DEFINITION 2.1 **(INTUITIVE DEFINITION OF LIMIT)** The equation

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{L},$$

where $\mathbf{f}\colon X \subseteq \mathbf{R}^n \to \mathbf{R}^m$, means that we can make $\|\mathbf{f}(\mathbf{x}) - \mathbf{L}\|$ arbitrarily small (i.e., near zero) by keeping $\|\mathbf{x} - \mathbf{a}\|$ sufficiently small (but nonzero).

THEOREM 2.5 (ALGEBRAIC PROPERTIES) Let $\mathbf{F}, \mathbf{G}\colon X \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be vector-valued functions, $f, g\colon X \subseteq \mathbf{R}^n \to \mathbf{R}$ be scalar-valued functions, and let $k \in \mathbf{R}$ be a scalar.

1. If $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{L}$ and $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{G}(\mathbf{x}) = \mathbf{M}$, then $\lim_{\mathbf{x}\to\mathbf{a}}(\mathbf{F}+\mathbf{G})(\mathbf{x}) = \mathbf{L}+\mathbf{M}$.
2. If $\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{F}(\mathbf{x}) = \mathbf{L}$, then $\lim_{\mathbf{x}\to\mathbf{a}} k\mathbf{F}(\mathbf{x}) = k\mathbf{L}$.
3. If $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$ and $\lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x}) = M$, then $\lim_{\mathbf{x}\to\mathbf{a}}(fg)(\mathbf{x}) = LM$.
4. If $\lim_{\mathbf{x}\to\mathbf{a}} f(\mathbf{x}) = L$, $g(\mathbf{x}) \neq 0$ for $\mathbf{x} \in X$, and $\lim_{\mathbf{x}\to\mathbf{a}} g(\mathbf{x}) = M \neq 0$, then $\lim_{\mathbf{x}\to\mathbf{a}}(f/g)(\mathbf{x}) = L/M$.

DEFINITION 2.7 Let $\mathbf{f}: X \subseteq \mathbf{R}^n \to \mathbf{R}^m$ and let $\mathbf{a} \in X$. Then $\mathbf{f}$ is said to be **continuous at a** if either $\mathbf{a}$ is an isolated point of X or if

$$\lim_{\mathbf{x} \to \mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}).$$

If $\mathbf{f}$ is continuous at all points of its domain X, then we simply say that $\mathbf{f}$ is **continuous**.

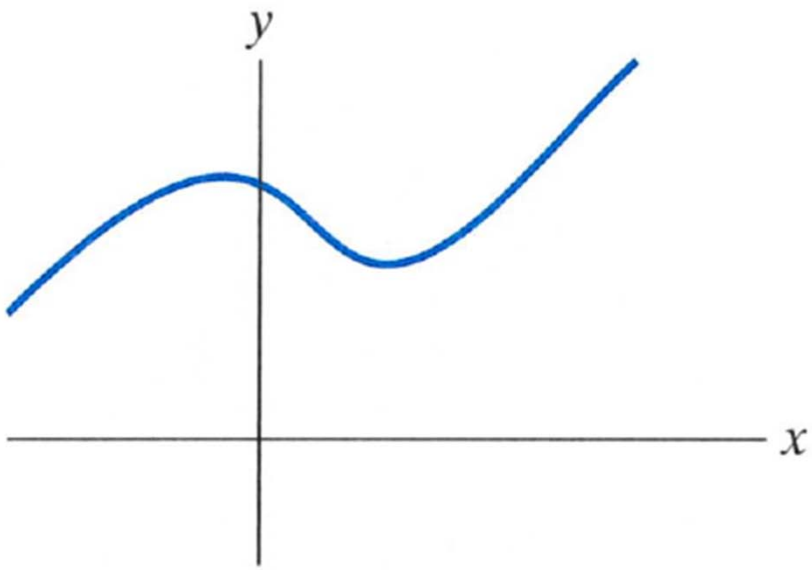

Figure 2.41 The graph of a continuous function.

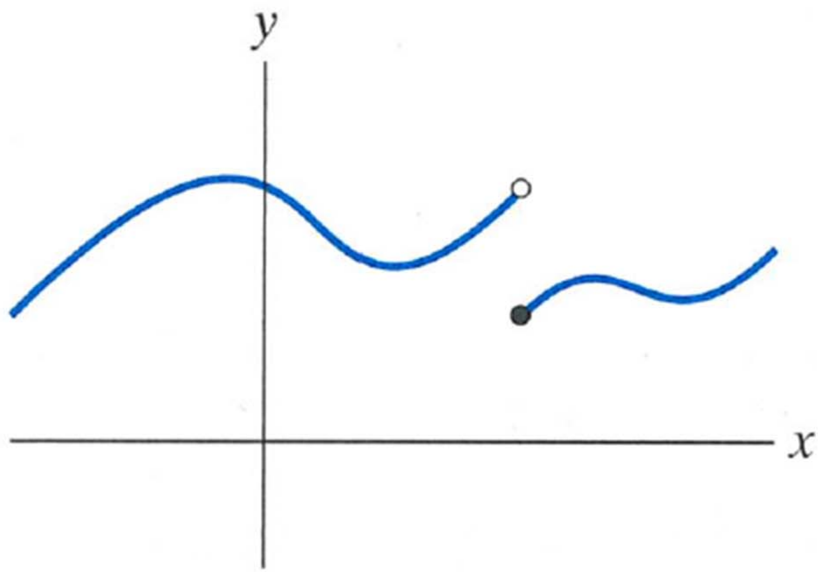

Figure 2.42 The graph of a function that is not continuous.

Recall that if $F\colon X \subseteq \mathbf{R} \to \mathbf{R}$ is a scalar-valued function of one variable, then the **derivative** of F at a number $a \in X$ is

$$F'(a) = \lim_{h \to 0} \frac{F(a+h) - F(a)}{h}. \tag{1}$$

Moreover, F is said to be **differentiable at** a precisely when the limit in equation (1) exists.

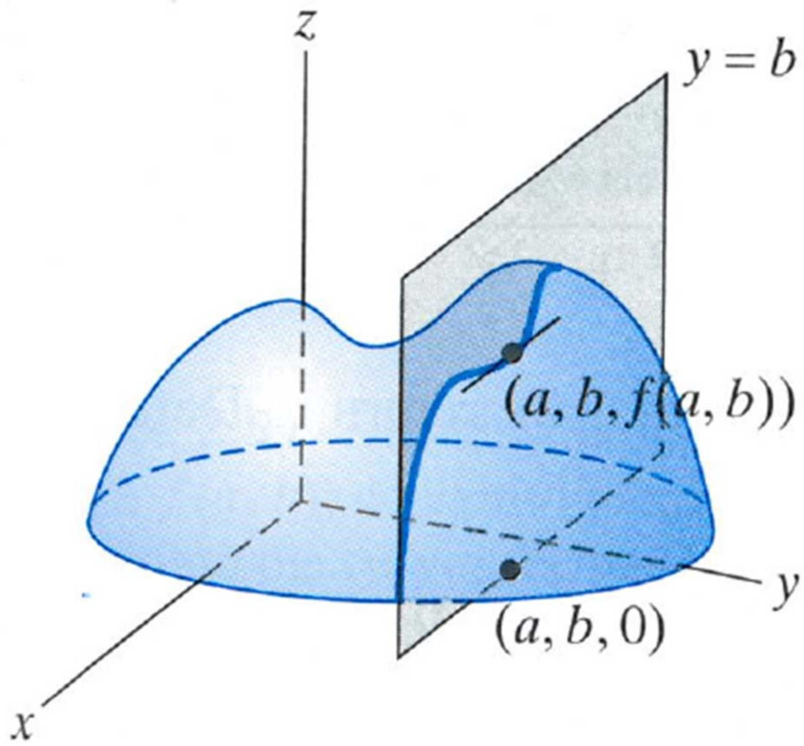

Figure 2.47 Visualizing the partial derivative $\frac{\partial f}{\partial x}(a, b)$.

