Chapter 14

Vectors

Vectors are sometimes represented as arrows.

Section 14.1 introduces vectors and their arithmetic. Section 14.2 examines the dot product, which is a number. This includes the geometry of the dot product and its role in projections. (A projection is related to the shadow cast by parallel rays of light.)

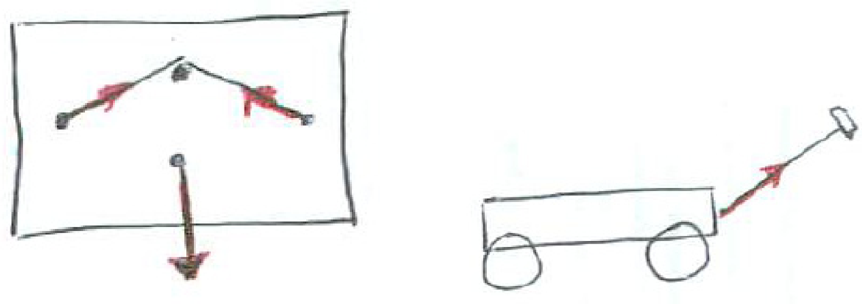
Section 14.3 examines the cross product, which is a vector. Determinants are reviewed, and the scalar triple product (a number) is introduced and used to find the volume of a parallelepiped.

Section 14.4 develops a number of fundamental properties of lines and planes, in terms of vectors. The distance from a point to a line or plane is developed, a parametric description of a line is given, using the dot and cross product. These ideas are used to talk about flows.

This algebra was developed primarily in response to James Clerk Maxwell's *Treatise on Electricity and Magnetism*, published in 1873. Josiah Gibbs, who in 1863 earned the first doctorate in engineering awarded in the United States and became a mathematical physicist, put vector analysis in its present form. His *Elements of Vector Analysis*, published in 1881, introduced the notation used in this chapter. Maxwell's contributions will be studied in greater detail in Chapter 18.

14.1 The Algebra of Vectors

You have lived with vectors all your life. When you hanged a picture on wire you dealt with three vectors: one describes the downward force of gravity and two describe the force of the wires pulling up to oppose gravity, as in Figure 14.1.1(a)



(a) (b)

Figure 14.1.1:

When you pull a wagon the force you use is represented by a vector, as in Figure 14.1.1(b). The harder you pull, the larger the vector.

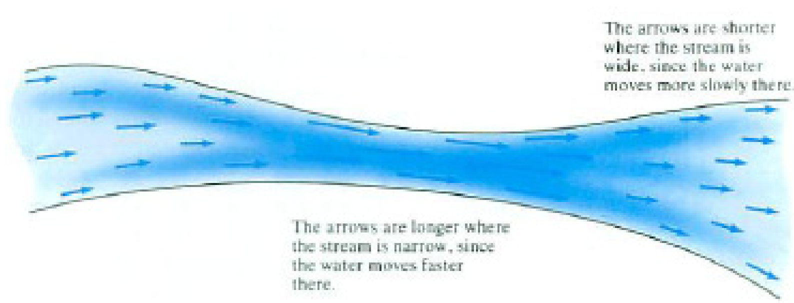


Figure 14.1.2:

A vector has a direction and a magnitude. You may think of it as an arrow, whose length and direction carry information. Vectors are of use in describing the flow of a fluid, as in Figure 14.1.2, or the wind, or the strength and direction of a magnetic field.

Vectors in the Plane

A vector in the xy plane is an ordered pair of numbers x and y, denoted $\langle x, y\rangle$. Its magnitude, or length, is $\sqrt{x^2 + y^2}$. Though the notation resembles that for a point, (x, y), we treat vectors quite differently. We can add them, subtract them and multiply them by a number. Two additional products of vectors are introduced in Sections 14.2 and 14.3.

We represent a vector by an arrow whose tail is at $(0, 0)$ and whose head (or "tip") is at (x, y), as in Figure 14.1.3.

More generally, we represent $\langle x, y\rangle$ by any pair of points $P = (a_1, a_2)$ and $Q = (b_1, b_2)$ if $b_1 - a_1 = x$ and $b_2 - a_2 = y$, as in Figure 14.1.4.

We speak then of "the vector from P to Q" and denote it $\overrightarrow{PQ}$. A vector $\langle x, y\rangle$ will be denoted by bold face letters, such as **A**, **B**, **r**, **v**, and **a**. In handwriting or on the blackboard they are decorated with a bar or arrow on top, for instance $\overrightarrow{A}$ or $\overline{A}$. A vector of length 1 is called a unit vector and is topped with a little hat, as in $\widehat{r}$, which is read aloud as "r hat".

Here is how we operate on vectors. Let $\mathbf{A} = \langle a_1, a_2\rangle$ and $\mathbf{B} = \langle b_1, b_2\rangle$ be vectors and let c be a number.

Figure 14.1.3: The arrow represents the vector $\langle x, y\rangle$.

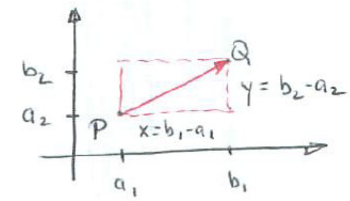
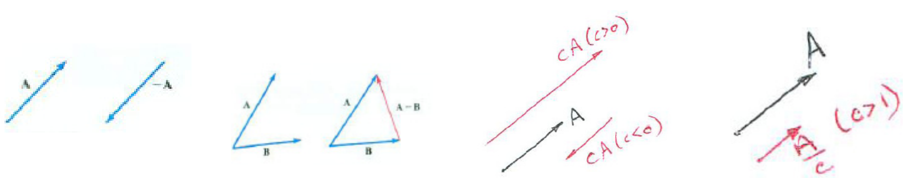


Figure 14.1.4:

(a) (b) (c) (d)

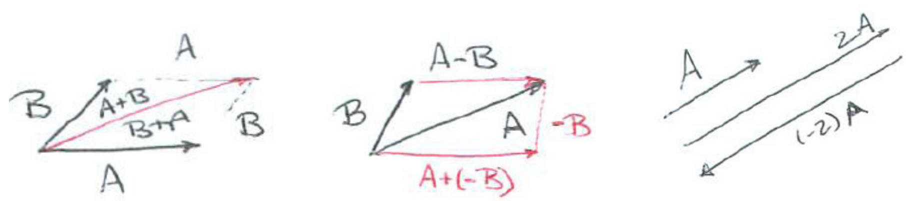
Figure 14.1.5:



(a) (b) (c) (d)

Figure 14.1.6:

Operation	Definition	Geometry	Comment
$\mathbf{A}+\mathbf{B}$	$\langle a_1+b_1, a_2+b_2\rangle$	Figure 14.1.5	The tail of **B** is placed at the head of **A**
$-\mathbf{A}$	$\langle -a_1, -a_2\rangle$	Figure 14.1.6(a)	$-\mathbf{A}$ points in opposite direction of **A**
$\mathbf{A}-\mathbf{B}$	$\langle a_1-b_1, a_2-b_2\rangle$	Figure 14.1.6(b)	What you add to **B** to get **A**
$c\mathbf{A}$	$\langle ca_1, ca_2\rangle$	Figure 14.1.6(c)	Parallel to **A** and $\lvert c\rvert$ times as long as **A**
$\frac{\mathbf{A}}{c}$	$\left\langle \frac{a_1}{c}, \frac{a_2}{c}\right\rangle$	Figure 14.1.6(d)	Parallel to **A** and $\frac{1}{c}$ times as long as **A** ($c \neq 0$)



(a) (b) (c)

Figure 14.1.7:

The operation of addition obeys the usual rules of addition of numbers. For instance, $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$ and $\mathbf{A}+(\mathbf{B}+\mathbf{C}) = (\mathbf{A}+\mathbf{B})+\mathbf{C}$. Also $\mathbf{A}-\mathbf{B} = \mathbf{A}+-\mathbf{B}$. This is easy to establish using the definitions. In terms of arrows it makes sense; see Figure 14.1.7(a).

$\mathbf{A}-\mathbf{B}$ and $\mathbf{A}+(-\mathbf{B})$ appears as opposite sides of a parallelogram. Figure 14.1.7(a) shows both $\mathbf{A}+\mathbf{B}$ and $\mathbf{B}+\mathbf{A}$; they are equal.

