# CHAPTER 13

# Partial Derivatives

This chapter is at the center of multidimensional calculus. Other chapters and other topics may be optional; this chapter and these topics are required. We are back to the basic idea of calculus—***the derivative***. There is a function $f$, the variables move a little bit, and $f$ moves. The question is how much $f$ moves and how fast. Chapters 1–4 answered this question for $f(x)$, a function of one variable. Now we have $f(x, y)$ or $f(x, y, z)$—with two or three or more variables that move independently. As $x$ and $y$ change, $f$ changes. The fundamental problem of differential calculus is to connect $\Delta x$ and $\Delta y$ to $\Delta f$.

Calculus solves that problem in the limit. ***It connects dx and dy to df.*** In using this language I am building on the work already done. You know that $df/dx$ is the limit of $\Delta f/\Delta x$. Calculus computes the rate of change—which is the slope of the tangent line. The goal is to extend those ideas to

$$f(x, y) = x^2 - y^2 \qquad \text{or} \qquad f(x, y) = \sqrt{x^2 + y^2} \qquad \text{or} \qquad f(x, y, z) = 2x + 3y + 4z.$$

These functions have graphs, they have derivatives, and they must have tangents.

The heart of this chapter is summarized in six lines. The subject is *differential* calculus—small changes in a short time. Still to come is *integral* calculus—adding up those small changes. We give the words and symbols for $f(x, y)$, matched with the words and symbols for $f(x)$. Please use this summary as a guide, to know where calculus is going.

***Curve*** $y = f(x)$ *vs.* ***Surface*** $z = f(x, y)$

| | | |
|---|---|---|
| $\dfrac{df}{dx}$ | becomes two partial derivatives | $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ |
| $\dfrac{d^2f}{dx^2}$ | becomes four second derivatives | $\dfrac{\partial^2 f}{\partial x^2}, \dfrac{\partial^2 f}{\partial x \partial y}, \dfrac{\partial^2 f}{\partial y \partial x}, \dfrac{\partial^2 f}{\partial y^2}$ |
| $\Delta f \approx \dfrac{df}{dx}\Delta x$ | becomes the linear approximation | $\Delta f \approx \dfrac{\partial f}{\partial x}\Delta x + \dfrac{\partial f}{\partial y}\Delta y$ |
| tangent line | becomes the tangent plane | $z - z_0 = \dfrac{\partial f}{\partial x}(x - x_0) + \dfrac{\partial f}{\partial y}(y - y_0)$ |
| $\dfrac{dy}{dt} = \dfrac{dy}{dx}\dfrac{dx}{dt}$ | becomes the chain rule | $\dfrac{dz}{dt} = \dfrac{\partial z}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial z}{\partial y}\dfrac{dy}{dt}$ |
| $\dfrac{df}{dx} = 0$ | becomes two maximum-minimum equations | $\dfrac{\partial f}{\partial x} = 0$ and $\dfrac{\partial f}{\partial y} = 0$. |

## 13.1 Surfaces and Level Curves

The graph of $y=f(x)$ is a curve in the $xy$ plane. There are two variables—$x$ is independent and free, $y$ is dependent on $x$. Above $x$ on the base line is the point $(x, y)$ on the curve. The curve can be displayed on a two-dimensional printed page.

The graph of $z=f(x, y)$ is a ***surface in*** $xyz$ ***space***. There are *three variables*—$x$ and $y$ are independent, $z$ is dependent. Above $(x, y)$ in the base plane is the point $(x, y, z)$ on the surface (Figure 13.1). Since the printed page remains two-dimensional, we shade or color or project the surface. The eyes are extremely good at converting two-dimensional images into three-dimensional understanding—they get a lot of practice. The mathematical part of our brain also has something new to work on—***two partial derivatives***.

This section uses examples and figures to illustrate surfaces and their level curves. The next section is also short. Then the work begins.

**EXAMPLE 1** Describe the ***surface*** and the ***level curves*** for $z=f(x, y)=\sqrt{x^2+y^2}$.

*The surface is a cone.* Reason: $\sqrt{x^2+y^2}$ is the distance in the base plane from $(0, 0)$ to $(x, y)$. When we go out a distance 5 in the base plane, we go up the same distance 5 to the surface. The cone climbs with slope 1. The distance out to $(x, y)$ equals the distance up to $z$ (this is a 45° cone).

*The level curves are circles.* At height 5, the cone contains a circle of points—all at the same "level" on the surface. The plane $z=5$ meets the surface $z=\sqrt{x^2+y^2}$ at those points (Figure 13.1b). The circle below them (in the base plane) is the level curve.

**DEFINITION** A ***level curve*** or ***contour line*** of $z=f(x, y)$ contains all points $(x, y)$ that share the *same value* $f(x, y)=c$. Above those points, the surface is at the height $z=c$.

There are different level curves for different $c$. To see the curve for $c=2$, cut through the surface with the horizontal plane $z=2$. The plane meets the surface above the points where $f(x, y)=2$. ***The level curve in the base plane has the equation*** $f(x, y)=2$. Above it are all the points at "level 2" or "level $c$" on the surface.

Every curve $f(x, y)=c$ is labeled by its constant $c$. This produces a ***contour map*** (the base plane is full of curves). For the cone, the level curves are given by $\sqrt{x^2+y^2}=c$, and the contour map consists of circles of radius $c$.

**Question** What are the level curves of $z=f(x, y)=x^2+y^2$?
Answer *Still circles.* But the surface is not a cone (it bends up like a parabola). The circle of radius 3 is the level curve $x^2+y^2=9$. On the surface above, the height is 9.

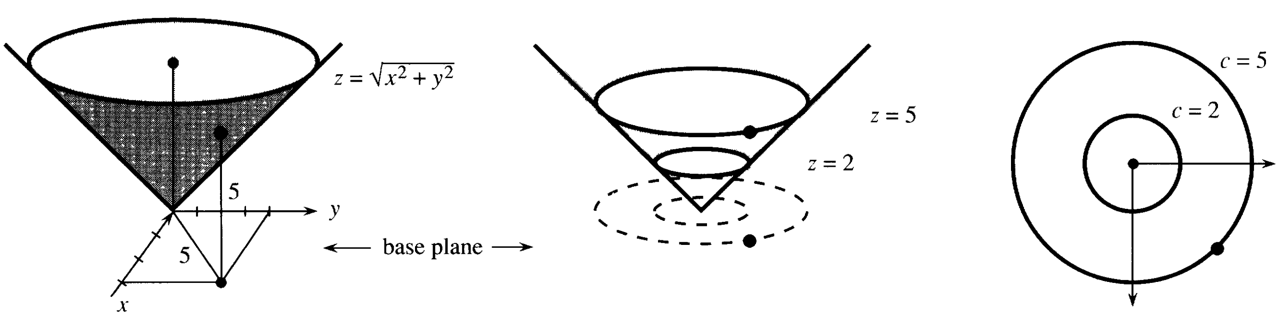



**Fig. 13.1** The surface for $z=f(x, y)=\sqrt{x^2+y^2}$ is a cone. The level curves are circles.

**EXAMPLE 2** For the *linear function* $f(x, y) = 2x + y$, the surface is a plane. Its level curves are straight lines. ***The surface*** $z = 2x + y$ ***meets the plane*** $z = c$ ***in the line*** $2x + y = c$. That line is above the base plane when $c$ is positive, and below when $c$ is negative. The contour lines are *in* the base plane. Figure 13.2b labels these parallel lines according to their height in the surface.

**Question** If the level curves are all straight lines, must they be parallel?
**Answer** No. The surface $z = y/x$ has level curves $y/x = c$. Those lines $y = cx$ swing around the origin, as the surface climbs like a spiral playground slide.

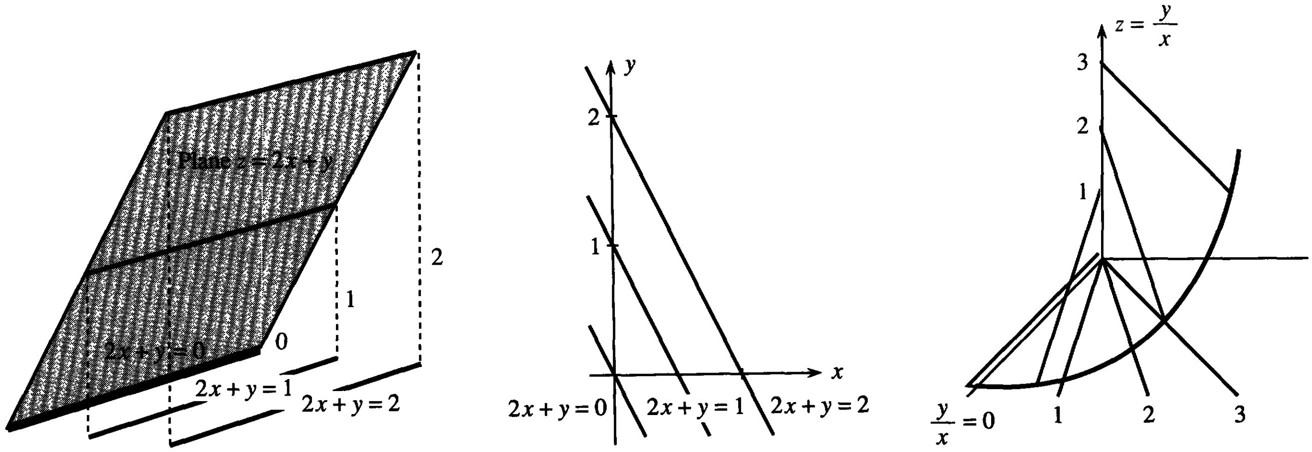



**Fig. 13.2** A plane has parallel level lines. The spiral slide $z = y/x$ has lines $y/x = c$.

**EXAMPLE 3** The weather map shows contour lines of the *temperature function*. Each level curve connects points at a constant temperature. One line runs from Seattle to Omaha to Cincinnati to Washington. In winter it is painful even to think about the line through L.A. and Texas and Florida. *USA Today* separates the contours by color, which is better. We had never seen a map of universities.

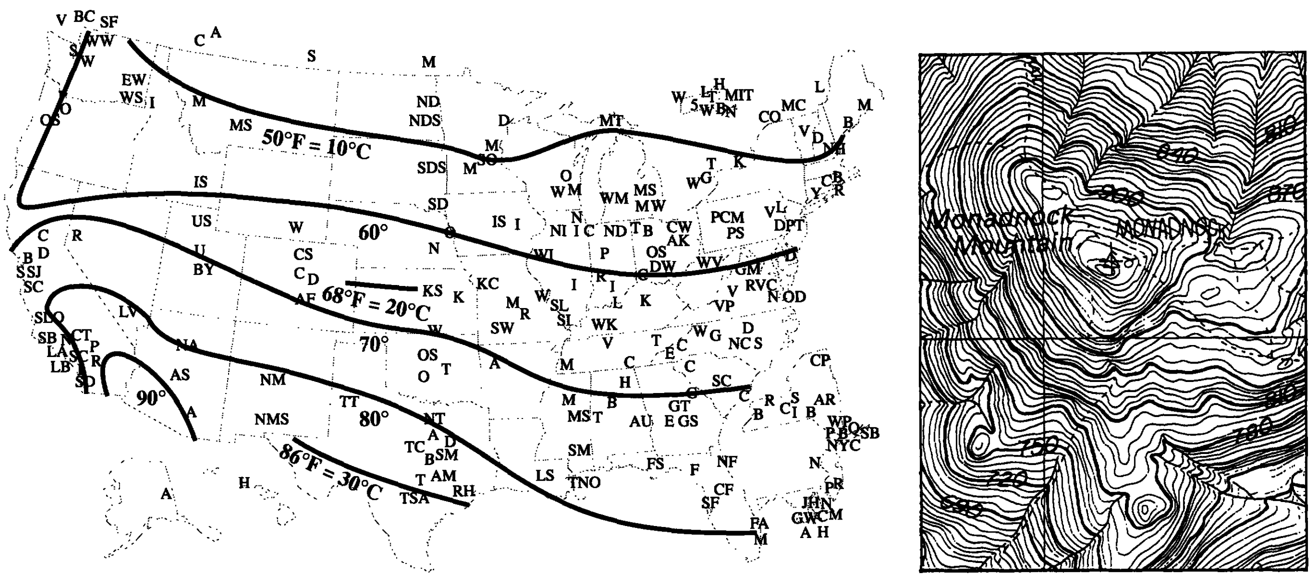



**Fig. 13.3** The temperature at many U.S. and Canadian universities. Mt. Monadnock in New Hampshire is said to be the most climbed mountain (except Fuji?) at 125,000/year. Contour lines every 6 meters.

**Question** From a contour map, how do you find the highest point?
Answer The level curves form *loops* around the maximum point. As $c$ increases the loops become tighter. Similarly the curves squeeze to the lowest point as $c$ decreases.

**EXAMPLE 4** A contour map of a mountain may be the best example of all. Normally the level curves are separated by 100 feet in height. On a steep trail those curves are bunched together—the trail climbs quickly. In a flat region the contour lines are far apart. Water runs perpendicular to the level curves. On my map of New Hampshire that is true of creeks but looks doubtful for rivers.

**Question** Which direction in the base plane is uphill on the surface?
Answer The steepest direction is perpendicular to the level curves. This is important. Proof to come.

**EXAMPLE 5** In economics $x^2y$ is a *utility function* and $x^2y=c$ is an *indifference curve.*

The utility function $x^2y$ gives the value of $x$ hours awake and $y$ hours asleep. Two hours awake and fifteen minutes asleep have the value $f=(2^2)(\frac{1}{4})$. This is the same as one hour of each: $f=(1^2)(1)$. Those lie on the same level curve in Figure 13.4a. We are indifferent, and willing to exchange any two points on a level curve.

The indifference curve is "*convex.*" ***We prefer the average of any two points.*** The line between two points is up on higher level curves.

Figure 13.4b shows an extreme case. The level curves are straight lines $4x+y=c$. Four quarters are freely substituted for one dollar. The value is $f=4x+y$ dollars.

Figure 13.4c shows the other extreme. Extra left shoes or extra right shoes are useless. The value (or utility) is the *smaller* of $x$ and $y$. That counts *pairs* of shoes.

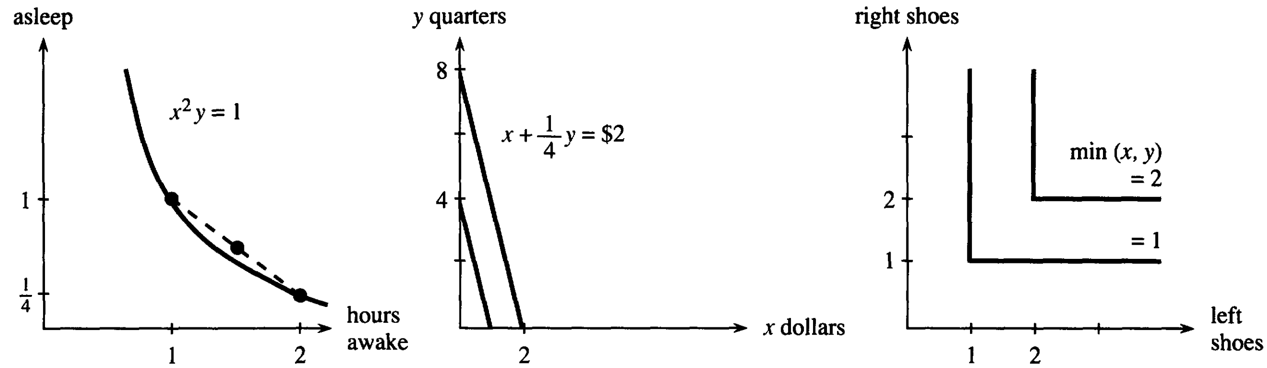



**Fig. 13.4** Utility functions $x^2y$, $4x+y$, $\min(x,y)$. Convex, straight substitution, complements.

## 13.1 EXERCISES

**Read-through questions**

The graph of $z=f(x,y)$ is a __a__ in __b__-dimensional space. The __c__ curve $f(x,y)=7$ lies down in the base plane. Above this level curve are all points at height __d__ in the surface. The __e__ $z=7$ cuts through the surface at those points. The level curves $f(x,y)=$ __f__ are drawn in the $xy$ plane and labeled by __g__. The family of labeled curves is a __h__ map.

For $z=f(x,y)=x^2-y^2$, the equation for a level curve is __i__. This curve is a __j__. For $z=x-y$ the curves are __k__. *Level curves never cross because* __l__. They crowd together when the surface is __m__. The curves tighten to a point when __n__. The steepest direction on a mountain is __o__ to the __p__.

1 Draw the surface $z = f(x, y)$ for these four functions:

$$f_1 = \sqrt{4 - x^2 - y^2} \qquad f_2 = 2 - \sqrt{x^2 + y^2}$$
$$f_3 = 2 - \tfrac{1}{2}(x^2 + y^2) \qquad f_4 = 1 + e^{-x^2-y^2}$$

2 The level curves of all four functions are ________. They enclose the maximum at ________. Draw the four curves $f(x, y) = 1$ and rank them by increasing radius.

3 Set $y = 0$ and compute the $x$ derivative of each function at $x = 2$. Which mountain is flattest and which is steepest at that point?

4 Set $y = 1$ and compute the $x$ derivative of each function at $x = 1$.

**For $f_5$ to $f_{10}$ draw the level curves $f = 0, 1, 2$. Also $f = -4$.**

5 $f_5 = x - y$ 6 $f_6 = (x + y)^2$

7 $f_7 = xe^{-y}$ 8 $f_8 = \sin(x - y)$

9 $f_9 = y - x^2$ 10 $f_{10} = y/x^2$

11 Suppose the level curves are parallel straight lines. Does the surface have to be a plane?

12 Construct a function whose level curve $f = 0$ is in two separate pieces.

13 Construct a function for which $f = 0$ is a circle and $f = 1$ is not.

14 Find a function for which $f = 0$ has infinitely many pieces.

15 Draw the contour map for $f = xy$ with level curves $f = -2, -1, 0, 1, 2$. Describe the surface.

16 Find a function $f(x, y)$ whose level curve $f = 0$ consists of a circle and *all points inside it*.

**Draw two level curves in 17–20. Are they ellipses, parabolas, or hyperbolas? Write $\sqrt{\phantom{xx}} - 2x = c$ as $\sqrt{\phantom{xx}} = c + 2x$ before squaring both sides.**

17 $f = \sqrt{4x^2 + y^2}$ 18 $f = \sqrt{4x^2 + y^2} - 2x$

19 $f = \sqrt{5x^2 + y^2} - 2x$ 20 $f = \sqrt{3x^2 + y^2} - 2x$

21 The level curves of $f = (y - 2)/(x - 1)$ are ________ through the point (1, 2) except that this point is not ________.

22 Sketch a map of the US with lines of constant temperature (isotherms) based on today's paper.

23 (a) The contour lines of $z = x^2 + y^2 - 2x - 2y$ are circles around the point ________, where $z$ is a minimum.
(b) The contour lines of $f =$ ________ are the circles $x^2 + y^2 = c + 1$ on which $f = c$.

24 Draw a contour map of any state or country (lines of constant height above sea level). Florida may be too flat.

25 The graph of $w = F(x, y, z)$ is a ________-dimensional surface in $xyzw$ space. Its level sets $F(x, y, z) = c$ are ________-dimensional surfaces in $xyz$ space. For $w = x - 2y + z$ those level sets are ________. For $w = x^2 + y^2 + z^2$ those level sets are ________.

26 The surface $x^2 + y^2 - z^2 = -1$ is in Figure 13.8. There is empty space when $z^2$ is smaller than 1 because ________.

27 The level sets of $F = x^2 + y^2 + qz^2$ look like footballs when $q$ is ________, like basketballs when $q$ is ________, and like frisbees when $q$ is ________.

28 Let $T(x, y)$ be the driving time from your home at (0, 0) to nearby towns at $(x, y)$. Draw the level curves.

29 (a) The level curves of $f(x, y) = \sin(x - y)$ are ________.
(b) The level curves of $g(x, y) = \sin(x^2 - y^2)$ are ________.
(c) The level curves of $h(x, y) = \sin(x - y^2)$ are ________.

30 Prove that if $x_1y_1 = 1$ and $x_2y_2 = 1$ then their average $x = \tfrac{1}{2}(x_1 + x_2)$, $y = \tfrac{1}{2}(y_1 + y_2)$ has $xy \geqslant 1$. The function $f = xy$ has convex level curves (hyperbolas).

31 The hours in a day are limited by $x + y = 24$. Write $x^2y$ as $x^2(24 - x)$ and maximize to find the optimal number of hours to stay awake.

32 Near $x = 16$ draw the level curve $x^2y = 2048$ and the line $x + y = 24$. Show that the curve is convex and the line is tangent.

33 The surface $z = 4x + y$ is a ________. The surface $z = \min(x, y)$ is formed from two ________. We are willing to exchange 6 left and 2 right shoes for 2 left and 4 right shoes but better is the average ________.

34 Draw a contour map of the top of your shoe.

## 13.2 Partial Derivatives

The central idea of differential calculus is the derivative. A change in $x$ produces a change in $f$. The ratio $\Delta f/\Delta x$ approaches the derivative, or slope, or rate of change. What to do if $f$ depends on both $x$ and $y$?

The new idea is to vary $x$ and $y$ one at a time. First, only $x$ moves. If the function is $x + xy$, then $\Delta f$ is $\Delta x + y\Delta x$. The ratio $\Delta f/\Delta x$ is $1 + y$. The "$x$ derivative" of $x + xy$

is $1 + y$. For all functions the method is the same: ***Keep y constant, change x, take the limit of*** $\Delta f/\Delta x$:

**DEFINITION**
$$\frac{\partial f}{\partial x}(x, y) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}. \tag{1}$$

On the left is a new symbol $\partial f/\partial x$. It signals that only $x$ is allowed to vary—$\partial f/\partial x$ is a ***partial derivative***. The different form $\partial$ of the same letter (still say "$d$") is a reminder that $x$ is not the only variable. Another variable $y$ is present but not moving.

**EXAMPLE 1** $f(x, y) = x^2y^2 + xy + y \qquad \frac{\partial f}{\partial x}(x, y) = 2xy^2 + y + 0.$

***Do not treat y as zero!*** Treat it as a constant, like 6. Its $x$ *derivative* is zero. If $f(x) = \sin 6x$ then $df/dx = 6 \cos 6x$. If $f(x, y) = \sin xy$ then $\partial f/\partial x = y \cos xy$.

Spoken aloud, $\partial f/\partial x$ is still "$d f d x$." It is a function of $x$ and $y$. When more is needed, call it "the partial of $f$ with respect to $x$." The symbol $f'$ is no longer available, since it gives no special indication about $x$. Its replacement $f_x$ is pronounced "$f x$" or "$f$ sub $x$," which is shorter than $\partial f/\partial x$ and means the same thing.

We may also want to indicate the point $(x_0, y_0)$ where the derivative is computed:

$$\frac{\partial f}{\partial x}(x_0, y_0) \quad \text{or} \quad f_x(x_0, y_0) \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{(x_0, y_0)} \quad \text{or just} \quad \left(\frac{\partial f}{\partial x}\right)_0.$$

**EXAMPLE 2** $f(x, y) = \sin 2x \cos y \qquad f_x = 2 \cos 2x \cos y \qquad (\cos y$ is constant for $\partial/\partial x)$

The particular point $(x_0, y_0)$ is $(0, 0)$. The height of the surface is $f(0, 0) = 0$. The slope in the $x$ direction is $f_x = 2$. At a different point $x_0 = \pi$, $y_0 = \pi$ we find $f_x(\pi, \pi) = -2$.

***Now keep x constant and vary y.*** The ratio $\Delta f/\Delta y$ approaches $\partial f/\partial y$:

$$f_y(x, y) = \lim_{\Delta y \to 0} \frac{\Delta f}{\Delta y} = \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}. \tag{2}$$

This is the slope in the $y$ direction. Please realize that a surface can go up in the $x$ direction and down in the $y$ direction. The plane $f(x, y) = 3x - 4y$ has $f_x = 3$ (up) and $f_y = -4$ (down). We will soon ask what happens in the 45° direction.

**EXAMPLE 3** $f(x, y) = \sqrt{x^2 + y^2} \qquad \frac{\partial f}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}} = \frac{x}{f} \qquad \frac{\partial f}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}} = \frac{y}{f}.$

The $x$ derivative of $\sqrt{x^2 + y^2}$ is really one-variable calculus, because $y$ is constant. The exponent drops from $\frac{1}{2}$ to $-\frac{1}{2}$, and there is $2x$ from the chain rule. ***This distance function has the curious derivative*** $\partial f/\partial x = x/f$.

The graph is a cone. Above the point $(0, 2)$ the height is $\sqrt{0^2 + 2^2} = 2$. The partial derivatives are $f_x = 0/2$ and $f_y = 2/2$. At that point, Figure 13.5 climbs in the $y$ direction. It is level in the $x$ direction. An actual step $\Delta x$ will increase $0^2 + 2^2$ to $(\Delta x)^2 + 2^2$. But this change is of order $(\Delta x)^2$ and the $x$ derivative is zero.

Figure 13.5 is rather important. It shows how $\partial f/\partial x$ and $\partial f/\partial y$ are the ordinary derivatives of $f(x, y_0)$ and $f(x_0, y)$. It is natural to call these ***partial functions***. The first has $y$ fixed at $y_0$ while $x$ varies. The second has $x$ fixed at $x_0$ while $y$ varies. Their graphs are ***cross sections down the surface***—cut out by the vertical planes $y = y_0$ and $x = x_0$. Remember that the level curve is cut out by the horizontal plane $z = c$.

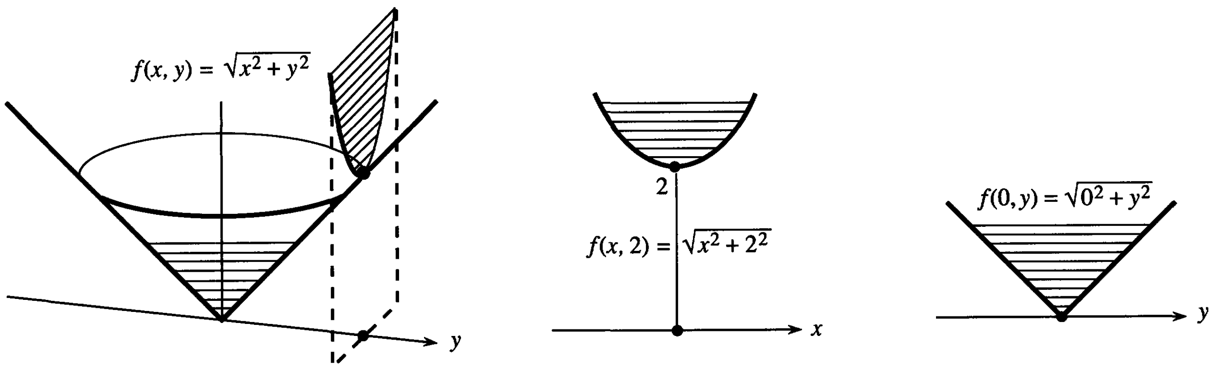



**Fig. 13.5** Partial functions $\sqrt{x^2+2^2}$ and $\sqrt{0^2+y^2}$ of the distance function $f = \sqrt{x^2+y^2}$.

The limits of $\Delta f/\Delta x$ and $\Delta f/\Delta y$ are computed as always. With partial functions we are back to a single variable. ***The partial derivative is the ordinary derivative of a partial function*** (constant $y$ or constant $x$). For the cone, $\partial f/\partial y$ exists at all points except (0, 0). The figure shows how the cross section down the middle of the cone produces the absolute value function: $f(0, y) = |y|$. It has one-sided derivatives but not a two-sided derivative.

Similarly $\partial f/\partial x$ will not exist at the sharp point of the cone. We develop the idea of a *continuous function* $f(x, y)$ as needed (the definition is in the exercises). Each partial derivative involves one direction, but limits and continuity involve all directions. The distance function is continuous at (0, 0), where it is not differentiable.

**EXAMPLE 4** $f(x, y) = y^2 - x^2 \qquad \partial f/\partial x = -2x \qquad \partial f/\partial y = 2y$

Move in the $x$ direction from (1, 3). Then $y^2 - x^2$ has the partial function $9 - x^2$. With $y$ fixed at 3, a parabola opens downward. In the $y$ direction (along $x = 1$) the partial function $y^2 - 1$ opens upward. The surface in Figure 13.6 is called a *hyperbolic paraboloid*, because the level curves $y^2 - x^2 = c$ are hyperbolas. Most people call it a saddle, and the special point at the origin is a ***saddle point***.