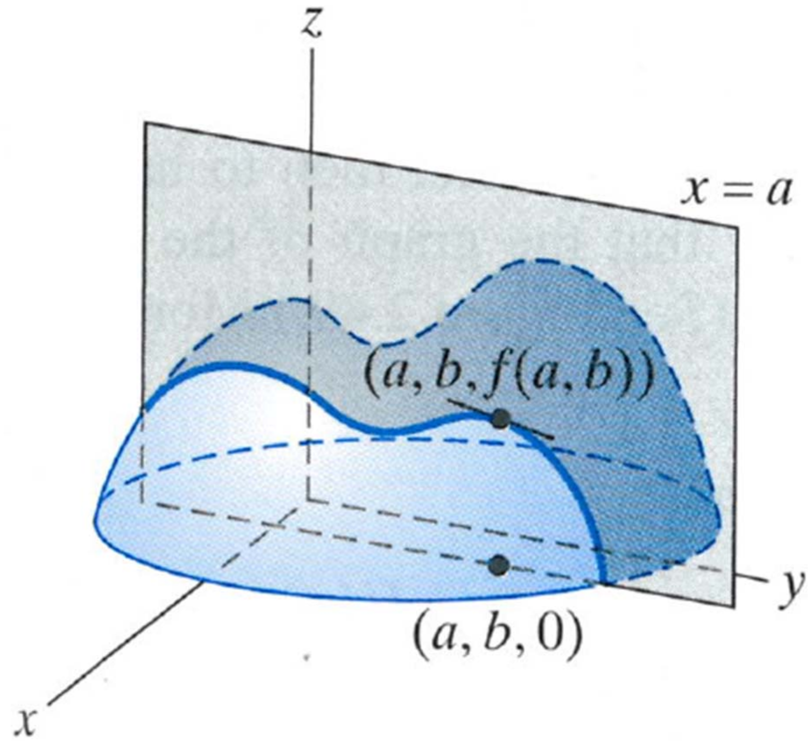

Figure 2.48 Visualizing the partial derivative $\frac{\partial f}{\partial y}(a, b)$.

DEFINITION 3.2 The **partial derivative of f with respect to x_i** is the (ordinary) derivative of the partial function with respect to x_i. That is, the partial derivative with respect to x_i is $F'(x_i)$, in the notation of Definition 3.1. Standard notations for the partial derivative of f with respect to x_i are

$$\frac{\partial f}{\partial x_i}, \qquad D_{x_i} f(x_1, \ldots, x_n), \qquad \text{and} \qquad f_{x_i}(x_1, \ldots, x_n).$$

Symbolically, we have

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1, \ldots, x_i + h, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}. \tag{2}$$

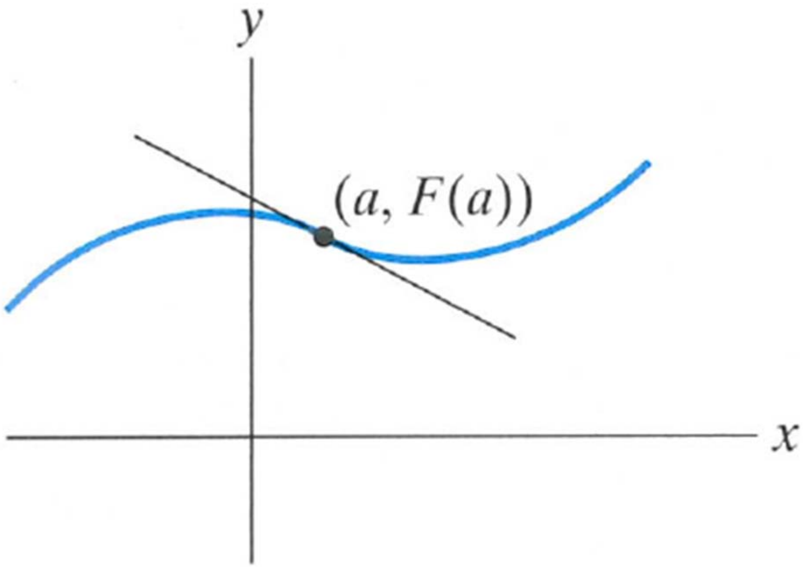

Figure 2.49 The tangent line to $y = F(x)$ at $x = a$ has equation $y = F(a) + F'(a)(x - a)$.

The derivative $F'(a)$ is the slope of the tangent line to $y = F(x)$ at $x = a$.

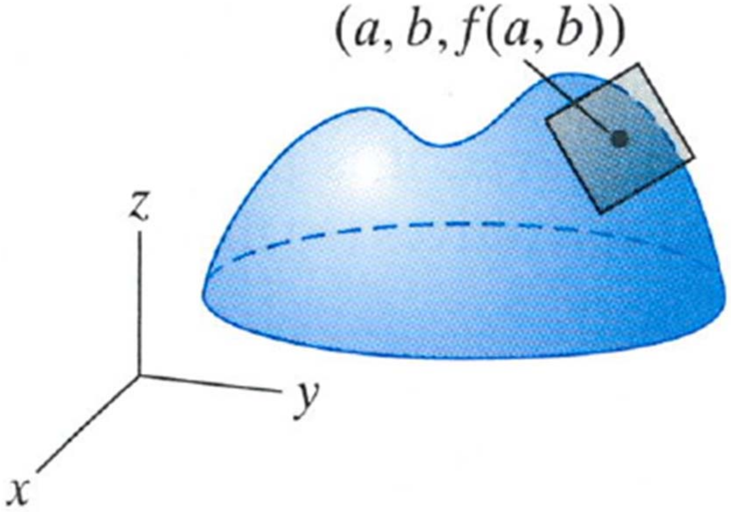

Figure 2.50 The plane tangent to $z = f(x, y)$ at $(a, b, f(a, b))$.