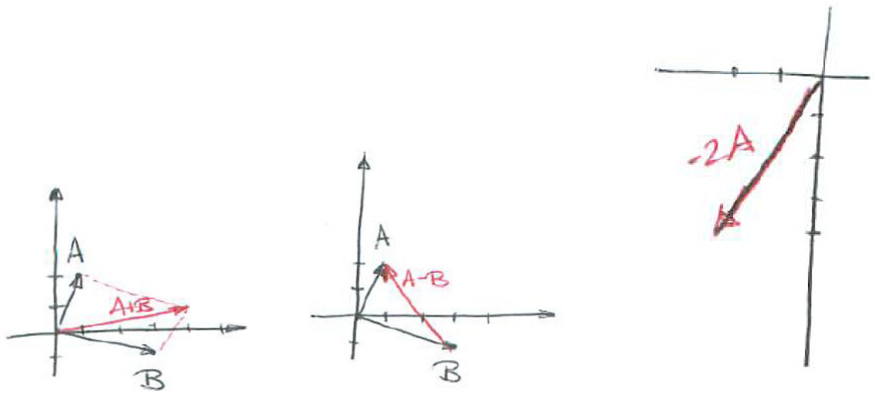
The magnitude of $\langle x, y\rangle$ is $\sqrt{(cx)^2+(cy)^2} = \sqrt{c^2}\sqrt{x^2+y^2}$, that is, $|c|$ times the magnitude of $\langle x, y\rangle$. If c is positive $\langle cx, cy\rangle$ and $\langle x, y\rangle$ point in the same direction. If c is negative they point in opposite direction, as the arrows in Figure 14.1.7(c) illustrate for $c = 2$ or -2.

When talking about numbers, such as c, x, and y, in the context of vectors, we call them **scalars**. Thus in $c\overrightarrow{A}$ the scalar c is multiplying the vector **A**.

The vector $\langle 0, 0\rangle$ is denoted **0** and is called the **zero vector**.

EXAMPLE 1 Let $\mathbf{A} = \langle 1, 2\rangle$, $\mathbf{B} = \langle 3, -1\rangle$ and $c = -2$. Complete $\mathbf{A}+\mathbf{B}$, $\mathbf{A}-\mathbf{B}$ and $c\mathbf{A}$. Then draw the corresponding arrows.
SOLUTION

$$\begin{aligned} A+B &= \langle 1,2\rangle + \langle 3,-1\rangle = \langle 1+3, 2+(-1)\rangle = \langle 4,1\rangle \\ A-B &= \langle 1,2\rangle - \langle 3,-1\rangle = \langle 1-3, 2-(-1)\rangle = \langle -2,3\rangle \\ cA &= -2\langle 1,2\rangle = \langle -2,(1), -2(2)\rangle = \langle -2,-4\rangle \end{aligned}$$



(a) (b) (c)

Figure 14.1.8:

Note that $\mathbf{A} - \mathbf{B}$ and $\mathbf{A} + \mathbf{B}$ lie on the two diagonals of a parallelogram. (See Figure 14.1.8.) ◇

Before we can make the similar definition for vectors in space, we must introduce an appropriate coordinate system.

Coordinates in Space

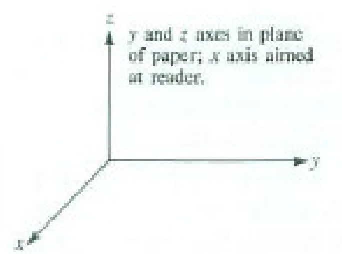


Figure 14.1.9: ARTIST: A "right hand" should be added to this figure.

First, pick a pair of perpendicular intersecting lines to serve as the x and y axes. The positive parts of these axes are indicated by arrows. These two lines determine the xy **plane**. The line perpendicular to the xy plane and meeting the x and y axes will be called the z-axis. The point where the three axes meet is called the **origin**. The 0 of the z-axis will be put at the origin. But which half of the z-axis will have positive numbers and which half will have the negative numbers? It is customary to determine this by the **right-hand rule**. Moving in the xy plane through a right angle from the positive x-axis to the positive y-axis determines a sense of rotation around the z-axis. If the fingers of the right hand curl in that sense, the thumb points in the direction of the *positive* z-axis, as shown in Figure 14.1.9.

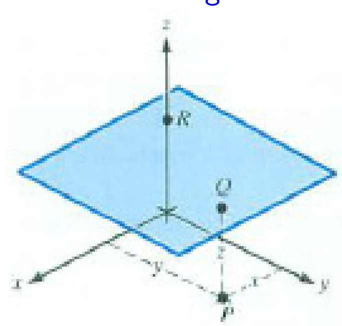
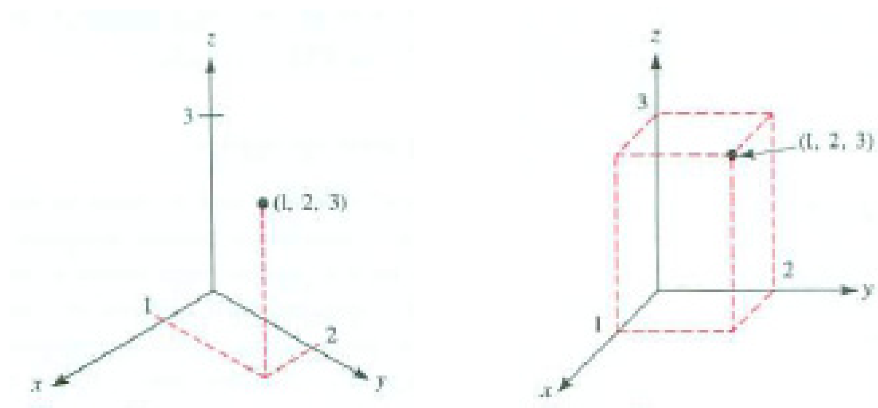


Figure 14.1.10:

Any point Q in space is now described by three numbers: First, two numbers specify the x and y coordinates of the point P in the xy plane directly below (or above) Q; then the height of Q above (or below) the xy plane is recorded by the z coordinate of the point R where the plane through Q and parallel to the xy plane meets the z-axis. The point Q is then denoted (x, y, z). See Figure 14.1.10.

The points (x, y, z) for which $z = 0$ lie in the xy plane. There are an infinite number of these points. The points (x, y, z) for which $x = 0$ lie entirely in the plane determined by the y and x axes, which is called the yz **plane**. Similarly, the equation $y-0$ describe the xz **plane**. The xy, xz and yz planes are called the **coordinate planes**.

EXAMPLE 2 Plot the point $(1,2,3)$.



(a) (b)

Figure 14.1.11:

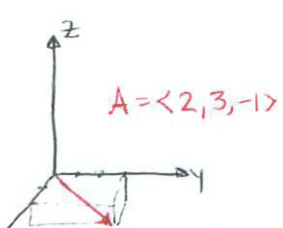
SOLUTION One way is to first plot the point $(1,2)$ in the xy plane. Then, on a line perpendicular to the xy plane at that point, show the point $(1,2,3)$ as done in Figure 14.1.11(a).

Another way is to draw a box whose edges are parallel to the axes and which has the origin $(0,0,0)$ and $(1,2,3)$ as done in Figure 14.1.11(b). (This time, the y and z axes make a right angle.) ◇

Just as the axes in the xy plane divide the plane with four quadrants, the three coordinate planes divide space with eight octants.

Vectors in Space

The only difference between a vector in space and a vector in the xy plane is that it has three components, x, y, and z, and is written $\langle x,y,z\rangle$. Its length or magnitude is defined as $\sqrt{x^2+y^2+z^2}$. The definition of the sum and difference of such vectors is so similar to the definition for planar vectors that we will not list them. For instance, $\langle a_1,a_2,a_3\rangle+\langle b_1,b_2,b_3\rangle$ is $\langle a_1+b_1,a_2+b_2,a_3+b_3\rangle$. The biggest difference is that they are harder to draw, even though each can be suggested by our arrow. It may help visualize such a vector by drawing a box in which it is a main diagonal. For instance, to draw the vector $\langle 2,3,-1\rangle$ you may draw the box shown in Figure 14.1.12

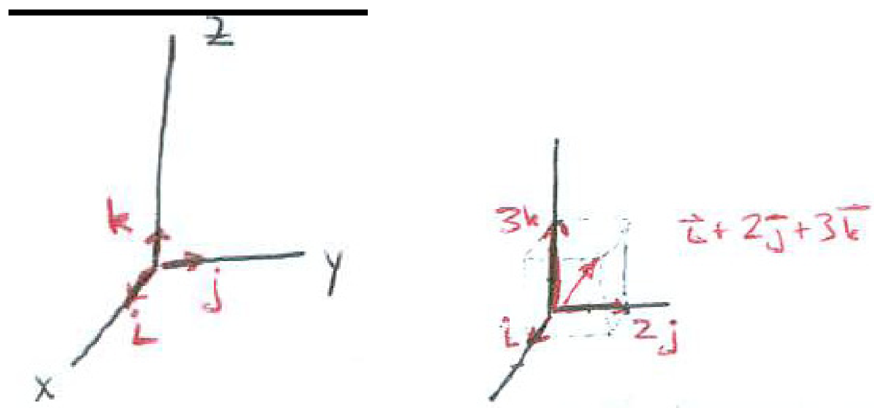


This representation of **A** has its tail at the arrow. Of course the arrow and box could be drawn with the tail of the arrow anywhere else.