The origin is special for $y^2 - x^2$ because both derivatives are zero. ***The bottom of the y parabola at*** (0, 0) ***is the top of the x parabola***. The surface is momentarily flat in all directions. It is the top of a hill and the bottom of a mountain range at the same

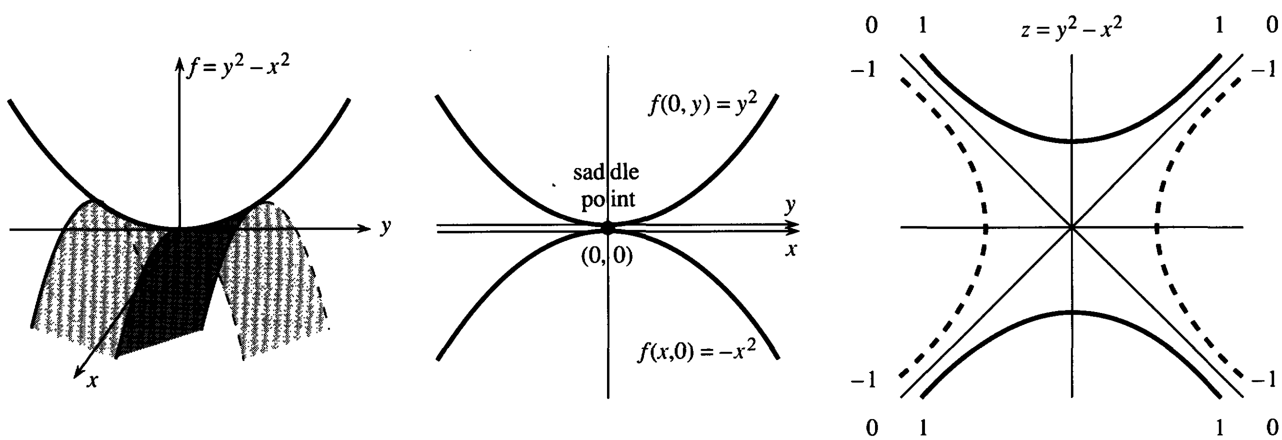



**Fig. 13.6** A saddle function, its partial functions, and its level curves.

time. A saddle point is neither a maximum nor a minimum, although both derivatives are zero.

*Note* Do not think that $f(x, y)$ must contain $y^2$ and $x^2$ to have a saddle point. The function $2xy$ does just as well. The level curves $2xy = c$ are still hyperbolas. The partial functions $2xy_0$ and $2x_0y$ now give straight lines—which is remarkable. Along the 45° line $x = y$, the function is $2x^2$ and climbing. Along the $-45°$ line $x = -y$, the function is $-2x^2$ and falling. The graph of $2xy$ is Figure 13.6 rotated by 45°.

**EXAMPLES 5–6** $f(x, y, z) = x^2 + y^2 + z^2 \qquad P(T, V) = nRT/V$

Example 5 shows more variables. Example 6 shows that the variables may not be named $x$ and $y$. Also, the function may not be named $f$! Pressure and temperature and volume are $P$ and $T$ and $V$. The letters change but nothing else:

$$\partial P/\partial T = nR/V \qquad \partial P/\partial V = -nRT/V^2 \qquad \text{(note the derivative of } 1/V).$$

There is no $\partial P/\partial R$ because $R$ is a constant from chemistry—not a variable.

Physics produces six variables for a moving body—the coordinates $x, y, z$ and the momenta $p_x, p_y, p_z$. Economics and the social sciences do better than that. If there are 26 products there are 26 variables—sometimes 52, to show prices as well as amounts. The profit can be a complicated function of these variables. ***The partial derivatives are the marginal profits***, as one of the 52 variables is changed. A spreadsheet shows the 52 values and the effect of a change. An infinitesimal spreadsheet shows the derivative.

## SECOND DERIVATIVE

Genius is not essential, to move to second derivatives. The only difficulty is that *two* first derivatives $f_x$ and $f_y$ lead to *four* second derivatives $f_{xx}$ and $f_{xy}$ and $f_{yx}$ and $f_{yy}$. (Two subscripts: $f_{xx}$ is the $x$ derivative of the $x$ derivative. Other notations are $\partial^2 f/\partial x^2$ and $\partial^2 f/\partial x \partial y$ and $\partial^2 f/\partial y \partial x$ and $\partial^2 f/\partial y^2$.) Fortunately $f_{xy}$ equals $f_{yx}$, as we see first by example.

**EXAMPLE 7** $f = x/y$ has $f_x = 1/y$, which has $f_{xx} = 0$ and $f_{xy} = -1/y^2$.

The function $x/y$ is linear in $x$ (which explains $f_{xx} = 0$). Its $y$ derivative is $f_y = -x/y^2$. This has the $x$ derivative $f_{yx} = -1/y^2$. ***The mixed derivatives $f_{xy}$ and $f_{yx}$ are equal.***

In the pure $y$ direction, the second derivative is $f_{yy} = 2x/y^3$. One-variable calculus is sufficient for all these derivatives, because only one variable is moving.

**EXAMPLE 8** $f = 4x^2 + 3xy + y^2$ has $f_x = 8x + 3y$ and $f_y = 3x + 2y$.

Both "*cross derivatives*" $f_{xy}$ and $f_{yx}$ equal 3. The second derivative in the $x$ direction is $\partial^2 f/\partial x^2 = 8$ or $f_{xx} = 8$. Thus "$f\,x\,x$" is "$d$ second $f\,d\,x$ squared." Similarly $\partial^2 f/\partial y^2 = 2$. The only change is from $d$ to $\partial$.

***If $f(x, y)$ has continuous second derivatives then $f_{xy} = f_{yx}$.*** Problem 43 sketches a proof based on the Mean Value Theorem. For third derivatives almost any example shows that $f_{xxy} = f_{xyx} = f_{yxx}$ is different from $f_{yyx} = f_{yxy} = f_{xyy}$.

**Question** ***How do you plot a space curve $x(t), y(t), z(t)$ in a plane?*** One way is to look parallel to the direction (1, 1, 1). On your $XY$ screen, plot $X = (y - x)/\sqrt{2}$ and $Y = (2z - x - y)/\sqrt{6}$. The line $x = y = z$ goes to the point (0, 0)!

***How do you graph a surface*** $z = f(x, y)$? Use the same $X$ and $Y$. Fix $x$ and let $y$ vary, for curves one way in the surface. Then fix $y$ and vary $x$, for the other partial function. For a parametric surface like $x = (2 + v\sin\frac{1}{2}u)\cos u$, $y = (2 + v\sin\frac{1}{2}u)\sin u$, $z = v\cos\frac{1}{2}u$, vary $u$ and then $v$. Dick Williamson showed how this draws a one-sided "Möbius strip."

## 13.2 EXERCISES

**Read-through questions**

The __a__ derivative $\partial f/\partial y$ comes from fixing __b__ and moving __c__. It is the limit of __d__. If $f = e^{2x}\sin y$ then $\partial f/\partial x =$ __e__ and $\partial f/\partial y =$ __f__. If $f = (x^2 + y^2)^{1/2}$ then $f_x =$ __g__ and $f_y =$ __h__. At $(x_0, y_0)$ the partial derivative $f_x$ is the ordinary derivative of the __i__ function $f(x, y_0)$. Similarly $f_y$ comes from $f($__j__$)$. Those functions are cut out by vertical planes $x = x_0$ and __k__, while the level curves are cut out by __l__ planes.

The four second derivatives are $f_{xx}$, __m__, __n__, __o__. For $f = xy$ they are __p__. For $f = \cos 2x \cos 3y$ they are __q__. In those examples the derivatives __r__ and __s__ are the same. That is always true when the second derivatives are __t__. At the origin, $\cos 2x \cos 3y$ is curving __u__ in the $x$ and $y$ directions, while $xy$ goes __v__ in the 45° direction and __w__ in the $-45°$ direction.

**Find $\partial f/\partial x$ and $\partial f/\partial y$ for the functions in 1–12.**

**1** $3x - y + x^2y^2$
**2** $\sin(3x - y) + y$
**3** $x^3y^2 - x^2 - e^y$
**4** $xe^{x+4}$
**5** $(x + y)/(x - y)$
**6** $1/\sqrt{x^2 + y^2}$
**7** $(x^2 + y^2)^{-1}$
**8** $\ln(x + 2y)$
**9** $\ln\sqrt{x^2 + y^2}$
**10** $y^x$
**11** $\tan^{-1}(y/x)$
**12** $\ln(xy)$

**Compute $f_{xx}$, $f_{xy} = f_{yx}$, and $f_{yy}$ for the functions in 13–20.**

**13** $x^2 + 3xy + 2y^2$
**14** $(x + 3y)^2$
**15** $(x + iy)^3$
**16** $e^{ax+by}$
**17** $1/\sqrt{x^2 + y^2}$
**18** $(x + y)^n$
**19** $\cos ax \cos by$
**20** $1/(x + iy)$

**Find the domain and range (all inputs and outputs) for the functions 21–26. Then compute $f_x, f_y, f_z, f_t$.**

**21** $1/(x - y)^2$
**22** $\sqrt{x^2 + y^2 - t^2}$
**23** $(y - x)/(z - t)$
**24** $\ln(x + t)$
**25** $x^{\ln t}$ *Why does this equal* $t^{\ln x}$?
**26** $\cos x \cos^{-1} y$

**27** Verify $f_{xy} = f_{yx}$ for $f = x^my^n$. If $f_{xy} = 0$ then $f_x$ does not depend on ______ and $f_y$ is independent of ______. The function must have the form $f(x, y) = G(x) +$ ______.

**28** In terms of $v$, compute $f_x$ and $f_y$ for $f(x, y) = \int_x^y v(t)\,dt$. First vary $x$. Then vary $y$.

**29** Compute $\partial f/\partial x$ for $f = \int_0^{xy} v(t)dt$. Keep $y$ constant.

**30** What is $f(x, y) = \int_x^y dt/t$ and what are $f_x$ and $f_y$?

**31** Calculate all eight third derivatives $f_{xxx}, f_{xxy}, \ldots$ of $f = x^3y^3$. How many are different?

**In 32–35, choose $g(y)$ so that $f(x, y) = e^{cx}g(y)$ satisfies the equation.**

**32** $f_x + f_y = 0$
**33** $f_x = 7f_y$
**34** $f_y = f_{xx}$
**35** $f_{xx} = 4f_{yy}$

**36** Show that $t^{-1/2}e^{-x^2/4t}$ satisfies the ***heat equation*** $f_t = f_{xx}$. This $f(x, t)$ is the temperature at position $x$ and time $t$ due to a point source of heat at $x = 0, t = 0$.

**37** The equation for heat flow in the $xy$ plane is $f_t = f_{xx} + f_{yy}$. Show that $f(x, y, t) = e^{-2t}\sin x \sin y$ is a solution. What exponent in $f = e^{\text{———}}\sin 2x \sin 3y$ gives a solution?

**38** Find solutions $f(x, y) = e^{\text{———}}\sin mx \cos ny$ of the heat equation $f_t = f_{xx} + f_{yy}$. Show that $t^{-1}e^{-x^2/4t}e^{-y^2/4t}$ is also a solution.

**39** The basic ***wave equation*** is $f_{tt} = f_{xx}$. Verify that $f(x, t) = \sin(x + t)$ and $f(x, t) = \sin(x - t)$ are solutions. Draw both graphs at $t = \pi/4$. Which wave moved to the left and which moved to the right?

**40** Continuing 39, the peaks of the waves moved a distance $\Delta x =$ ______ in the time step $\Delta t = \pi/4$. The wave velocity is $\Delta x/\Delta t =$ ______.

**41** Which of these satisfy the wave equation $f_{tt} = c^2 f_{xx}$?

$\sin(x - ct)$, $\cos(x + ct)$, $e^{x-ct}$, $e^x - e^{ct}$, $e^x \cos ct$.

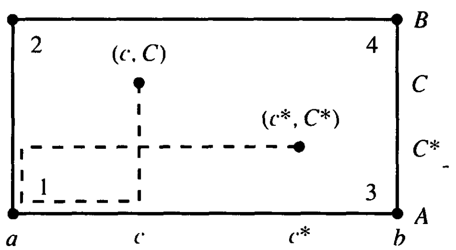
**42** Suppose $\partial f/\partial t = \partial f/\partial x$. Show that $\partial^2 f/\partial t^2 = \partial^2 f/\partial x^2$.

**43** The proof of $f_{xy}=f_{yx}$ studies $f(x, y)$ in a small rectangle. The top-bottom difference is $g(x)=f(x, B)-f(x, A)$. The difference at the corners 1, 2, 3, 4 is:

$$\begin{aligned} Q &= [f_4-f_3]-[f_2-f_1] \\ &= g(b)-g(a) \quad \text{(definition of } g) \\ &= (b-a)g_x(c) \quad \text{(Mean Value Theorem)} \\ &= (b-a)[f_x(c, B)-f_x(c, A)] \quad \text{(compute } g_x) \\ &= (b-a)(B-A)f_{xy}(c, C) \quad \text{(MVT again)} \end{aligned}$$

(a) The right-left difference is $h(y)=f(b, y)-f(a, y)$. The same $Q$ is $h(B)-h(A)$. Change the steps to reach $Q=(B-A)(b-a)f_{yx}(c^*, C^*)$.

(b) The two forms of $Q$ make $f_{xy}$ at $(c, C)$ equal to $f_{yx}$ at $(c^*, C^*)$. Shrink the rectangle toward $(a, A)$. What assumption yields $f_{xy}=f_{yx}$ at that typical point?




**44** Find $\partial f/\partial x$ and $\partial f/\partial y$ where they exist, based on equations (1) and (2).

(a) $f=|xy|$ (b) $f=x^2+y^2$ if $x \neq 0$, $f=0$ if $x=0$

**Questions 45–52 are about limits in two dimensions.**

**45** Complete these *four correct definitions of limit*: **1** The points $(x_n, y_n)$ approach the point $(a, b)$ if $x_n$ converges to $a$ and ______. **2** For any circle around $(a, b)$, the points $(x_n, y_n)$ eventually go ______ the circle and stay ______. **3** The distance from $(x_n, y_n)$ to $(a, b)$ is ______ and it approaches ______. **4** For any $\varepsilon > 0$ there is an $N$ such that the distance ______ $< \varepsilon$ for all $n >$ ______.

**46** Find $(x_2, y_2)$ and $(x_4, y_4)$ and the limit $(a, b)$ if it exists. Start from $(x_0, y_0)=(1, 0)$.

(a) $(x_n, y_n)=(1/(n+1), n/(n+1))$
(b) $(x_n, y_n)=(x_{n-1}, y_{n-1})$
(c) $(x_n, y_n)=(y_{n-1}, x_{n-1})$
(d) $(x_n, y_n)=(x_{n-1}+y_{n-1}, x_{n-1}-y_{n-1})$

**47** (*Limit of* $f(x, y)$) **1** Informal definition: the numbers $f(x_n, y_n)$ approach $L$ when the points $(x_n, y_n)$ approach $(a, b)$. **2** *Epsilon-delta definition*: For each $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x, y)-L|$ is less than ______ when the distance from $(x, y)$ to $(a, b)$ is ______. The value of $f$ *at* $(a, b)$ is not involved.

**48** Write down the limit $L$ as $(x, y) \to (a, b)$. At which points $(a, b)$ does $f(x, y)$ have no limit?

(a) $f(x, y)=\sqrt{x^2+y^2}$ (b) $f(x, y)=x/y$
(c) $f(x, y)=1/(x+y)$ (d) $f(x, y)=xy/(x^2+y^2)$

In (d) find the limit at (0, 0) along the line $y=mx$. The limit changes with $m$, so $L$ does not exist at (0, 0). Same for $x/y$.

**49** *Definition of continuity*: $f(x, y)$ is continuous at $(a, b)$ if $f(a, b)$ is defined and $f(x, y)$ approaches the limit ______ as $(x, y)$ approaches $(a, b)$. Construct a function that is *not* continuous at (1, 2).

**50** Show that $x^2y/(x^4+y^2) \to 0$ along every straight line $y=mx$ to the origin. But traveling down the parabola $y=x^2$, the ratio equals ______.

**51** Can you define $f(0, 0)$ so that $f(x, y)$ is continuous at (0, 0)?

(a) $f=|x|+|y-1|$ (b) $f=(1+x)^y$ (c) $f=x^{1+y}$.

**52** Which functions approach zero as $(x, y) \to (0, 0)$ and why?

(a) $\dfrac{xy^2}{x^2+y^2}$ (b) $\dfrac{x^2y^2}{x^4+y^4}$ (c) $\dfrac{x^my^n}{x^m+y^n}$.

## 13.3 Tangent Planes and Linear Approximations

Over a short range, a smooth curve $y=f(x)$ is almost straight. The curve changes direction, but the tangent line $y-y_0=f'(x_0)(x-x_0)$ keeps the same slope forever. The tangent line immediately gives the linear approximation to $y=f(x)$: $y \approx y_0+f'(x_0)(x-x_0)$.

***What happens with two variables***? The function is $z=f(x, y)$, and its graph is a ***surface***. We are at a point on that surface, and we are near-sighted. We don't see far away. The surface may curve out of sight at the horizon, or it may be a bowl or a saddle. To our myopic vision, the surface looks flat. We believe we are on a plane (not necessarily horizontal), and we want the equation of this ***tangent plane***.

*Notation* The basepoint has coordinates $x_0$ and $y_0$. The height on the surface is $z_0 = f(x_0, y_0)$. Other letters are possible: the point can be $(a, b)$ with height $w$. The subscript $_0$ indicates the value of $x$ or $y$ or $z$ or $\partial f/\partial x$ or $\partial f/\partial y$ *at the point.*

With one variable the tangent line has slope $df/dx$. With two variables there are two derivatives $\partial f/\partial x$ and $\partial f/\partial y$. At the particular point, they are $(\partial f/\partial x)_0$ and $(\partial f/\partial y)_0$. ***Those are the slopes of the tangent plane.*** Its equation is the key to this chapter:

**13A** The tangent plane at $(x_0, y_0, z_0)$ has the same slopes as the surface $z = f(x, y)$. The equation of the tangent plane (a linear equation) is

$$z - z_0 = \left(\frac{\partial f}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial f}{\partial y}\right)_0 (y - y_0). \tag{1}$$

The normal vector $\mathbf{N}$ to that plane has components $(\partial f/\partial x)_0, (\partial f/\partial y)_0, -1$.

**EXAMPLE 1** Find the tangent plane to $z = 14 - x^2 - y^2$ at $(x_0, y_0, z_0) = (1, 2, 9)$.

Solution The derivatives are $\partial f/\partial x = -2x$ and $\partial f/\partial y = -2y$. When $x = 1$ and $y = 2$ those are $(\partial f/\partial x)_0 = -2$ and $(\partial f/\partial y)_0 = -4$. The equation of the tangent plane is

$$z - 9 = -2(x - 1) - 4(y - 2) \quad \text{or} \quad z + 2x + 4y = 19. \tag{2}$$

This $z(x, y)$ has derivatives $-2$ and $-4$, just like the surface. So the plane is tangent.

The normal vector $\mathbf{N}$ has components $-2, -4, -1$. ***The equation of the normal line is*** $(x, y, z) = (1, 2, 9) + t(-2, -4, -1)$. Starting from $(1, 2, 9)$ the line goes out along $\mathbf{N}$—perpendicular to the plane and the surface.

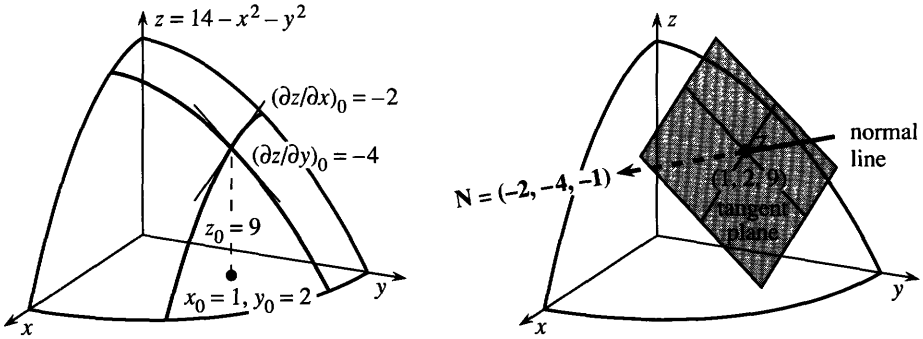



**Fig. 13.7** The tangent plane contains the $x$ and $y$ tangent lines, perpendicular to $\mathbf{N}$.

Figure 13.7 shows more detail about the tangent plane. The dotted lines are the $x$ and $y$ tangent lines. They lie in the plane. All tangent lines lie in the tangent plane! These particular lines are tangent to the "partial functions"—where $y$ is fixed at $y_0 = 2$ or $x$ is fixed at $x_0 = 1$. The plane is balancing on the surface and touching at the tangent point.

More is true. In the surface, *every curve through the point is tangent to the plane.* Geometrically, the curve goes up to the point and "kisses" the plane.† The tangent $\mathbf{T}$ to the curve and the normal $\mathbf{N}$ to the surface are perpendicular: $\mathbf{T} \cdot \mathbf{N} = 0$.

†A safer word is "osculate." At saddle points the plane is kissed from both sides.

**EXAMPLE 2** Find the tangent plane to the sphere $z^2 = 14 - x^2 - y^2$ at (1, 2, 3).

Solution Instead of $z = 14 - x^2 - y^2$ we have $z = \sqrt{14 - x^2 - y^2}$. At $x_0 = 1$, $y_0 = 2$ the height is now $z_0 = 3$. The surface is a sphere with radius $\sqrt{14}$. The only trouble from the square root is its derivatives:

$$\frac{\partial z}{\partial x} = \frac{\partial}{\partial x}\sqrt{14 - x^2 - y^2} = \frac{\frac{1}{2}(-2x)}{\sqrt{14 - x^2 - y^2}} \quad \text{and} \quad \frac{\partial z}{\partial y} = \frac{\frac{1}{2}(-2y)}{\sqrt{14 - x^2 - y^2}} \tag{3}$$

At (1, 2) those slopes are $-\frac{1}{3}$ and $-\frac{2}{3}$. The equation of the tangent plane is linear: $z - 3 = -\frac{1}{3}(x - 1) - \frac{2}{3}(y - 2)$. I cannot resist improving the equation, by multiplying through by 3 and moving all terms to the left side:

$$\textbf{\textit{tangent plane to sphere}}: \quad 1(x - 1) + 2(y - 2) + 3(z - 3) = 0. \tag{4}$$

If mathematics is the "science of patterns," equation (4) is a prime candidate for study. The numbers 1, 2, 3 appear twice. The coordinates are $(x_0, y_0, z_0) = (1, 2, 3)$. The normal vector is $1\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$. The tangent equation is $1x + 2y + 3z = 14$. None of this can be an accident, but the square root of $14 - x^2 - y^2$ made a simple pattern look complicated.

***This square root is not necessary***. Calculus offers a direct way to find $dz/dx$—*implicit differentiation*. Just differentiate every term as it stands:

$$x^2 + y^2 + z^2 = 14 \quad \text{leads to} \quad 2x + 2z\,\partial z/\partial x = 0 \quad \text{and} \quad 2y + 2z\,\partial z/\partial y = 0. \tag{5}$$

Canceling the 2's, the derivatives on a sphere are $-x/z$ and $-y/z$. Those are the same as in (3). The equation for the tangent plane has an extremely symmetric form:

$$z - z_0 = -\frac{x_0}{z_0}(x - x_0) - \frac{y_0}{z_0}(y - y_0) \quad \text{or} \quad x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0) = 0. \tag{6}$$

Reading off $\mathbf{N} = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}$ from the last equation, calculus proves something we already knew: **The normal vector to a sphere points outward along the radius.**

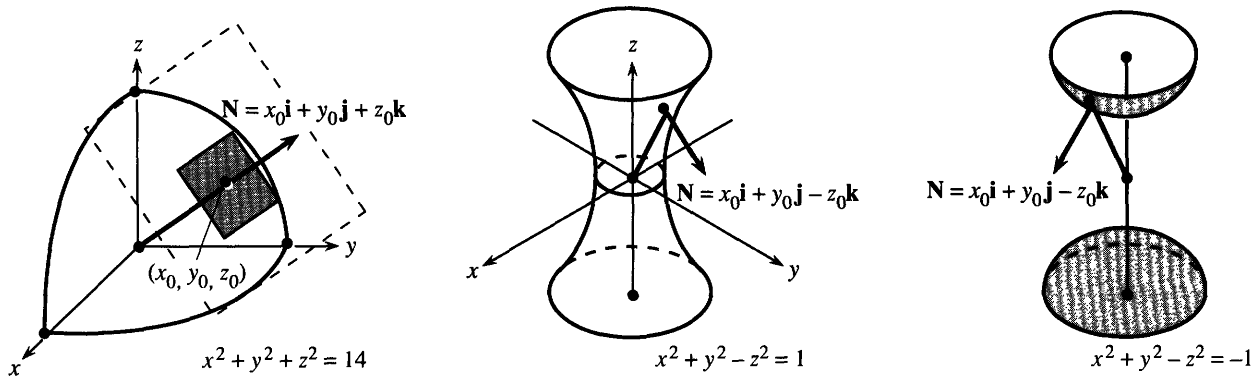



**Fig. 13.8** Tangent plane and normal N for a sphere. Hyperboloids of 1 and 2 sheets.

## THE TANGENT PLANE TO $F(x, y, z) = c$

The sphere suggests a question that is important for other surfaces. Suppose the equation is $F(x, y, z) = c$ instead of $z = f(x, y)$. Can the partial derivatives and tangent plane be found directly from $F$?

The answer is *yes*. It is not necessary to solve first for $z$. The derivatives of $F$, computed at $(x_0, y_0, z_0)$, give a second formula for the tangent plane and normal vector.

**13B** The tangent plane to the surface $F(x, y, z) = c$ has the linear equation

$$\left(\frac{\partial F}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial F}{\partial y}\right)_0 (y - y_0) + \left(\frac{\partial F}{\partial z}\right)_0 (z - z_0) = 0. \qquad (7)$$

The normal vector is $\mathbf{N} = \left(\frac{\partial F}{\partial x}\right)_0 \mathbf{i} + \left(\frac{\partial F}{\partial y}\right)_0 \mathbf{j} + \left(\frac{\partial F}{\partial z}\right)_0 \mathbf{k}$.

Notice how this includes the original case $z = f(x, y)$. ***The function F becomes*** $f(x, y) - z$. Its partial derivatives are $\partial f/\partial x$ and $\partial f/\partial y$ and $-1$. (The $-1$ is from the derivative of $-z$.) Then equation (7) is the same as our original tangent equation (1).

**EXAMPLE 3** The surface $F = x^2 + y^2 - z^2 = c$ is a ***hyperboloid***. Find its tangent plane.

Solution The partial derivatives are $F_x = 2x$, $F_y = 2y$, $F_z = -2z$. Equation (7) is

$$\textbf{\textit{tangent plane}}: \quad 2x_0(x - x_0) + 2y_0(y - y_0) - 2z_0(z - z_0) = 0. \qquad (8)$$

We can cancel the 2's. The normal vector is $\mathbf{N} = x_0\mathbf{i} + y_0\mathbf{j} - z_0\mathbf{k}$. For $c > 0$ *this hyperboloid has one sheet* (Figure 13.8). For $c = 0$ it is a cone and for $c < 0$ it breaks into two sheets (Problem 13.1.26).

## DIFFERENTIALS

Come back to the linear equation $z - z_0 = (\partial z/\partial x)_0(x - x_0) + (\partial z/\partial y)_0(y - y_0)$ for the tangent plane. That may be the most important formula in this chapter. Move along the tangent plane instead of the curved surface. Movements in the plane are $dx$ and $dy$ and $dz$—while $\Delta x$ and $\Delta y$ and $\Delta z$ are movements in the surface. The $d$'s are governed by the tangent equation—the $\Delta$'s are governed by $z = f(x, y)$. In Chapter 2 the $d$'s were ***differentials*** along the tangent line:

$$dy = (dy/dx)dx \ \ \text{(straight line)} \quad \text{and} \quad \Delta y \approx (dy/dx)\Delta x \ \ \text{(on the curve)}. \qquad (9)$$

Now $y$ is independent like $x$. The dependent variable is $z$. The idea is the same. The distances $x - x_0$ and $y - y_0$ and $z - z_0$ (on the tangent plane) are $dx$ and $dy$ and $dz$. *The equation of the plane is*

$$dz = (\partial z/\partial x)_0 dx + (\partial z/\partial y)_0 dy \quad \text{or} \quad df = f_x dx + f_y dy. \qquad (10)$$

This is the ***total differential***. All letters $dz$ and $df$ and $dw$ can be used, but $\partial z$ and $\partial f$ are not used. Differentials suggest small movements in $x$ and $y$; then $dz$ is the resulting movement in $z$. On the tangent plane, equation (10) holds exactly.