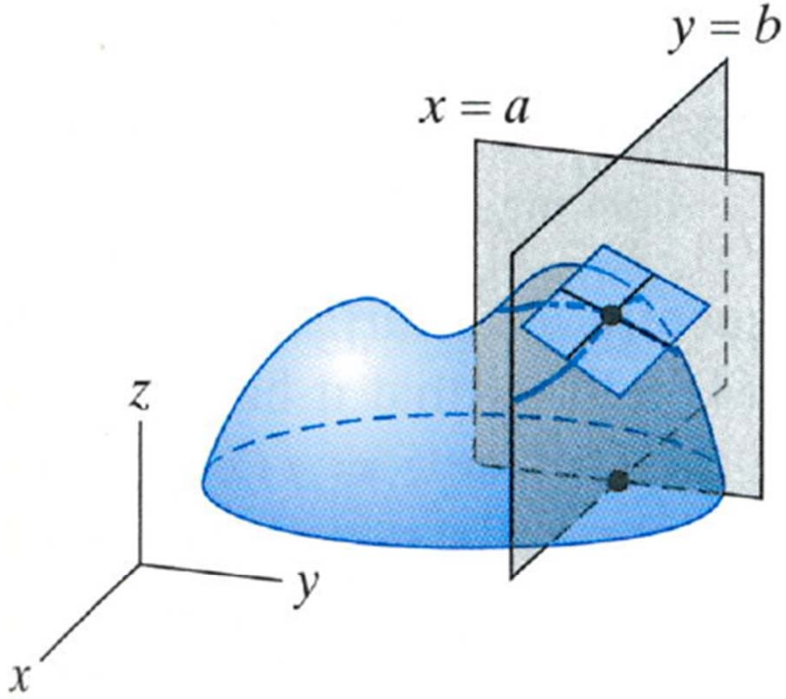

Figure 2.51 The tangent plane at $(a, b, f(a, b))$ contains the lines tangent to the curves formed by intersecting the surface $z = f(x, y)$ by the planes $x = a$ and $y = b$.

THEOREM 3.3 If the graph of $z = f(x, y)$ has a tangent plane at $(a, b, f(a, b))$, then that tangent plane has equation

$$z = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b). \tag{4}$$

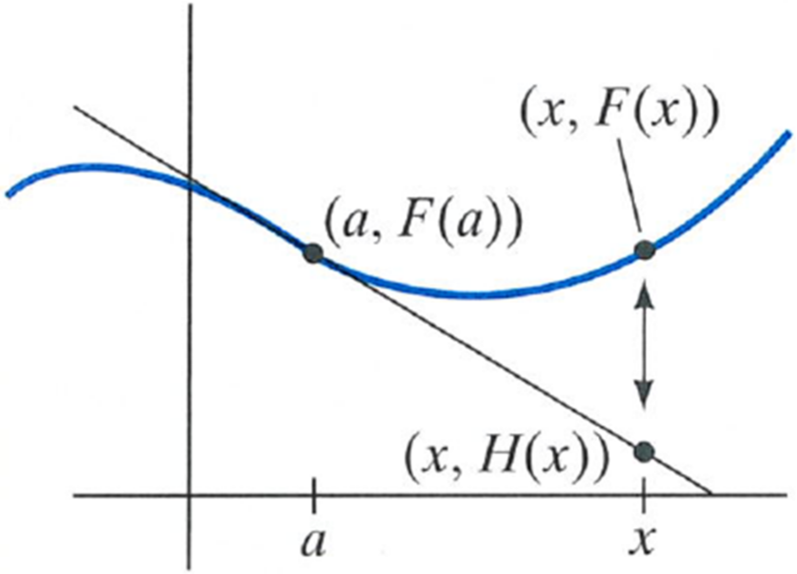

Figure 2.53 If F is differentiable at a, the vertical distance between $F(x)$ and $H(x)$ must approach zero faster than the horizontal distance between x and a does.

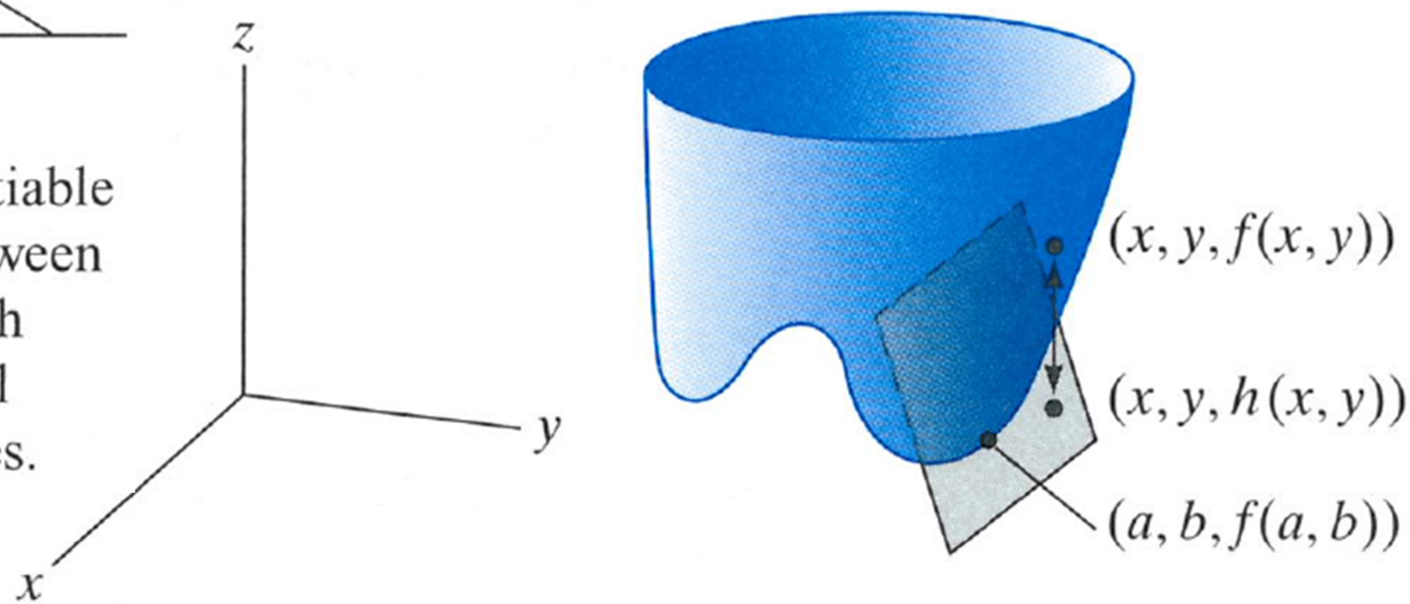

Figure 2.54 If f is differentiable at (a, b), the distance between $f(x, y)$ and $h(x, y)$ must approach zero faster than the distance between (x, y) and (a, b) does.

DEFINITION 3.4 Let X be open in $\mathbf{R}^2$ and $f: X \subseteq \mathbf{R}^2 \to \mathbf{R}$ be a scalar-valued function of two variables. We say that f is **differentiable at** $(a, b) \in X$ if the partial derivatives $f_x(a, b)$ and $f_y(a, b)$ exist and if the function

$$h(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is a good linear approximation to f near (a, b)—that is, if

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - h(x, y)}{\|(x, y) - (a, b)\|} = 0.$$

Moreover, if f is differentiable at (a, b), then the equation $z = h(x, y)$ defines the **tangent plane** to the graph of f at the point $(a, b, f(a, b))$. If f is differentiable at all points of its domain, then we simply say that f is **differentiable**.

