction as $\mathbf{u}$, but of unit length.

(b) Find the rate of change of $e^{xy} \sin(xyz)$ in the direction $\mathbf{u}$ at $(0, 1, 1)$.

31. Let $h(x, y) = 2e^{-x^2} + e^{-3y^2}$ denote the height on a mountain at position (x, y). In what direction from $(1, 0)$ should one begin walking in order to climb the fastest?

32. Compute an equation for the plane tangent to the graph of

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

at $x = 1$, $y = 2$.

33. (a) Give a careful statement of the general form of the chain rule.

(b) Let $f(x, y) = x^2 + y$ and $\mathbf{h}(u) = (\sin 3u, \cos 8u)$. Let $g(u) = f(\mathbf{h}(u))$. Compute dg/du at $u = 0$ both directly *and* by using the chain rule.

34. (a) Sketch the level curves of $f(x, y) = -x^2 - 9y^2$ for $c = 0, -1, -10$.

(b) On your sketch, draw in ∇f at $(1, 1)$. Discuss.

35. At time $t = 0$, a particle is ejected from the surface $x^2 + 2y^2 + 3z^2 = 6$ at the point $(1, 1, 1)$ in a direction normal to the surface at a speed of 10 units per second. At what time does it cross the sphere $x^2 + y^2 + z^2 = 103$?

36. At what point(s) on the surface in Exercise 35 is the normal vector parallel to the line $x = y = z$?

37. Compute $\partial z/\partial x$ and $\partial z/\partial y$ if

$$z = \frac{u^2 + v^2}{u^2 - v^2}, \qquad u = e^{-x-y}, \qquad v = e^{xy}$$

(a) by substitution and direct calculation, and (b) by the chain rule.

38. Compute the partial derivatives as in Exercise 37 if $z = uv$, $u = x + y$, and $v = x - y$.

39. What is wrong with the following argument? Suppose that $w = f(x, y)$ and $y = x^2$. By the chain rule,

$$\frac{\partial w}{\partial x} = \frac{\partial w}{\partial x}\frac{\partial x}{\partial x} + \frac{\partial w}{\partial y}\frac{\partial y}{\partial x} = \frac{\partial w}{\partial x} + 2x\frac{\partial w}{\partial y}.$$

Hence, $0 = 2x(\partial w/\partial y)$, and so $\partial w/\partial y = 0$. Choose an explicit example to really see that this is incorrect.

40. A boat is sailing northeast at 20 km/h. Assuming that the temperature drops at a rate of 0.2°C/km in the northerly direction and 0.3°C/km in the easterly direction, what is the time rate of change of temperature as observed on the boat?

41. Use the chain rule to find a formula for $(d/dt)\exp[f(t)g(t)]$.

42. Use the chain rule to find a formula for $(d/dt)(f(t)^{g(t)})$.

43. Verify the chain rule for the function $f(x, y, z) = [\ln(1 + x^2 + 2z^2)]/(1 + y^2)$ and the path $\mathbf{c}(t) = (t, 1 - t^2, \cos t)$.

44. Verify the chain rule for the function $f(x, y) = x^2/(2 + \cos y)$ and the path $x = e^t$, $y = e^{-t}$.

45. Suppose that $u(x, t)$ satisfies the differential equation $u_t + uu_x = 0$ and that x, as a function $x = f(t)$ of t, satisfies $dx/dt = u(x, t)$. Prove that $u(f(t), t)$ is constant in t.

46. The displacement at time t and horizontal position on a line x of a certain violin string is given by $u = \sin(x - 6t) + \sin(x + 6t)$. Calculate the velocity of the string at $x = 1$ when $t = \frac{1}{3}$.

47. The ***ideal gas law*** $PV = nRT$ involves a constant R, the number n of moles of the gas, the volume V, the Kelvin temperature T, and the pressure P.

(a) Show that each of n, P, T, V is a function of the remaining variables, and determine explicitly the defining equations.
(b) Calculate $\partial V/\partial T$, $\partial T/\partial P$, $\partial P/\partial V$ and show that their product equals -1.

48. The ***potential temperature*** θ is defined in terms of temperature T and pressure p by

$$\theta = T\left(\frac{1000}{p}\right)^{0.286}.$$

The temperature and pressure may be thought of as functions of position (x, y, z) in the atmosphere and also of time t.

(a) Find formulas for $\partial\theta/\partial x$, $\partial\theta/\partial y$, $\partial\theta/\partial z$, $\partial\theta/\partial t$ in terms of partial derivatives of T and p.

(b) The condition $\partial\theta/\partial z < 0$ is regarded as an unstable atmosphere, for it leads to large vertical excursions of air parcels from a single upward or downward impetus. Meteorologists use the formula

$$\frac{\partial\theta}{\partial z} = \frac{\theta}{T}\left(\frac{\partial T}{\partial z} + \frac{g}{C_p}\right)$$

where $g = 32.2$ and C_p is a positive constant. How does the temperature change in the upward direction for an unstable atmosphere?

49. The specific volume V, pressure P, and temperature T of a van der Waals gas are related by $P = RT/(V - \beta) - \alpha/V^2$, where α, β, and R are constants.

(a) Explain why any two of V, P, and T can be considered independent variables that determine the third variable.

(b) Find $\partial T/\partial P$, $\partial P/\partial V$, $\partial V/\partial T$. Identify which variables are constant, and interpret each partial derivative physically.

(c) Verify that $(\partial T/\partial P)(\partial P/\partial V)(\partial V/\partial T) = -1$(not $+1$!).

50. The height h of the Hawaiian volcano Mauna Loa is (roughly) described by the function $h(x, y) = 2.59 - 0.00024y^2 - 0.00065x^2$, where h is the height above sea level in miles and x and y measure east-west and north-south distances in miles from the top of the mountain. At $(x, y) = (-2, -4)$:

(a) How fast is the height increasing in the direction $(1, 1)$ (that is, northeastward)? Express your answer in miles of height per mile of horizontal distance traveled.

(b) In what direction is the steepest upward path?

51. (a) In what direction is the directional derivative of $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$ at $(1, 1)$ equal to zero?

(b) How about at an arbitrary point (x_0, y_0) in the first quadrant?

(c) Describe the level curves of f. In particular, discuss them in terms of the result of part (b).

52. (a) Show that the curve $x^2 - y^2 = c$, for *any* value of c, satisfies the differential equation $dy/dx = x/y$.

(b) Draw in a few of the curves $x^2 - y^2 = c$, say for $c = \pm 1$. At several points (x, y) along each of these curves, draw a short segment of slope x/y; check that these segments appear to be tangent to the curve. What happens when $y = 0$? What happens when $c = 0$?

53. Suppose that f is a differentiable function of one variable and that a function $u = g(x, y)$ is defined by

$$u = g(x, y) = xyf\left(\frac{x+y}{xy}\right).$$

Show that u satisfies a (partial) differential equation of the form

$$x^2\frac{\partial u}{\partial x} - y^2\frac{\partial u}{\partial y} = G(x, y)u$$

and find the function $G(x, y)$.

54. (a) Let F be a function of one variable and f a function of two variables. Show that the gradient vector of $g(x, y) = F(f(x, y))$ is parallel to the gradient vector of $f(x, y)$.

(b) Let $f(x, y)$ and $g(x, y)$ be functions such that $\nabla f = \lambda \nabla g$ for some function $\lambda(x, y)$. What is the relation between the level curves of f and g? Explain why there might be a function F such that $g(x, y) = F(f(x, y))$.