A "centering transform" has put $x_0, y_0, z_0$ at the center of coordinates. Then the "zoom transform" stretches the surface into its tangent plane.

**EXAMPLE 4** The area of a triangle is $A = \frac{1}{2}ab \sin\theta$. Find the total differential $dA$.

Solution The base has length $b$ and the sloping side has length $a$. The angle between them is $\theta$. You may prefer $A = \frac{1}{2}bh$, where $h$ is the perpendicular height $a \sin\theta$. Either way we need the partial derivatives. If $A = \frac{1}{2}ab \sin\theta$, then

$$\frac{\partial A}{\partial a} = \frac{1}{2} b \sin\theta \qquad \frac{\partial A}{\partial b} = \frac{1}{2} a \sin\theta \qquad \frac{\partial A}{\partial \theta} = \frac{1}{2} ab \cos\theta. \qquad (11)$$

These lead immediately to the total differential $dA$ (like a product rule):

$$dA=\left(\frac{\partial A}{\partial a}\right)da+\left(\frac{\partial A}{\partial b}\right)db+\left(\frac{\partial A}{\partial \theta}\right)d\theta=\frac{1}{2}b\sin\theta\,da+\frac{1}{2}a\sin\theta\,db+\frac{1}{2}ab\cos\theta\,d\theta.$$

**EXAMPLE 5** The volume of a cylinder is $V=\pi r^2h$. Decide whether $V$ is more sensitive to a change from $r=1.0$ to $r=1.1$ or from $h=1.0$ to $h=1.1$.

Solution The partial derivatives are $\partial V/\partial r=2\pi rh$ and $\partial V/\partial h=\pi r^2$. ***They measure the sensitivity to change.*** Physically, they are the side area and base area of the cylinder. The volume differential $dV$ comes from a shell around the side plus a layer on top:

$$dV=\text{shell}+\text{layer}=2\pi rh\,dr+\pi r^2dh. \tag{12}$$

Starting from $r=h=1$, that differential is $dV=2\pi dr+\pi dh$. With $dr=dh=.1$, the shell volume is $.2\pi$ and the layer volume is only $.1\pi$. So $V$ is sensitive to $dr$.

For a short cylinder like a penny, the layer has greater volume. $V$ is more sensitive to $dh$. In our case $V=\pi r^2h$ increases from $\pi(1)^3$ to $\pi(1.1)^3$. ***Compare $\Delta V$ to $dV$:***

$$\Delta V=\pi(1.1)^3-\pi(1)^3=.331\pi \qquad \text{and} \qquad dV=2\pi(.1)+\pi(.1)=.300\pi.$$

The difference is $\Delta V-dV=.031\pi$. The shell and layer missed a small volume in Figure 13.9, just above the shell and around the layer. The mistake is of order $(dr)^2+(dh)^2$. For $V=\pi r^2h$, the differential $dV=2\pi rh\,dr+\pi r^2dh$ is a ***linear approximation*** to the true change $\Delta V$. We now explain that properly.

## LINEAR APPROXIMATION

***Tangents lead immediately to linear approximations.*** That is true of tangent planes as it was of tangent lines. The plane stays close to the surface, as the line stayed close to the curve. Linear functions are simpler than $f(x)$ or $f(x,y)$ or $F(x,y,z)$. All we need are first derivatives *at the point*. Then the approximation is good *near the point*.

This key idea of calculus is already present in differentials. On the plane, $df$ equals $f_xdx+f_ydy$. On the curved surface that is a linear approximation to $\Delta f$:

**13C** The linear approximation to $f(x,y)$ near the point $(x_0,y_0)$ is

$$f(x,y)\approx f(x_0,y_0)+\left(\frac{\partial f}{\partial x}\right)_0(x-x_0)+\left(\frac{\partial f}{\partial y}\right)_0(y-y_0). \tag{13}$$

In other words $\Delta f\approx f_x\Delta x+f_y\Delta y$, as proved in Problem 24. The right side of (13) is a linear function $f_L(x,y)$. At $(x_0,y_0)$, the functions $f$ and $f_L$ have the same slopes. Then $f(x,y)$ curves away from $f_L$ with an error of "second order:"

$$|f(x,y)-f_L(x,y)|\leqslant M[(x-x_0)^2+(y-y_0)^2]. \tag{14}$$

This assumes that $f_{xx}, f_{xy}$, and $f_{yy}$ are continuous and bounded by $M$ along the line from $(x_0,y_0)$ to $(x,y)$. Example 3 of Section 13.5 shows that $|f_{tt}|\leqslant 2M$ along that line. A factor $\frac{1}{2}$ comes from equation 3.8.12, for the error $f-f_L$ with one variable.

For the volume of a cylinder, $r$ and $h$ went from 1.0 to 1.1. The second derivatives of $V=\pi r^2h$ are $V_{rr}=2\pi h$ and $V_{rh}=2\pi r$ and $V_{hh}=0$. They are below $M=2.2\pi$. Then (14) gives the error bound $2.2\pi(.1^2+.1^2)=.044\pi$, not far above the actual error $.031\pi$. The main point is that ***the error in linear approximation comes from the quadratic terms***—those are the first terms to be ignored by $f_L$.

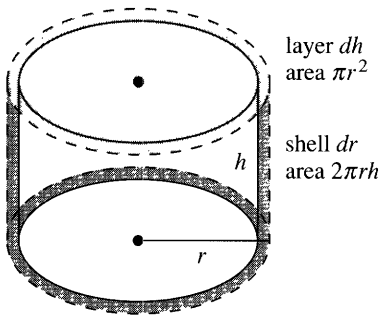



**Fig. 13.9** Shell plus layer gives $dV = .300\pi$. Including top ring gives $\Delta V = .331\pi$.

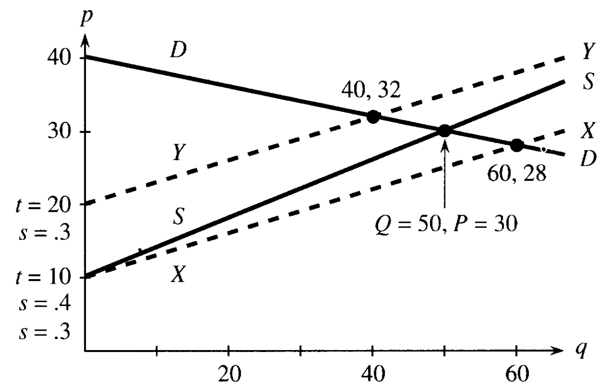



**Fig. 13.10** Quantity $Q$ and price $P$ move with the lines.

**EXAMPLE 6** Find a linear approximation to the distance function $r = \sqrt{x^2 + y^2}$.

Solution The partial derivatives are $x/r$ and $y/r$. Then $\Delta r \approx (x/r)\Delta x + (y/r)\Delta y$.

For $(x, y, r)$ near $(1, 2, \sqrt{5})$: $\sqrt{x^2 + y^2} \approx \sqrt{1^2 + 2^2} + (x - 1)/\sqrt{5} + 2(y - 2)/\sqrt{5}$.

If $y$ is fixed at 2, this is a one-variable approximation to $\sqrt{x^2 + 2^2}$. If $x$ is fixed at 1, it is a linear approximation in $y$. Moving both variables, you might think $dr$ would involve $dx$ and $dy$ in a square root. It doesn't. Distance involves $x$ and $y$ in a square root, but *change of distance* is linear in $\Delta x$ and $\Delta y$—to a first approximation.

There is a rough point at $x = 0$, $y = 0$. Any movement from $(0, 0)$ gives $\Delta r = \sqrt{(\Delta x)^2 + (\Delta y)^2}$. The square root has returned. The reason is that ***the partial derivatives*** $x/r$ ***and*** $y/r$ ***are not continuous at*** $(0, 0)$. The cone has a sharp point with no tangent plane. *Linear approximation breaks down.*

The next example shows how to approximate $\Delta z$ from $\Delta x$ and $\Delta y$, when the equation is $F(x, y, z) = c$. We use the implicit derivatives in (7) instead of the explicit derivatives in (1). The idea is the same: Look at the tangent equation as a way to find $\Delta z$, instead of an equation for $z$. Here is Example 6 with new letters.

**EXAMPLE 7** From $F = -x^2 - y^2 + z^2 = 0$ find a linear approximation to $\Delta z$.

Solution (implicit derivatives) Use the derivatives of $F$: $-2x\Delta x - 2y\Delta y + 2z\Delta z \approx 0$. Then solve for $\Delta z$, which gives $\Delta z \approx (x/z)\Delta x + (y/z)\Delta y$—the same as Example 6.

**EXAMPLE 8** How does the equilibrium price change when the supply curve changes?

The equilibrium price is at the intersection of the supply and demand curves (***supply = demand***). As the price $p$ rises, the demand $q$ drops (the slope is $-.2$):

$$\text{demand line } DD\text{: } p = -.2q + 40. \qquad (15)$$

The supply (also $q$) goes *up* with the price. The slope $s$ is positive (here $s = .4$):

$$\text{supply line } SS\text{: } p = sq + t = .4q + 10.$$

Those lines are in Figure 13.10. They meet at the ***equilibrium price*** $P = \$30$. The quantity $Q = 50$ is available at $P$ (on $SS$) and demanded at $P$ (on $DD$). So it is sold.

Where do partial derivatives come in? The reality is that those lines $DD$ and $SS$ are not fixed for all time. Technology changes, and competition changes, and the value of money changes. Therefore the lines move. Therefore the crossing point $(Q, P)$ also moves. Please recognize that derivatives are hiding in those sentences.

Main point: ***The equilibrium price $P$ is a function of $s$ and $t$.*** Reducing $s$ by better technology lowers the supply line to $p = .3q + 10$. The demand line has not changed. The customer is as eager or stingy as ever. But the price $P$ and quantity $Q$ are different. The new equilibrium is at $Q = 60$ and $P = \$28$, where the new line $XX$ crosses $DD$.

If the technology is expensive, the supplier will raise $t$ when reducing $s$. Line $YY$ is $p = .3q + 20$. That gives a higher equilibrium $P = \$32$ at a lower quantity $Q = 40$—the demand was too weak for the technology.

**Calculus question** Find $\partial P/\partial s$ and $\partial P/\partial t$. The difficulty is that $P$ is not given as a function of $s$ and $t$. So take implicit derivatives of the supply = demand equations:

$$\begin{aligned} \text{supply} = \text{demand:} \quad & P = -.2Q + 40 = sQ + t && (16)\\ s \text{ derivative:} \quad & P_s = -.2Q_s = sQ_s + Q \quad (\text{note } t_s = 0) &&\\ t \text{ derivative:} \quad & P_t = -.2Q_t = sQ_t + 1 \quad (\text{note } t_t = 1) && \end{aligned}$$

Now substitute $s = .4$, $t = 10$, $P = 30$, $Q = 50$. That is the starting point, around which we are finding a linear approximation. The last two equations give $P_s = 50/3$ and $P_t = 1/3$ (Problem 25). The linear approximation is

$$P = 30 + 50(s - .4)/3 + (t - 10)/3. \qquad (17)$$

*Comment* This example turned out to be subtle (so is economics). I hesitated before including it. The equations are linear and their derivatives are easy, but something in the problem is hard—there is no explicit formula for $P$. The function $P(s, t)$ is not known. Instead of a point on a surface, we are following the intersection of two lines. *The solution changes as the equation changes.* ***The derivative of the solution comes from the derivative of the equation.***

*Summary* The foundation of this section is equation (1) for the tangent plane. Everything builds on that—total differential, linear approximation, sensitivity to small change. Later sections go on to the chain rule and "directional derivatives" and "gradients." The central idea of differential calculus is $\Delta f \approx f_x \Delta x + f_y \Delta y$.

## NEWTON'S METHOD FOR TWO EQUATIONS

Linear approximation is used ***to solve equations***. To find out where a function is zero, look first to see where its approximation is zero. To find out where a graph crosses the $xy$ plane, look to see where its tangent plane crosses.

Remember Newton's method for $f(x) = 0$. The current guess is $x_n$. Around that point, $f(x)$ is close to $f(x_n) + (x - x_n)f'(x_n)$. This is zero at the next guess $x_{n+1} = x_n - f(x_n)/f'(x_n)$. That is where the tangent line crosses the $x$ axis.

With two variables the idea is the same— but two unknowns $x$ and $y$ require *two equations*. We solve $g(x, y) = 0$ and $h(x, y) = 0$. Both functions have linear approximations that start from the current point $(x_n, y_n)$—where derivatives are computed:

$$\begin{aligned} g(x, y) &\approx g(x_n, y_n) + (\partial g/\partial x)(x - x_n) + (\partial g/\partial y)(y - y_n) \\ h(x, y) &\approx h(x_n, y_n) + (\partial h/\partial x)(x - x_n) + (\partial h/\partial y)(y - y_n). \end{aligned} \qquad (18)$$

The natural idea is to ***set these approximations to zero***. That gives linear equations for $x - x_n$ and $y - y_n$. Those are the steps $\Delta x$ and $\Delta y$ that take us to the next guess in Newton's method:

> **13D** Newton's method to solve $g(x, y) = 0$ and $h(x, y) = 0$ has linear equations for the steps $\Delta x$ and $\Delta y$ that go from $(x_n, y_n)$ to $(x_{n+1}, y_{n+1})$:
> $$\left(\frac{\partial g}{\partial x}\right)\Delta x + \left(\frac{\partial g}{\partial y}\right)\Delta y = -g(x_n, y_n) \quad \text{and} \quad \left(\frac{\partial h}{\partial x}\right)\Delta x + \left(\frac{\partial h}{\partial y}\right)\Delta y = -h(x_n, y_n). \qquad (19)$$

**EXAMPLE 9** $g = x^3 - y = 0$ and $h = y^3 - x = 0$ have 3 solutions $(1, 1), (0, 0), (-1, -1)$.

I will start at different points $(x_0, y_0)$. The next guess is $x_1 = x_0 + \Delta x$, $y_1 = y_0 + \Delta y$. It is of extreme interest to know which solution Newton's method will choose—if it converges at all. I made three small experiments.

1. Suppose $(x_0, y_0) = (2, 1)$. At that point $g = 2^3 - 1 = 7$ and $h = 1^3 - 2 = -1$. The derivatives are $g_x = 3x^2 = 12$, $g_y = -1$, $h_x = -1$, $h_y = 3y^2 = 3$. The steps $\Delta x$ and $\Delta y$ come from solving (19):

$$\begin{array}{rcl} 12\Delta x - \Delta y = -7 & \Rightarrow & \Delta x = -4/7 \\ -\Delta x + 3\Delta y = +1 & & \Delta y = +1/7 \end{array} \Rightarrow \begin{array}{l} x_1 = x_0 + \Delta x = 10/7 \\ y_1 = y_0 + \Delta y = 8/7. \end{array}$$

This new point $(10/7, 8/7)$ is closer to the solution at $(1, 1)$. The next point is $(1.1, 1.05)$ and convergence is clear. Soon convergence is fast.

2. Start at $(x_0, y_0) = (\frac{1}{2}, 0)$. There we find $g = 1/8$ and $h = -1/2$:

$$\begin{array}{rcl} (3/4)\Delta x - \Delta y = -1/8 & \Rightarrow & \Delta x = -1/2 \\ -\Delta x + 0\Delta y = +1/2 & & \Delta y = +1/4 \end{array} \Rightarrow \begin{array}{l} x_1 = x_0 + \Delta x = 0 \\ y_1 = y_0 + \Delta y = -1/4. \end{array}$$

Newton has jumped from $(\frac{1}{2}, 0)$ on the $x$ axis to $(0, -\frac{1}{4})$ on the $y$ axis. The next step goes to $(1/32, 0)$, back on the $x$ axis. We are in the "basin of attraction" of $(0, 0)$.

3. Now start further out the axis at $(1, 0)$, where $g = 1$ and $h = -1$:

$$\begin{array}{rcl} 3\Delta x - \Delta y = -1 & \Rightarrow & \Delta x = -1 \\ -\Delta x + 0\Delta y = +1 & & \Delta y = -2 \end{array} \Rightarrow \begin{array}{l} x_1 = x_0 + \Delta x = 0 \\ y_1 = y_0 + \Delta y = -2. \end{array}$$

Newton moves from $(1, 0)$ to $(0, -2)$ to $(16, 0)$. Convergence breaks down—the method blows up. This danger is ever-present, when we start far from a solution.

Please recognize that even a small computer will uncover amazing patterns. It can start from hundreds of points $(x_0, y_0)$, and follow Newton's method. Each solution has a ***basin of attraction***, containing all $(x_0, y_0)$ leading to that solution. There is also a basin leading to infinity. The basins in Figure 13.11 are completely mixed together—a color figure shows them as ***fractals***. The most extreme behavior is on the borderline between basins, when Newton can't decide which way to go. Frequently we see chaos.

Chaos is irregular movement that follows a definite rule. Newton's method determines an *iteration* from each point $(x_n, y_n)$ to the next. In scientific problems it normally converges to the solution we want. (We start close enough.) But the computer makes it posible to study iterations from faraway points. This has created a new part of mathematics—so new that any experiments you do are likely to be original.

Section 3.7 found chaos when trying to solve $x^2 + 1 = 0$. But don't think Newton's method is a failure. On the contrary, it is the best method to solve nonlinear equations. The error is squared as the algorithm converges, because linear approximations have errors of order $(\Delta x)^2 + (\Delta y)^2$. Each step doubles the number of correct digits, *near the solution*. The example shows why it is important to be near.

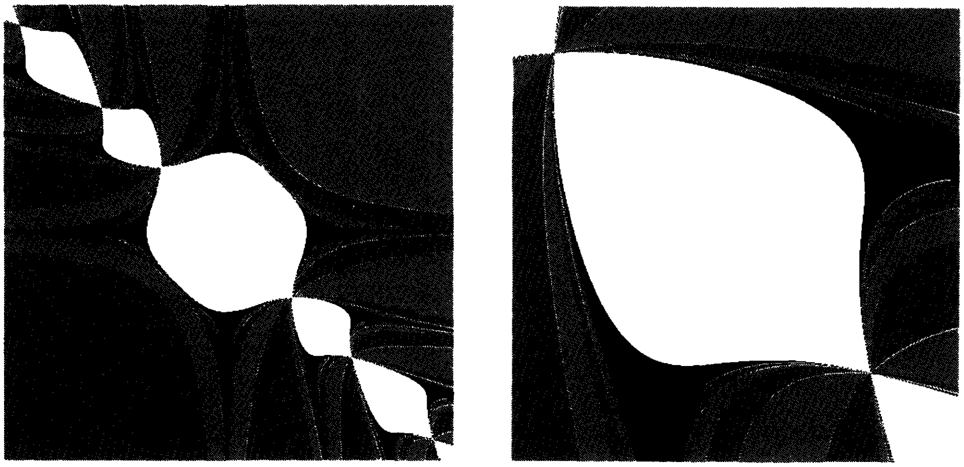


**Fig. 13.11** The basins of attraction to (1, 1), (0, 0), (−1, −1), and infinity.

## 13.3 EXERCISES

**Read-through questions**

The tangent line to $y=f(x)$ is $y-y_0=$ __a__. The tangent plane to $w=f(x, y)$ is $w-w_0=$ __b__. The normal vector is $\mathbf{N}=$ __c__. For $w=x^3+y^3$ the tangent equation at (1, 1, 2) is __d__. The normal vector is $\mathbf{N}=$ __e__. For a sphere, the direction of $\mathbf{N}$ is __f__.

The surface given implicitly by $F(x, y, z)=c$ has tangent equation $(\partial F/\partial x)_0(x-x_0)+$ __g__. For $xyz=6$ at (1, 2, 3) the tangent plane is __h__. On that plane the differentials satisfy __i__ $dx+$ __j__ $dy+$ __k__ $dz=0$. The differential of $z=f(x, y)$ is $dz=$ __l__. This holds exactly on the tangent plane, while $\Delta z \approx$ __m__ holds approximately on the __n__. The height $z=3x+7y$ is more sensitive to a change in __o__ than in $x$, because the partial derivative __p__ is larger than __q__.

The linear approximation to $f(x, y)$ is $f(x_0, y_0)+$ __r__. This is the same as $\Delta f \approx$ __s__ $\Delta x+$ __t__ $\Delta y$. The error is of order __u__. For $f=\sin xy$ the linear approximation around (0, 0) is $f_L=$ __v__. We are moving along the __w__ instead of the __x__. When the equation is given as $F(x, y, z)=c$, the linear approximation is __y__ $\Delta x+$ __z__ $\Delta y+$ __A__ $\Delta z=0$.

Newton's method solves $g(x, y)=0$ and $h(x, y)=0$ by a __B__ approximation. Starting from $x_n$, $y_n$ the equations are replaced by __C__ and __D__. The steps $\Delta x$ and $\Delta y$ go to the next point __E__. Each solution has a basin of __F__. Those basins are likely to be __G__.

**In 1–8 find the tangent plane and the normal vector at $P$.**

**1** $z=\sqrt{x^2+y^2}$, $P=(0, 1, 1)$

**2** $x+y+z=17$, $P=(3, 4, 10)$

**3** $z=x/y$, $P=(6, 3, 2)$

**4** $z=e^{x+2y}$, $P=(0, 0, 1)$

**5** $x^2+y^2+z^2=6$, $P=(1, 2, 1)$

**6** $x^2+y^2+2z^2=7$, $P=(1, 2, 1)$

**7** $z=x^y$, $P=(1, 1, 1)$

**8** $V=\pi r^2 h$, $P=(2, 2, 8\pi)$.

**9** Show that the tangent plane to $z^2-x^2-y^2=0$ goes through the origin and makes a 45° angle with the $z$ axis.

**10** The planes $z=x+4y$ and $z=2x+3y$ meet at (1, 1, 5). The whole line of intersection is $(x, y, z)=(1, 1, 5)+\mathbf{v}t$. Find $\mathbf{v}=\mathbf{N}_1 \times \mathbf{N}_2$.

**11** If $z=3x-2y$ find $dz$ from $dx$ and $dy$. If $z=x^3/y^2$ find $dz$ from $dx$ and $dy$ at $x_0=1$, $y_0=1$. If $x$ moves to 1.02 and $y$ moves to 1.03, find the approximate $dz$ and exact $\Delta z$ for both functions. The first surface is the ________ to the second surface.

**12** The surfaces $z = x^2 + 4y$ and $z = 2x + 3y^2$ meet at (1, 1, 5). Find the normals $\mathbf{N}_1$ and $\mathbf{N}_2$ and also $\mathbf{v} = \mathbf{N}_1 \times \mathbf{N}_2$. The line in this direction $\mathbf{v}$ is tangent to what curve?

**13** The normal $\mathbf{N}$ to the surface $F(x, y, z) = 0$ has components $F_x, F_y, F_z$. The *normal line* has $x = x_0 + F_x t$, $y = y_0 + F_y t$, $z =$ ________. For the surface $xyz - 24 = 0$, find the tangent plane and normal line at (4, 2, 3).

**14** For the surface $x^2y^2 - z = 0$, the normal line at (1, 2, 4) has $x =$ ________, $y =$ ________, $z =$ ________.

**15** For the sphere $x^2 + y^2 + z^2 = 9$, find the equation of the tangent plane through (2, 1, 2). Also find the equation of the normal line and show that it goes through (0, 0, 0).

**16** If the normal line at every point on $F(x, y, z) = 0$ goes through (0, 0, 0), show that $F_x = cx$, $F_y = cy$, $F_z = cz$. The surface must be a sphere.

**17** For $w = xy$ near $(x_0, y_0)$, the linear approximation is $dw =$ ________. This looks like the ________ rule for derivatives. The difference between $\Delta w = xy - x_0y_0$ and this approximation is ________.

**18** If $f = xyz$ (3 independent variables) what is $df$?

**19** You invest $P = \$4000$ at $R = 8\%$ to make $I = \$320$ per year. If the numbers change by $dP$ and $dR$ what is $dI$? If the rate drops by $dR = .002$ (to 7.8%) what change $dP$ keeps $dI = 0$? Find the exact interest $I$ after those changes in $R$ and $P$.

**20** Resistances $R_1$ and $R_2$ have parallel resistance $R$, where $1/R = 1/R_1 + 1/R_2$. Is $R$ more sensitive to $\Delta R_1$ or $\Delta R_2$ if $R_1 = 1$ and $R_2 = 2$?

**21** (a) If your batting average is $A = (25 \text{ hits})/(100 \text{ at bats}) = .250$, compute the increase (to 26/101) with a hit and the decrease (to 25/101) with an out.

(b) If $A = x/y$ then $dA =$ ________ $dx +$ ________ $dy$. A hit ($dx = dy = 1$) gives $dA = (1 - A)/y$. An out ($dy = 1$) gives $dA = -A/y$. So at $A = .250$ a hit has ________ times the effect of an out.

**22** (a) 2 hits and 3 outs ($dx = 2$, $dy = 5$) will raise your average ($dA > 0$) provided $A$ is less than ________.

(b) A player batting $A = .500$ with $y = 400$ at bats needs $dx =$ ________ hits to raise his average to .505.

**23** If $x$ and $y$ change by $\Delta x$ and $\Delta y$, find the approximate change $\Delta\theta$ in the angle $\theta = \tan^{-1}(y/x)$.

**24** The *Fundamental Lemma* behind equation (13) writes $\Delta f = a\Delta x + b\Delta y$. The Lemma says that $a \to f_x(x_0, y_0)$ and $b \to f_y(x_0, y_0)$ when $\Delta x \to 0$ and $\Delta y \to 0$. The proof takes $\Delta x$ first and then $\Delta y$:

(1) $f(x_0 + \Delta x, y_0) - f(x_0, y_0) = \Delta x f_x(c, y_0)$ where $c$ is between ________ and ________ (by which theorem?)

(2) $f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0 + \Delta x, y_0) = \Delta y f_y(x_0 + \Delta x, C)$ where $C$ is between ________ and ________.

(3) $a = f_x(c, y_0) \to f_x(x_0, y_0)$ provided $f_x$ is ________.

(4) $b = f_y(x_0 + \Delta x, C) \to f_y(x_0, y_0)$ provided $f_y$ is ________.

**25** If the supplier reduces $s$, Figure 13.10 shows that $P$ decreases and $Q$ ________.

(a) Find $P_s = 50/3$ and $P_t = 1/3$ in the economics equation (17) by solving the equations above it for $Q_s$ and $Q_t$.

(b) What is the linear approximation to $Q$ around $s = .4$, $t = 10$, $P = 30$, $Q = 50$?

**26** Solve the equations $P = -.2Q + 40$ and $P = sQ + t$ for $P$ and $Q$. Then find $\partial P/\partial s$ and $\partial P/\partial t$ explicitly. At the same $s, t, P, Q$ check 50/3 and 1/3.

**27** If the supply = demand equation (16) changes to $P = sQ + t = -Q + 50$, find $P_s$ and $P_t$ at $s = 1$, $t = 10$.

**28** To find out how the roots of $x^2 + bx + c = 0$ vary with $b$, take partial derivatives of the equation with respect to ________. Compare $\partial x/\partial b$ with $\partial x/\partial c$ to show that a root at $x = 2$ is more sensitive to $b$.

**29** Find the tangent planes to $z = xy$ and $z = x^2 - y^2$ at $x = 2$, $y = 1$. Find the Newton point where those planes meet the $xy$ plane (set $z = 0$ in the tangent equations).

**30** (a) To solve $g(x, y) = 0$ and $h(x, y) = 0$ is to find the meeting point of three surfaces: $z = g(x, y)$ and $z = h(x, y)$ and ________.

(b) Newton finds the meeting point of three planes: the tangent plane to the graph of $g$, ________, and ________.

**Problems 31–36 go further with Newton's method for $g = x^3 - y$ and $h = y^3 - x$. This is Example 9 with solutions (1, 1), (0, 0), (−1, −1).**

**31** Start from $x_0 = 1$, $y_0 = 1$ and find $\Delta x$ and $\Delta y$. Where are $x_1$ and $y_1$, and what line is Newton's method moving on?

**32** Start from $(\frac{1}{2}, \frac{1}{2})$ and find the next point. This is in the basin of attraction of which solution?

**33** Starting from $(a, -a)$ find $\Delta y$ which is also $-\Delta x$. Newton goes toward (0, 0). But can you find the sharp point in Figure 13.11 where the lemon meets the spade?