DEFINITION 3.7 Let X be open in $\mathbf{R}^n$ and $f\colon X \to \mathbf{R}$ be a scalar-valued function; let $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in X$. We say that f is **differentiable at a** if all the partial derivatives $f_{x_i}(\mathbf{a})$, $i = 1, \ldots, n$, exist and if the function $h\colon \mathbf{R}^n \to \mathbf{R}$ defined by

$$\begin{aligned} h(\mathbf{x}) = f(\mathbf{a}) &+ f_{x_1}(\mathbf{a})(x_1 - a_1) + f_{x_2}(\mathbf{a})(x_2 - a_2) \\ &+ \cdots + f_{x_n}(\mathbf{a})(x_n - a_n) \end{aligned} \tag{6}$$

is a good linear approximation to f near $\mathbf{a}$, meaning that

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

DEFINITION 3.8 (GRAND DEFINITION OF DIFFERENTIABILITY) Let $X \subseteq \mathbf{R}^n$ be open, let $\mathbf{f}\colon X \to \mathbf{R}^m$, and let $\mathbf{a} \in X$. We say that $\mathbf{f}$ is **differentiable at a** if $D\mathbf{f}(\mathbf{a})$ exists and if the function $\mathbf{h}\colon \mathbf{R}^n \to \mathbf{R}^m$ defined by

$$\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})$$

is a good linear approximation to $\mathbf{f}$ near $\mathbf{a}$. That is, we must have

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - \mathbf{h}(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{a}\|} = \lim_{\mathbf{x}\to\mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x}) - [\mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a})]\|}{\|\mathbf{x} - \mathbf{a}\|} = 0.$$

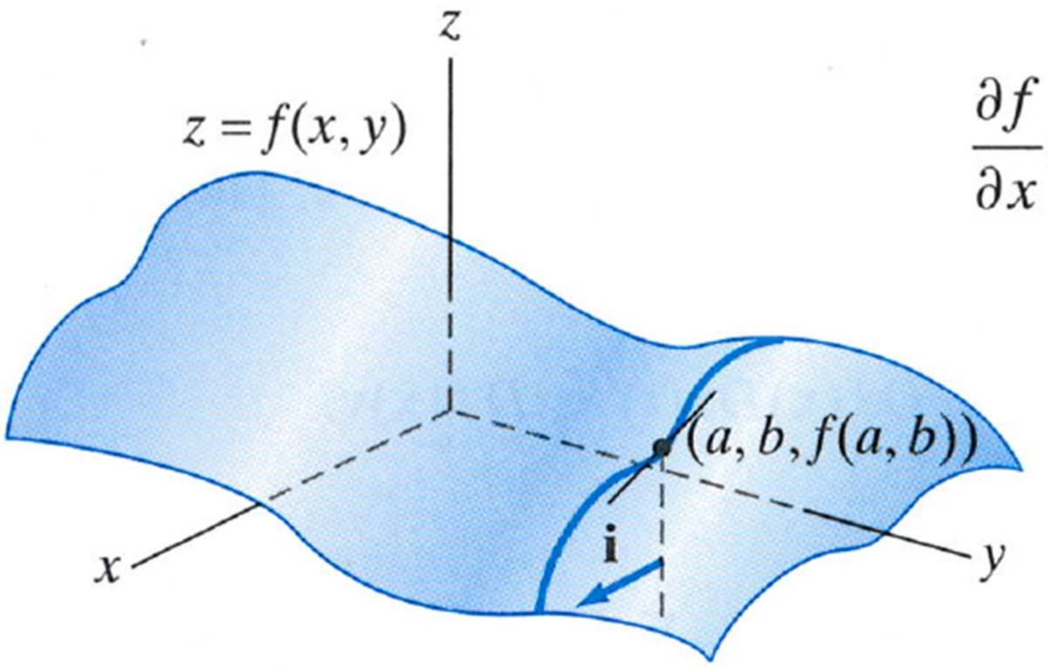

Figure 2.67 Another way to view the partial derivative $\partial f/\partial x$ at a point.

$$\frac{\partial f}{\partial x}(a, b) = \lim_{h\to 0} \frac{f(a+h, b) - f(a, b)}{h}$$
$$= \lim_{h\to 0} \frac{f(\mathbf{a} + h\mathbf{i}) - f(\mathbf{a})}{h}.$$

where $\mathbf{a} = (a, b) = a\mathbf{i} + b\mathbf{j}$.

"the directional derivative of f at $\mathbf{a}$ in the direction of $\mathbf{v}$"