The Standard Unit Vectors

The three most important unit vectors indicate the positive directions of the positive x, y, and z axes. They will be denoted **i**, **j** and **k**, respectively. For instance, $\mathbf{i} = \langle 1, 0, 0\rangle$. The vectors $\langle x, y, z\rangle$ can also be written $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

EXAMPLE 3 Draw **i**, **j**, **k** and $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.
SOLUTION Figure 14.1.13(a) shows **i**, **j**, **k** and Figure 14.1.13(b) shows



(a) (b)

Figure 14.1.13:

$\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$. ◇

The magnitude of **A** is indicated by $\|\mathbf{A}\|$. $\|\mathbf{A}\|$ is a scalar and $\mathbf{A}/\|\mathbf{A}\|$ is a vector.

The vector $\frac{\mathbf{A}}{\|\mathbf{A}\|}$ is a unit vector for any non-zero vector **A**. To see this, we let $\mathbf{A} = \langle x, y, z\rangle$ and compute $\mathbf{A}/\|\mathbf{A}\|$:

$$\frac{\mathbf{A}}{\|\mathbf{A}\|} = \frac{\langle x, y, z\rangle}{\sqrt{x^2 + y^2 + z^2}} = \langle \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}}\rangle.$$

The square of the length of $\mathbf{A}/\|\mathbf{A}\|$ is

$$\left(\frac{x}{\sqrt{x^2 + y^2 + z^2}}\right)^2 + \left(\frac{y}{\sqrt{x^2 + y^2 + z^2}}\right)^2 + \left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)^2 = \frac{x^2 + y^2 + z^2}{x^2 + y^2 + z^2} = 1.$$

Thus $\mathbf{A}/\|\mathbf{A}\|$ is a unit vector.

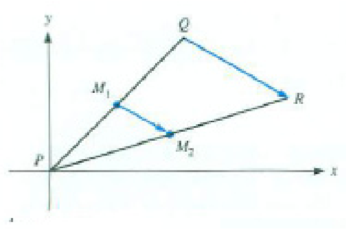


Figure 14.1.15:
Figure 14.1.14:

Example 4 shows how vectors can be used to establish geometric properties.

EXAMPLE 4 Prove that the line which joins the midpoints of two sides of a triangle is parallel to the third side and half as long.
SOLUTION Let the triangle have vertices P, Q, and R. Let the midpoint of side PQ be M and the midpoint of side PR be N as in Figure 14.1.14.

Introduce an xy coordinate system in the plane of the triangle. Through its origin could be anywhere in the plane, we should put it at P in order to simplify the calculations. (See Figure 14.1.15.)

We wish to show that the vector $\overrightarrow{MN}$ is $\frac{1}{2}\overrightarrow{QR}$. To do so, we compute $\overrightarrow{MN}$ and $\overrightarrow{QR}$ in terms of vectors involving P, Q, and R.

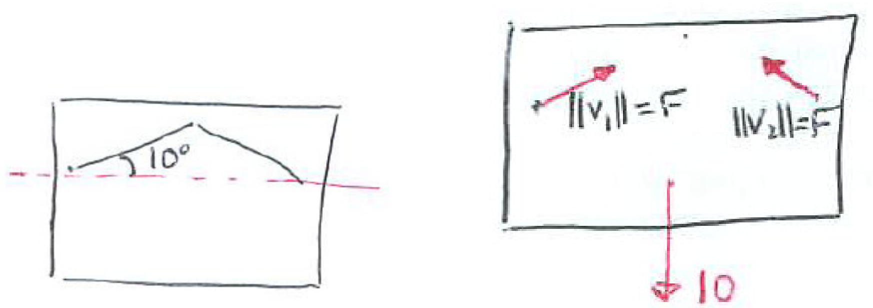
First of all, $\overrightarrow{PM} = \frac{1}{2}\overrightarrow{PQ}$ and $\overrightarrow{PN} = \frac{1}{2}\overrightarrow{PR}$. Thus

$$\overrightarrow{MN} = \frac{1}{2}\overrightarrow{PR} - \frac{1}{2}\overrightarrow{PQ} = \frac{1}{2}(\overrightarrow{PR} - \overrightarrow{PQ}) = \frac{1}{2}(\overrightarrow{QR}).$$

◇

The next example shows the importance of thinking vectorally. Not thinking that way, one of the other had a picture fall and break a vase.

EXAMPLE 5 A picture weighing 10 pounds has a wire on the back, which rests on a picture hook, as shown in Figure 14.1.16(a). Find the force (tension) on the wire.
SOLUTION There are three vectors involved. One is straight down, with magnitude 10 lbs. and two are along the wire, with unknown magnitude F: $\|\mathbf{v}_1\| = F = \|\mathbf{v}_2\|$.



(a) (b)

Figure 14.1.16:

To balance the downward force of gravity, each end of the wire must have a vertical component of 5 lbs. Since the angle with the horizontal is 10° we must have $F\sin(10°) = 5$ or $F = 5/\sin(10°) \approx 29$ pounds. That is much greater than the weight of the painting and creates quite a pull on the screws at the bases of the wire. This force can (sadly, we learned) eventually pull a screw out of the wall. $\diamond$

Summary

We introduced the notion of a vector $\langle x, y\rangle$ in the xy plane or $\langle x, y, z\rangle$ in space and defined their addition and subtraction. Furthermore we defined the operation of a scalar c as a vector $\langle x, y, z\rangle$, as $\langle cx, cy, cz\rangle$.

We visualized vectors with the aid of arrows, which could be drawn anywhere in the xy plane or in space.

Each vector in the xy-plane can be written as $x\mathbf{i} + y\mathbf{j}$. Vector in space can be written as $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

EXERCISES for Section 14.1 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 and 2 use the plane of your paper as the xy plane.

1.[R] Draw the vector $2\mathbf{i}+3\mathbf{j}$, placing its tail at (a) $(0,0)$, (b) $(-1,2)$, (c) $(1,1)$.

2.[R] Draw the vector $-\mathbf{i}+2\mathbf{j}$, placing its tail at (a) $(0,0)$, (b) $(3,0)$, (c) $(-2,2)$.

In Exercises 3 to 6 draw the vector $\mathbf{A}$ and enough extra lines to show how it is situated in space.

3.[R] $\mathbf{A}=2\mathbf{i}+\mathbf{j}+3\mathbf{k}$,

(a) tail at $(0,0,0)$,

(b) tail at $(1,1,1)$.

4.[R] $\mathbf{A}=\mathbf{i}+\mathbf{j}+\mathbf{k}$,

(a) tail at $(0,0,0)$,

(b) tail at $(2,3,4)$.

5.[R] $\mathbf{A}=-\mathbf{i}-2\mathbf{j}+2\mathbf{k}$,

(a) tail at $(0,0,0)$,

(b) tail at $(1,1,-1)$.

6.[R] $\mathbf{A}=\mathbf{j}+\mathbf{k}$,

(a) tail at $(0,0,0)$,

(b) tail at $(-1,-1,-1)$.

In Exercises 7 to 10 plot the points P and Q, draw the vector $\overrightarrow{PQ}$, express it in the form $x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, and find its length.

7.[R] $P=(0,0,0), Q=(1,3,4)$

8.[R] $P=(1,2,3), Q=(2,5,4)$

9.[R] $P=(2,5,4), Q=(1,2,2)$

10.[R] $P = (1, 1, 1), Q = (-1, 3, -2)$

In Exercises 11 and 12 express the vector $\mathbf{A}$ in the form $x\mathbf{i} + y\mathbf{j}$. North is along the positive y-axis and east is along the positive x-axis.

11.[R]

(a) $\|\mathbf{A}\| = 10$ and $\mathbf{A}$ points northwest;

(b) $\|\mathbf{A}\| = 6$ and $\mathbf{A}$ points south;

(c) $\|\mathbf{A}\| = 9$ and $\mathbf{A}$ points southeast;

(d) $\|\mathbf{A}\| = 5$ and $\mathbf{A}$ points east.

12.[R]

(a) $\|\mathbf{A}\| = 1$ and $\mathbf{A}$ points southwest;

(b) $\|\mathbf{A}\| = 2$ and $\mathbf{A}$ points west;

(c) $\|\mathbf{A}\| = \sqrt{8}$ and $\mathbf{A}$ points northeast;

(d) $\|\mathbf{A}\| = 1/2$ and $\mathbf{A}$ points south.