**34** Starting from $(a, 0)$ show that Newton's method goes to $(0, -2a^3)$ and find the next point $(x_2, y_2)$. Which numbers $a$ lead to convergence? Which special number $a$ leads to a cycle, in which $(x_2, y_2)$ is the same as the starting point $(a, 0)$?

**35** Show that $x^3 = y$, $y^3 = x$ has exactly three solutions.

**36** Locate a point from which Newton's method diverges.

**37** Apply Newton's method to a linear problem: $g = x + 2y - 5 = 0$, $h = 3x - 3 = 0$. From any starting point show that $(x_1, y_1)$ is the exact solution (convergence in *one step*).

**38** The complex equation $(x+iy)^3 = 1$ contains two real equations, $x^3 - 3xy^2 = 1$ from the real part and $3x^2y - y^3 = 0$ from the imaginary part. Search by computer for the basins of attraction of the three solutions $(1, 0)$, $(-1/2, \sqrt{3}/2)$, and $(-1/2, -\sqrt{3}/2)$—which give the cube roots of 1.

**39** In Newton's method the new guess comes from $(x_n, y_n)$ by an *iteration*: $x_{n+1} = G(x_n, y_n)$ and $y_{n+1} = H(x_n, y_n)$. What are $G$ and $H$ for $g = x^2 - y = 0$, $h = x - y = 0$? First find $\Delta x$ and $\Delta y$; then $x_n + \Delta x$ gives $G$ and $y_n + \Delta y$ gives $H$.

**40** In Problem 39 find the basins of attraction of the solution $(0, 0)$ and $(1, 1)$.

**41** The matrix in Newton's method is the ***Jacobian***:

$$J = \begin{bmatrix} \partial g/\partial x & \partial g/\partial y \\ \partial h/\partial x & \partial h/\partial y \end{bmatrix} \quad \text{and} \quad J\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} -g_n \\ -h_n \end{bmatrix}.$$

Find $J$ and $\Delta x$ and $\Delta y$ for $g = e^x - 1$, $h = e^y + x$.

**42** Find the Jacobian matrix at $(1, 1)$ when $g = x^2 + y^2$ and $h = xy$. This matrix is ________ and Newton's method fails. The graphs of $g$ and $h$ have ________ tangent planes.

**43** Solve $g = x^2 - y^2 + 1 = 0$ and $h = 2xy = 0$ by Newton's method from three starting points: $(0, 2)$ and $(-1, 1)$ and $(2, 0)$. Take ten steps by computer or one by hand. The solution $(0, 1)$ attracts when $y_0 > 0$. If $y_0 = 0$ you should find the chaos iteration $x_{n+1} = \frac{1}{2}(x_n - x_n^{-1})$.

## 13.4 Directional Derivatives and Gradients

As $x$ changes, we know how $f(x, y)$ changes. The partial derivative $\partial f/\partial x$ treats $y$ as constant. Similarly $\partial f/\partial y$ keeps $x$ constant, and gives the slope in the $y$ direction. But east-west and north-south are not the only directions to move. We could go along a 45° line, where $\Delta x = \Delta y$. In principle, before we draw axes, no direction is preferred. The graph is a surface with slopes in *all* directions.

On that surface, calculus looks for the rate of change (or the slope). There is a ***directional derivative***, whatever the direction. In the 45° case we are inclined to divide $\Delta f$ by $\Delta x$, but we would be wrong.

Let me state the problem. We are given $f(x, y)$ around a point $P = (x_0, y_0)$. We are also given a direction $\mathbf{u}$ (a unit vector). There must be a natural definition of $D_{\mathbf{u}}f$—***the derivative of f in the direction*** $\mathbf{u}$. To compute this slope at $P$, we need a formula. Preferably the formula is based on $\partial f/\partial x$ and $\partial f/\partial y$, which we already know.

Note that the 45° direction has $\mathbf{u} = \mathbf{i}/\sqrt{2} + \mathbf{j}/\sqrt{2}$. The square root of 2 is going to enter the derivative. This shows that dividing $\Delta f$ by $\Delta x$ is wrong. We should divide by the step length $\Delta s$.

**EXAMPLE 1** Stay on the surface $z = xy$. When $(x, y)$ moves a distance $\Delta s$ in the 45° direction from $(1, 1)$, what is $\Delta z/\Delta s$?

Solution The step is $\Delta s$ times the unit vector $\mathbf{u}$. Starting from $x = y = 1$ the step ends at $x = y = 1 + \Delta s/\sqrt{2}$. (The components of $\mathbf{u}\Delta s$ are $\Delta s/\sqrt{2}$.) Then $z = xy$ is

$$z = (1 + \Delta s/\sqrt{2})^2 = 1 + \sqrt{2}\Delta s + \tfrac{1}{2}(\Delta s)^2, \text{ which means } \Delta z = \sqrt{2}\Delta s + \tfrac{1}{2}(\Delta s)^2.$$

The ratio $\Delta z/\Delta s$ approaches $\sqrt{2}$ as $\Delta s \to 0$. That is the slope in the 45° direction.

**DEFINITION** *The derivative of f in the direction* $\mathbf{u}$ *at the point P is* $D_{\mathbf{u}}f(P)$:

$$D_{\mathbf{u}}f(P) = \lim_{\Delta s \to 0} \frac{\Delta f}{\Delta s} = \lim_{\Delta s \to 0} \frac{f(P + \mathbf{u}\Delta s) - f(P)}{\Delta s}. \tag{1}$$

The step from $P = (x_0, y_0)$ has length $\Delta s$. It takes us to $(x_0 + u_1\Delta s, y_0 + u_2\Delta s)$. We compute the change $\Delta f$ and divide by $\Delta s$. But formula (2) below saves time.

The $x$ direction is $\mathbf{u}=(1,0)$. Then $\mathbf{u}\Delta s$ is $(\Delta s, 0)$ and we recover $\partial f/\partial x$:

$$\frac{\Delta f}{\Delta s}=\frac{f(x_0+\Delta s, y_0)-f(x_0, y_0)}{\Delta s} \text{ approaches } D_{(1,0)}f=\frac{\partial f}{\partial x}.$$

Similarly $D_{\mathbf{u}}f=\partial f/\partial y$, when $\mathbf{u}=(0,1)$ is in the $y$ direction. *What is $D_{\mathbf{u}}f$ when $\mathbf{u}=(0,-1)$?* That is the negative $y$ direction, so $D_{\mathbf{u}}f=-\partial f/\partial y$.

## CALCULATING THE DIRECTIONAL DERIVATIVE

$D_{\mathbf{u}}f$ is the slope of the surface $z=f(x,y)$ in the direction $\mathbf{u}$. How do you compute it? From $\partial f/\partial x$ and $\partial f/\partial y$, in two special directions, there is a quick way to find $D_{\mathbf{u}}f$ in all directions. **Remember that u is a unit vector.**

**13E** The ***directional derivative*** $D_{\mathbf{u}}f$ in the direction $\mathbf{u}=(u_1,u_2)$ equals

$$D_{\mathbf{u}}f=\frac{\partial f}{\partial x}u_1+\frac{\partial f}{\partial y}u_2. \qquad (2)$$

The reasoning goes back to the linear approximation of $\Delta f$:

$$\Delta f\approx\frac{\partial f}{\partial x}\Delta x+\frac{\partial f}{\partial y}\Delta y=\frac{\partial f}{\partial x}u_1\Delta s+\frac{\partial f}{\partial y}u_2\Delta s.$$

Divide by $\Delta s$ and let $\Delta s$ approach zero. Formula (2) is the limit of $\Delta f/\Delta s$, as the approximation becomes exact. A more careful argument guarantees this limit provided $f_x$ and $f_y$ are continuous at the basepoint $(x_0, y_0)$.

Main point: ***Slopes in all directions are known from slopes in two directions.***

**EXAMPLE 1 (repeated)** $f=xy$ and $P=(1,1)$ and $\mathbf{u}=(1/\sqrt{2}, 1/\sqrt{2})$. Find $D_{\mathbf{u}}f(P)$.

The derivatives $f_x=y$ and $f_y=x$ equal 1 at $P$. The 45° derivative is

$$D_{\mathbf{u}}f(P)=f_xu_1+f_yu_2=1(1/\sqrt{2})+1(1/\sqrt{2})=\sqrt{2} \text{ as before.}$$

**EXAMPLE 2** The linear function $f=3x+y+1$ has slope $D_{\mathbf{u}}f=3u_1+u_2$.

The $x$ direction is $u=(1,0)$, and $D_{\mathbf{u}}f=3$. That is $\partial f/\partial x$. In the $y$ direction $D_{\mathbf{u}}f=1$. ***Two other directions are special—along the level lines of $f(x,y)$ and perpendicular:***

*Level direction*: $D_{\mathbf{u}}f$ is zero because $f$ is constant

*Steepest direction*: $D_{\mathbf{u}}f$ is as large as possible (with $u_1^2+u_2^2=1$).

To find those directions, look at $D_{\mathbf{u}}f=3u_1+u_2$. The level direction has $3u_1+u_2=0$. Then $u$ is proportional to $(1,-3)$. Changing $x$ by 1 and $y$ by $-3$ produces no change in $f=3x+y+1$.

In the steepest direction $\mathbf{u}$ is proportional to (3, 1). Note the partial derivatives $f_x=3$ and $f_y=1$. The dot product of (3, 1) and (1, $-3$) is zero—*steepest direction is perpendicular to level direction.* To make (3, 1) a unit vector, divide by $\sqrt{10}$.

*Steepest climb*: $D_{\mathbf{u}}f=3(3/\sqrt{10})+1(1/\sqrt{10})=10/\sqrt{10}=\sqrt{10}$

*Steepest descent*: Reverse to $\mathbf{u}=(-3/\sqrt{10}, -1/\sqrt{10})$ and $D_{\mathbf{u}}f=-\sqrt{10}$.

The contour lines around a mountain follow $D_{\mathbf{u}}f=0$. The creeks are perpendicular. On a plane like $f=3x+y+1$, those directions stay the same at all points (Figure 13.12). On a mountain the steepest direction changes as the slopes change.

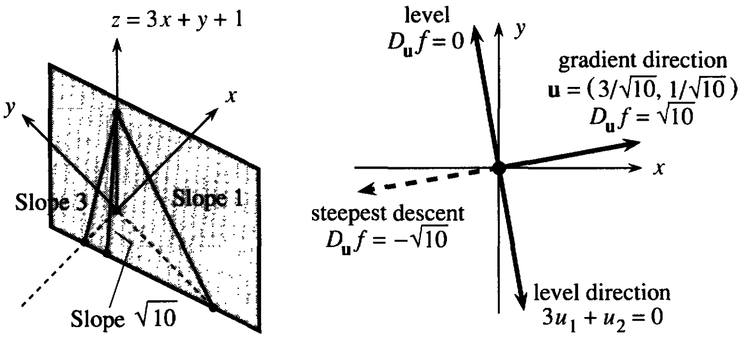



**Fig. 13.12** Steepest direction is along the gradient. Level direction is perpendicular.

## THE GRADIENT VECTOR

Look again at $f_x u_1 + f_y u_2$, which is the directional derivative $D_u f$. *This is the dot product of two vectors.* One vector is $\mathbf{u} = (u_1, u_2)$, which sets the direction. The other vector is $(f_x, f_y)$, which comes from the function. This second vector is the ***gradient***.

**DEFINITION** The ***gradient*** of $f(x, y)$ is the vector whose components are $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$:

$$\operatorname{grad} f = \nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} \qquad \left(\text{add } \frac{\partial f}{\partial z}\mathbf{k} \text{ in three dimensions}\right).$$

The space-saving symbol $\nabla$ is read as "grad." In Chapter 15 it becomes "del."

For the linear function $3x + y + 1$, the gradient is the constant vector (3, 1). It is the way to climb the plane. For the nonlinear function $x^2 + xy$, the gradient is the non-constant vector $(2x + y, x)$. Notice that grad $f$ shares the two derivatives in $\mathbf{N} = (f_x, f_y, -1)$. But the gradient is not the normal vector. $\mathbf{N}$ is in three dimensions, pointing away from the surface $z = f(x, y)$. ***The gradient vector is in the** xy **plane**!* The gradient tells which way on the surface is up, but it does that from down in the base.

The level curve is also in the $xy$ plane, perpendicular to the gradient. The contour map is a projection on the base plane of what the hiker sees on the mountain. The vector grad $f$ tells the ***direction*** of climb, and its length $|\operatorname{grad} f|$ gives the ***steepness***.

> **13F** The directional derivative is $D_u f = (\operatorname{grad} f) \cdot \mathbf{u}$. The level direction is perpendicular to grad $f$, since $D_u f = 0$. ***The slope** $D_u f$ **is largest when u is parallel to*** grad $f$. That maximum slope is the length $|\operatorname{grad} f| = \sqrt{f_x^2 + f_y^2}$:
>
> $$\textit{for } \mathbf{u} = \frac{\operatorname{grad} f}{|\operatorname{grad} f|} \quad \textit{the slope is} \quad (\operatorname{grad} f) \cdot \mathbf{u} = \frac{|\operatorname{grad} f|^2}{|\operatorname{grad} f|} = |\operatorname{grad} f|.$$

The example $f = 3x + y + 1$ had grad $f = (3, 1)$. Its steepest slope was in the direction $\mathbf{u} = (3, 1)/\sqrt{10}$. The maximum slope was $\sqrt{10}$. That is $|\operatorname{grad} f| = \sqrt{9 + 1}$.

Important point: ***The maximum of*** $(\operatorname{grad} f) \cdot \mathbf{u}$ ***is the length*** $|\operatorname{grad} f|$. In nonlinear examples, the gradient and steepest direction and slope will vary. But look at one particular point in Figure 13.13. Near that point, and near any point, the linear picture takes over.

On the graph of $f$, the special vectors are the level direction $\mathbf{L} = (f_y, -f_x, 0)$ and the uphill direction $\mathbf{U} = (f_x, f_y, f_x^2 + f_y^2)$ and the normal $\mathbf{N} = (f_x, f_y, -1)$. Problem 18 checks that those are perpendicular.

**EXAMPLE 3** The gradient of $f(x, y) = (14 - x^2 - y^2)/3$ is $\nabla f = (-2x/3, -2y/3)$.

On the surface, the normal vector is $\mathbf{N} = (-2x/3, -2y/3, -1)$. At the point $(1, 2, 3)$, this perpendicular is $\mathbf{N} = (-2/3, -4/3, -1)$. At the point $(1, 2)$ down in the base, the gradient is $(-2/3, -4/3)$. The length of grad $f$ is the slope $\sqrt{20}/3$.

Probably a hiker does not go straight up. A "grade" of $\sqrt{20}/3$ is fairly steep (almost 150%). To estimate the slope in other directions, measure the distance along the path between two contour lines. If $\Delta f = 1$ in a distance $\Delta s = 3$ the slope is about $1/3$. This calculation is not exact until the limit of $\Delta f/\Delta s$, which is $D_{\mathbf{u}}f$.

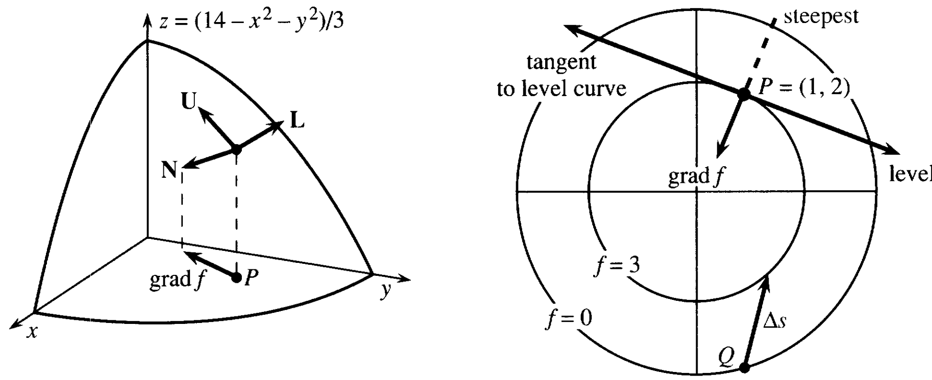



**Fig. 13.13** N perpendicular to surface and grad $f$ perpendicular to level line (in the base).

**EXAMPLE 4** The gradient of $f(x, y, z) = xy + yz + xz$ has three components.

The pattern extends from $f(x, y)$ to $f(x, y, z)$. The gradient is now the three-dimensional vector $(f_x, f_y, f_z)$. For this function grad $f$ is $(y + z, x + z, x + y)$. To draw the graph of $w = f(x, y, z)$ would require a four-dimensional picture, with axes in the $xyzw$ directions.

Notice the dimensions. The graph is a 3-dimensional "surface" in 4-dimensional space. The gradient is down below in the 3-dimensional base. The level sets of $f$ come from $xy + yz + zx = c$—they are 2-dimensional. The gradient is perpendicular to that level set (still down in 3 dimensions). The gradient is not **N**! The normal vector is $(f_x, f_y, f_z, -1)$, perpendicular to the surface up in 4-dimensional space.

**EXAMPLE 5** Find grad $z$ when $z(x, y)$ is given implicitly: $F(x, y, z) = x^2 + y^2 - z^2 = 0$.

In this case we find $z = \pm\sqrt{x^2 + y^2}$. The derivatives are $\pm x/\sqrt{x^2 + y^2}$ and $\pm y/\sqrt{x^2 + y^2}$, which go into grad $z$. But the point is this: To find that gradient faster, differentiate $F(x, y, z)$ as it stands. Then divide by $F_z$:

$$F_x dx + F_y dy + F_z dz = 0 \quad \text{or} \quad dz = (-F_x dx - F_y dy)/F_z. \tag{3}$$

***The gradient is*** $(-F_x/F_z, -F_y/F_z)$. Those derivatives are evaluated at $(x_0, y_0)$. The computation does not need the explicit function $z = f(x, y)$:

$$F = x^2 + y^2 - z^2 \Rightarrow F_x = 2x,\ F_y = 2y,\ F_z = -2z \Rightarrow \text{grad } z = (x/z, y/z).$$

To go uphill on the cone, move in the direction $(x/z, y/z)$. That gradient direction goes radially outward. The steepness of the cone is the length of the gradient vector:

$$|\text{grad } z| = \sqrt{(x/z)^2 + (y/z)^2} = 1 \text{ because } z^2 = x^2 + y^2 \text{ on the cone.}$$

### DERIVATIVES ALONG CURVED PATHS

On a straight path the derivative of $f$ is $D_{\mathbf{u}}f = (\text{grad} f) \cdot \mathbf{u}$. What is the derivative on a curved path? ***The path direction* u *is the tangent vector* T**. So replace **u** by **T**, which gives the "direction" of the curve.

The path is given by the position vector $\mathbf{R}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$. The velocity is $\mathbf{v} = (dx/dt)\mathbf{i} + (dy/dt)\mathbf{j}$. The tangent vector is $\mathbf{T} = \mathbf{v}/|\mathbf{v}|$. Notice the choice—to move at any speed (with $\mathbf{v}$) or to go at unit speed (with $\mathbf{T}$). There is the same choice for the derivative of $f(x, y)$ along this curve:

$$\textbf{\textit{rate of change}} \quad \frac{df}{dt} = (\text{grad} f) \cdot \mathbf{v} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \tag{4}$$

$$\textbf{\textit{slope}} \quad \frac{df}{ds} = (\text{grad} f) \cdot \mathbf{T} = \frac{\partial f}{\partial x}\frac{dx}{ds} + \frac{\partial f}{\partial y}\frac{dy}{ds} \tag{5}$$

The first involves *time*. If we move faster, $df/dt$ increases. The second involves *distance*. If we move a distance $ds$, at any speed, the function changes by $df$. So the slope in that direction is $df/ds$. Chapter 1 introduced velocity as $df/dt$ and slope as $dy/dx$ and mixed them up. Finally we see the difference.

Uniform motion on a straight line has $\mathbf{R} = \mathbf{R}_0 + \mathbf{v}t$. The velocity $\mathbf{v}$ is constant. The direction $\mathbf{T} = \mathbf{u} = \mathbf{v}/|\mathbf{v}|$ is also constant. The directional derivative is $(\text{grad} f) \cdot \mathbf{u}$, but the rate of change is $(\text{grad} f) \cdot \mathbf{v}$.

Equations (4) and (5) look like chain rules. They are chain rules. The next section extends $df/dt = (df/dx)(dx/dt)$ to more variables, proving (4) and (5). Here we focus on the meaning: ***df/ds is the derivative of f in the direction* u = T *along the curve.***

**EXAMPLE 7** Find $df/dt$ and $df/ds$ for $f = r$. The curve is $x = t^2$, $y = t$ in Figure 13.14a.

Solution The velocity along the curve is $\mathbf{v} = 2t\mathbf{i} + \mathbf{j}$. At the typical point $t = 1$ it is $\mathbf{v} = 2\mathbf{i} + \mathbf{j}$. The unit tangent is $\mathbf{T} = \mathbf{v}/\sqrt{5}$. The gradient is a unit vector $\mathbf{i}/\sqrt{2} + \mathbf{j}/\sqrt{2}$ pointing outward, when $f(x, y)$ is the distance $r$ from the center. The dot product with $\mathbf{v}$ is $df/dt = 3/\sqrt{2}$. The dot product with $\mathbf{T}$ is $df/ds = 3/\sqrt{10}$.

When we slow down to speed 1 (with $\mathbf{T}$), the changes in $f(x, y)$ slow down too.

**EXAMPLE 8** Find $df/ds$ for $f = xy$ along the circular path $x = \cos t$, $y = \sin t$.

First take a direct approach. On the circle, $xy$ equals $(\cos t)(\sin t)$. Its derivative comes from the product rule: $df/dt = \cos^2 t - \sin^2 t$. Normally this is different from $df/ds$, because the time $t$ need not equal the arc length $s$. There is a speed factor $ds/dt$ to divide by—but here the speed is 1. (A circle of length $s = 2\pi$ is completed at $t = 2\pi$.) Thus the slope $df/ds$ along the roller-coaster in Figure 13.14 is $\cos^2 t - \sin^2 t$.

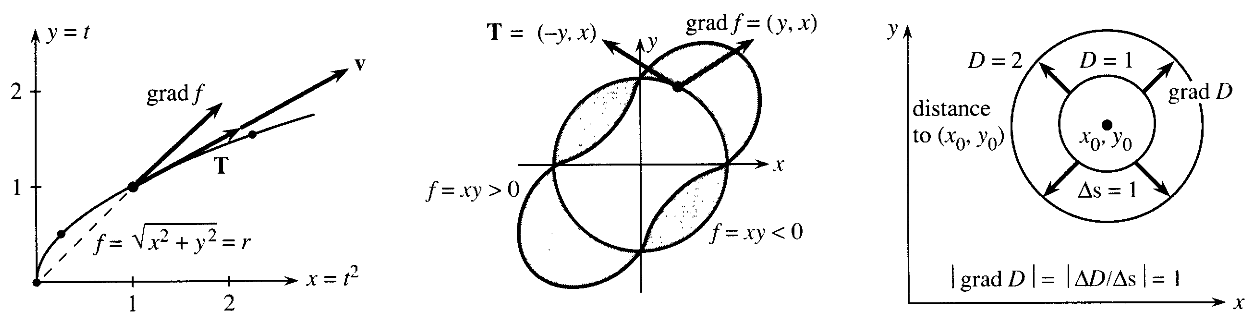



**Fig. 13.14** The distance $f = r$ changes along the curve. The slope of the roller-coaster is $(\text{grad} f) \cdot \mathbf{T}$. The distance $D$ from $(x_0, y_0)$ has grad $D$ = unit vector.

The second approach uses the vectors grad $f$ and $\mathbf{T}$. The gradient of $f = xy$ is $(y, x) = (\sin t, \cos t)$. The unit tangent vector to the path is $\mathbf{T} = (-\sin t, \cos t)$. Their dot product is the same $df/ds$:

$$\textit{slope along path} = (\text{grad } f) \cdot \mathbf{T} = -\sin^2 t + \cos^2 t.$$

### GRADIENTS WITHOUT COORDINATES

This section ends with a little "philosophy." What is the *coordinate-free definition* of the gradient? Up to now, grad $f = (f_x, f_y)$ depended totally on the choice of $x$ and $y$ axes. But the steepness of a surface is independent of the axes. Those are added later, to help us compute.

The steepness $df/ds$ involves only $f$ and the direction, nothing else. The gradient should be a "tensor"—its meaning does not depend on the coordinate system. The gradient has different formulas in different systems ($xy$ or $r\theta$ or ...), but the direction and length of grad $f$ are determined by $df/ds$—without any axes:

The ***direction*** of grad $f$ is the one in which $df/ds$ is largest.
The ***length*** $|\text{grad } f|$ is that largest slope.

The key equation is (***change in*** $f$) $\approx$ (***gradient of*** $f$) $\cdot$ (***change in position***). That is another way to write $\Delta f \approx f_x \Delta x + f_y \Delta y$. It is the multivariable form—we used two variables—of the basic linear approximation $\Delta y \approx (dy/dx)\Delta x$.

**EXAMPLE 9** $D(x, y) =$ distance from $(x, y)$ to $(x_0, y_0)$. Without derivatives prove $|\text{grad } D| = 1$. The graph of $D(x, y)$ is a cone with slope 1 and sharp point $(x_0, y_0)$.

**First question** In which direction does the distance $D(x, y)$ increase fastest?
Answer Going directly away from $(x_0, y_0)$. Therefore this is the direction of grad $D$.

**Second question** How quickly does $D$ increase in that steepest direction?
Answer A step of length $\Delta s$ increases $D$ by $\Delta s$. Therefore $|\text{grad } D| = \Delta s/\Delta s = 1$.

***Conclusion*** grad $D$ ***is a unit vector***. The derivatives of $D$ in Problem 48 are $(x - x_0)/D$ and $(y - y_0)/D$. The sum of their squares is 1, because $(x - x_0)^2 + (y - y_0)^2$ equals $D^2$.

## 13.4 EXERCISES

**Read-through questions**

$D_{\mathbf{u}}f$ gives the rate of change of __a__ in the direction __b__. It can be computed from the two derivatives __c__ in the special directions __d__. In terms of $u_1, u_2$ the formula is $D_{\mathbf{u}}f =$ __e__. This is a __f__ product of $\mathbf{u}$ with the vector __g__, which is called the __h__. For the linear function $f = ax + by$, the gradient is grad $f =$ __i__ and the directional derivative is $D_{\mathbf{u}}f =$ __j__ $\cdot$ __k__.

The gradient $\nabla f = (f_x, f_y)$ is not a vector in __l__ dimensions, it is a vector in the __m__. It is perpendicular to the __n__ lines. It points in the direction of __o__ climb. Its magnitude $|\text{grad } f|$ is __p__. For $f = x^2 + y^2$ the gradient points __q__ and the slope in that steepest direction is __r__.

The gradient of $f(x, y, z)$ is __s__. This is different from the gradient on the surface $F(x, y, z) = 0$, which is $-(F_x/F_z)\mathbf{i} +$ __t__. Traveling with velocity $\mathbf{v}$ on a curved path, the rate of change of $f$ is $df/dt =$ __u__. When the tangent direction is $\mathbf{T}$, the slope of $f$ is $df/ds =$ __v__. In a straight direction $\mathbf{u}$, $df/ds$ is the same as __w__.

**Compute grad $f$, then $D_{\mathbf{u}}f = (\text{grad } f) \cdot \mathbf{u}$, then $D_{\mathbf{u}}f$ at $P$.**

| | | |
|---|---|---|
| **1** $f(x, y) = x^2 - y^2$ | $\mathbf{u} = (\sqrt{3}/2, 1/2)$ | $P = (1, 0)$ |
| **2** $f(x, y) = 3x + 4y + 7$ | $\mathbf{u} = (3/5, 4/5)$ | $P = (0, \pi/2)$ |
| **3** $f(x, y) = e^x \cos y$ | $\mathbf{u} = (0, 1)$ | $P = (0, \pi/2)$ |
| **4** $f(x, y) = y^{10}$ | $\mathbf{u} = (0, -1)$ | $P = (1, -1)$ |

5 $f(x, y) =$ distance to $(0, 3)$ $\mathbf{u} = (1, 0)$ $P = (1, 1)$

**Find grad $f = (f_x, f_y, f_z)$ for the functions 6–8 from physics.**

6 $1/\sqrt{x^2 + y^2 + z^2}$ (point source at the origin)

7 $\ln(x^2 + y^2)$ (line source along $z$ axis)

8 $1/\sqrt{(x-1)^2 + y^2 + z^2} - 1/\sqrt{(x+1)^2 + y^2 + z^2}$ (dipole)

9 For $f = 3x^2 + 2y^2$ find the steepest direction and the level direction at (1, 2). Compute $D_{\mathbf{u}}f$ in those directions.