$$\lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h}$$

where $\mathbf{v} = (A, B) = A\mathbf{i} + B\mathbf{j}$

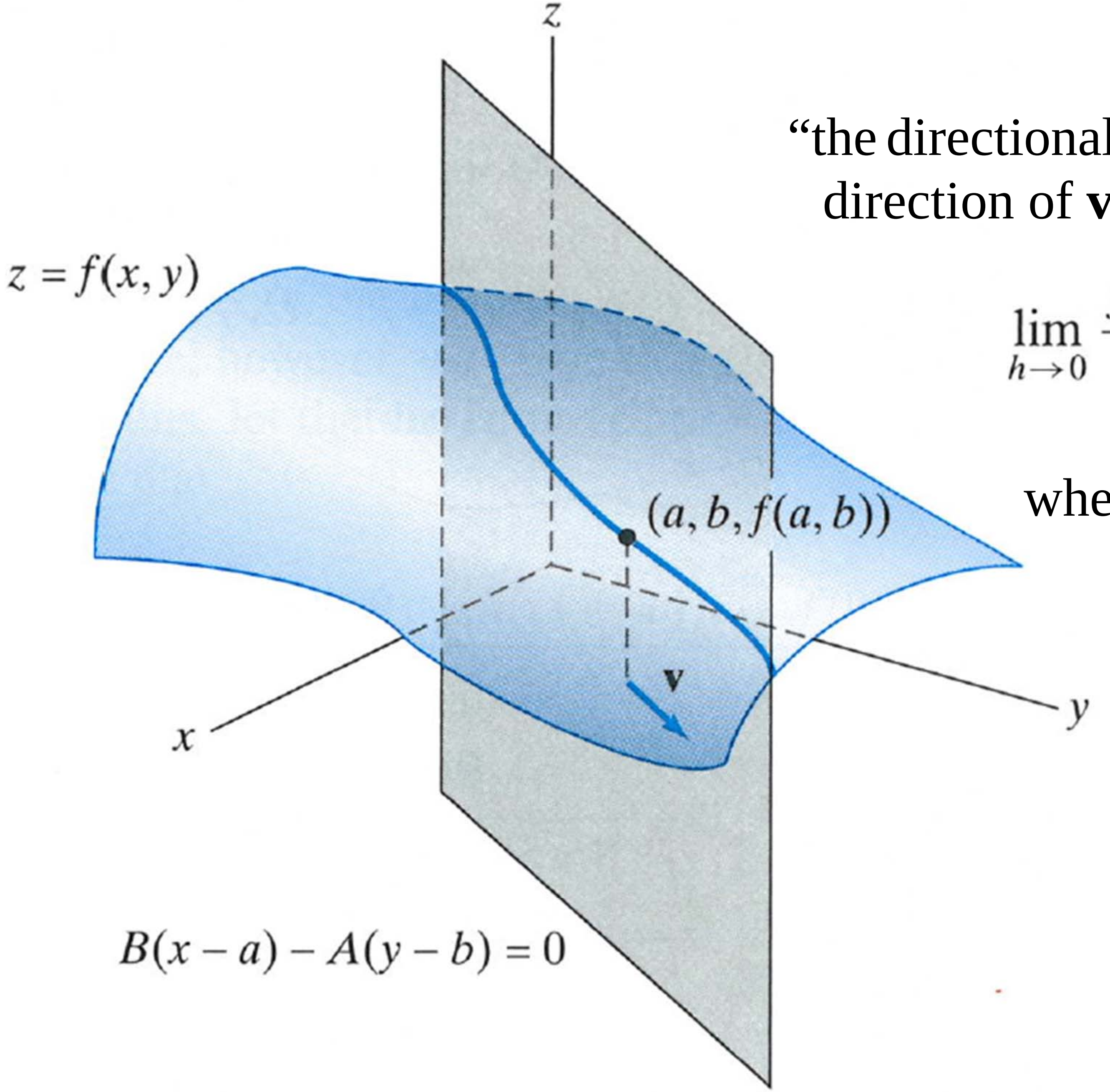

Figure 2.68 The directional derivative.

DEFINITION 6.1 Let X be open in $\mathbf{R}^n$, $f: X \subseteq \mathbf{R}^n \to \mathbf{R}$ a scalar-valued function, and $\mathbf{a} \in X$. If $\mathbf{v} \in \mathbf{R}^n$ is any unit vector, then the **directional derivative of f at $\mathbf{a}$ in the direction of $\mathbf{v}$**, denoted $D_{\mathbf{v}} f(\mathbf{a})$, is

$$D_{\mathbf{v}} f(\mathbf{a}) = \lim_{h \to 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h}$$

(provided that this limit exists).

THEOREM 6.2 Let $X \subseteq \mathbf{R}^n$ be open and suppose $f\colon X \to \mathbf{R}$ is differentiable at $\mathbf{a} \in X$. Then the directional derivative $D_{\mathbf{v}} f(\mathbf{a})$ exists for all directions (unit vectors) $\mathbf{v} \in \mathbf{R}^n$ and, moreover, we have

$$D_{\mathbf{v}} f(\mathbf{a}) = \nabla f(\mathbf{a}) \cdot \mathbf{v}.$$

THEOREM 6.3 The directional derivative $D_{\mathbf{u}} f(\mathbf{a})$ is maximized, with respect to direction, when $\mathbf{u}$ points in the *same* direction as $\nabla f(\mathbf{a})$ and is minimized when $\mathbf{u}$ points in the *opposite* direction. Furthermore, the maximum and minimum values of $D_{\mathbf{u}} f(\mathbf{a})$ are $\|\nabla f(\mathbf{a})\|$ and $-\|\nabla f(\mathbf{a})\|$, respectively.

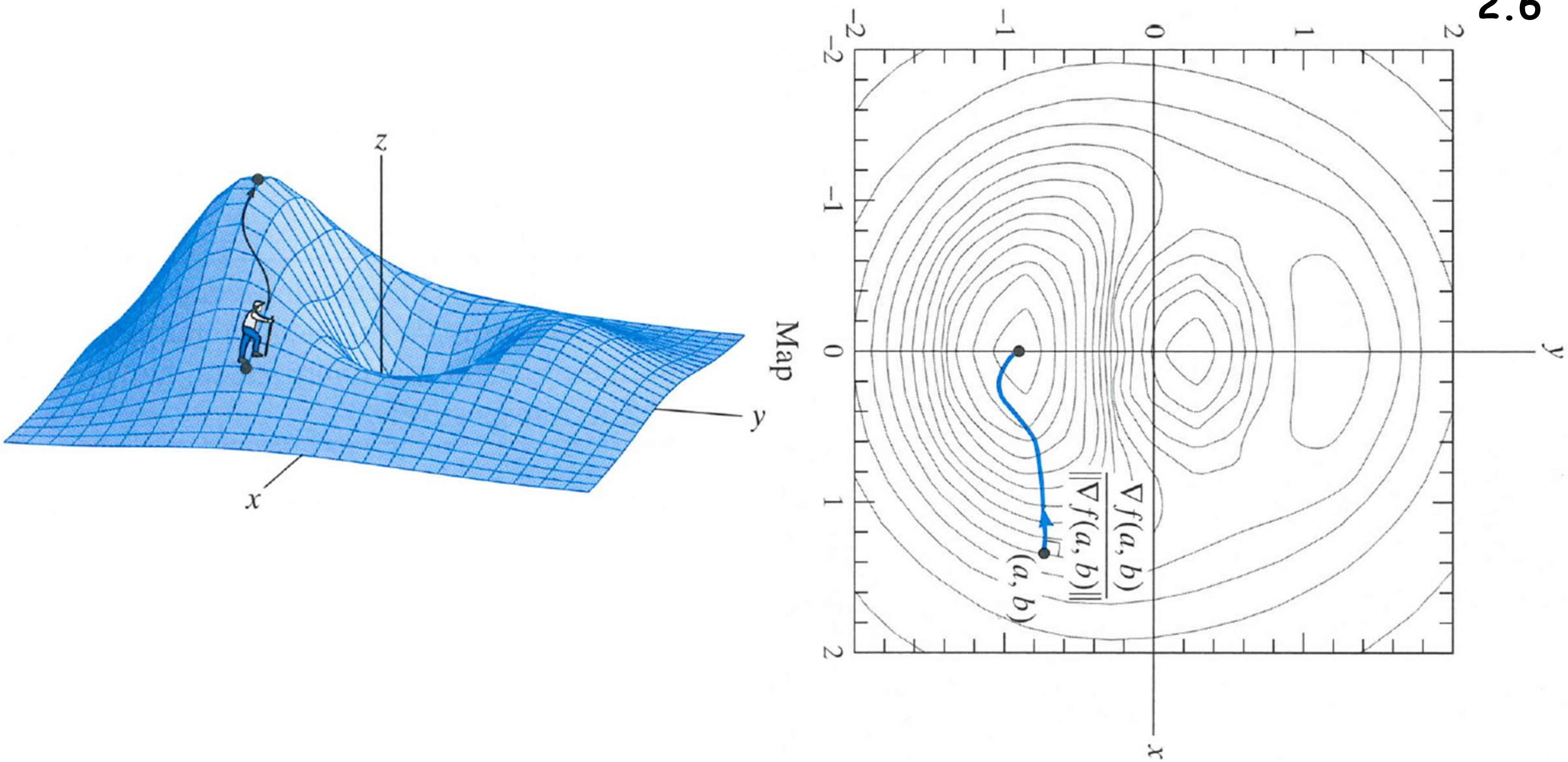

Figure 2.69 Select $\nabla f(a, b)/\|\nabla f(a, b)\|$ for direction of steepest ascent.