13.[M] The wind is 30 miles per hour to the northeast. An airplane is traveling 100 miles per hour relative to the air, and the vector from the tail of the plane to its front tip points to the southeast. (See Figure 14.1.17.)

(a) What is the speed of the plane relative to the ground?

(b) What is the direction of the flight relative to the ground?

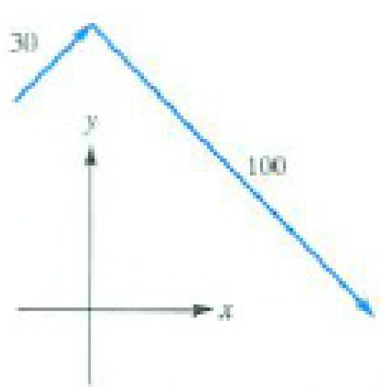


Figure 37

Figure 14.1.17:

14.[M] (See Exercise 13.) The jet stream is moving 200 miles per hour to the southeast. A plane with a speed of 550 miles per hour relative to the air is aimed to the northwest.

(a) Draw the vectors representing the wind and the plane relative to the air. (Choose a scale and make an accurate drawing.)

(b) Using your drawing, estimate the speed of the plane relative to the ground.

(c) Compute the speed in (b) exactly.

15.[R] Compute $\mathbf{A}+\mathbf{B}$ and $\mathbf{A}-\mathbf{B}$ if

(a) $\mathbf{A}=\langle -1,2,3\rangle$ and $\mathbf{B}=\langle 7,0,2\rangle$.

(b) $\mathbf{A}=3\mathbf{j}+4\mathbf{k}$ and $\mathbf{B}=6\mathbf{i}+7\mathbf{j}$.

16.[R] Compute $\mathbf{A}+\mathbf{B}$ and $\mathbf{A}-\mathbf{B}$ if

(a) $\mathbf{A}=\langle 1/2,1/3,1/6\rangle$ and $\mathbf{B}=\langle 2,3,-1/3\rangle$.

(b) $\mathbf{A}=2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$ and $\mathbf{B}=-\mathbf{i}+5\mathbf{j}+6\mathbf{k}$.

17.[R] Compute and sketch $c\mathbf{A}$ if $\mathbf{A}-2\mathbf{i}+3\mathbf{j}+\mathbf{k}$ and c is

(a) 2,

(b) -2,

(c) $\frac{1}{2}$,

(d) $-\frac{1}{2}$.

18.[R] Express each of the following vectors in the form $c(2\mathbf{i}+3\mathbf{j}+4\mathbf{k})$ for suitable c:

(a) $\langle 4,6,8\rangle$

(b) $-2\mathbf{i}-3\mathbf{j}-4\mathbf{k}$

(c) $\mathbf{0}$

(d) $\frac{2}{11}\mathbf{i} + \frac{3}{11}\mathbf{j} + \frac{4}{11}\mathbf{k}$

19.[R] If $\|\mathbf{A}\| = 6$, find the length of the following vectors

(a) $-2\mathbf{A}$

(b) $\mathbf{A}/3$

(c) $\mathbf{A}/\|\mathbf{A}\|$

(d) $-\mathbf{A}$

(e) $\mathbf{A} + 2\mathbf{A}$.

20.[R] If $\|\mathbf{A}\| = 3$, find the length of the following vecrps

(a) $-4\mathbf{A}$

(b) $13\mathbf{A} - 7\mathbf{A}$

(c) $\mathbf{A}/\|\mathbf{A}\|$

(d) $\mathbf{A}/0.05$

(e) $\mathbf{A} - \mathbf{A}$.

21.[R]

(a) Find a unit vector $\mathbf{u}$ that has the same direction as $\mathbf{A} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.

(b) Draw $\mathbf{A}$ and $\mathbf{u}$, with their tails at the origin.

22.[R]

(a) Find a unit vector $\mathbf{u}$ that has the same direction as $\mathbf{A} = 2\mathbf{i} - 2\mathbf{j} + 3\mathbf{k}$.

(b) Draw $\mathbf{A}$ and $\mathbf{u}$, with their tails at the origin.

23.[R] Using the definition of addition of vectors $\mathbf{A} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3 \rangle$, show the $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$ and $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$.

24.[R] Using the definition of addition of vectors show that $\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}$.

25.[R] Which unit vector points in the same direction as $2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$?

26.[R] Sketch a unit vector pointing in the same direction as $3\mathbf{i} + 4\mathbf{j}$.

27.[M] *(Midpoint formula)* Let A and B be two points in space. Let M be their midpoint. Let $\mathbf{A} = \overrightarrow{OA}$, $\mathbf{B} = \overrightarrow{OB}$, and $\mathbf{M} = \overrightarrow{OM}$.

(a) Show that $\mathbf{M} = \mathbf{A} + \frac{1}{2}(\mathbf{B} - \mathbf{A})$.

(b) Deduce that $\mathbf{M} = (\mathbf{A} + \mathbf{B})/2$. *Hint:* Draw a picture.

28.[M] Let A and B be two distinct points in space. Let C be the point on the line segment AB that is twice as far from A as it is from B. Let $\mathbf{A} = \overrightarrow{OA}$, $\mathbf{B} = \overrightarrow{OB}$, and $\mathbf{C} = \overrightarrow{OC}$. Show that $\mathbf{C} = \frac{1}{3}\mathbf{A} + \frac{2}{3}\mathbf{B}$. *Hint:* Draw a picture.

29.[M] Show that $2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$ and $6\mathbf{i} + 9\mathbf{j} + 12\mathbf{k}$ are parallel.

30.[M] Show that $\mathbf{i} - 3\mathbf{j} + 6\mathbf{k}$ and $-2\mathbf{i} + 6\mathbf{j} - 12\mathbf{k}$ are parallel.

31.[M] This exercise outlines a proof of the distributive rule: $c(\mathbf{A}+\mathbf{B}) = c\mathbf{A}+c\mathbf{B}$. Write $\mathbf{A}$ and $\mathbf{B}$ in components, and obtain the rule by expressing both $c(\mathbf{A} + \mathbf{B})$ and $c\mathbf{A} + c\mathbf{B}$ in components.

32.[M]

(a) Show that the vectors $\mathbf{u}_1 = \frac{1}{2}\mathbf{i} + (\sqrt{3}/2\mathbf{j}$ and $\mathbf{u}_2 = (\sqrt{3}/2\mathbf{i} - \frac{1}{2}\mathbf{j}$ are perpendicular unit vectors. *Hint:* What angles do they make with the x-axis?

(b) Find scalars x and y such that $\mathbf{i} = x\mathbf{u}_1 + y\mathbf{u}_2$.

33.[M]

(a) Show that the vectors $\mathbf{u}_1 = (\sqrt{2}/2\mathbf{i}) + (\sqrt{2}/2\mathbf{j})$ and $\mathbf{u}_2 = (-\sqrt{2}/2\mathbf{i}) + (\sqrt{2}/2\mathbf{j})$ are perpendicular unit vectors. *Hint:* Draw them.

(b) Express $\mathbf{i}$ in the form of $x\mathbf{u}_1 + y\mathbf{u}_2$. *Hint:* Draw $\mathbf{i}$,$\mathbf{u}_1$, and $\mathbf{u}_2$.

(c) Express $\mathbf{j}$ in the form $x\mathbf{u}_1 + y\mathbf{u}_2$.

(d) Express $-2\mathbf{i} + 3\mathbf{j}$ in the form $x\mathbf{u}_1 + y\mathbf{u}_2$.

34.[M]

(a) Draw a unit vector **u** tangent to the curve $y = \sin x$ at $(0, 0)$.

(b) Express **u** in the form $x\mathbf{i} + y\mathbf{j}$.

35.[M]

(a) Draw a unit vector **u** tangent to the curve $y = x^3$ at $(1, 1)$.

(b) Express **u** in the form $x\mathbf{i} + y\mathbf{j}$.

36.[M]

(a) What is the sum of the five vectors shown in Figure 14.1.18?

(b) Sketch the figure corresponding to the sum $\mathbf{A} + \mathbf{C} + \mathbf{D} + \mathbf{E} + \mathbf{B}$.