10 Example 2 claimed that $f = 3x + y + 1$ has steepest slope $\sqrt{10}$. Maximize $D_{\mathbf{u}}f = 3u_1 + u_2 = 3u_1 + \sqrt{1 - u_1^2}$.

11 **True or false**, when $f(x, y)$ is any smooth function:

(a) There is a direction $\mathbf{u}$ at $P$ in which $D_{\mathbf{u}}f = 0$.

(b) There is a direction $\mathbf{u}$ in which $D_{\mathbf{u}}f = \text{grad}\, f$.

(c) There is a direction $\mathbf{u}$ in which $D_{\mathbf{u}}f = 1$.

(d) The gradient of $f(x)g(x)$ equals $g$ grad $f + f$ grad $g$.

12 What is the gradient of $f(x)$? (One component only.) What are the two possible directions $\mathbf{u}$ and the derivatives $D_{\mathbf{u}}f$? What is the normal vector $\mathbf{N}$ to the curve $y = f(x)$? (Two components.)

**In 13–16 find the direction u in which $f$ increases fastest at $P =$ (1, 2). How fast?**

13 $f(x, y) = ax + by$ 14 $f(x, y) =$ smaller of $2x$ and $y$

15 $f(x, y) = e^{x-y}$ 16 $f(x, y) = \sqrt{5 - x^2 - y^2}$ (careful)

17 (Looking ahead) At a point where $f(x, y)$ is a maximum, what is grad $f$? Describe the level curve containing the maximum point $(x, y)$.

18 (a) Check by dot products that the normal and uphill and level directions on the graph are perpendicular: $\mathbf{N} = (f_x, f_y, -1), \mathbf{U} = (f_x, f_y, f_x^2 + f_y^2), \mathbf{L} = (f_y, -f_x, 0)$.

(b) $\mathbf{N}$ is ________ to the tangent plane, $\mathbf{U}$ and $\mathbf{L}$ are ________ to the tangent plane.

(c) The gradient is the $xy$ projection of ________ and also of ________. The projection of $\mathbf{L}$ points along the ________.

19 Compute the $\mathbf{N}, \mathbf{U}, \mathbf{L}$ vectors for $f = 1 - x + y$ and draw them at a point on the flat surface.

20 Compute $\mathbf{N}, \mathbf{U}, \mathbf{L}$ for $x^2 + y^2 - z^2 = 0$ and draw them at a typical point on the cone.

**With gravity in the negative $z$ direction, in what direction $-\mathbf{U}$ will water flow down the roofs 21–24?**

21 $z = 2x$ (flat roof) 22 $z = 4x - 3y$ (flat roof)

23 $z = \sqrt{1 - x^2 - y^2}$ (sphere) 24 $z = -\sqrt{x^2 + y^2}$ (cone)

25 Choose two functions $f(x, y)$ that depend only on $x + 2y$. Their gradients at (1, 1) are in the direction ________. Their level curves are ________.

26 The level curve of $f = y/x$ through (1, 1) is ________. The direction of the gradient must be ________. Check grad $f$.

27 Grad $f$ is perpendicular to $2\mathbf{i} + \mathbf{j}$ with length 1, and grad $g$ is parallel to $2\mathbf{i} + \mathbf{j}$ with length 5. Find grad $f$, grad $g$, $f$, and $g$.

28 **True or false**:

(a) If we know grad $f$, we know $f$.

(b) The line $x = y = -z$ is perpendicular to the plane $z = x + y$.

(c) The gradient of $z = x + y$ lies along that line.

29 Write down the level direction $\mathbf{u}$ for $\theta = \tan^{-1}(y/x)$ at the point (3, 4). Then compute grad $\theta$ and check $D_{\mathbf{u}}\theta = 0$.

30 On a circle around the origin, distance is $\Delta s = r\Delta\theta$. Then $d\theta/ds = 1/r$. Verify by computing grad $\theta$ and $\mathbf{T}$ and (grad $\theta$)$\cdot\mathbf{T}$.

31 At the point (2, 1, 6) on the mountain $z = 9 - x - y^2$, which way is up? On the roof $z = x + 2y + 2$, which way is down? The roof is ________ to the mountain.

32 Around the point $(1, -2)$ the temperature $T = e^{-x^2 - y^2}$ has $\Delta T \approx$ ________ $\Delta x +$ ________ $\Delta y$. In what direction $\mathbf{u}$ does it get hot fastest?

33 Figure A shows level curves of $z = f(x, y)$.

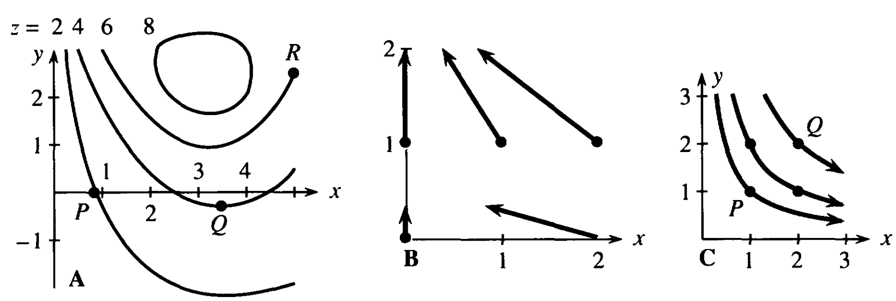
(a) Estimate the direction and length of grad $f$ at $P, Q, R$.

(b) Locate two points where grad $f$ is parallel to $\mathbf{i} + \mathbf{j}$.

(c) Where is $|\text{grad}\, f|$ largest? Where is it smallest?

(d) What is your estimate of $z_{max}$ on this figure?

(e) On the straight line from $P$ to $R$, describe $z$ and estimate its maximum.



34 A quadratic function $ax^2 + by^2 + cx + dy$ has the gradients shown in Figure B. Estimate $a, b, c, d$ and sketch two level curves.

35 The level curves of $f(x, y)$ are circles around (1, 1). The curve $f = c$ has radius $2c$. What is $f$? What is grad $f$ at (0, 0)?

36 Suppose grad $f$ is tangent to the hyperbolas $xy$ = constant in Figure C. Draw three level curves of $f(x, y)$. Is $|\text{grad } f|$ larger at $P$ or $Q$? Is $|\text{grad } f|$ constant along the hyperbolas? Choose a function that could be $f$: $x^2 + y^2$, $x^2 - y^2$, $xy$, $x^2y^2$.

37 Repeat Problem 36, if grad $f$ is *perpendicular* to the hyperbolas in Figure C.

38 If $f = 0, 1, 2$ at the points (0, 1), (1, 0), (2, 1), estimate grad $f$ by assuming $f = Ax + By + C$.

39 What functions have the following gradients?

(a) $(2x + y, x)$ (b) $(e^{x-y}, -e^{x-y})$ (c) $(y, -x)$ (careful)

40 Draw level curves of $f(x, y)$ if grad $f = (y, x)$.

**In 41–46 find the velocity v and the tangent vector T. Then compute the rate of change $df/dt = \text{grad} f \cdot \mathbf{v}$ and the slope $df/ds = \text{grad} f \cdot \mathbf{T}$.**

41 $f = x^2 + y^2$ $x = t$ $y = t^2$

42 $f = x$ $x = \cos 2t$ $y = \sin 2t$

43 $f = x^2 - y^2$ $x = x_0 + 2t$ $y = y_0 + 3t$

44 $f = xy$ $x = t^2 + 1$ $y = 3$

45 $f = \ln xyz$ $x = e^t$ $y = e^{2t}$ $z = e^{-t}$

46 $f = 2x^2 + 3y^2 + z^2$ $x = t$ $y = t^2$ $z = t^3$

47 (a) Find $df/ds$ and $df/dt$ for the roller-coaster $f = xy$ along the path $x = \cos 2t$, $y = \sin 2t$. (b) Change to $f = x^2 + y^2$ and explain why the slope is zero.

48 The distance $D$ from $(x, y)$ to (1, 2) has $D^2 = (x-1)^2 + (y-2)^2$. Show that $\partial D/\partial x = (x-1)/D$ and $\partial D/\partial y = (y-2)/D$ and $|\text{grad } D| = 1$. The graph of $D(x, y)$ is a ________ with its vertex at ________.

49 If $f = 1$ and grad $f = (2, 3)$ at the point (4, 5), find the tangent plane at (4, 5). If $f$ is a linear function, find $f(x, y)$.

50 Define the derivative of $f(x, y)$ in the direction $\mathbf{u} = (u_1, u_2)$ at the point $P = (x_0, y_0)$. What is $\Delta f$ (approximately)? What is $D_{\mathbf{u}} f$ (exactly)?

51 The slope of $f$ along a level curve is $df/ds$ = ________ = 0. This says that grad $f$ is perpendicular to the vector ________ in the level direction.

## 13.5 The Chain Rule

Calculus goes back and forth between solving problems and getting ready for harder problems. The first is "*application*," the second looks like "*theory*." If we minimize $f$ to save time or money or energy, that is an application. If we don't take derivatives to find the minimum—maybe because $f$ is a function of other functions, and we don't have a chain rule—then it is time for more theory. The chain rule is a fundamental working tool, because $f(g(x))$ appears all the time in applications. So do $f(g(x, y))$ and $f(x(t), y(t))$ and worse. We have to know their derivatives. Otherwise calculus can't continue with the applications.

You may instinctively say: Don't bother with the theory, just teach me the formulas. That is not possible. You now regard the derivative of sin $2x$ as a trivial problem, unworthy of an answer. That was not always so. Before the chain rule, the slopes of $\sin 2x$ and $\sin x^2$ and $\sin^2 x^2$ were hard to compute from $\Delta f/\Delta x$. We are now at the same point for $f(x, y)$. We know the *meaning* of $\partial f/\partial x$, but if $f = r \tan \theta$ and $x = r \cos \theta$ and $y = r \sin \theta$, we need a way to *compute* $\partial f/\partial x$. A little theory is unavoidable, if the problem-solving part of calculus is to keep going.

To repeat: ***The chain rule applies to a function of a function.*** In one variable that was $f(g(x))$. With two variables there are more possibilities:

| | | |
|---|---|---|
| **1.** $f(z)$ | with $z = g(x, y)$ | Find $\partial f/\partial x$ and $\partial f/\partial y$ |
| **2.** $f(x, y)$ | with $x = x(t)$, $y = y(t)$ | Find $df/dt$ |
| **3.** $f(x, y)$ | with $x = x(t, u)$, $y = y(t, u)$ | Find $\partial f/\partial t$ and $\partial f/\partial u$ |

All derivatives are assumed continuous. More exactly, the *input* derivatives like $\partial g/\partial x$ and $dx/dt$ and $\partial x/\partial u$ are continuous. Then the output derivatives like $\partial f/\partial x$ and $df/dt$ and $\partial f/\partial u$ will be continuous from the chain rule. We avoid points like $r = 0$ in polar coordinates—where $\partial r/\partial x = x/r$ has a division by zero.

***A Typical Problem*** Start with a function of $x$ and $y$, for example $x$ times $y$. Thus $f(x, y) = xy$. Change $x$ to $r \cos \theta$ and $y$ to $r \sin \theta$. The function becomes $(r \cos \theta)$ times $(r \sin \theta)$. We want its derivatives with respect to $r$ and $\theta$. First we have to decide on its *name*.

To be correct, we should not reuse the letter $f$. The new function can be $F$:

$$f(x, y) = xy \qquad f(r \cos \theta, r \sin \theta) = (r \cos \theta)(r \sin \theta) = F(r, \theta).$$

*Why not call it $f(r, \theta)$?* Because strictly speaking that is $r$ times $\theta$! If we follow the rules, then $f(x, y)$ is $xy$ and $f(r, \theta)$ should be $r\theta$. The new function $F$ does the right thing—it multiplies $(r \cos \theta)(r \sin \theta)$. But in many cases, the rules get bent and the letter $F$ is changed back to $f$.

This crime has already occurred. The end of the last page ought to say $\partial F/\partial t$. Instead the printer put $\partial f/\partial t$. The purpose of the chain rule is to find derivatives in the new variables $t$ and $u$ (or $r$ and $\theta$). In our example we want ***the derivative of $F$ with respect to $r$***. Here is the chain rule:

$$\frac{\partial F}{\partial r} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} = (y)(\cos \theta) + (x)(\sin \theta) = 2r \sin \theta \cos \theta.$$

I substituted $r \sin \theta$ and $r \cos \theta$ for $y$ and $x$. You immediately check the answer: $F(r, \theta) = r^2 \cos \theta \sin \theta$ does lead to $\partial F/\partial r = 2r \cos \theta \sin \theta$. The derivative is correct. The only incorrect thing—but we do it anyway—is to write $f$ instead of $F$.

**Question** What is $\dfrac{\partial f}{\partial \theta}$? Answer It is $\dfrac{\partial f}{\partial x}\dfrac{\partial x}{\partial \theta} + \dfrac{\partial f}{\partial y}\dfrac{\partial y}{\partial \theta}$.

### THE DERIVATIVES OF $f(g(x, y))$

Here $g$ depends on $x$ and $y$, and $f$ depends on $g$. Suppose $x$ moves by $dx$, while $y$ stays constant. Then $g$ moves by $dg = (\partial g/\partial x)dx$. When $g$ changes, $f$ also changes: $df = (df/dg)dg$. Now substitute for $dg$ to make the chain: $df = (df/dg)(\partial g/\partial x)dx$. This is the first rule:

> **13G** ***Chain rule for $f(g(x, y))$:*** $\dfrac{\partial f}{\partial x} = \dfrac{df}{dg}\dfrac{\partial g}{\partial x}$ and $\dfrac{\partial f}{\partial y} = \dfrac{df}{dg}\dfrac{\partial g}{\partial y}$. (1)

**EXAMPLE 1** Every $f(x + cy)$ satisfies the 1-way wave equation $\partial f/\partial y = c\, \partial f/\partial x$.

The inside function is $g = x + cy$. The outside function can be anything, $g^2$ or $\sin g$ or $e^g$. The composite function is $(x + cy)^2$ or $\sin(x + cy)$ or $e^{x+cy}$. In each separate case we could check that $\partial f/\partial y = c\, \partial f/\partial x$. The chain rule produces this equation in all cases at once, from $\partial g/\partial x = 1$ and $\partial g/\partial y = c$:

$$\frac{\partial f}{\partial x} = \frac{df}{dg}\frac{\partial g}{\partial x} = 1\frac{df}{dg} \quad \text{and} \quad \frac{\partial f}{\partial y} = \frac{df}{dg}\frac{\partial g}{\partial y} = c\frac{df}{dg} \quad \text{so} \quad \frac{\partial f}{\partial y} = c\frac{\partial f}{\partial x}. \tag{2}$$

This is important: $\partial f/\partial y = c\, \partial f/\partial x$ is our first example of a ***partial differential equation***. The unknown $f(x, y)$ has two variables. Two partial derivatives enter the equation.

Up to now we have worked with $dy/dt$ and ***ordinary differential equations***. The independent variable was time *or* space (and only one dimension in space). For partial differential equations the variables are time *and* space (possibly several dimensions in space). The great equations of mathematical physics—heat equation, wave equation, Laplace's equation—are partial differential equations.

Notice how the chain rule applies to $f = \sin xy$. Its $x$ derivative is $y \cos xy$. A patient reader would check that $f$ is $\sin g$ and $g$ is $xy$ and $f_x$ is $f_g g_x$. Probably you are not so patient—you know the derivative of $\sin xy$. Therefore we pass quickly to the next chain rule. Its outside function depends on *two* inside functions, and each of those depends on $t$. We want $df/dt$.

### THE DERIVATIVE OF $f(x(t), y(t))$

Before the formula, here is the idea. Suppose $t$ changes by $\Delta t$. That affects $x$ and $y$; they change by $\Delta x$ and $\Delta y$. There is a domino effect on $f$; it changes by $\Delta f$. Tracing backwards,

$$\Delta f \approx \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y \quad \text{and} \quad \Delta x \approx \frac{dx}{dt}\Delta t \quad \text{and} \quad \Delta y \approx \frac{dy}{dt}\Delta t.$$

Substitute the last two into the first, connecting $\Delta f$ to $\Delta t$. Then let $\Delta t \to 0$:

**13H** ***Chain rule for*** $f(x(t), y(t))$: $\displaystyle \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$ (3)

This is close to the one-variable rule $dz/dx = (dz/dy)(dy/dx)$. There we could "cancel" $dy$. (We actually canceled $\Delta y$ in $(\Delta z/\Delta y)(\Delta y/\Delta x)$, and then approached the limit.) Now $\Delta t$ affects $\Delta f$ in two ways, through $x$ and through $y$. *The chain rule has two terms*. If we cancel in $(\partial f/\partial x)(dx/dt)$ we only get one of the terms!

We mention again that the true name for $f(x(t), y(t))$ is $F(t)$ not $f(t)$. For $f(x, y, z)$ the rule has three terms: $f_x x_t + f_y y_t + f_z z_t$ is $f_t$ (or better $dF/dt$).

**EXAMPLE 2** How quickly does the temperature change when you drive to Florida?

Suppose the Midwest is at 30°F and Florida is at 80°F. Going 1000 miles south increases the temperature $f(x, y)$ by 50°, or .05 degrees per mile. Driving straight south at 70 miles per hour, the rate of increase is $(.05)(70) = 3.5$ degrees per hour. ***Note how*** **(degrees/mile)** ***times*** **(miles/hour)** ***equals*** **(degrees/hour)**. That is the ordinary chain rule $(df/dx)(dx/dt) = (df/dt)$—there is no $y$ variable going south.

If the road goes southeast, the temperature is $f = 30 + .05x + .01y$. Now $x(t)$ is distance south and $y(t)$ is distance east. What is $df/dt$ if the speed is still 70?

Solution $\displaystyle \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = .05\frac{70}{\sqrt{2}} + .01\frac{70}{\sqrt{2}} \approx 3$ degrees/hour.

In reality there is another term. The temperature also depends *directly on t*, because of night and day. The factor $\cos(2\pi t/24)$ has a period of 24 hours, and it brings an extra term into the chain rule:

$$\textbf{\textit{For }} f(x, y, t) \textbf{\textit{ the chain rule is }} \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial t}. \qquad (4)$$

This is the ***total*** derivative $df/dt$, from all causes. Changes in $x$, $y$, $t$ all affect $f$. The partial derivative $\partial f/\partial t$ is only one part of $df/dt$. (Note that $dt/dt = 1$.) If night and day add 12 cos$(2\pi t/24)$ to $f$, the extra term is $\partial f/\partial t = -\pi \sin(2\pi t/24)$. At nightfall that is $-\pi$ degrees per hour. You have to drive faster than 70 mph to get warm.

## SECOND DERIVATIVES

What is $d^2f/dt^2$? We need the derivative of (4), which is painful. It is like acceleration in Chapter 12, with many terms. So start with movement in a straight line.

Suppose $x = x_0 + t\cos\theta$ and $y = y_0 + t\sin\theta$. We are moving at the fixed angle $\theta$, with speed 1. The derivatives are $x_t = \cos\theta$ and $y_t = \sin\theta$ and $\cos^2\theta + \sin^2\theta = 1$. Then $df/dt$ is immediate from the chain rule:

$$f_t = f_x x_t + f_y y_t = f_x \cos\theta + f_y \sin\theta. \tag{5}$$

For the second derivative $f_{tt}$, apply this rule to $f_t$. Then $f_{tt}$ is

$$(f_t)_x \cos\theta + (f_t)_y \sin\theta = (f_{xx}\cos\theta + f_{yx}\sin\theta)\cos\theta + (f_{xy}\cos\theta + f_{yy}\sin\theta)\sin\theta.$$

Collect terms: $$f_{tt} = f_{xx}\cos^2\theta + 2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta. \tag{6}$$

In polar coordinates change $t$ to $r$. When we move in the $r$ direction, $\theta$ is fixed. Equation (6) gives $f_{rr}$ from $f_{xx}, f_{xy}, f_{yy}$. Second derivatives on curved paths (with new terms from the curving) are saved for the exercises.

**EXAMPLE 3** If $f_{xx}, f_{xy}, f_{yy}$ are all continuous and bounded by $M$, find a bound on $f_{tt}$. This is the second derivative along any line.

Solution Equation (6) gives $|f_{tt}| \leqslant M\cos^2\theta + M\sin 2\theta + M\sin^2\theta \leqslant 2M$. This upper bound $2M$ was needed in equation 13.3.14, for the error in linear approximation.

## THE DERIVATIVES OF $f(x(t, u), y(t, u))$

Suppose there are two inside functions $x$ and $y$, each depending on $t$ and $u$. When $t$ moves, $x$ and $y$ both move: $dx = x_t dt$ and $dy = y_t dt$. Then $dx$ and $dy$ force a change in $f$: $df = f_x dx + f_y dy$. The chain rule for $\partial f/\partial t$ is no surprise:

**13I** Chain rule for $f(x(t, u), y(t, u))$: $$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}. \tag{7}$$

This rule has $\partial/\partial t$ instead of $d/dt$, because of the extra variable $u$. The symbols remind us that $u$ is constant. Similarly $t$ is constant while $u$ moves, and there is a second chain rule for $\partial f/\partial u$: $f_u = f_x x_u + f_y y_u$.

**EXAMPLE 4** In polar coordinates find $f_\theta$ and $f_{\theta\theta}$. Start from $f(x, y) = f(r\cos\theta, r\sin\theta)$.

The chain rule uses the $\theta$ derivatives of $x$ and $y$:

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} = \left(\frac{\partial f}{\partial x}\right)(-r\sin\theta) + \left(\frac{\partial f}{\partial y}\right)(r\cos\theta). \tag{8}$$

The second $\theta$ derivative is harder, because (8) has four terms that depend on $\theta$. Apply the chain rule to the first term $\partial f/\partial x$. It is a function of $x$ and $y$, and $x$ and $y$ are functions of $\theta$:

$$\frac{\partial}{\partial \theta}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)\frac{\partial x}{\partial \theta} + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\frac{\partial y}{\partial \theta} = f_{xx}(-r\sin\theta) + f_{xy}(r\cos\theta).$$

The $\theta$ derivative of $\partial f/\partial y$ is similar. So apply the product rule to (8):

$$f_{\theta\theta} = [f_{xx}(-r\sin\theta) + f_{xy}(r\cos\theta)](-r\sin\theta) + f_x(-r\cos\theta)$$
$$+ [f_{yx}(-r\sin\theta) + f_{yy}(r\cos\theta)](r\cos\theta) + f_y(-r\sin\theta). \qquad (9)$$

This formula is not attractive. In mathematics, a messy formula is almost always a signal of asking the wrong question. In fact the combination $f_{xx} + f_{yy}$ is much more special than the separate derivatives. We might hope the same for $f_{rr} + f_{\theta\theta}$, but dimensionally that is impossible—since $r$ is a length and $\theta$ is an angle. The dimensions of $f_{xx}$ and $f_{yy}$ are matched by $f_{rr}$ and $f_r/r$ and $f_{\theta\theta}/r^2$. We could even hope that

$$f_{xx} + f_{yy} = f_{rr} + \frac{1}{r} f_r + \frac{1}{r^2} f_{\theta\theta}. \qquad (10)$$

This equation is true. Add (5) + (6) + (9) with $t$ changed to $r$. ***Laplace's equation*** $f_{xx} + f_{yy} = 0$ ***is now expressed in polar coordinates***: $f_{rr} + f_r/r + f_{\theta\theta}/r^2 = 0$.

## A PARADOX

Before leaving polar coordinates there is one more question. It goes back to $\partial r/\partial x$, which was practically the first example of partial derivatives:

$$\frac{\partial r}{\partial x} = \frac{\partial}{\partial x}\sqrt{x^2+y^2} = x/\sqrt{x^2+y^2} = x/r. \qquad (11)$$

My problem is this. We know that $x$ is $r\cos\theta$. So $x/r$ on the right side is $\cos\theta$. On the other hand $r$ is $x/\cos\theta$. So $\partial r/\partial x$ is also $1/\cos\theta$. ***How can*** $\partial r/\partial x$ ***lead to*** $\cos\theta$ ***one way and*** $1/\cos\theta$ ***the other way***?

I will admit that this cost me a sleepless night. There must be an explanation—we cannot end with $\cos\theta = 1/\cos\theta$. This paradox brings a new respect for partial derivatives. May I tell you what I finally noticed? You could cover up the next paragraph and think about the puzzle first.

The key to partial derivatives is to ask: ***Which variable is held constant***? In equation (11), $y$ is constant. But when $r = x/\cos\theta$ gave $\partial r/\partial x = 1/\cos\theta$, $\theta$ was constant. In both cases we change $x$ and look at the effect on $r$. The movement is on a horizontal line (constant $y$) or on a radial line (constant $\theta$). Figure 13.15 shows the difference.

***Remark*** This example shows that $\partial r/\partial x$ is different from $1/(\partial x/\partial r)$. The neat formula $(\partial r/\partial x)(\partial x/\partial r) = 1$ is not generally true. May I tell you what takes its place? We have to include $(\partial r/\partial y)(\partial y/\partial r)$. With two variables $xy$ and two variables $r\theta$, we need 2 by 2 matrices! Section 14.4 gives the details:

$$\begin{bmatrix} \partial r/\partial x & \partial r/\partial y \\ \partial\theta/\partial x & \partial\theta/\partial y \end{bmatrix}\begin{bmatrix} \partial x/\partial r & \partial x/\partial\theta \\ \partial y/\partial r & \partial y/\partial\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

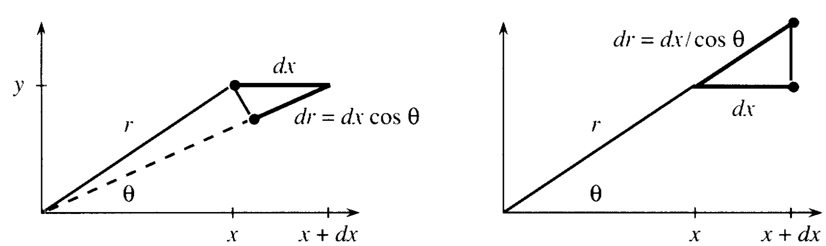



**Fig. 13.15** $dr = dx\cos\theta$ when $y$ is constant, $dr = dx/\cos\theta$ when $\theta$ is constant.

### NON-INDEPENDENT VARIABLES

This paradox points to a serious problem. In computing partial derivatives of $f(x, y, z)$, we assumed that $x, y, z$ were independent. Up to now, $x$ could move while $y$ and $z$ were fixed. In physics and chemistry and economics that may not be possible. If there is a relation between $x, y, z$, then $x$ can't move by itself.

**EXAMPLE 5** The gas law $PV = nRT$ relates pressure to volume and temperature. $P, V, T$ *are not independent*. What is the meaning of $\partial V/\partial P$? Does it equal $1/(\partial P/\partial V)$?

Those questions have no answers, until we say what is held constant. In the paradox, $\partial r/\partial x$ had one meaning for fixed $y$ and another meaning for fixed $\theta$. ***To indicate what is held constant, use an extra subscript*** (not denoting a derivative):

$$(\partial r/\partial x)_y = \cos\theta \qquad (\partial r/\partial x)_\theta = 1/\cos\theta. \tag{12}$$

$(\partial f/\partial P)_V$ has constant volume and $(\partial f/\partial P)_T$ has constant temperature. The usual $\partial f/\partial P$ has both $V$ and $T$ constant. But then the gas law won't let us change $P$.

**EXAMPLE 6** Let $f = 3x + 2y + z$. Compute $\partial f/\partial x$ on the plane $z = 4x + y$.

Solution 1 Think of $x$ and $y$ as independent. Replace $z$ by $4x + y$:

$$f = 3x + 2y + (4x + y) \quad \text{so} \quad (\partial f/\partial x)_y = 7.$$

Solution 2 Keep $x$ and $y$ independent. Deal with $z$ by the chain rule:

$$(\partial f/\partial x)_y = \partial f/\partial x + (\partial f/\partial z)(\partial z/\partial x) = 3 + (1)(4) = 7.$$

Solution 3 (***different***) Make $x$ and $z$ independent. Then $y = z - 4x$:

$$(\partial f/\partial x)_z = \partial f/\partial x + (\partial f/\partial y)(\partial y/\partial x) = 3 + (2)(-4) = -5.$$

Without a subscript, $\partial f/\partial x$ means: Take the $x$ derivative the usual way. The answer is $\partial f/\partial x = 3$, when $y$ and $z$ don't move. But on the plane $z = 4x + y$, one of them must move! 3 is only part of the total answer, which is $(\partial f/\partial x)_y = 7$ or $(\partial f/\partial x)_z = -5$.