THEOREM 6.4 Let $X \subseteq \mathbf{R}^n$ be open and $f: X \to \mathbf{R}$ be a function of class C^1. If $\mathbf{x}_0$ is a point on the level set $S = \{\mathbf{x} \in X \mid f(\mathbf{x}) = c\}$, then the vector $\nabla f(\mathbf{x}_0)$ is perpendicular to S.

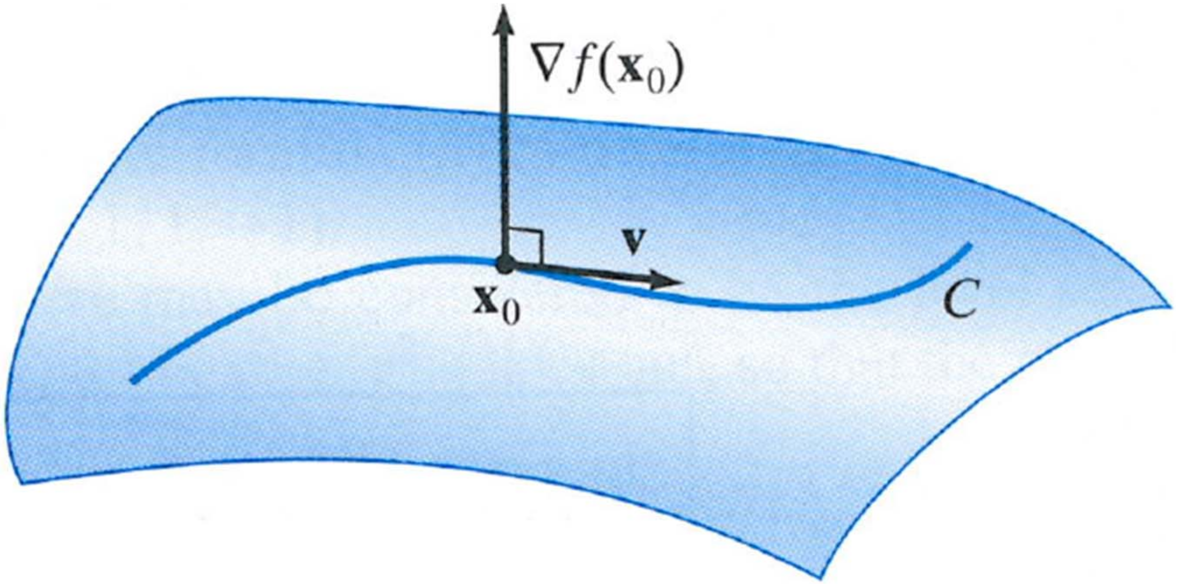

Figure 2.71 The level set surface $S = \{\mathbf{x} \mid f(\mathbf{x}) = c\}$.

In general, if S is a surface in $\mathbf{R}^3$ defined by an equation of the form

$$f(x, y, z) = c,$$

then if $\mathbf{x}_0 \in X$, the gradient vector $\nabla f(\mathbf{x}_0)$ is perpendicular to S and, consequently, if nonzero, is a vector normal to the plane tangent to S at $\mathbf{x}_0$. Thus, the equation

$$\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = 0 \tag{5}$$

or, equivalently,

$$\begin{aligned} f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) \\ + f_z(x_0, y_0, z_0)(z - z_0) = 0 \end{aligned} \tag{6}$$

is an equation for the tangent plane to S at $\mathbf{x}_0$.

PROPOSITION 4.1 (LINEARITY OF DIFFERENTIATION) Let $\mathbf{f}, \mathbf{g}\colon X \subseteq \mathbf{R}^n \to \mathbf{R}^m$ be two functions that are both differentiable at a point $\mathbf{a} \in X$, and let $c \in \mathbf{R}$ be any scalar. Then

1. The function $\mathbf{h} = \mathbf{f} + \mathbf{g}$ is also differentiable at $\mathbf{a}$, and we have
$$D\mathbf{h}(\mathbf{a}) = D(\mathbf{f} + \mathbf{g})(\mathbf{a}) = D\mathbf{f}(\mathbf{a}) + D\mathbf{g}(\mathbf{a}).$$
2. The function $\mathbf{k} = c\mathbf{f}$ is differentiable at $\mathbf{a}$ and
$$D\mathbf{k}(\mathbf{a}) = D(c\mathbf{f})(\mathbf{a}) = cD\mathbf{f}(\mathbf{a}).$$

PROPOSITION 4.2 Let $f, g\colon X \subseteq \mathbf{R}^n \to \mathbf{R}$ be differentiable at $\mathbf{a} \in X$. Then

1. The product function fg is also differentiable at $\mathbf{a}$, and
$$D(fg)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a}).$$
2. If $g(\mathbf{a}) \neq 0$, then the quotient function f/g is differentiable at $\mathbf{a}$, and
$$D(f/g)(\mathbf{a}) = \frac{g(\mathbf{a})Df(\mathbf{a}) - f(\mathbf{a})Dg(\mathbf{a})}{g(\mathbf{a})^2}.$$

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} (\text{or } f_{x_1}, f_{x_2}, \ldots, f_{x_n}).$$

$$\frac{\partial^2 f}{\partial x_i^2} = \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right). \quad \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right). \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