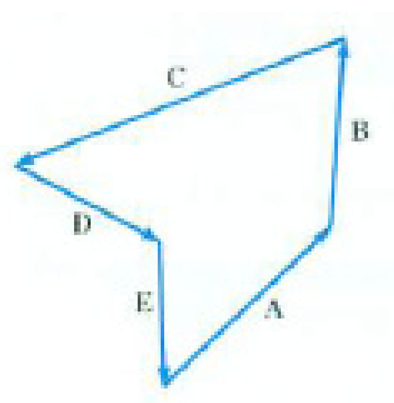


Figure 14.1.18:

37.[M] A rectangular box has sides of length x, y, and z. Show that the length of a longest diagonal (arc joining opposite corner) is $\sqrt{x^2 + y^2 + z^2}$. HINT: Use the Pythogorean Theorem, twice.

38.[M] See Example 5 concerning hanging a picture. What would be the tension in the wire if it were at an angle of

(a) $60°$ instead of $10°$ to the horizontal,

(b) $5°$ instead of $10°$ to the horizontal?

39.[C]

(a) Draw the vectors $\mathbf{A} = 2\mathbf{i} + \mathbf{j}$, $\mathbf{B} = 4\mathbf{i} - \mathbf{j}$, and $\mathbf{C} = 5\mathbf{i} + 2\mathbf{j}$.

(b) With the aid of the drawing show that there are scalars x and y such that $\mathbf{C} = x\mathbf{A} + y\mathbf{B}$.

(c) Using the drawing in (a), estimate x and y.

(d) Find x and y exactly.

40.[C] (See Exercise 13.) Let $\mathbf{A}$ and $\mathbf{B}$ be two nonzero and nonparallel vectors in the xy plane. Let $\mathbf{C}$ be any vector in the xy plane. Show with the aid of a sketch that there are scalars x and y such that $\mathbf{C} = x\mathbf{A} + y\mathbf{B}$.

41.[C] Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be three vectors that do not all lie in one plane. Let $\mathbf{D}$ be any vector in space. Show with the aid of a sketch that there are scalars x, y, and z such that $\mathbf{D} = x\mathbf{A} + y\mathbf{B} + z\mathbf{C}$.

42.[C] Let A, B and C be the vertices of a triangle. Let $\mathbf{A} = \overrightarrow{OA}$, $\mathbf{B} = \overrightarrow{OB}$, and $\mathbf{C} = \overrightarrow{OC}$.

(a) Let P be the point that is on the line segment joining A to the midpoint of the edge BC and twice as far from A as from the midpoint. Show that $\overrightarrow{OP} = (\mathbf{A} + \mathbf{B} + \mathbf{C})/3$.

(b) Use (a) to show that the three medians of a triangle are concurrent.

43.[C] The midpoints of a quadrilateral in space are joined to form another quadrilateral. Prove that this second quadrilateral is a parallelogram.

44.[C]

(a) Using an appropriate diagram, explain why $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$. (This is called the **triangle inequality**.

(b) For which pairs of vectors $\mathbf{A}$ and $\mathbf{B}$ is $\|\mathbf{A} + \mathbf{B}\| = \|\mathbf{A}\| + \|\mathbf{B}\|$?

45.[C] From Exercise 44 deduce that for any four real numbers x_1, y_1, x_2, and y_2,

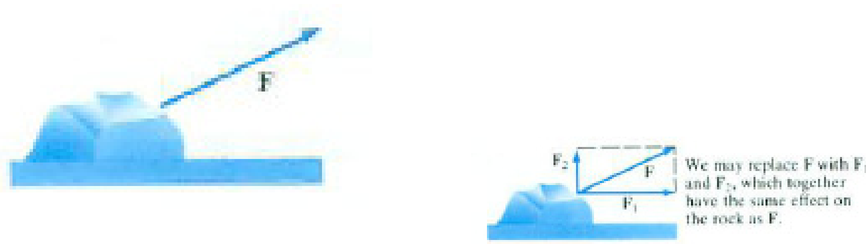
$$x_1x_2 + y_1y_2 \leq \sqrt{x_1^2 + y_1^2}\sqrt{x_2^2 + y_2^2}.$$

When does equality hold?

14.2 The Dot Product of Two Vectors

The dot product is a number, or scalar.

The "dot product" or "scalar product" is a number that is defined for every pair of vectors. Consider a rock being pulled along level ground by a



(a) (b)

Figure 14.2.1:

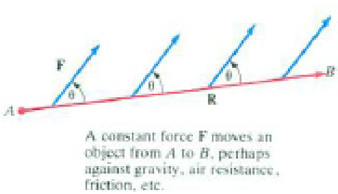


Figure 14.2.2:

rope inclined at at fixed angle to the ground. Let the force applied to the rock be represented by the vector $\mathbf{F}$. The force $\mathbf{F}$ can be expressed as the sum of a vertical force $\mathbf{F}_2$ and a horizontal force $\mathbf{F}_1$, as shown in Figure 14.2.1(b).

How much work is done by the force $\mathbf{F}$ in moving the rock along the ground? The physicist defines the work accomplished by a constant force $\mathbf{F}$ (whatever direction it may have) as the product of the component of $\mathbf{F}$ in the direction of motion and the distance traveled. Say that the force $\mathbf{F}$, as shown in Figure 14.2.2, moves an object along a straight line from the tail to the head of $\mathbf{R}$.

By definition

$$\text{Work} = \underbrace{\|\mathbf{F}\|\cos(\theta)}_{\text{Force in Direction of } \mathbf{R}} \cdot \underbrace{\|\mathbf{R}\|}_{\text{Distance traveled}}$$

where θ is the angle between $\mathbf{R}$ and $\mathbf{F}$.

The force $\mathbf{F}_2$ in Figure 14.2.1 accomplishes no work. The work accomplished by $\mathbf{F}$ in pulling the rock is the same as that accomplished by $\mathbf{F}_1$.

The Dot Product

This important physical concept illustrates the dot product of two vectors, which will be introduced after the following definition.

> **DEFINITION** *(Angle between two nonzero vectors.)* Let **A** and **B** be two nonparallel and nonzero vectors. They determine a triangle and an angle θ, shown in Figure 14.2.3. The **angle between A and B** is θ. Note that
>
> $$0 < \theta < \pi$$
>
> If **A** and **B** are parallel, the angle between them is 0 (if they have the same direction) or π (if they have opposite directions). The angle between 0 and any other vector is not defined.

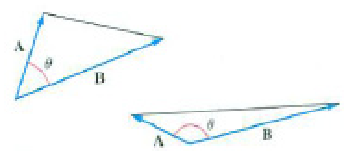


Figure 14.2.3:

The angle between **i** and **j** is $\pi/2$. The angle between $\mathbf{A} = -\mathbf{i} - \mathbf{j}$ and $\mathbf{B} = 3\mathbf{i}$ is $3\pi/4$, as Figure 14.2.4 shows. The angle between **k** and $-\mathbf{k}$ is π; the angle between $2\mathbf{i}$ and $5\mathbf{i}$ is 0.

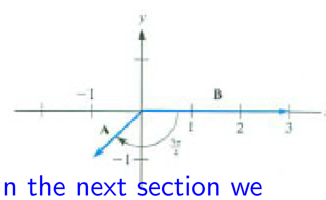


Figure 14.2.4:

In the next section we define another product of **A** and **B**, which will be a vector.

Sometimes we write $\cos(\mathbf{A}, \mathbf{B})$ instead of $\cos(\theta)$.

> **DEFINITION** *(Dot product)* Let **A** and **B** be two nonzero vectors. Their **dot product** is the number
>
> $$\|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta),$$
>
> where θ is the angle between **A** and **B**. If **A** or **B** is 0, their dot product is 0. The dot product is denoted $\mathbf{A} \cdot \mathbf{B}$. It is a scalar and is also called the **scalar product** of **A** and **B**.

The dot product satisfies several useful identities, which follow from the definition:

$$\begin{aligned}
\mathbf{A} \cdot \mathbf{B} &= \mathbf{B} \cdot \mathbf{A} && \text{(the dot product is commutative)} \\
\mathbf{A} \cdot \mathbf{A} &= \|\mathbf{A}\|^2 && \\
(c\mathbf{A}) \cdot \mathbf{B} &= c(\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \cdot (c\mathbf{B}) && (c \text{ is a scalar}) \\
\text{and} \quad \mathbf{0} \cdot \mathbf{A} &= 0. &&
\end{aligned}$$

For instance, to establish that $\mathbf{A} \cdot \mathbf{A} = \|\mathbf{A}\|^2$, we calculate $\mathbf{A} \cdot \mathbf{A}$:

$$\mathbf{A} \cdot \mathbf{A} = \|\mathbf{A}\|\|\mathbf{A}\|\cos(\theta) = \|\mathbf{A}\|^2,$$

since the angle θ between **A** and **A** is 0, and $\cos(0) = 1$.