Here is the geometrical meaning. We are on the plane $z = 4x + y$. The derivative $(\partial f/\partial x)_y$ moves $x$ but not $y$. To stay on the plane, $dz$ is $4dx$. The change in $f = 3x + 2y + z$ is $df = 3dx + 0 + dz = 7dx$.

**EXAMPLE 7** On the world line $x^2 + y^2 + z^2 = t^2$ find $(\partial f/\partial y)_{x,z}$ for $f = xyzt$.

The subscripts $x, z$ mean that $x$ and $z$ are fixed. The chain rule skips $\partial f/\partial x$ and $\partial f/\partial z$:

$$(\partial f/\partial y)_{x,z} = \partial f/\partial y + (\partial f/\partial t)(\partial t/\partial y) = xzt + (xyz)(y/t). \ \textit{Why } y/t?$$

**EXAMPLE 8** From the law $PV = T$, compute the product $(\partial P/\partial V)_T(\partial V/\partial T)_P(\partial T/\partial P)_V$.

Any intelligent person cancels $\partial V$'s, $\partial T$'s, and $\partial P$'s to get 1. The right answer is $-1$:

$$(\partial P/\partial V)_T = -T/V^2 \qquad (\partial V/\partial T)_P = 1/P \qquad (\partial T/\partial P)_V = V.$$

The product is $-T/PV$. ***This is*** $-1$ ***not*** $+1$! The chain rule is tricky (Problem 42).

**EXAMPLE 9** Implicit differentiation was used in Chapter 4. The chain rule explains it:

$$\textit{If } F(x, y) = 0 \textit{ then } F_x + F_y y_x = 0 \textit{ so } dy/dx = -F_x/F_y. \tag{13}$$

## 13.5 EXERCISES

### Read-through questions

The chain rule applies to a function of a __a__. The $x$ derivative of $f(g(x, y))$ is $\partial f/\partial x =$ __b__. The $y$ derivative is $\partial f/\partial y =$ __c__. The example $f=(x+y)^n$ has $g=$ __d__. Because $\partial g/\partial x = \partial g/\partial y$ we know that __e__ = __f__. This __g__ differential equation is satisfied by any function of $x+y$.

Along a path, the derivative of $f(x(t), y(t))$ is $df/dt =$ __h__. The derivative of $f(x(t), y(t), z(t))$ is __i__. If $f=xy$ then the chain rule gives $df/dt =$ __j__. That is the same as the __k__ rule! When $x=u_1t$ and $y=u_2t$ the path is __l__. The chain rule for $f(x, y)$ gives $df/dt =$ __m__. That is the __n__ derivative $D_{\mathbf{u}}f$.

The chain rule for $f(x(t, u), y(t, u))$ is $\partial f/\partial t =$ __o__. We don't write $df/dt$ because __p__. If $x = r\cos\theta$ and $y = r\sin\theta$, the variables $t, u$ change to __q__. In this case $\partial f/\partial r =$ __r__ and $\partial f/\partial\theta =$ __s__. That connects the derivatives in __t__ and __u__ coordinates. The difference between $\partial r/\partial x = x/r$ and $\partial r/\partial x = 1/\cos\theta$ is because __v__ is constant in the first and __w__ is constant in the second.

With a relation like $xyz=1$, the three variables are __x__ independent. The derivatives $(\partial f/\partial x)_y$ and $(\partial f/\partial x)_z$ and $(\partial f/\partial x)$ mean __y__ and __z__ and __A__. For $f=x^2+y^2+z^2$ with $xyz=1$, we compute $(\partial f/\partial x)_z$ from the chain rule __B__. In that rule $\partial z/\partial x =$ __C__ from the relation $xyz=1$.

**Find $f_x$ and $f_y$ in Problems 1–4. What equation connects them?**

**1** $f(x, y) = \sin(x+cy)$ **2** $f(x, y) = (ax+by)^{10}$

**3** $f(x, y) = e^{x+7y}$ **4** $f(x, y) = \ln(x+7y)$

**5** Find both terms in the $t$ derivative of $(g(x(t), y(t)))^3$.

**6** If $f(x, y) = xy$ and $x = u(t)$ and $y = v(t)$, what is $df/dt$? Probably all other rules for derivatives follow from the chain rule.

**7** The step function $f(x)$ is zero for $x<0$ and one for $x>0$. Graph $f(x)$ and $g(x) = f(x+2)$ and $h(x) = f(x+4)$. If $f(x+2t)$ represents a wall of water (a tidal wave), which way is it moving and how fast?

**8** The wave equation is $f_{tt} = c^2 f_{xx}$. (a) Show that $(x+ct)^n$ is a solution. (b) Find $C$ different from $c$ so that $(x+Ct)^n$ is also a solution.

**9** If $f = \sin(x-t)$, draw two lines in the $xt$ plane along which $f=0$. Between those lines sketch a sine wave. Skiing on top of the sine wave, what is your speed $dx/dt$?

**10** If you float at $x=0$ in Problem 9, do you go up first or down first? At time $t=4$ what is your height and upward velocity?

**11** *Laplace's equation* is $f_{xx}+f_{yy}=0$. Show from the chain rule that any function $f(x+iy)$ satisfies this equation if $i^2 = -1$. Check that $f=(x+iy)^2$ and its real part ________ and its imaginary part ________ all satisfy Laplace's equation.

**12** Equation (10) gave the polar form $f_{rr}+f_r/r+f_{\theta\theta}/r^2=0$ of Laplace's equation. (a) Check that $f=r^2e^{2i\theta}$ and its real part $r^2\cos 2\theta$ and its imaginary part $r^2\sin 2\theta$ all satisfy Laplace's equation. (b) Show from the chain rule that any function $f(re^{i\theta})$ satisfies this equation if $i^2=-1$.

**In Problems 13–18 find $df/dt$ from the chain rule (3).**

**13** $f=x^2+y^2,\ x=t,\ y=t^2$

**14** $f=\sqrt{x^2+y^2},\ x=t,\ y=t^2$

**15** $f=xy,\ x=1-\sqrt{t},\ y=1+\sqrt{t}$

**16** $f=x/y,\ x=e^t,\ y=2e^t$

**17** $f=\ln(x+y),\ x=e^t,\ y=e^t$

**18** $f=x^4,\ x=t,\ y=t$

**19** If a cone grows in height by $dh/dt=1$ and in radius by $dr/dt=2$, starting from zero, how fast is its volume growing at $t=3$?

**20** If a rocket has speed $dx/dt=6$ down range and $dy/dt=2t$ upward, how fast is it moving away from the launch point at (0, 0)? How fast is the angle $\theta$ changing, if $\tan\theta = y/x$?

**21** If a train approaches a crossing at 60 mph and a car approaches (at right angles) at 45 mph, how fast are they coming together? (a) Assume they are both 90 miles from the crossing. (b) Assume they are going to hit.

**22** In Example 2 does the temperature increase faster if you drive due south at 70 mph or southeast at 80 mph?

**23** On the line $x=u_1t,\ y=u_2t,\ z=u_3t$, what combination of $f_x, f_y, f_z$ gives $df/dt$? This is the *directional derivative* in 3D.

**24** On the same line $x=u_1t,\ y=u_2t,\ z=u_3t$, find a formula for $d^2f/dt^2$. Apply it to $f=xyz$.

**25** For $f(x, y, t) = x+y+t$ find $\partial f/\partial t$ and $df/dt$ when $x=2t$ and $y=3t$. Explain the difference.

**26** If $z=(x+y)^2$ then $x=\sqrt{z}-y$. Does $(\partial z/\partial x)(\partial x/\partial z)=1$?

**27** Suppose $x_t=t$ and $y_t=2t$, not constant as in (5–6). For $f(x, y)$ find $f_t$ and $f_{tt}$. The answer involves $f_x, f_y, f_{xx}, f_{xy}, f_{yy}$.

**28** Suppose $x_t=t$ and $y_t=t^2$. For $f=(x+y)^3$ find $f_t$ and then $f_{tt}$ from the chain rule.

**29** Derive $\partial f/\partial r = (\partial f/\partial x)\cos\theta + (\partial f/\partial y)\sin\theta$ from the chain rule. Why do we take $\partial x/\partial r$ as $\cos\theta$ and not $1/\cos\theta$?

**30** Compute $f_{xx}$ for $f(x, y) = (ax+by+c)^{10}$. If $x=t$ and $y=t$ compute $f_{tt}$. True or false: $(\partial f/\partial x)(\partial x/\partial t) = \partial f/\partial t$.

**31** Show that $\partial^2 r/\partial x^2 = y^2/r^3$ in two ways:

(1) Find the $x$ derivative of $\partial r/\partial x = x/\sqrt{x^2+y^2}$

(2) Find the $x$ derivative of $\partial r/\partial x = x/r$ by the chain rule.

**32** Reversing $x$ and $y$ in Problem 31 gives $r_{yy} = x^2/r^3$. But show that $r_{xy} = -xy/r^3$.

**33** If $\sin z = x + y$ find $(\partial z/\partial x)_y$ in two ways:

(1) Write $z = \sin^{-1}(x + y)$ and compute its derivative.

(2) Take $x$ derivatives of $\sin z = x + y$. Verify that these answers, explicit and implicit, are equal.

**34** By direct computation find $f_x$ and $f_{xx}$ and $f_{xy}$ for $f = \sqrt{x^2 + y^2}$.

**35** Find a formula for $\partial^2 f/\partial r \partial \theta$ in terms of the $x$ and $y$ derivatives of $f(x, y)$.

**36** Suppose $z = f(x, y)$ is solved for $x$ to give $x = g(y, z)$. Is it true that $\partial z/\partial x = 1/(\partial x/\partial z)$? Test on examples.

**37** Suppose $z = e^{xy}$ and therefore $x = (\ln z)/y$. Is it true or not that $(\partial z/\partial x) = 1/(\partial x/\partial z)$?

**38** If $x = x(t, u, v)$ and $y = y(t, u, v)$ and $z = z(t, u, v)$, find the $t$ derivative of $f(x, y, z)$.

**39** The $t$ derivative of $f(x(t, u), y(t, u))$ is in equation (7). What is $f_{tt}$?

**40** (a) For $f = x^2 + y^2 + z^2$ compute $\partial f/\partial x$ (no subscript, $x, y, z$ all independent).

(b) When there is a further relation $z = x^2 + y^2$, use it to remove $z$ and compute $(\partial f/\partial x)_y$.

(c) Compute $(\partial f/\partial x)_y$ using the chain rule $(\partial f/\partial x) + (\partial f/\partial z)(\partial z/\partial x)$.

(d) Why doesn't that chain rule contain $(\partial f/\partial y)(\partial y/\partial x)$?

**41** For $f = ax + by$ on the plane $z = 3x + 5y$, find $(\partial f/\partial x)_z$ and $(\partial f/\partial x)_y$ and $(\partial f/\partial z)_x$.

**42** The gas law in physics is $PV = nRT$ or a more general relation $F(P, V, T) = 0$. Show that the three derivatives in Example 8 still multiply to give $-1$. First find $(\partial P/\partial V)_T$ from $\partial F/\partial V + (\partial F/\partial P)(\partial P/\partial V)_T = 0$.

**43** If Problem 42 changes to four variables related by $F(x, y, z, t) = 0$, what is the corresponding product of four derivatives?

**44** Suppose $x = t + u$ and $y = tu$. Find the $t$ and $u$ derivatives of $f(x, y)$. Check when $f(x, y) = x^2 - 2y$.

**45** (a) For $f = r^2 \sin^2\theta$ find $f_x$ and $f_y$.

(b) For $f = x^2 + y^2$ find $f_r$ and $f_\theta$.

**46** On the curve $\sin x + \sin y = 0$, find $dy/dx$ and $d^2y/dx^2$ by implicit differentiation.

**47** (horrible) Suppose $f_{xx} + f_{yy} = 0$. If $x = u + v$ and $y = u - v$ and $f(x, y) = g(u, v)$, find $g_u$ and $g_v$. Show that $g_{uu} + g_{vv} = 0$.

**48** A function has *constant returns to scale* if $f(cx, cy) = cf(x, y)$. When $x$ and $y$ are doubled so are $f = \sqrt{x^2 + y^2}$ and $f = \sqrt{xy}$. In economics, input/output is constant. In mathematics $f$ is *homogeneous* of degree one.

Prove that $x\, \partial f/\partial x + y\, \partial f/\partial y = f(x, y)$, by computing the $c$ derivative at $c = 1$. Test this equation on the two examples and find a third example.

**49** ***True or false***: The directional derivative of $f(r, \theta)$ in the direction of $\mathbf{u}_\theta$ is $\partial f/\partial \theta$.

## 13.6 Maxima, Minima, and Saddle Points

The outstanding equation of differential calculus is also the simplest: $df/dx = 0$. The slope is zero and the tangent line is horizontal. Most likely we are at the top or bottom of the graph—a maximum or a minimum. This is the point that the engineer or manager or scientist or investor is looking for—maximum stress or production or velocity or profit. With more variables in $f(x, y)$ and $f(x, y, z)$, the problem becomes more realistic. The question still is: ***How to locate the maximum and minimum***?

The answer is in the *partial derivatives*. When the graph is level, they are zero. Deriving the equations $f_x = 0$ and $f_y = 0$ is pure mathematics and pure pleasure. Applying them is the serious part. We watch out for saddle points, and also for a minimum at a boundary point—this section takes extra time. Remember the steps for $f(x)$ in one-variable calculus:

1. The leading candidates are ***stationary*** points (where $df/dx = 0$).
2. The other candidates are ***rough points*** (no derivative) and ***endpoints*** ($a$ or $b$).
3. Maximum vs. minimum is decided by the sign of the ***second derivative***.

In two dimensions, a stationary point requires $\partial f/\partial x = 0$ and $\partial f/\partial y = 0$. The tangent line becomes a tangent plane. The endpoints $a$ and $b$ are replaced by a ***boundary curve***. In practice boundaries contain about 40% of the minima and 80% of the work.

Finally there are three second derivatives $f_{xx}, f_{xy}$, and $f_{yy}$. They tell how the graph bends away from the tangent plane—up at a *minimum*, down at a *maximum*, both ways at a *saddle point*. This will be determined by comparing $(f_{xx})(f_{yy})$ with $(f_{xy})^2$.

### STATIONARY POINT → HORIZONTAL TANGENT → ZERO DERIVATIVES

Suppose $f$ has a minimum at the point $(x_0, y_0)$. This may be an ***absolute minimum*** or only a ***local minimum***. In both cases $f(x_0, y_0) \leqslant f(x, y)$ near the point. For an absolute minimum, this inequality holds wherever $f$ is defined. For a local minimum, the inequality can fail far away from $(x_0, y_0)$. The bottom of your foot is an absolute minimum, the end of your finger is a local minimum.

We assume for now that $(x_0, y_0)$ is an *interior point* of the domain of $f$. At a boundary point, we cannot expect a horizontal tangent and zero derivatives.

Main conclusion: At a minimum or maximum (absolute or local) a nonzero derivative is impossible. The tangent plane would tilt. In some direction $f$ would decrease. Note that the minimum *point* is $(x_0, y_0)$, and the minimum *value* is $f(x_0, y_0)$.

> **13J** If derivatives exist at an interior minimum or maximum, they are zero:
>
> $$\partial f/\partial x = 0 \quad \text{and} \quad \partial f/\partial y = 0 \quad (\textit{together this is } \operatorname{grad} f = \mathbf{0}). \tag{1}$$
>
> For a function $f(x, y, z)$ of three variables, add the third equation $\partial f/\partial z = 0$.

The reasoning goes back to the one-variable case. That is because we look along the lines $x = x_0$ and $y = y_0$. The minimum of $f(x, y)$ is at the point where the lines meet. So this is also the minimum ***along each line separately***.

Moving in the $x$ direction along $y = y_0$, we find $\partial f/\partial x = 0$. Moving in the $y$ direction, $\partial f/\partial y = 0$ at the same point. ***The slope in every direction is zero***. In other words $\operatorname{grad} f = \mathbf{0}$.

Graphically, $(x_0, y_0)$ is the low point of the surface $z = f(x, y)$. Both cross sections in Figure 13.16 touch bottom. The phrase "if derivatives exist" rules out the vertex of a cone, which is a ***rough point***. The absolute value $f = |x|$ has a minimum without $df/dx = 0$, and so does the distance $f = r$. The rough point is $(0, 0)$.

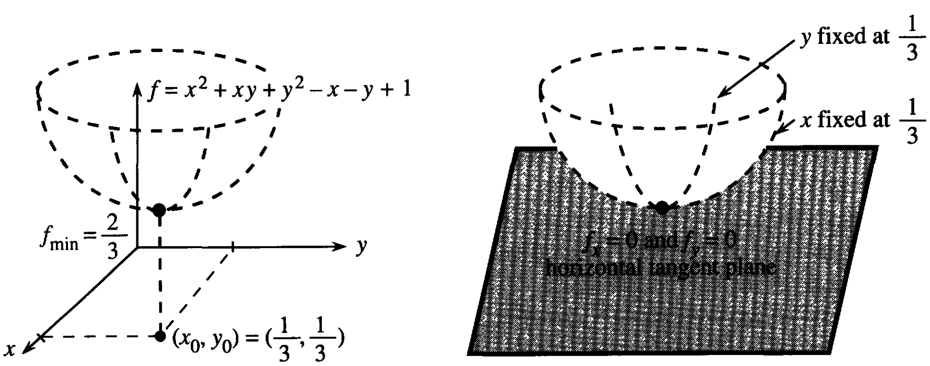



**Fig. 13.16** $\partial f/\partial x = 0$ and $\partial f/\partial y = 0$ at the minimum. Quadratic $f$ has linear derivatives.

**EXAMPLE 1** Minimize the quadratic $f(x, y) = x^2 + xy + y^2 - x - y + 1$.

To locate the minimum (or maximum), set $f_x = 0$ and $f_y = 0$:

$$f_x = 2x + y - 1 = 0 \qquad \text{and} \qquad f_y = x + 2y - 1 = 0. \tag{2}$$

Notice what's important: ***There are two equations for two unknowns*** $x$ ***and*** $y$. Since $f$ is quadratic, the equations are linear. Their solution is $x_0 = \frac{1}{3}$, $y_0 = \frac{1}{3}$ (the stationary point). This is actually a minimum, but to prove that you need to read further.

The constant 1 affects the minimum value $f = \frac{2}{3}$—but not the minimum point. The graph shifts up by 1. The linear terms $-x - y$ affect $f_x$ and $f_y$. They move the minimum away from (0, 0). The quadratic part $x^2 + xy + y^2$ makes the surface curve upwards. Without that curving part, a plane has its minimum and maximum at boundary points.

**EXAMPLE 2** (*Steiner's problem*) ***Find the point that is nearest to three given points.***

This example is worth your attention. We are locating an airport close to three cities. Or we are choosing a house close to three jobs. The problem is to get as near as possible to the corners of a triangle. The best point depends on the meaning of "*near.*"

The distance to the first corner $(x_1, y_1)$ is $d_1 = \sqrt{(x - x_1)^2 + (y - y_1)^2}$. The distances to the other corners $(x_2, y_2)$ and $(x_3, y_3)$ are $d_2$ and $d_3$. Depending on whether cost equals (*distance*) or (*distance*)$^2$ or (*distance*)$^p$, our problem will be:

$$\textit{Minimize} \quad d_1 + d_2 + d_3 \quad \textit{or} \quad d_1^2 + d_2^2 + d_3^2 \quad \textit{or even} \quad d_1^p + d_2^p + d_3^p.$$

The second problem is the easiest, when $d_1^2$ and $d_2^2$ and $d_3^2$ are quadratics:

$$f(x, y) = (x - x_1)^2 + (y - y_1)^2 + (x - x_2)^2 + (y - y_2)^2 + (x - x_3)^2 + (y - y_3)^2$$

$$\partial f/\partial x = 2[x - x_1 + x - x_2 + x - x_3] = 0 \qquad \partial f/\partial y = 2[y - y_1 + y - y_2 + y - y_3] = 0.$$

Solving $\partial f/\partial x = 0$ gives $x = \frac{1}{3}(x_1 + x_2 + x_3)$. Then $\partial f/\partial y = 0$ gives $y = \frac{1}{3}(y_1 + y_2 + y_3)$. The best point is the ***centroid of the triangle*** (Figure 13.17a). It is the nearest point to the corners when the cost is (distance)$^2$. Note how squaring makes the derivatives linear. ***Least squares*** dominates an enormous part of applied mathematics.

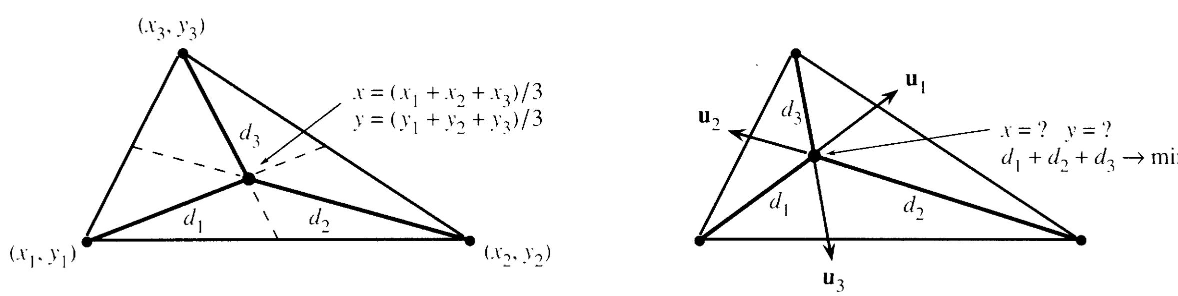



**Fig. 13.17** The centroid minimizes $d_1^2 + d_2^2 + d_3^2$. The Steiner point minimizes $d_1 + d_2 + d_3$.

The real "*Steiner problem*" is to minimize $f(x, y) = d_1 + d_2 + d_3$. We are laying down roads from the corners, with cost proportional to length. The equations $f_x = 0$ and $f_y = 0$ look complicated because of square roots. But the nearest point in Figure 13.17b has a remarkable property, which you will appreciate.

Calculus takes derivatives of $d_1^2 = (x - x_1)^2 + (y - y_1)^2$. The $x$ derivative leaves $2d_1(\partial d_1/\partial x) = 2(x - x_1)$. Divide both sides by $2d_1$:

$$\frac{\partial d_1}{\partial x} = \frac{x - x_1}{d_1} \quad \text{and} \quad \frac{\partial d_1}{\partial y} = \frac{y - y_1}{d_1} \quad \text{so grad } d_1 = \left(\frac{x - x_1}{d_1}, \frac{y - y_1}{d_1}\right). \tag{3}$$

***This gradient is a unit vector.*** The sum of $(x - x_1)^2/d_1^2$ and $(y - y_1)^2/d_1^2$ is $d_1^2/d_1^2 = 1$. This was already in Section 13.4: Distance increases with slope 1 away from the center. The gradient of $d_1$ (call it $\mathbf{u}_1$) is a unit vector from the center point $(x_1, y_1)$.

Similarly the gradients of $d_2$ and $d_3$ are unit vectors $\mathbf{u}_2$ and $\mathbf{u}_3$. They point directly away from the other corners of the triangle. The total cost is $f(x, y) = d_1 + d_2 + d_3$, so we add the gradients. The equations $f_x = 0$ and $f_y = 0$ combine into the vector equation

$$\text{grad } f = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 = \mathbf{0} \textit{ at the minimum.}$$

***The three unit vectors add to zero***! Moving away from one corner brings us closer to another. The nearest point to the three corners is where those movements cancel. This is the meaning of "grad $f = \mathbf{0}$ at the minimum."

It is unusual for three unit vectors to add to zero—this can only happen in one way. ***The three directions must form angles of*** 120°. The best point has this property, which is repeated in Figure 13.18a. The unit vectors cancel each other. At the "Steiner point," the roads to the corners make 120° angles. This optimal point solves the problem, except for one more possibility.

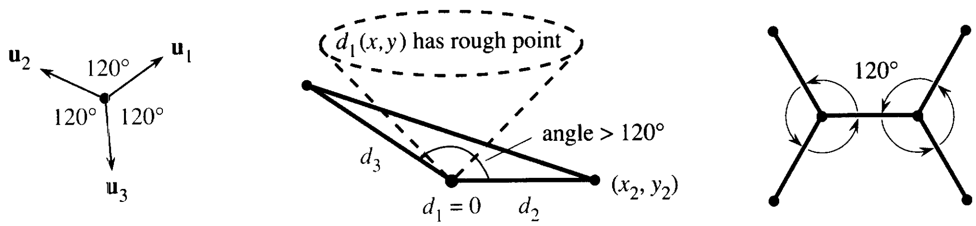



Fig. 13.18 Gradients $\mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3 = \mathbf{0}$ for 120° angles. Corner wins at wide angle. *Four corners.* In this case two branchpoints are better—still 120°.

The other possibility is a minimum at a ***rough point***. The graph of the distance function $d_1(x, y)$ is a cone. It has a sharp point at the center $(x_1, y_1)$. All three corners of the triangle are rough points for $d_1 + d_2 + d_3$, so all of them are possible minimizers.

***Suppose the angle at a corner exceeds*** 120°. Then there is no Steiner point. Inside the triangle, the angle would become even wider. The best point must be a rough point—one of the corners. The winner is the corner with the wide angle. In the figure that means $d_1 = 0$. Then the sum $d_2 + d_3$ comes from the two shortest edges.

***Summary*** The solution is at a 120° point or a wide-angle corner. That is the theory. The real problem is to compute the Steiner point—which I hope you will do.

***Remark* 1** Steiner's problem for *four points* is surprising. We don't minimize $d_1 + d_2 + d_3 + d_4$—there is a better problem. Connect the four points with roads, minimizing their total length, ***and allow the roads to branch***. A typical solution is in Figure 13.18c. The angles at the branch points are 120°. There are at most two branch points (two less than the number of corners).

***Remark* 2** For other powers $p$, the cost is $(d_1)^p + (d_2)^p + (d_3)^p$. The $x$ derivative is

$$\partial f/\partial x = p(d_1)^{p-2}(x - x_1) + p(d_2)^{p-2}(x - x_2) + p(d_3)^{p-2}(x - x_3). \tag{4}$$

The key equations are still $\partial f/\partial x = 0$ and $\partial f/\partial y = 0$. Solving them requires a computer and an algorithm. To share the work fairly, I will supply the algorithm (Newton's method) if you supply the computer. Seriously, this is a terrific example. It is typical of real problems—we know $\partial f/\partial x$ and $\partial f/\partial y$ but not the point where they are zero. You can calculate that nearest point, which changes as $p$ changes. You can also discover new mathematics, about how that point moves. I will collect all replies I receive to Problems 38 and 39.

## MINIMUM OR MAXIMUM ON THE BOUNDARY

Steiner's problem had no boundaries. The roads could go anywhere. But most applications have restrictions on $x$ and $y$, like $x \geqslant 0$ or $y \leqslant 0$ or $x^2 + y^2 \geqslant 1$. The minimum with these restrictions is probably higher than the absolute minimum. There are three possibilities:

(1) ***stationary point*** $f_x = 0, f_y = 0$ (2) ***rough point*** (3) ***boundary point***

That third possibility requires us to maximize or minimize $f(x, y)$ along the boundary.

**EXAMPLE 3** Minimize $f(x, y) = x^2 + xy + y^2 - x - y + 1$ in the *half-plane* $x \geqslant 0$.

The minimum in Example 1 was $\frac{2}{3}$. It occurred at $x_0 = \frac{1}{3}$, $y_0 = \frac{1}{3}$. *This point is still allowed*. It satisfies the restriction $x \geqslant 0$. So the minimum is not moved.

**EXAMPLE 4** Minimize the same $f(x, y)$ restricted to the *lower half-plane* $y \leqslant 0$.

Now the absolute minimum at $(\frac{1}{3}, \frac{1}{3})$ is not allowed. There are no rough points. We look for a minimum on the boundary line $y = 0$ in Figure 13.19a. Set $y = 0$, so $f$ depends only on $x$. Then choose the best $x$:

$$f(x, 0) = x^2 + 0 - x - 0 + 1 \qquad \text{and} \qquad f_x = 2x - 1 = 0.$$

The minimum is at $x = \frac{1}{2}$ and $y = 0$, where $f = \frac{3}{4}$. This is up from $\frac{2}{3}$.