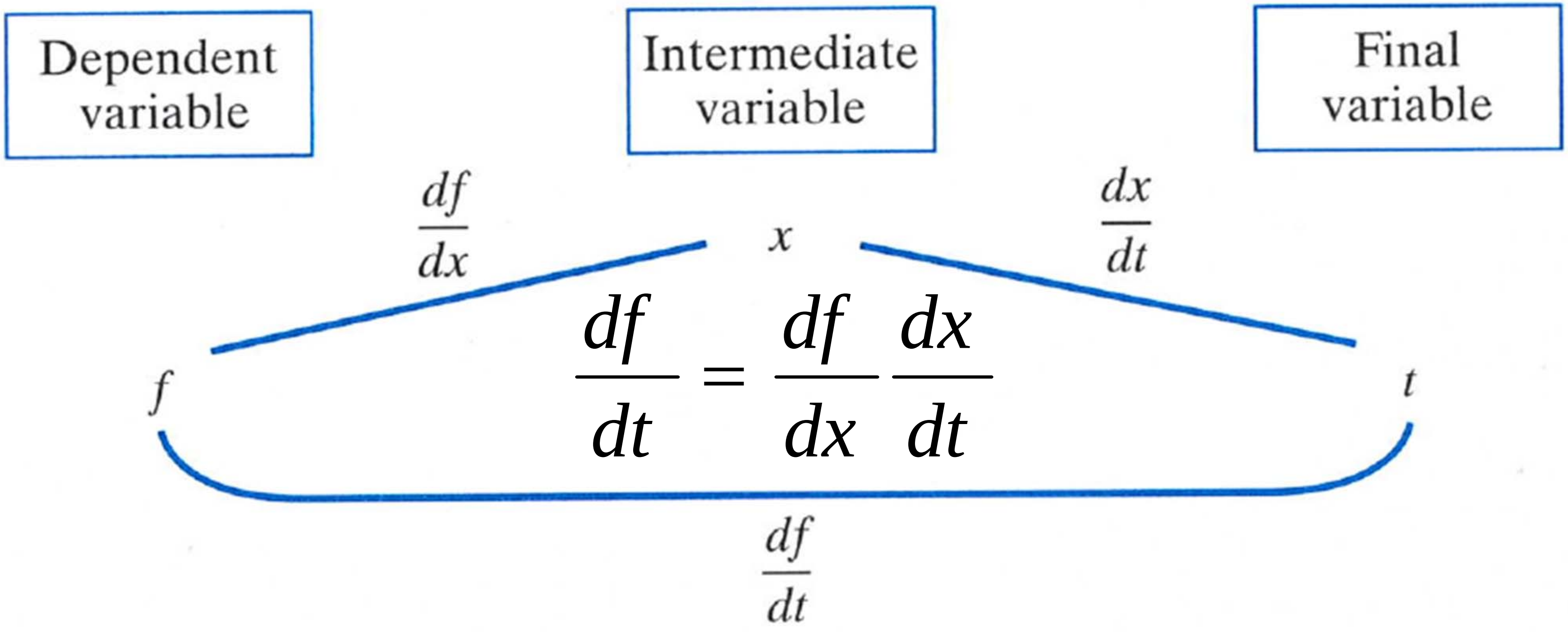

Figure 2.58 The chain rule for functions of a single variable.

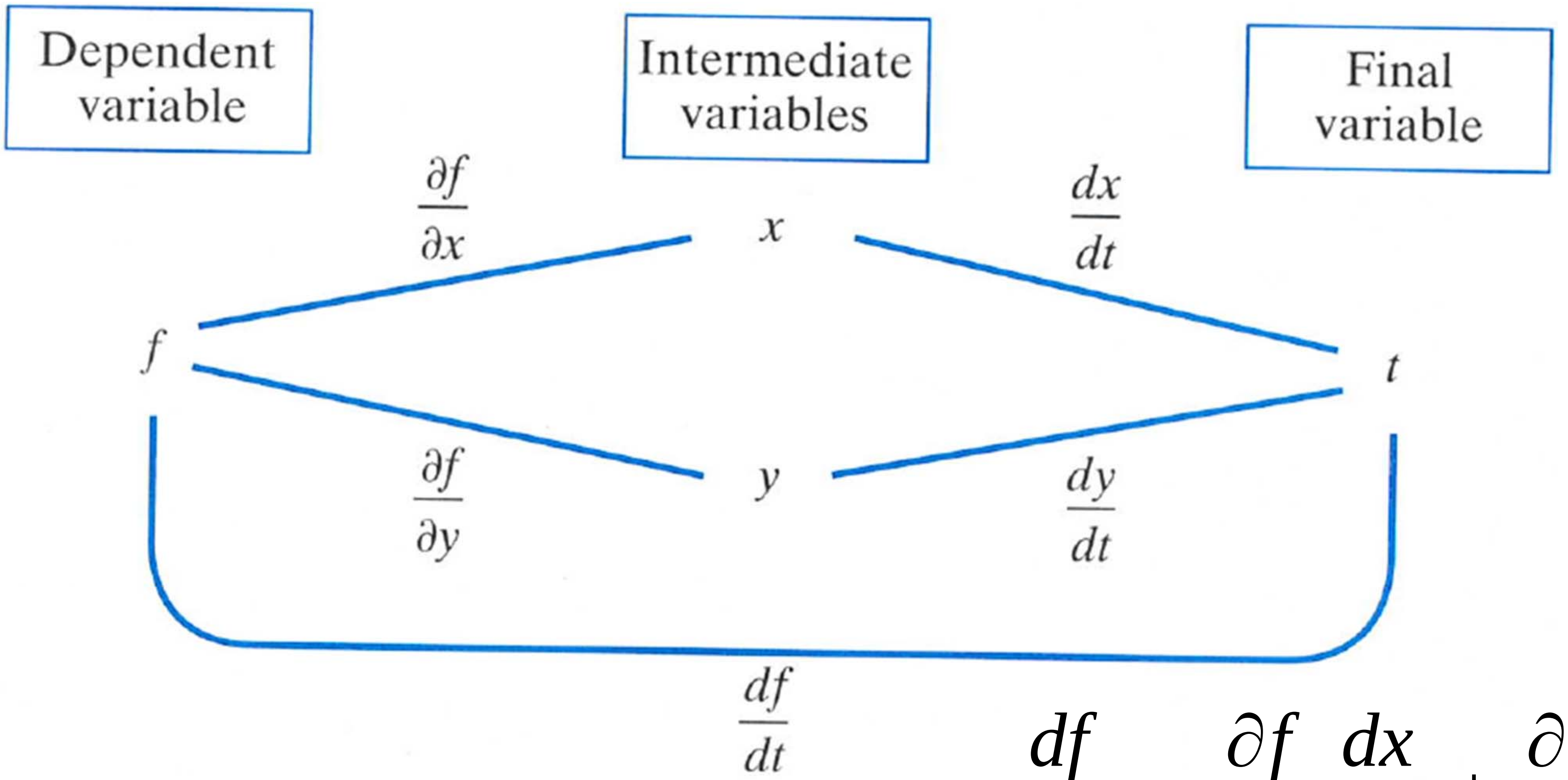

Figure 2.60 The chain rule of Proposition 5.2.

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

$$\frac{df}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial f}{\partial x_2}\frac{dx_2}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}$$