EXAMPLE 1 Find the dot product $\mathbf{A} \cdot \mathbf{B}$ if $\mathbf{A} = 3\mathbf{i} + 3\mathbf{j}$ and $\mathbf{B} = -5\mathbf{i}$.
SOLUTION Inspection of Figure 14.2.5 shows that θ, the angle between **A** and **B**, is $3\pi/4$. Also,

$$\|\mathbf{A}\| = \sqrt{3^2 + 3^2} = \sqrt{18} \text{ and } \|\mathbf{B}\| = \sqrt{5^2 + 0^2} = 5.$$

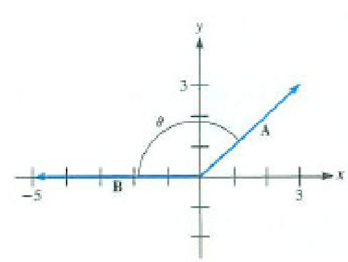


Figure 14.2.5:

Thus

$$\mathbf{A}\cdot\mathbf{B} = \|\mathbf{A}\|\|\mathbf{B}\|\cos\theta = \sqrt{18}\cdot\left(\frac{-\sqrt{2}}{2}\right) = -15.$$

◇

EXAMPLE 2 Find

1. $\mathbf{i}\cdot\mathbf{j}$,
2. $\mathbf{i}\cdot\mathbf{i}$,
3. $2\mathbf{k}\cdot(-3\mathbf{k})$.

Recall that $\mathbf{i}$ and $\mathbf{j}$ are perpendicular, be definition.

SOLUTION

1. The angle between $\mathbf{i}$ and $\mathbf{j}$ is $\pi/2$. Thus
$$\mathbf{i}\cdot\mathbf{j} \quad = \quad \|\mathbf{i}\|\|\mathbf{j}\|\cos\left(\frac{\pi}{2}\right) = 1\cdot 1\cdot 0 = 0.$$

This is a special case of the fact that $\mathbf{A}\cdot\mathbf{A} = \|\mathbf{A}\|^2$.

2. The angle between $\mathbf{i}$ and $\mathbf{i}$ is 0. Thus
$$\mathbf{i}\cdot\mathbf{i} \quad = \quad \|\mathbf{i}\|\|\mathbf{i}\|\cos(0) = 1\cdot 1\cdot 1 = 1.$$

3. The angle between $2\mathbf{k}$ and $-3\mathbf{k}$ is π. Thus
$$2\mathbf{k}\cdot(-3\mathbf{k}) \quad = \quad \|2\mathbf{k}\|\|-3\mathbf{k}\|\cos(\pi) = 2\cdot 3\cdot(-1) = -6.$$

◇ Computations like those in Example 2 show that $a\mathbf{i}\cdot b\mathbf{i} = ab$, $a\mathbf{j}\cdot b\mathbf{j} = ab$, and $a\mathbf{k}\cdot b\mathbf{k} = ab$, while $a\mathbf{i}\cdot b\mathbf{j} = 0$, $a\mathbf{i}\cdot b\mathbf{k} = 0$, and $a\mathbf{j}\cdot b\mathbf{k} = 0$.

In particular, $\mathbf{i}\cdot\mathbf{i} = \mathbf{j}\cdot\mathbf{j} = \mathbf{k}\cdot\mathbf{k} = 1$, while $\mathbf{i}\cdot\mathbf{j} = \mathbf{i}\cdot\mathbf{k} = \mathbf{j}\cdot\mathbf{k} = 0$.

The Geometry of the Dot Product

Let $\mathbf{A}$ and $\mathbf{B}$ be nonzero vectors and θ the angle between them. Their dot product is

$$\mathbf{A}\cdot\mathbf{B} = \|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta).$$

Observe that, by definition, the zero vector, $\mathbf{0}$, is perpendicular to every vector in the xy plane.

The quantities $\|\mathbf{A}\|$ and $\|\mathbf{B}\|$, being the lengths of vectors, are positive. However, $\cos(\theta)$ can be positive, zero, or negative. Note that $\cos(\theta) = 0$ only when $\theta = \pi/2$, that is when $\mathbf{A}$ and $\mathbf{B}$ are perpendicular. So the dot product provides a way of telling whether $\mathbf{A}$ and $\mathbf{B}$ are perpendicular:

A test for perpendicularity

Let $\mathbf{A}$ and $\mathbf{B}$ be nonzero vectors. If $\mathbf{A}\cdot\mathbf{B} = 0$, then $\mathbf{A}$ and $\mathbf{B}$ are perpendicular. Conversely, if $\mathbf{A}$ and $\mathbf{B}$ are perpendicular, then $\mathbf{A}\cdot\mathbf{B} = 0$.

As Figure 14.2.6 shows, $\mathbf{A}$ can be expressed as the sum of a vector parallel to $\mathbf{B}$ and a vector perpendicular to $\mathbf{B}$.

The vector parallel to $\mathbf{B}$ we call the **projection** of $\mathbf{A}$ on $\mathbf{B}$, denoted $\text{proj}_{\mathbf{B}}\,\mathbf{A}$. The vector perpendicular to $\mathbf{B}$ is then $\mathbf{A} - \text{proj}_{\mathbf{B}}\,\mathbf{A}$.

The length of $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ is $\|\mathbf{A}\||\cos\theta|$, which equals $\frac{|\mathbf{A}\cdot\mathbf{B}|}{\|\mathbf{B}\|}$. If θ is less than $\pi/2$, $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ points in the same direction as $\mathbf{B}$.

If $\pi/2 < \theta \leq \pi$, then $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ points in the direction opposite to that of $\mathbf{B}$. In either case, since $\mathbf{B}/\|\mathbf{B}\|$ is the unit vector in the direction of $\mathbf{B}$, we have

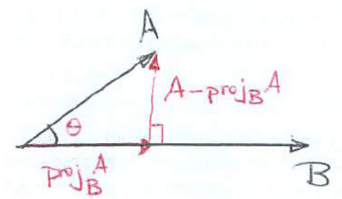


Figure 14.2.6:

Let $\mathbf{A}$ and $\mathbf{B}$ be vectors. $\text{proj}_{\mathbf{B}}\,\mathbf{A} = \frac{\mathbf{A}\cdot\mathbf{B}}{\|B\|}\frac{\mathbf{B}}{\|\mathbf{B}\|}$

- If $\mathbf{A}\cdot\mathbf{B}$ is positive, then the angle between the vectors is less than $\pi/2$. In this case $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ points in the same direction as $\mathbf{B}$.
- If $\mathbf{A}\cdot\mathbf{B}$ is negative, then the angle between the vectors is greater than $\pi/2$. In this case $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ points in the opposite direction as $\mathbf{B}$.

If $\mathbf{A}\cdot\mathbf{B}$ is negative, then the angle between $\mathbf{A}$ and $\mathbf{B}$ is obtuse (greater than $\pi/2$). Figure 14.2.7 shows this situation. As Figure 14.2.7 illustrates, $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ points in the direction opposite that of $\mathbf{B}$.

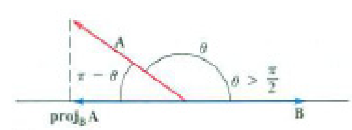


Figure 14.2.7:

Computing $\mathbf{A}\cdot\mathbf{B}$ in Terms of Their Components

We defined $\mathbf{A}\cdot\mathbf{B}$, using the geometric interpretation of $\mathbf{A}$ and $\mathbf{B}$. But what if $\mathbf{A}$ and $\mathbf{B}$ are given in terms of their components, $\mathbf{A} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3\rangle$? How would we find $\mathbf{A}\cdot\mathbf{B}$ in that case?

The answer turns out to be quite simple:

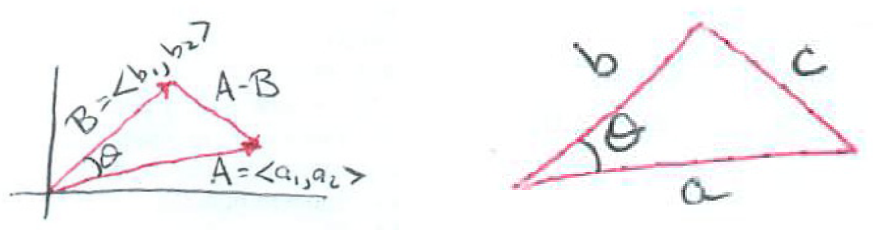
If $\mathbf{A} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{B} = \langle b_1, b_2, b_3\rangle$, then $\mathbf{A}\cdot\mathbf{B} = a_1b_1 + a_2b_2 + a_3b_3$.

The dot product is the sum of three numbers. Each number is a product of corresponding components.