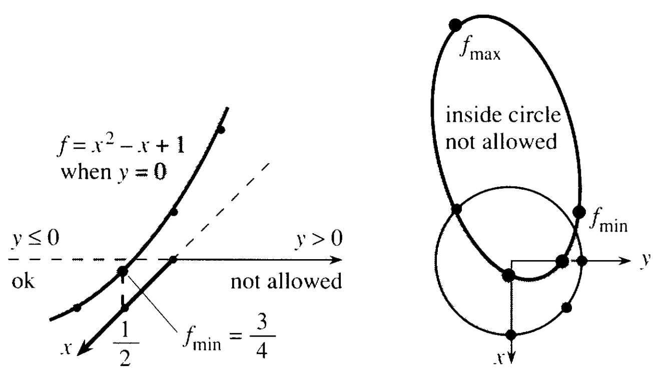



**Fig. 13.19** The boundaries $y = 0$ and $x^2 + y^2 = 1$ contain the minimum points.

**EXAMPLE 5** ***Minimize the same*** $f(x, y)$ ***on or outside the circle*** $x^2 + y^2 = 1$.

One possibility is $f_x = 0$ and $f_y = 0$. But this is at $(\frac{1}{3}, \frac{1}{3})$, inside the circle. The other possibility is a minimum at a boundary point, ***on the circle***.

To follow this boundary we can set $y = \sqrt{1 - x^2}$. The function $f$ gets complicated, and $df/dx$ is worse. There is a way to avoid square roots: Set $x = \cos t$ and $y = \sin t$. Then $f = x^2 + xy + y^2 - x - y + 1$ is a function of the angle $t$:

$$f(t) = 1 + \cos t \sin t - \cos t - \sin t + 1$$

$$df/dt = \cos^2 t - \sin^2 t + \sin t - \cos t = (\cos t - \sin t)(\cos t + \sin t - 1).$$

Now $df/dt = 0$ locates a minimum or maximum along the boundary. The first factor $(\cos t - \sin t)$ is zero when $x = y$. The second factor is zero when $\cos t + \sin t = 1$, or $x + y = 1$. Those points *on the circle* are the candidates. Problem 24 sorts them out, and Section 13.7 finds the minimum in a new way—using "***Lagrange multipliers***."

Minimization on a boundary is a serious problem—it gets difficult quickly—and multipliers are ultimately the best solution.

## MAXIMUM VS. MINIMUM VS. SADDLE POINT

How to separate the maximum from the minimum? When possible, try all candidates and decide. Compute $f$ at every stationary point and other critical point (maybe also out at infinity), and compare. Calculus offers another approach, based on ***second derivatives***.

With one variable the second derivative test was simple: $f_{xx} > 0$ at a minimum, $f_{xx} = 0$ at an inflection point, $f_{xx} < 0$ at a maximum. This is a local test, which may not give a global answer. But it decides whether the slope is increasing (bottom of the graph) or decreasing (top of the graph). We now find a similar test for $f(x, y)$.

The new test involves all three second derivatives. It applies where $f_x = 0$ and $f_y = 0$. The tangent plane is horizontal. ***We ask whether the graph of $f$ goes above or below that plane.*** The tests $f_{xx} > 0$ and $f_{yy} > 0$ guarantee a minimum in the $x$ and $y$ directions, but there are other directions.

**EXAMPLE 6** $f(x, y) = x^2 + 10xy + y^2$ has $f_{xx} = 2$, $f_{xy} = 10$, $f_{yy} = 2$ (***minimum or not?***)

All second derivatives are positive—but wait and see. The stationary point is (0, 0), where $\partial f/\partial x$ and $\partial f/\partial y$ are both zero. Our function is the sum of $x^2 + y^2$, which goes upward, and $10xy$ which has a saddle. The second derivatives must decide whether $x^2 + y^2$ or $10xy$ is stronger.

Along the $x$ axis, where $y = 0$ and $f = x^2$, our point is at the bottom. The minimum in the $x$ direction is at (0, 0). Similarly for the $y$ direction. But (0, 0) is ***not a minimum point*** for the whole function, because of $10xy$.

Try $x = 1$, $y = -1$. Then $f = 1 - 10 + 1$, which is negative. The graph goes *below* the $xy$ plane in that direction. The stationary point at $x = y = 0$ is a ***saddle point***.

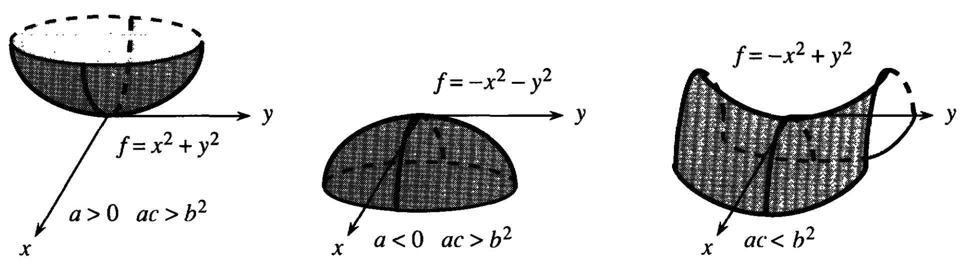



**Fig. 13.20** Minimum, maximum, saddle point based on the signs of $a$ and $ac - b^2$.

**EXAMPLE 7** $f(x, y) = x^2 + xy + y^2$ has $f_{xx} = 2$, $f_{xy} = 1$, $f_{yy} = 2$ (***minimum or not?***)

The second derivatives 2, 1, 2 are again positive. The graph curves up in the $x$ and $y$ directions. But there is a big difference from Example 6: $f_{xy}$ is reduced from 10 to 1. ***It is the size of $f_{xy}$ (not its sign!) that makes the difference.*** The extra terms $-x - y + 4$ in Example 1 moved the stationary point to $(\frac{1}{3}, \frac{1}{3})$. The second derivatives are still 2, 1, 2, and they pass the test for a minimum:

> **13K** At (0, 0) the quadratic function $f(x, y) = ax^2 + 2bxy + cy^2$ has a
>
> ***minimum*** if $\begin{matrix} a > 0 \\ ac > b^2 \end{matrix}$ ***maximum*** if $\begin{matrix} a < 0 \\ ac > b^2 \end{matrix}$ ***saddle point*** if $ac < b^2$.

For a direct proof, split $f(x, y)$ into two parts by "completing the square:"

$$ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2.$$

That algebra can be checked (notice the $2b$). It is the conclusion that's important:

if $a > 0$ and $ac > b^2$, both parts are positive: ***minimum*** at (0, 0)

if $a < 0$ and $ac > b^2$, both parts are negative: ***maximum*** at (0, 0)

if $ac < b^2$, the parts have opposite signs: ***saddle point*** at (0, 0).

Since the test involves the *square* of $b$, its sign has no importance. Example 6 had $b = 5$ and a saddle point. Example 7 had $b = \frac{1}{2}$ and a minimum. Reversing to $-x^2 - xy - y^2$ yields a maximum. So does $-x^2 + xy - y^2$.

Now comes the final step, from $ax^2 + 2bxy + cy^2$ to a general function $f(x, y)$. For all functions, quadratics or not, it is the ***second order terms*** that we test.

**EXAMPLE 8** $f(x, y) = e^x - x - \cos y$ has a stationary point at $x = 0$, $y = 0$.

The first derivatives are $e^x - 1$ and $\sin y$, both zero. The second derivatives are $f_{xx} = e^x = 1$ and $f_{yy} = \cos y = 1$ and $f_{xy} = 0$. We only use the derivatives *at the stationary point*. The first derivatives are zero, so the second order terms come to the front in the series for $e^x - x - \cos y$:

$$(1 + x + \tfrac{1}{2}x^2 + \cdots) - x - (1 - \tfrac{1}{2}y^2 + \cdots) = \tfrac{1}{2}x^2 + \tfrac{1}{2}y^2 + \text{higher order terms.} \quad (7)$$

There is a *minimum* at the origin. The quadratic part $\frac{1}{2}x^2 + \frac{1}{2}y^2$ goes upward. The $x^3$ and $y^4$ terms are too small to protest. Eventually those terms get large, but near a stationary point it is the quadratic that counts. We didn't need the whole series, because from $f_{xx} = f_{yy} = 1$ and $f_{xy} = 0$ we knew it would start with $\frac{1}{2}x^2 + \frac{1}{2}y^2$.

> **13L** The test in **13K** applies to the second derivatives $a = f_{xx}$, $b = f_{xy}$, $c = f_{yy}$ of any $f(x, y)$ at any stationary point. At all points the test decides whether the graph is concave up, concave down, or "indefinite."

**EXAMPLE 9** $f(x, y) = e^{xy}$ has $f_x = ye^{xy}$ and $f_y = xe^{xy}$. The stationary point is (0, 0).

The second derivatives at that point are $a = f_{xx} = 0$, $b = f_{xy} = 1$, and $c = f_{yy} = 0$. The test $b^2 > ac$ makes this a saddle point. Look at the infinite series:

$$e^{xy} = 1 + xy + \tfrac{1}{2}x^2y^2 + \cdots.$$

No linear term because $f_x = f_y = 0$: The origin is a ***stationary point***. No $x^2$ or $y^2$ term (only $xy$): The stationary point is a ***saddle point***.

At $x = 2$, $y = -2$ we find $f_{xx}f_{yy} > (f_{xy})^2$. The graph is concave up at that point—but it's not a minimum since the first derivatives are not zero.

The series begins with the constant term—not important. Then come the linear terms—extremely important. Those terms are decided by *first* derivatives, and they give the tangent plane. It is only at stationary points—when the linear part disappears and the tangent plane is horizontal—that second derivatives take over. Around any basepoint, ***these constant-linear-quadratic terms are the start of the Taylor series.***

### THE TAYLOR SERIES

We now put together the whole infinite series. It is a "Taylor series"—which means it is ***a power series that matches all derivatives of*** $f$ (at the basepoint). For one variable, the powers were $x^n$ when the basepoint was 0. For two variables, the powers are $x^n$ times $y^m$ when the basepoint is (0, 0). Chapter 10 multiplied the $n$th derivative $d^nf/dx^n$ by $x^n/n!$ *Now every mixed derivative* $(\partial/\partial x)^n(\partial/\partial y)^m f(x, y)$ *is computed at the basepoint* (subscript $_0$).

We multiply those numbers by $x^ny^m/n!m!$ to match each derivative of $f(x, y)$:

> **13M** When the basepoint is (0, 0), the Taylor series is a double sum $\Sigma\Sigma a_{nm}x^ny^m$. The term $a_{nm}x^ny^m$ has the same mixed derivative at (0, 0) as $f(x, y)$. The series is
>
> $$f(0, 0) + x\left(\frac{\partial f}{\partial x}\right)_0 + y\left(\frac{\partial f}{\partial y}\right)_0 + \frac{x^2}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)_0 + xy\left(\frac{\partial^2 f}{\partial x \partial y}\right)_0 + \frac{y^2}{2}\left(\frac{\partial^2 f}{\partial y^2}\right)_0$$
> $$+ \sum\sum_{n+m>2} \frac{x^ny^m}{n!m!}\left(\frac{\partial^{n+m} f}{\partial x^n \partial y^m}\right)_0.$$
>
> The derivatives of this series agree with the derivatives of $f(x, y)$ at the basepoint.

The first three terms are the linear approximation to $f(x, y)$. They give the tangent plane at the basepoint. The $x^2$ term has $n = 2$ and $m = 0$, so $n!m! = 2$. The $xy$ term has $n = m = 1$, and $n!m! = 1$. ***The quadratic part*** $\frac{1}{2}(ax^2 + 2bxy + cy^2)$ ***is in control when the linear part is zero.***

**EXAMPLE 10** All derivatives of $e^{x+y}$ equal one at the origin. The Taylor series is

$$e^{x+y} = 1 + x + y + \frac{x^2}{2} + xy + \frac{y^2}{2} + \cdots = \sum\sum \frac{x^ny^m}{n!m!}$$

This happens to have $ac = b^2$, the special case that was omitted in **13M** and **13N**. *It is the two-dimensional version of an inflection point.* The second derivatives fail to decide the concavity. When $f_{xx}f_{yy} = (f_{xy})^2$, the decision is passed up to the higher derivatives. But in ordinary practice, the Taylor series is stopped after the quadratics.

If the basepoint moves to $(x_0, y_0)$, the powers become $(x - x_0)^n(y - y_0)^m$—and all derivatives are computed at this new basepoint.

***Final question***: ***How would you compute a minimum numerically***? One good way is to solve $f_x = 0$ and $f_y = 0$. These are the functions $g$ and $h$ of Newton's method (Section 13.3). At the current point $(x_n, y_n)$, the derivatives of $g = f_x$ and $h = f_y$ give linear equations for the steps $\Delta x$ and $\Delta y$. Then the next point $x_{n+1} = x_n + \Delta x$, $y_{n+1} = y_n + \Delta y$ comes from those steps. The input is $(x_n, y_n)$, the output is the new point, and the linear equations are

$$\begin{aligned}(g_x)\Delta x + (g_y)\Delta y &= -g(x_n, y_n)\\ (h_x)\Delta x + (h_y)\Delta y &= -h(x_n, y_n)\end{aligned} \quad \text{or} \quad \begin{aligned}(f_{xx})\Delta x + (f_{xy})\Delta y &= -f_x(x_n, y_n)\\ (f_{xy})\Delta x + (f_{yy})\Delta y &= -f_y(x_n, y_n).\end{aligned} \qquad (5)$$

When the second derivatives of $f$ are available, use Newton's method.

When the problem is too complicated to go beyond first derivatives, here is an alternative—***steepest descent***. The goal is to move down the graph of $f(x, y)$, like a boulder rolling down a mountain. The steepest direction at any point is given by the *gradient*, with a minus sign to go down instead of up. So move in the direction $\Delta x = -s\,\partial f/\partial x$ and $\Delta y = -s\,\partial f/\partial y$.

The question is: How far to move? Like a boulder, a steep start may not aim directly toward the minimum. The stepsize $s$ is monitored, to end the step when the function $f$ starts upward again (Problem 54). At the end of each step, compute first derivatives and start again in the new steepest direction.

## 13.6 EXERCISES

**Read-through questions**

A minimum occurs at a __a__ point (where $f_x = f_y = 0$) or a __b__ point (no derivative) or a __c__ point. Since $f = x^2 - xy + 2y$ has $f_x =$ __d__ and $f_y =$ __e__, the stationary point is $x =$ __f__, $y =$ __g__. This is not a minimum, because $f$ decreases when __h__.

The minimum of $d^2 = (x - x_1)^2 + (y - y_1)^2$ occurs at the rough point __i__. The graph of $d$ is a __j__ and grad $d$ is a __k__ vector that points __l__. The graph of $f = |xy|$ touches bottom along the lines __m__. Those are "rough lines" because the derivative __n__. The maximum of $d$ and $f$ must occur on the __o__ of the allowed region because it doesn't occur __p__.

When the boundary curve is $x = x(t)$, $y = y(t)$, the derivative of $f(x, y)$ along the boundary is __q__ (chain rule). If $f = x^2 + 2y^2$ and the boundary is $x = \cos t$, $y = \sin t$, then $df/dt =$ __r__. It is zero at the points __s__. The maximum is at __t__ and the minimum is at __u__. Inside the circle $f$ has an absolute minimum at __v__.

To separate maximum from minimum from __w__, compute the __x__ derivatives at a __y__ point. The tests for a minimum are __z__. The tests for a maximum are __A__. In case $ac <$ __B__ or $f_{xx}f_{yy} <$ __C__, we have a __D__. At all points these tests decide between concave up and __E__ and "indefinite." For $f = 8x^2 - 6xy + y^2$, the origin is a __F__. The signs of $f$ at (1, 0) and (1, 3) are __G__.

The Taylor series for $f(x, y)$ begins with the six terms __H__. The coefficient of $x^n y^m$ is __I__. To find a stationary point numerically, use __J__ or __K__.

**Find all stationary points ($f_x = f_y = 0$) in 1–16. Separate minimum from maximum from saddle point. Test 13K applies to $a = f_{xx}$, $b = f_{xy}$, $c = f_{yy}$.**

**1** $x^2 + 2xy + 3y^2$ **2** $xy - x + y$

**3** $x^2 + 4xy + 3y^2 - 6x - 12y$ **4** $x^2 - y^2 + 4y$

**5** $x^2y^2 - x$ **6** $xe^y - e^x$

**7** $-x^2 + 2xy - 3y^2$ **8** $(x + y)^2 + (x + 2y - 6)^2$

**9** $x^2 + y^2 + z^2 - 4z$ **10** $(x + y)(x + 2y - 6)$

**11** $(x - y)^2$ **12** $(1 + x^2)/(1 + y^2)$

**13** $(x + y)^2 - (x + 2y)^2$ **14** $\sin x - \cos y$

**15** $x^3 + y^3 - 3x^2 + 3y^2$ **16** $8xy - x^4 - y^4$

**17** A rectangle has sides on the $x$ and $y$ axes and a corner on the line $x + 3y = 12$. Find its maximum area.

**18** A box has a corner at (0, 0, 0) and all edges parallel to the axes. If the opposite corner $(x, y, z)$ is on the plane $3x + 2y + z = 1$, what position gives maximum volume? Show first that the problem maximizes $xy - 3x^2y - 2xy^2$.

**19** (Straight line fit, Section 11.4) Find $x$ and $y$ to minimize the error

$$E = (x + y)^2 + (x + 2y - 5)^2 + (x + 3y - 4)^2.$$

Show that this gives a minimum not a saddle point.

**20** (*Least squares*) What numbers $x, y$ come closest to satisfying the three equations $x - y = 1$, $2x + y = -1$, $x + 2y = 1$? Square and add the errors, $(x - y - 1)^2 +$ ________ $+$ ________. Then minimize.

**21** Minimize $f = x^2 + xy + y^2 - x - y$ restricted by
(a) $x \leqslant 0$ (b) $y \geqslant 1$ (c) $x \leqslant 0$ and $y \geqslant 1$.

**22** Minimize $f = x^2 + y^2 + 2x + 4y$ in the regions
(a) all $x, y$ (b) $y \geqslant 0$ (c) $x \geqslant 0, y \geqslant 0$

**23** Maximize and minimize $f = x + \sqrt{3}y$ on the circle $x = \cos t$, $y = \sin t$.

**24** Example 5 followed $f = x^2 + xy + y^2 - x - y + 1$ around the circle $x^2 + y^2 = 1$. The four stationary points have $x = y$ or $x + y = 1$. Compute $f$ at those points and locate the minimum.

**25** (a) Maximize $f = ax + by$ on the circle $x^2 + y^2 = 1$.
(b) Minimize $x^2 + y^2$ on the line $ax + by = 1$.

**26** For $f(x, y) = \frac{1}{4}x^4 - xy + \frac{1}{4}y^4$, what are the equations $f_x = 0$ and $f_y = 0$? What are their solutions? What is $f_{\min}$?

**27** Choose $c > 0$ so that $f = x^2 + xy + cy^2$ has a saddle point at (0, 0). Note that $f > 0$ on the lines $x = 0$ and $y = 0$ and $y = x$ and $y = -x$, so checking four directions does not confirm a minimum.

**Problems 28–42 minimize the Steiner distance $f = d_1 + d_2 + d_3$ and related functions. A computer is needed for 33 and 36–39.**

**28** Draw the triangle with corners at (0, 0), (1, 1), and (1, −1). By symmetry the Steiner point will be on the $x$ axis. Write down the distances $d_1, d_2, d_3$ to $(x, 0)$ and find the $x$ that minimizes $d_1 + d_2 + d_3$. Check the 120° angles.

**29** Suppose three unit vectors add to zero. Prove that the angles between them must be 120°.

**30** In three dimensions, Steiner minimizes the total distance $f(x, y, z) = d_1 + d_2 + d_3 + d_4$ from four points. Show that grad $d_1$ is still a unit vector (in which direction?) At what angles do four unit vectors add to zero?

**31** With four points in a plane, the Steiner problem allows branches (Figure 13.18c). Find the shortest network connecting the corners of a rectangle, if the side lengths are (a) 1 and 2 (b) 1 and 1 (two solutions for a square) (c) 1 and 0.1.

**32** Show that a Steiner point (120° angles) can never be outside the triangle.

**33** Write a program to minimize $f(x, y) = d_1 + d_2 + d_3$ by Newton's method in equation (5). Fix two corners at (0, 0), (3, 0), vary the third from (1, 1) to (2, 1) to (3, 1) to (4, 1), and compute Steiner points.

**34** Suppose one side of the triangle goes from (−1, 0) to (1, 0). Above that side are points from which the lines to (−1, 0) and (1, 0) meet at a 120° angle. Those points lie on a circular arc—draw it and find its center and its radius.

**35** Continuing Problem 34, there are circular arcs for all three sides of the triangle. On the arcs, every point sees one side of the triangle at a 120° angle. Where is the Steiner point? (Sketch three sides with their arcs.)

**36** Invent an algorithm to converge to the Steiner point based on Problem 35. Test it on the triangles of Problem 33.

**37** Write a code to minimize $f = d_1^4 + d_2^4 + d_3^4$ by solving $f_x = 0$ and $f_y = 0$. Use Newton's method in equation (5).

**38** Extend the code to allow all powers $p \geqslant 1$, not only $p = 4$. Follow the minimizing point from the centroid at $p = 2$ to the Steiner point at $p = 1$ (try $p = 1.8, 1.6, 1.4, 1.2$).

**39** Follow the minimizing point with your code as $p$ increases: $p = 2$, $p = 4$, $p = 8$, $p = 16$. Guess the limit at $p = \infty$ and test whether it is equally distant from the three corners.

**40** At $p = \infty$ we are making the largest of the distances $d_1, d_2, d_3$ as small as possible. The best point for a 1, 1, $\sqrt{2}$ right triangle is ________.

**41** Suppose the road from corner 1 is wider than the others, and the total cost is $f(x, y) = \sqrt{2}\, d_1 + d_2 + d_3$. Find the gradient of $f$ and the angles at which the best roads meet.

**42** Solve Steiner's problem for *two* points. Where is $d_1 + d_2$ a minimum? Solve also for three points if only the three corners are allowed.

**Find all derivatives at (0, 0). Construct the Taylor series:**

**43** $f(x, y) = (x + y)^3$ **44** $f(x, y) = xe^y$

**45** $f(x, y) = \ln(1 - xy)$

**Find $f_x, f_y, f_{xx}, f_{xy}, f_{yy}$ at the basepoint. Write the quadratic approximation to $f(x, y)$ — the Taylor series through second-order terms:**

**46** $f = e^{x+y}$ at (0, 0) **47** $f = e^{x+y}$ at (1, 1)

**48** $f = \sin x \cos y$ at (0, 0) **49** $f = x^2 + y^2$ at (1, −1)

**50** The Taylor series around $(x, y)$ is also written with steps $h$ and $k$: $f(x + h, y + k) = f(x, y) + h$ ________ $+ k$ ________ $+ \frac{1}{2}h^2$ ________ $+ hk$ ________ $+ \cdots$. Fill in those four blanks.

**51** Find lines along which $f(x, y)$ is constant (these functions have $f_{xx}f_{yy} = f_{xy}^2$ or $ac = b^2$):
(a) $f = x^2 - 4xy + 4y^2$ (b) $f = e^x e^y$

**52** For $f(x, y, z)$ the first three terms after $f(0, 0, 0)$ in the Taylor series are ________. The next six terms are ________.

**53** (a) For the error $f - f_L$ in linear approximation, the Taylor series at (0, 0) starts with the quadratic terms ________.
(b) The graph of $f$ goes up from its tangent plane (and $f > f_L$) if ________. Then $f$ is concave upward.
(c) For (0, 0) to be a minimum we also need ________

**54** The gradient of $x^2 + 2y^2$ at the point (1, 1) is (2, 4). Steepest descent is along the line $x = 1 - 2s$, $y = 1 - 4s$ (minus sign to go downward). Minimize $x^2 + 2y^2$ with respect to the stepsize $s$. That locates the next point ________, where steepest descent begins again.

**55** Newton's method minimizes $x^2 + 2y^2$ in one step. Starting at $(x_0, y_0) = (1, 1)$, find $\Delta x$ and $\Delta y$ from equation (5).

**56** If $f_{xx} + f_{yy} = 0$, show that $f(x, y)$ cannot have an interior maximum or minimum (only saddle points).

**57** The value of $x$ theorems and $y$ exercises is $f = x^2 y$ (maybe). The most that a student or author can deal with is $4x + y = 12$. Substitute $y = 12 - 4x$ and maximize $f$. Show that the line $4x + y = 12$ is tangent to the level curve $x^2 y = f_{\max}$.

**58** The desirability of $x$ houses and $y$ yachts is $f(x, y)$. The constraint $px + qy = k$ limits the money available. The cost of a house is ________, the cost of a yacht is ________. Substitute $y = (k - px)/q$ into $f(x, y) = F(x)$ and use the chain rule for $dF/dx$. Show that the slope $-f_x/f_y$ at the best $x$ is $-p/q$.

**59** At the farthest point in a baseball field, explain why the fence is perpendicular to the line from home plate. Assume it is not a rough point (corner) or endpoint (foul line).

## 13.7 Constraints and Lagrange Multipliers

This section faces up to a practical problem. We often minimize one function $f(x, y)$ while another function $g(x, y)$ is fixed. There is a ***constraint*** on $x$ and $y$, given by $g(x, y) = k$. This restricts the material available or the funds available or the energy available. With this constraint, the problem is to do the best possible ($f_{max}$ or $f_{min}$).

At the absolute minimum of $f(x, y)$, the requirement $g(x, y) = k$ is probably violated. In that case the minimum point is not allowed. We cannot use $f_x = 0$ and $f_y = 0$— those equations don't account for $g$.

***Step* 1** Find equations for the ***constrained minimum*** or ***constrained maximum***. They will involve $f_x$ and $f_y$ and also $g_x$ and $g_y$, which give local information about $f$ and $g$. To see the equations, look at two examples.

**EXAMPLE 1** Minimize $f = x^2 + y^2$ subject to the constraint $g = 2x + y = k$.

*Trial runs* The constraint allows $x = 0$, $y = k$, where $f = k^2$. Also $(\frac{1}{2}k, 0)$ satisfies the constraint, and $f = \frac{1}{4}k^2$ is smaller. Also $x = y = \frac{1}{3}k$ gives $f = \frac{2}{9}k^2$ (best so far).

*Idea of solution* Look at the level curves of $f(x, y)$ in Figure 13.21. They are circles $x^2 + y^2 = c$. When $c$ is small, the circles do not touch the line $2x + y = k$. There are no points that satisfy the constraint, when $c$ is too small. *Now increase $c$.*

Eventually the growing circles $x^2 + y^2 = c$ will *just touch the line* $x + 2y = k$. The point where they touch is the winner. It gives the smallest value of $c$ that can be achieved on the line. The touching point is $(x_{min}, y_{min})$, and the value of $c$ is $f_{min}$.

What equation describes that point? When the circle touches the line, they are *tangent*. They have the same slope. *The perpendiculars to the circle and the line go in the same direction.* That is the key fact, which you see in Figure 13.21a. The direction perpendicular to $f = c$ is given by grad $f = (f_x, f_y)$. The direction perpendicular to $g = k$ is given by grad $g = (g_x, g_y)$. The key equation says that those two vectors are parallel. One gradient vector is a multiple of the other gradient vector, with a multiplier $\lambda$ (called lambda) that is unknown:

> **13N** At the minimum of $f(x, y)$ subject to $g(x, y) = k$, the gradient of $f$ is parallel to the gradient of $g$—with an unknown number $\lambda$ as the multiplier:
>
> $$\text{grad}\, f = \lambda\, \text{grad}\, g \quad \text{so} \quad \frac{\partial f}{\partial x} = \lambda \frac{\partial g}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y} = \lambda \frac{\partial g}{\partial y}. \tag{1}$$

***Step* 2** There are now three unknowns $x$, $y$, $\lambda$. There are also *three equations*:

$$\begin{aligned} \partial f/\partial x = \lambda \partial g/\partial x \text{ is } & \quad 2x = 2\lambda \\ \partial f/\partial y = \lambda \partial g/\partial y \text{ is } & \quad 2y = \lambda \\ g(x, y) = k \text{ is } & \quad 2x + y = k. \end{aligned} \tag{2}$$

In the third equation, substitute $2\lambda$ for $2x$ and $\frac{1}{2}\lambda$ for $y$. Then $2x + y$ equals $\frac{5}{2}\lambda$ equals $k$. Knowing $\lambda = \frac{2}{5}k$, go back to the first two equations for $x$, $y$, and $f_{min}$:

$$x = \lambda = \frac{2}{5}k, \quad y = \frac{1}{2}\lambda = \frac{1}{5}k, \quad f_{min} = \left(\frac{2}{5}k\right)^2 + \left(\frac{1}{5}k\right)^2 = \frac{5}{25}k^2 = \frac{1}{5}k^2.$$

The winning point $(x_{min}, y_{min})$ is $(\frac{2}{5}k, \frac{1}{5}k)$. It minimizes the "distance squared," $f = x^2 + y^2 = \frac{1}{5}k^2$, from the origin to the line.