$$= DfDx = \nabla f \cdot \vec{x}\,'$$

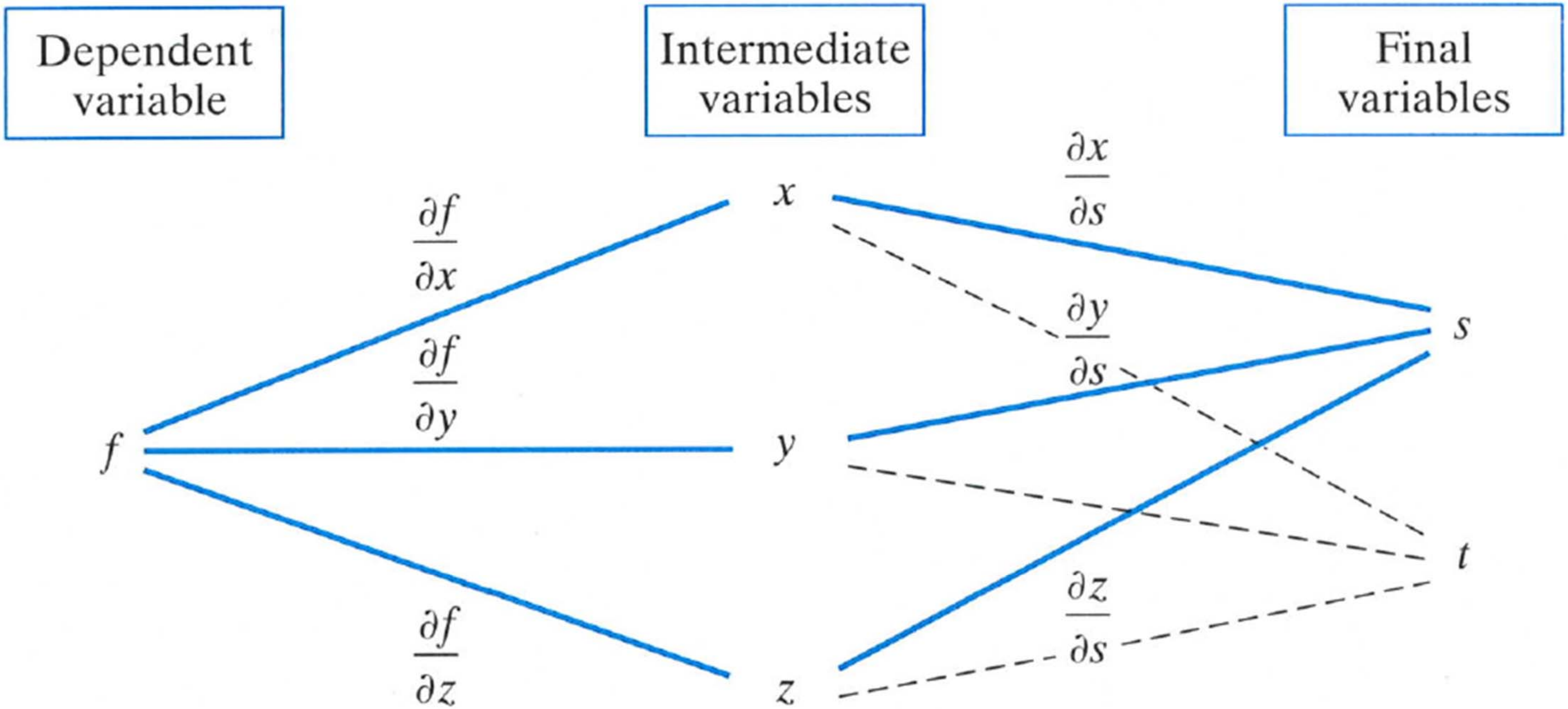

Figure 2.63 The chain rule for $f \circ \mathbf{x}$, where $f\colon X \subseteq \mathbf{R}^3 \to \mathbf{R}$ and $\mathbf{x}\colon T \subseteq \mathbf{R}^2 \to \mathbf{R}^3$.

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

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial t}$$

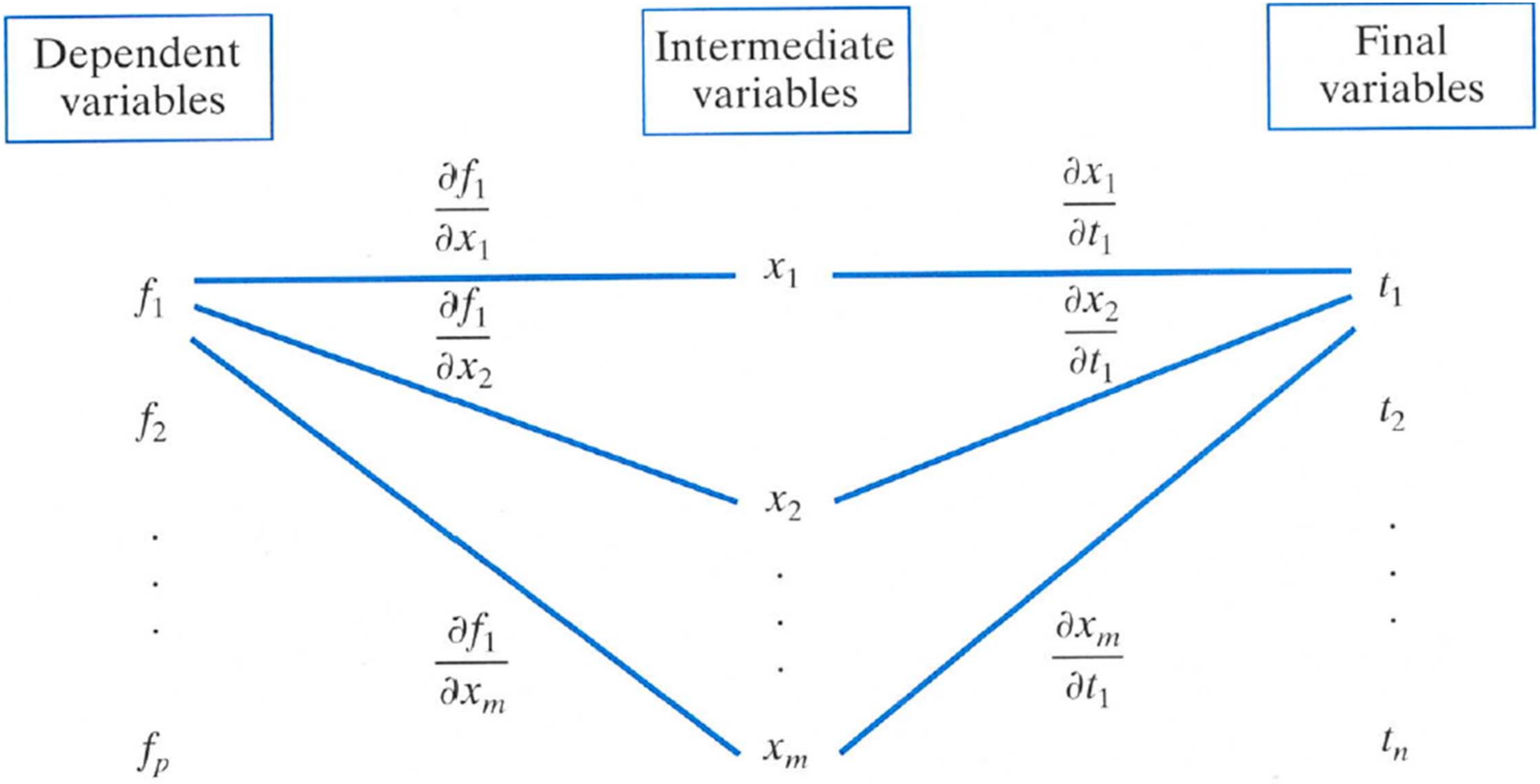

Figure 2.65 The chain rule diagram for $\mathbf{f} \circ \mathbf{x}$, where $\mathbf{f}: X \subseteq \mathbf{R}^m \to \mathbf{R}^p$ and $\mathbf{x}: T \subseteq \mathbf{R}^n \to \mathbf{R}^m$.

$$\frac{\partial f_1}{\partial t_2} = \frac{\partial f_1}{\partial x_1}\frac{\partial x_1}{\partial t_2} + \frac{\partial f_1}{\partial x_2}\frac{\partial x_2}{\partial t_2} + \cdots + \frac{\partial f_1}{\partial x_m}\frac{\partial x_m}{\partial t_2}$$

$$D\vec{h}(\vec{t}\,) = D(\vec{f} \circ \vec{x})(\vec{t}\,) = D\vec{f}\,(\vec{x})D\vec{x}(\vec{t}\,)$$