For vectors in the xy-plane, the result is a bit shorter:

If $\mathbf{A} = \langle a_1, a_2 \rangle$ and $\mathbf{B} = \langle b_1, b_2 \rangle$, then $\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2$.

A proof of the Law of Cosines is defined in Exercise 45



(a) (b)

Figure 14.2.8:

For convenience we establish the second result. Our reasoning rests on the Law of Cosines. It says that in a triangle where sides have lengths a, b, and c, and angle θ opposite the side with length c, as in Figure 14.2.8(b), $c^2 = a^2 + b^2 - 2ab\cos(\theta)$.

Then

$$\|\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{A}\|^2 + \|\mathbf{B}\|^2 - 2\|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta),$$

which tells us that

$$\|\mathbf{A} - \mathbf{B}\|^2 = \|\mathbf{A}\|^2 + \|\mathbf{B}\|^2 - 2\mathbf{A} \cdot \mathbf{B}, \tag{14.2.1}$$

All that's left is to complete the three squares and solve for $\mathbf{A} \cdot \mathbf{B}$.

Translating (14.2.1) into components, we have

$$(a_1 - b_1)^2 + (a_2 - b_2)^2 = a_1^2 + a_2^2 + b_1^2 + b_2^2 - 2\mathbf{A} \cdot \mathbf{B}$$

or

$$a_1^2 - 2a_1 b_1 + b_1^2 + a_2^2 - 2a_2 b_2 + b_2^2 = a_1^2 + a_2^2 + b_1^2 + b_2^2 - 2\mathbf{A} \cdot \mathbf{B}.$$

Thus

$$-2(a_1 b_1 + a_2 b_2) = -2\mathbf{A} \cdot \mathbf{B},$$

from which it follows, as the night follows the day, that

$$\mathbf{A} \cdot \mathbf{B} = a_1 b_1 + a_2 b_2.$$

The argument in the case of space vectors is practically the same, as doing Exercise 38 will show.

EXAMPLE 3 Find $\cos(\mathbf{A},\mathbf{B})$ when $\mathbf{A} = \langle 6, 3\rangle$ and $\mathbf{B} = \langle -1, 1\rangle$.
SOLUTION We know that $\mathbf{A}\cdot\mathbf{B} = \|\mathbf{A}\|\|\mathbf{B}\|\cos(\mathbf{A},\mathbf{B})$. Thus

$$6\cdot(-1) + 3\cdot(1) = \sqrt{2^2+3^2}\sqrt{(-1)+1^2}\cos(\mathbf{A},\mathbf{B})$$

or

$$-3 = \sqrt{26}\cos(\mathbf{A},\mathbf{B}),$$

e conclude that

$$\cos(\mathbf{A},\mathbf{B}) = -3/\sqrt{26}.$$

◇

Clearly θ is an obtuse angle. A calculator would estimate θ, if we were curious. Figure 14.2.9 shows that the answer is reasonable.

As Example 3 illustrates

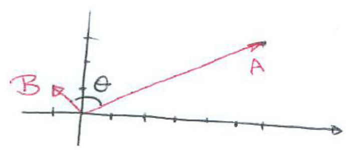


Figure 14.2.9:

$$\cos(\theta) = \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}$$

EXAMPLE 4

1. Find the projection of $\mathbf{A} = 2\mathbf{i} + \mathbf{j}$ on $\mathbf{B} = -3\mathbf{i} + 2\mathbf{j}$.
2. Express $\mathbf{A}$ as the sum of a vector parallel to $\mathbf{B}$ and a vector perpendicular to $\mathbf{B}$.

SOLUTION

1. In this case

$$\begin{aligned}\operatorname{proj}_{\mathbf{B}}\mathbf{A} &= \frac{\mathbf{A}\cdot\mathbf{B}}{\|\mathbf{B}\|}\frac{\mathbf{B}}{\|\mathbf{B}\|}\\ &= \frac{(2\mathbf{i}+\mathbf{j})\cdot(-3\mathbf{i}+2\mathbf{j})}{|-3\mathbf{i}+2\mathbf{j}|}\frac{-3\mathbf{i}+2\mathbf{j}}{|-2\mathbf{i}+2\mathbf{j}|}\\ &= \frac{(-6+2)}{\sqrt{13}}\frac{(-3\mathbf{i}+2\mathbf{j})}{\sqrt{13}}\\ &= \frac{-4}{13}(-3\mathbf{i}+2\mathbf{j}) = \frac{12}{13}\mathbf{i} - \frac{8}{13}\mathbf{j}.\end{aligned}$$

Figure 14.2.10 shows the vector $\mathbf{A}$, $\mathbf{B}$, and $\operatorname{proj}_{\mathbf{B}}\mathbf{A}$.

In this case $\mathbf{A}\cdot\mathbf{B}$ is negative, the angle between $\mathbf{A}$ and $\mathbf{B}$ is obtuse, and $\operatorname{proj}_{\mathbf{B}}\mathbf{A}$ points in the direction opposite to the direction of $\mathbf{B}$.

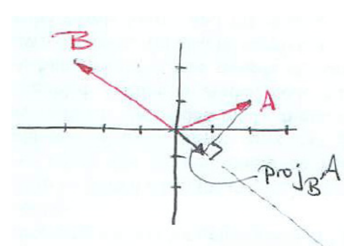


Figure 14.2.10:

2. The vector $\mathbf{A} - \text{proj}_{\mathbf{B}}\,\mathbf{A}$ is perpendicular to $\mathbf{B}$ and we have

$$\begin{aligned}\mathbf{A} &= (\text{proj}_{\mathbf{B}}\,\mathbf{A}) + (\mathbf{A} - \text{proj}_{\mathbf{B}}\,\mathbf{A}) \\ &= \left(\frac{12}{13}\mathbf{i} - \frac{8}{13}\mathbf{j}\right) + \left(2\mathbf{i} + \mathbf{j} - \left(\frac{12}{13}\mathbf{i} - \frac{8}{13}\mathbf{j}\right)\right) \\ &= \underbrace{\left(\frac{12}{13}\mathbf{i} - \frac{8}{13}\mathbf{j}\right)}_{\text{parallel to } \mathbf{B}} + \underbrace{\left(\frac{14}{13}\mathbf{i} + \frac{21}{13}\mathbf{j}\right)}_{\text{perpendicular to } \mathbf{B}}.\end{aligned}$$

◇

The scalar $\mathbf{A} \cdot (\mathbf{B}/\|\mathbf{B}\|)$ is the component of $\mathbf{A}$ in the direction of $\mathbf{B}$, denoted $\text{comp}_{\mathbf{B}}(\mathbf{A})$. It can be positive, negative, or zero. Its absolute value is the length of $\text{proj}_{\mathbf{B}}(\mathbf{A})$.

EXAMPLE 5 Find $\text{proj}_{\mathbf{B}}(\mathbf{A})$ and $\text{comp}_{\mathbf{B}}(\mathbf{A})$ when $\mathbf{A} = \mathbf{i} + 3\mathbf{j}$ and $\mathbf{B} = \mathbf{i} - \mathbf{j}$.

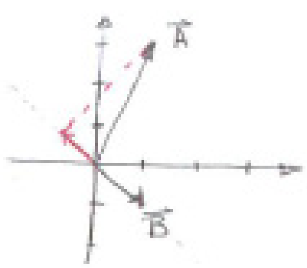


Figure 14.2.11:

SOLUTION Since $\|\mathbf{B}\| = \sqrt{1^2 + 1^2} = \sqrt{2}$ and $\mathbf{A} \cdot \mathbf{B} = 1 - 3 = -2$,

$$\text{proj}_{\mathbf{B}}(\mathbf{A}) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{B}\|} \frac{\mathbf{B}}{\|\mathbf{B}\|} = \frac{-2}{\sqrt{2}} \frac{\mathbf{i} - \mathbf{j}}{\sqrt{2}} = -\mathbf{i} + \mathbf{j}$$

and $\text{comp}_{\mathbf{B}}(\mathbf{A}) = (\mathbf{A} \cdot \mathbf{B})/\|\mathbf{B}\| = -2/\sqrt{2} = -\sqrt{2}$. This agrees with Figure 14.2.11. ◇

Properties of the Dot Product

With the aid of the formula for the dot product in terms of components, it is easy to establish the following properties:

$$\begin{aligned}\mathbf{A} \cdot \mathbf{B} &= \mathbf{B} \cdot \mathbf{A} && \text{commutative} \\ \mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) &= \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} && \text{distributive} \\ c\mathbf{A} \cdot \mathbf{B} &= c(\mathbf{A} \cdot \mathbf{B}) && c \text{ a scalar.}\end{aligned}$$

$$\cos(\theta) = \cos(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|}. \tag{14.2.2}$$

Equation (14.2.2) tells us how to find the cosine of the angle between two vectors. With the aid of a calculator, we then can find the angle itself. Note that if $\cos(\theta) > 0$, then $0 < \theta < \pi/2$, and when $\cos(\theta) < 0$, then $\pi/2 < \theta \leq \pi$.