**Question** *What is the meaning of the Lagrange multiplier $\lambda$?*

Mysterious answer ***The derivative of $\frac{1}{5}k^2$ is $\frac{2}{5}k$, which equals $\lambda$. The multiplier $\lambda$ is the derivative of $f_{\min}$ with respect to $k$.*** Move the line by $\Delta k$, and $f_{\min}$ changes by about $\lambda \Delta k$. Thus the Lagrange multiplier measures the ***sensitivity*** to $k$.

Pronounce his name "Lagronge" or better "Lagrongh" as if you are French.

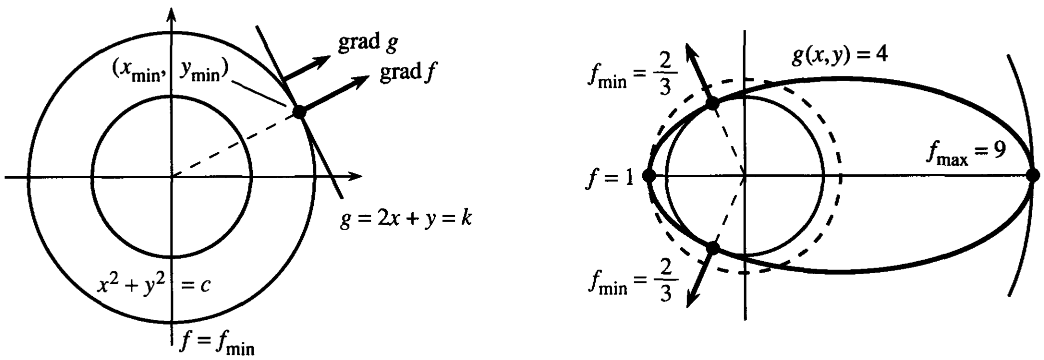



**Fig. 13.21** Circles $f = c$ tangent to line $g = k$ and ellipse $g = 4$: parallel gradients.

**EXAMPLE 2** Maximize and minimize $f = x^2 + y^2$ on the ellipse $g = (x-1)^2 + 4y^2 = 4$.

Idea and equations The circles $x^2 + y^2 = c$ grow until they touch the ellipse. The touching point is $(x_{\min}, y_{\min})$ and that smallest value of $c$ is $f_{\min}$. As the circles grow they cut through the ellipse. Finally there is a point $(x_{\max}, y_{\max})$ where the last circle touches. That largest value of $c$ is $f_{\max}$.

The minimum and maximum are described by the same rule: the circle is tangent to the ellipse (Figure 13.21b). ***The perpendiculars go in the same direction.*** Therefore $(f_x, f_y)$ is a multiple of $(g_x, g_y)$, and the unknown multiplier is $\lambda$:

$$\begin{aligned} f_x &= \lambda g_x: && 2x = \lambda 2(x-1) \\ f_y &= \lambda g_y: && 2y = \lambda 8y \\ g &= k: && (x-1)^2 + 4y^2 = 4. \end{aligned} \tag{3}$$

Solution The second equation allows two possibilities: $y = 0$ or $\lambda = \frac{1}{4}$. Following up $y = 0$, the last equation gives $(x-1)^2 = 4$. Thus $x = 3$ or $x = -1$. Then the first equation gives $\lambda = 3/2$ or $\lambda = 1/2$. The values of $f$ are $x^2 + y^2 = 3^2 + 0^2 = 9$ and $x^2 + y^2 = (-1)^2 + 0^2 = 1$.

Now follow $\lambda = 1/4$. The first equation yields $x = -1/3$. Then the last equation requires $y^2 = 5/9$. Since $x^2 = 1/9$ we find $x^2 + y^2 = 6/9 = 2/3$. This is $f_{\min}$.

Conclusion The equations (3) have four solutions, at which the circle and ellipse are tangent. The four points are $(3, 0)$, $(-1, 0)$, $(-1/3, \sqrt{5}/3)$, and $(-1/3, -\sqrt{5}/3)$. The four values of $f$ are $9, 1, \frac{2}{3}, \frac{2}{3}$.

***Summary*** The three equations are $f_x = \lambda g_x$ and $f_y = \lambda g_y$ and $g = k$. The unknowns are $x$, $y$, and $\lambda$. There is no absolute system for solving the equations (unless they are linear; then use elimination or Cramer's Rule). Often the first two equations yield $x$ and $y$ in terms of $\lambda$, and substituting into $g = k$ gives an equation for $\lambda$.

At the minimum, the level curve $f(x, y) = c$ is tangent to the constraint curve $g(x, y) = k$. If that constraint curve is given parametrically by $x(t)$ and $y(t)$, then minimizing $f(x(t), y(t))$ uses the chain rule:

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = 0 \quad \text{or} \quad (\text{grad} f)\cdot(\text{tangent to curve}) = 0.$$

This is the calculus proof that grad$f$ is perpendicular to the curve. *Thus* grad$f$ *is parallel to* grad $g$. This means $(f_x, f_y) = \lambda(g_x, g_y)$.

We have lost $f_x = 0$ and $f_y = 0$. But a new function $L$ has *three* zero derivatives:

**13O** The Lagrange function is $L(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - k)$. Its three derivatives are $L_x = L_y = L_\lambda = 0$ at the solution:

$$\frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - \lambda\frac{\partial g}{\partial x} = 0 \qquad \frac{\partial L}{\partial y} = \frac{\partial f}{\partial y} - \lambda\frac{\partial g}{\partial y} = 0 \qquad \frac{\partial L}{\partial \lambda} = -g + k = 0. \qquad (4)$$

Note that $\partial L/\partial\lambda = 0$ automatically produces $g = k$. The constraint is "***built in***" to $L$. *Lagrange has included a term* $\lambda(g - k)$, which is destined to be zero—but its derivatives are absolutely needed in the equations! At the solution, $g = k$ and $L = f$ and $\partial L/\partial k = \lambda$.

What is important is $f_x = \lambda g_x$ and $f_y = \lambda g_y$, coming from $L_x = L_y = 0$. In words: The constraint $g = k$ forces $dg = g_x dx + g_y dy = 0$. *This restricts the movements* $dx$ *and* $dy$. They must keep to the curve. The equations say that $df = f_x dx + f_y dy$ is equal to $\lambda dg$. Thus $df$ is zero *in the allowed direction*—which is the key point.

## MAXIMUM AND MINIMUM WITH TWO CONSTRAINTS

The whole subject of min(max)imization is called ***optimization***. Its applications to business decisions make up *operations research*. The special case of linear functions is always important—in this part of mathematics it is called ***linear programming***. A book about those subjects won't fit inside a calculus book, but we can take one more step—to allow a second constraint.

The function to minimize or maximize is now $f(x, y, z)$. The constraints are $g(x, y, z) = k_1$ and $h(x, y, z) = k_2$. The multipliers are $\lambda_1$ and $\lambda_2$. We need at least three variables $x, y, z$ because two constraints would completely determine $x$ and $y$.

**13P** To minimize $f(x, y, z)$ subject to $g(x, y, z) = k_1$ and $h(x, y, z) = k_2$, solve five equations for $x, y, z, \lambda_1, \lambda_2$. Combine $g = k_1$ and $h = k_2$ with

$$\frac{\partial f}{\partial x} = \lambda_1\frac{\partial g}{\partial x} + \lambda_2\frac{\partial h}{\partial x}, \qquad \frac{\partial f}{\partial y} = \lambda_1\frac{\partial g}{\partial y} + \lambda_2\frac{\partial h}{\partial y}, \qquad \frac{\partial f}{\partial z} = \lambda_1\frac{\partial g}{\partial z} + \lambda_2\frac{\partial h}{\partial z}. \qquad (5)$$

Figure 13.22a shows the geometry behind these equations. For convenience $f$ is $x^2 + y^2 + z^2$, so we are minimizing distance (squared). The constraints $g = x + y + z = 9$ and $h = x + 2y + 3z = 20$ are linear—their graphs are planes. The constraints keep $(x, y, z)$ on both planes—and therefore on the line where they meet. We are finding *the squared distance from* $(0, 0, 0)$ *to a line.*

What equation do we solve? The level surfaces $x^2 + y^2 + z^2 = c$ are spheres. They grow as $c$ increases. The first sphere to touch the line is tangent to it. That touching point gives the solution (the smallest $c$). ***All three vectors*** grad$f$, grad $g$, grad $h$ ***are perpendicular to the line***:

line tangent to sphere $\Rightarrow$ grad $f$ perpendicular to line

line in both planes $\Rightarrow$ grad $g$ and grad $h$ perpendicular to line.

Thus grad $f$, grad $g$, grad $h$ are *in the same plane*—perpendicular to the line. With three vectors in a plane, grad $f$ is a combination of grad $g$ and grad $h$:

$$(f_x, f_y, f_z) = \lambda_1(g_x, g_y, g_z) + \lambda_2(h_x, h_y, h_z). \tag{6}$$

This is the key equation (5). It applies to curved surfaces as well as planes.

**EXAMPLE 3** Minimize $x^2 + y^2 + z^2$ when $x + y + z = 9$ and $x + 2y + 3z = 20$.

In Figure 13.22b, the normals to those planes are grad $g = (1, 1, 1)$ and grad $h = (1, 2, 3)$. The gradient of $f = x^2 + y^2 + z^2$ is $(2x, 2y, 2z)$. The equations (5)–(6) are

$$2x = \lambda_1 + \lambda_2, \quad 2y = \lambda_1 + 2\lambda_2, \quad 2z = \lambda_1 + 3\lambda_2.$$

Substitute these $x, y, z$ into the other two equations $g = x + y + z = 9$ and $h = 20$:

$$\frac{\lambda_1 + \lambda_2}{2} + \frac{\lambda_1 + 2\lambda_2}{2} + \frac{\lambda_1 + 3\lambda_2}{2} = 9 \quad \text{and} \quad \frac{\lambda_1 + \lambda_2}{2} + 2\frac{\lambda_1 + 2\lambda_2}{2} + 3\frac{\lambda_1 + 3\lambda_2}{2} = 20.$$

After multiplying by 2, these simplify to $3\lambda_1 + 6\lambda_2 = 18$ and $6\lambda_1 + 14\lambda_2 = 40$. The solutions are $\lambda_1 = 2$ and $\lambda_2 = 2$. Now the previous equations give $(x, y, z) = (2, 3, 4)$.

The Lagrange function with two constraints is $L(x, y, z, \lambda_1, \lambda_2) = f - \lambda_1(g - k_1) - \lambda_2(h - k_2)$. Its five derivatives are zero—those are our five equations. Lagrange has increased the number of unknowns from 3 to 5, by adding $\lambda_1$ and $\lambda_2$. The best point $(2, 3, 4)$ gives $f_{\min} = 29$. The $\lambda$'s give $\partial f/\partial k$—the sensitivity to changes in 9 and 20.

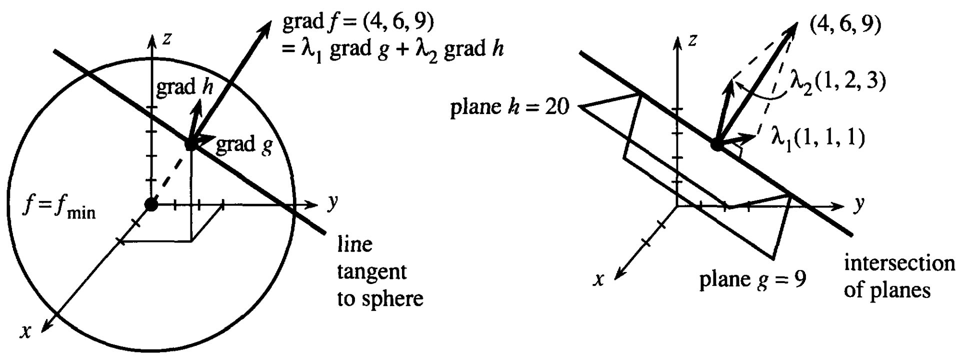



**Fig. 13.22** Perpendicular vector grad $f$ is a combination $\lambda_1$ grad $g + \lambda_2$ grad $h$.

## INEQUALITY CONSTRAINTS

In practice, applications involve ***inequalities*** as well as equations. The constraints might be $g \leqslant k$ and $h \geqslant 0$. The first means: It is not required to use the whole resource $k$, but you cannot use more. The second means: $h$ measures a quantity that cannot be negative. ***At the minimum point, the multipliers must satisfy the same inequalities***: $\lambda_1 \leqslant 0$ and $\lambda_2 \geqslant 0$. There are inequalities on the $\lambda$'s when there are inequalities in the constraints.

Brief reasoning: With $g \leqslant k$ the minimum can be *on or inside* the constraint curve. Inside the curve, where $g < k$, we are free to move in all directions. The constraint is not really constraining. This brings back $f_x = 0$ and $f_y = 0$ and $\lambda = 0$—an ordinary minimum. On the curve, where $g = k$ constrains the minimum from going lower, we have $\lambda < 0$. We don't know in advance which to expect.

For 100 constraints $g_i \leqslant k_i$, there are 100 $\lambda$'s. Some $\lambda$'s are zero (when $g_i < k_i$) and some are nonzero (when $g_i = k_i$). It is those $2^{100}$ possibilities that make optimization interesting. In ***linear programming*** with two variables, the constraints are $x \geqslant 0$, $y \geqslant 0$:

**EXAMPLE 4** Minimize $f = 5x + 6y$ with $g = x + y = 4$ and $h = x \geqslant 0$ and $H = y \geqslant 0$.

The constraint $g = 4$ is an equation, $h$ and $H$ yield inequalities. Each has its own Lagrange multiplier—and the inequalities require $\lambda_2 \geqslant 0$ and $\lambda_3 \geqslant 0$. The derivatives of $f, g, h, H$ are no problem to compute:

$$\frac{\partial f}{\partial x} = \lambda_1 \frac{\partial g}{\partial x} + \lambda_2 \frac{\partial h}{\partial x} + \lambda_3 \frac{\partial H}{\partial x} \quad \text{yields} \quad 5 = \lambda_1 + \lambda_2$$
$$\frac{\partial f}{\partial y} = \lambda_1 \frac{\partial g}{\partial y} + \lambda_2 \frac{\partial h}{\partial y} + \lambda_3 \frac{\partial H}{\partial y} \quad \text{yields} \quad 6 = \lambda_1 + \lambda_3. \tag{7}$$

Those equations make $\lambda_3$ larger than $\lambda_2$. Therefore $\lambda_3 > 0$, which means that the constraint on $H$ must be an equation. (Inequality for the multiplier means equality for the constraint.) In other words $H = y = 0$. Then $x + y = 4$ leads to $x = 4$. The solution is at $(x_{\min}, y_{\min}) = (4, 0)$, where $f_{\min} = 20$.

At this minimum, $h = x = 4$ is above zero. The multiplier for the constraint $h \geqslant 0$ must be $\lambda_2 = 0$. Then the first equation gives $\lambda_1 = 5$. As always, the multiplier measures sensitivity. When $g = 4$ is increased by $\Delta k$, the cost $f_{\min} = 20$ is increased by $5\Delta k$. In economics $\lambda_1 = 5$ is called a ***shadow price****—it is the cost of increasing the constraint.*

Behind this example is a nice problem in geometry. The constraint curve $x + y = 4$ is a line. The inequalities $x \geqslant 0$ and $y \geqslant 0$ leave a piece of that line—from $P$ to $Q$ in Figure 13.23. The level curves $f = 5x + 6y = c$ move out as $c$ increases, until they touch the line. ***The first touching point is*** $Q = (4, 0)$, ***which is the solution. It is always an endpoint***—or a corner of the triangle $PQR$. It gives the smallest cost $f_{\min}$, which is $c = 20$.

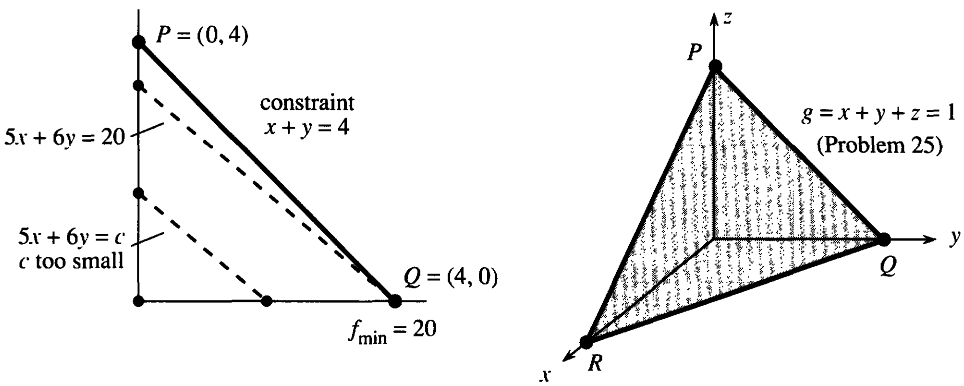



**Fig. 13.23** Linear programming: $f$ and $g$ are linear, inequalities cut off $x$ and $y$.

## 13.7 EXERCISES

**Read-through questions**

A restriction $g(x, y) = k$ is called a __a__. The minimizing equations for $f(x, y)$ subject to $g = k$ are __b__. The number $\lambda$ is the Lagrange __c__. Geometrically, grad $f$ is __d__ to grad $g$ at the minimum. That is because the __e__ curve $f = f_{\min}$ is __f__ to the constraint curve $g = k$. The number $\lambda$ turns out to be the derivative of __g__ with respect to __h__. The Lagrange function is $L =$ __i__ and the three equations for $x, y, \lambda$ are __j__ and __k__ and __l__.

To minimize $f=x^2-y$ subject to $g=x-y=0$, the three equations for $x, y, \lambda$ are __m__. The solution is __n__. In this example the curve $f(x, y)=f_{\min}=$ __o__ is a __p__ which is __q__ to the line $g=0$ at $(x_{\min}, y_{\min})$.

With two constraints $g(x, y, z)=k_1$ and $h(x, y, z)=k_2$ there are __r__ multipliers. The five unknowns are __s__. The five equations are __t__. The level surface $f=f_{\min}$ is __u__ to the curve where $g=k_1$ and $h=k_2$. Then grad $f$ is __v__ to this curve, and so are grad $g$ and __w__. Thus __x__ is a combination of grad $g$ and __y__. With nine variables and six constraints, there will be __z__ multipliers and eventually __A__ equations. If a constraint is an __B__ $g \leqslant k$, then its multiplier must satisfy $\lambda \leqslant 0$ at a minimum.

**1** Example 1 minimized $f=x^2+y^2$ subject to $2x+y=k$. Solve the constraint equation for $y=k-2x$, substitute into $f$, and minimize this function of $x$. The minimum is at $(x, y)=$ ________, where $f=$ ________.

**Note**: This direct approach *reduces* to one unknown $x$. Lagrange *increases* to $x, y, \lambda$. But Lagrange is better when the first step of solving for $y$ is difficult or impossible.

**Minimize and maximize $f(x, y)$ in 2–6. Find $x$, $y$, and $\lambda$.**

**2** $f=x^2y$ with $g=x^2+y^2=1$

**3** $f=x+y$ with $g=\dfrac{1}{x}+\dfrac{1}{y}=1$

**4** $f=3x+y$ with $g=x^2+9y^2=1$

**5** $f=x^2+y^2$ with $g=x^6+y^6=2$.

**6** $f=x+y$ with $g=x^{1/3}y^{2/3}=k$. With $x=$ capital and $y=$ labor, $g$ is a Cobb-Douglas function in economics. Draw two of its level curves.

**7** Find the point on the circle $x^2+y^2=13$ where $f=2x-3y$ is a maximum. Explain the answer.

**8** Maximize $ax+by+cz$ subject to $x^2+y^2+z^2=k^2$. Write your answer as the Schwarz inequality for dot products: $(a, b, c)\cdot(x, y, z) \leqslant$ ________$k$.

**9** Find the plane $z=ax+by+c$ that best fits the points $(x, y, z)=(0, 0, 1), (1, 0, 0), (1, 1, 2), (0, 1, 2)$. The answer $a, b, c$ minimizes the sum of $(z-ax-by-c)^2$ at the four points.

**10** The base of a triangle is the top of a rectangle (5 sides, combined area $=1$). What dimensions minimize the distance around?

**11** Draw the hyperbola $xy=-1$ touching the circle $g=x^2+y^2=2$. The minimum of $f=xy$ on the circle is reached at the points ________. The equations $f_x=\lambda g_x$ and $f_y=\lambda g_y$ are satisfied at those points with $\lambda=$ ________.

**12** Find the maximum of $f=xy$ on the circle $g=x^2+y^2=2$ by solving $f_x=\lambda g_x$ and $f_y=\lambda g_y$ and substituting $x$ and $y$ into $f$. Draw the level curve $f=f_{\max}$ that touches the circle.

**13** Draw the level curves of $f=x^2+y^2$ with a closed curve $C$ across them to represent $g(x, y)=k$. Mark a point where $C$ *crosses* a level curve. Why is that point not a minimum of $f$ on $C$? Mark a point where $C$ is *tangent* to a level curve. Is that the minimum of $f$ on $C$?

**14** On the circle $g=x^2+y^2=1$, Example 5 of 13.6 minimized $f=xy-x-y$. (a) Set up the three Lagrange equations for $x, y, \lambda$. (b) The first two equations give $x=y=$ ________. (c) There is another solution for the special value $\lambda=-\frac{1}{2}$, when the equations become ________. This is easy to miss but it gives $f_{\min}=-1$ at the point ________.

**Problems 15–18 develop the theory of Lagrange multipliers.**

**15** (*Sensitivity*) Certainly $L=f-\lambda(g-k)$ has $\partial L/\partial k=\lambda$. Since $L=f_{\min}$ and $g=k$ at the minimum point, this seems to prove the key formula $df_{\min}/dk=\lambda$. But $x_{\min}$, $y_{\min}$, $\lambda$, and $f_{\min}$ all change with $k$. We need the *total* derivative of $L(x, y, \lambda, k)$:

$$\frac{dL}{dk}=\frac{\partial L}{\partial x}\frac{dx}{dk}+\frac{\partial L}{\partial y}\frac{dy}{dk}+\frac{\partial L}{\partial \lambda}\frac{d\lambda}{dk}+\frac{\partial L}{\partial k}\frac{dk}{dk}.$$

Equation (1) at the minimum point should now yield the sensitivity formula $df_{\min}/dk=\lambda$.

**16** (Theory behind $\lambda$) When $g(x, y)=k$ is solved for $y$, it gives a curve $y=R(x)$. Then minimizing $f(x, y)$ along this curve yields

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\frac{dR}{dx}=0, \quad \frac{\partial g}{\partial x}+\frac{\partial g}{\partial y}\frac{dR}{dx}=0.$$

Those come from the ________ rule: $df/dx=0$ at the minimum and $dg/dx=0$ along the curve because $g=$ ________. Multiplying the second equation by $\lambda=(\partial f/\partial y)/(\partial g/\partial y)$ and subtracting from the first gives ________ $=0$. Also $\partial f/\partial y=\lambda \partial g/\partial y$. These are the equations (1) for $x, y, \lambda$.

**17** (*Example of failure*) $\lambda=f_y/g_y$ breaks down if $g_y=0$ at the minimum point.

(a) $g=x^2-y^3=0$ does not allow negative $y$ because ________.

(b) When $g=0$ the minimum of $f=x^2+y$ is at the point ________.

(c) At that point $f_y=\lambda g_y$ becomes ________ which is impossible.

(d) Draw the pointed curve $g=0$ to see why it is not tangent to a level curve of $f$.

**18** (*No maximum*) Find a point on the line $g=x+y=1$ where $f(x, y)=2x+y$ is greater than 100 (or 1000). Write out grad $f=\lambda$ grad $g$ to see that there is no solution.

**19** Find the minimum of $f=x^2+2y^2+z^2$ if $(x, y, z)$ is restricted to the planes $g=x+y+z=0$ and $h=x-z=1$.

**20** (a) Find by Lagrange multipliers the volume $V=xyz$ of the largest box with sides adding up to $x+y+z=k$. (b) Check that $\lambda=dV_{\max}/dk$. (c) United Airlines accepts baggage with $x+y+z=108''$. If it changes to $111''$, approximately how much (by $\lambda$) and exactly how much does $V_{\max}$ increase?

**21** The planes $x=0$ and $y=0$ intersect in the line $x=y=0$, which is the $z$ axis. Write down a vector perpendicular to the plane $x=0$ and a vector perpendicular to the plane $y=0$. Find $\lambda_1$ times the first vector plus $\lambda_2$ times the second. This combination is perpendicular to the line ______.

**22** Minimize $f=x^2+y^2+z^2$ on the plane $ax+by+cz=d$—one constraint and one multiplier. Compare $f_{\min}$ with the distance formula $|d|/\sqrt{a^2+b^2+c^2}$ in Section 11.2.

**23** At the absolute minimum of $f(x,y)$, the derivatives ______ are zero. If this point happens to fall on the curve $g(x,y)=k$ then the equations $f_x=\lambda g_x$ and $f_y=\lambda g_y$ hold with $\lambda=$ ______.

**Problems 24–33 allow inequality constraints, optional but good.**

**24** Find the minimum of $f=3x+5y$ with the constraints $g=x+2y=4$ and $h=x\geqslant 0$ and $H=y\geqslant 0$, using equations like (7). Which multiplier is zero?

**25** Figure 13.23 shows the constraint plane $g=x+y+z=1$ chopped off by the inequalities $x\geqslant 0$, $y\geqslant 0$, $z\geqslant 0$. What are the three "endpoints" of this triangle? Find the minimum and maximum of $f=4x-2y+5z$ on the triangle, by testing $f$ at the endpoints.

**26** With an inequality constraint $g\leqslant k$, the multiplier at the minimum satisfies $\lambda\leqslant 0$. If $k$ is increased, $f_{\min}$ goes down (since $\lambda=df_{\min}/dk$). *Explain the reasoning*: By increasing $k$, (more) (fewer) points satisfy the constraints. Therefore (more) (fewer) points are available to minimize $f$. Therefore $f_{\min}$ goes (up) (down).

**27** With an inequality constraint $g\leqslant k$, the multiplier at a *maximum* point satisfies $\lambda\geqslant 0$. Change the reasoning in **26**.

**28** When the constraint $h\geqslant k$ is a strict inequality $h>k$ at the minimum, the multiplier is $\lambda=0$. *Explain the reasoning*: For a small increase in $k$, the same minimizer is still available (since $h>k$ leaves room to move). Therefore $f_{\min}$ is (changed)(unchanged), and $\lambda=df_{\min}/dk$ is ______.

**29** Minimize $f=x^2+y^2$ subject to the inequality constraint $x+y\leqslant 4$. The minimum is obviously at ______, where $f_x$ and $f_y$ are zero. The multiplier is $\lambda=$ ______. A small change from 4 will leave $f_{\min}=$ ______ so the sensitivity $df_{\min}/dk$ still equals $\lambda$.

**30** Minimize $f=x^2+y^2$ subject to the inequality constraint $x+y\geqslant 4$. Now the minimum is at ______ and the multiplier is $\lambda=$ ______ and $f_{\min}=$ ______. A small change to $4+dk$ changes $f_{\min}$ by what multiple of $dk$?

**31** Minimize $f=5x+6y$ with $g=x+y=4$ and $h=x\geqslant 0$ and $H=y\leqslant 0$. Now $\lambda_3\leqslant 0$ and the sign change destroys Example 4. Show that equation (7) has no solution, and choose $x,y$ to make $5x+6y<-1000$.

**32** Minimize $f=2x+3y+4z$ subject to $g=x+y+z=1$ and $x,y,z\geqslant 0$. These constraints have multipliers $\lambda_2\geqslant 0$, $\lambda_3\geqslant 0$, $\lambda_4\geqslant 0$. The equations are $2=\lambda_1+\lambda_2$, ______, and $4=\lambda_1+\lambda_4$. Explain why $\lambda_3>0$ and $\lambda_4>0$ and $f_{\min}=2$.

**33** A wire 40″ long is used to enclose *one or two* squares (side $x$ and side $y$). Maximize the total area $x^2+y^2$ subject to $x\geqslant 0$, $y\geqslant 0$, $4x+4y=40$.