EXAMPLE 6 Show that the vectors $\langle 2, -3, 4\rangle$ and $\langle 1, 2, 1\rangle$ are perpendicular.

SOLUTION We want to show that the angle θ between the vector in $\pi/2$. To do this we show $\cos(\theta) = 0$. Now,

$$\cos(\theta) = \frac{\mathbf{A}\cdot\mathbf{B}}{|\mathbf{A}||\mathbf{B}|} = \frac{(1\cdot 2) + 2(-3) + 1\cdot 4}{|\mathbf{A}||\mathbf{B}|} = \frac{2-6+4}{|\mathbf{A}||\mathbf{B}|} = 0.$$

Therefore the vectors are perpendicular. ◇

Example 6 illustrates this test for two vectors being perpendicular to each other.

Two nonzero vectors are perpendicular if their dot product is 0.

The Dot Product in Business and Statistics

Imagine that a fast food restaurant sells 30 hamburgers, 20 salads, 15 soft drinks, and 13 orders of french fries. This is recorded by the four-dimensional "vector" $\langle 30, 20, 15, 13\rangle$. A hamburger sells for \$1.99, a salad for \$1.50, a soft drink for \$1.00, and an order of french fries for \$1.10. The "price vector" is $\langle 1.99, 1.50, 1.00, 1.10\rangle$. The dot product of these two vectors, $30(1.99) + 20(1.50) + 15(1.00) + 13(1.10)$, would be the total amount paid for all items. Descriptions of the economy use "production vectors," "cost vectors," "price vectors," and "profit vectors" with many more than the four componenets of our restaurant example.

In statistics the coefficient of correlation is defined in terms of a dot product. For instance, you may determine the height and weight of n persons. Let the height of the ith person be h_i and the weight be w_i. Let h be the average of the n heights and w be the average of the n weights. Let $\mathbf{H} = \langle h_1 - h, h_2 - h, \cdots, h_n - h\rangle$ and $\mathbf{W} = \langle w_1 - w, w_2 - w, \cdots, w_n - w\rangle$. Then coefficient of correlation between the heights and weights is defined to be

$$\frac{\mathbf{H}\cdot\mathbf{W}}{\|\mathbf{H}\|\|\mathbf{W}\|}.$$

In analogy with vectors in the plane or space,

$$\mathbf{HW} = \sum_{i=1}^{n}(h_i - h)(w_i - w), \|\mathbf{H}\| = \sqrt{\sum_{i=1}^{n}(h_i - h)^2}, \|\mathbf{W}\| = \sqrt{\sum_{i=1}^{n}(w_i - w)^2}.$$

It turns out that the coefficient of correlation is simply the cosine of the angle between the points $\mathbf{H} = \langle h_1 - h, h_2 - h, \cdots, h_n - h\rangle$ and $\mathbf{W}\langle w_1 - w, w_2 - w, \cdots, w_n - w\rangle$ in n-dimensional space.

Summary

We defined the dot (scalar) product of two vectors $\mathbf{A}$ and $\mathbf{B}$ geometrically as $\|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta)$, where θ is the angle between them. We then obtained a formula for $\mathbf{A}\cdot\mathbf{B}$ in terms of their components, as $\langle a_1, a_2\rangle\cdot\langle b_1, b_2\rangle = a_1b_1 + a_2b_2$ and a similar formula for the dot product of two space vectors.

The dot product enabled us to express a vector $\mathbf{A}$ as the sum of a vector parallel to $\mathbf{B}$ ($\text{proj}_{\mathbf{B}}\,\mathbf{A}$) and a vector perpendicular to $\mathbf{B}$ ($\mathbf{A} - \text{proj}_{\mathbf{B}}\,\mathbf{A}$).

When their dot product is 0, two non-zero vectors are perpendicular.

The zero-vector, $\mathbf{0}$, is considered to be perpendicular to every vector.

More generally, we can use the dot product to find the angle θ between two vectors:

$$\cos(\theta) = \cos(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \cdot \mathbf{B}}{|\mathbf{A}||\mathbf{B}|}.$$

EXERCISES for Section 14.2 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4 compute $\mathbf{A} \cdot \mathbf{B}$.

1.[R] $\mathbf{A}$ has length 3, $\mathbf{B}$ has length 4, and the angle between $\mathbf{A}$ and $\mathbf{B}$ is $\pi/4$.

2.[R] $\mathbf{A}$ has length 2, $\mathbf{B}$ has length 3, and the angle between $\mathbf{A}$ and $\mathbf{B}$ is $3\pi/4$.

3.[R] $\mathbf{A}$ has length 5, $\mathbf{B}$ has length $\frac{1}{2}$, and the angle between $\mathbf{A}$ and $\mathbf{B}$ is $\pi/2$.

4.[R] $\mathbf{A}$ is the zero vector $\mathbf{0}$, and $\mathbf{B}$ has length 5.

In Exercises 5 to 8 compute $\mathbf{A} \cdot \mathbf{B}$ using the formula in terms of components.

5.[R] $\mathbf{A} = -2\mathbf{i} + 3\mathbf{j}$, $\mathbf{B} = 4\mathbf{i} + 4\mathbf{j}$

6.[R] $\mathbf{A} = 0.3\mathbf{i} + 0.5\mathbf{j}$, $\mathbf{B} = 2\mathbf{i} - 1.5\mathbf{j}$

7.[R] $\mathbf{A} = 2\mathbf{i} - 3\mathbf{j} - \mathbf{k}$, $\mathbf{B} = 3\mathbf{i} + 4\mathbf{j} - \mathbf{k}$

8.[R] $\mathbf{A} = \mathbf{i} + \mathbf{j} + \mathbf{k}$,$\mathbf{B} = 2\mathbf{i} + +3\mathbf{j} - 5\mathbf{k}$

9.[R]

(a) Draw the vectors $7\mathbf{i} + 12\mathbf{j}$ and $9\mathbf{i} - 5\mathbf{j}$.

(b) Do they seem to be perpendicular?

(c) Determine whether they are perpendicular by examining their dot product.

10.[R]

(a) Draw the vectors $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$ and $\mathbf{i} + \mathbf{j} - \mathbf{k}$.

(b) Do they seem to be perpendicular?

(c) Determine whether they are perpendicular by examining their dot product.

11.[R]

(a) Estimate the angle between $\mathbf{A} = 3\mathbf{i} + 4\mathbf{j}$ and $\mathbf{B} = 5\mathbf{i} + 12\mathbf{j}$ by drawing them.

(b) Find the angle between $\mathbf{A}$ and $\mathbf{B}$.

12.[R] Let $P = (6, 1)$, $Q = (3, 2)$, $R = (1, 3)$, and $S = (4, 5)$.

(a) Draw the vectors $\overrightarrow{PQ}$ and $\overrightarrow{RS}$.

(b) Using the diagram in (a) estimate the angle between $\overrightarrow{PQ}$ and $\overrightarrow{RS}$.

(c) Using the dot product, find the $\cos(\overrightarrow{PQ}, \overrightarrow{RS})$, that is, the cosine of the angle between $\overrightarrow{PQ}$ and $\overrightarrow{RS}$.

(d) Using (c) and a calculator, find the angle in (b).

13.[R] Find the angle between $2\mathbf{i} - 4\mathbf{j} + 6\mathbf{k}$ and $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.

14.[R] Find the angle betwen $\mathbf{i} + \mathbf{j} + 3\mathbf{k}$ and $3\mathbf{i} + 6\mathbf{j} - 3\mathbf{k}$.

15.[R] Find the angle between $\overrightarrow{AB}$ and $\overrightarrow{CD}$ if $A = (1, 3)$, $B = (7, 4)$, $C = (2, 8)$, and $D = (1, -5)$.

16.[R] Find the angle between $\overrightarrow{AB}$ and $\overrightarrow{CD}$ if $A = (1, 2, -5)$, $B = (1, 0, 1)$, $C = (0, -1, 3)$, and $D = (2, 1, 4)$.

17.[R] Find the length of the projection of $-4\mathbf{i} + 5\mathbf{j}$ on the line through $(2, -1)$ and $(6, 1)$.

(a) By making a drawing and estimating the length by eye.

(b) By using the dot product.

18.[R]

(a) Find a vector $\mathbf{C}$ parallel to $\mathbf{i} + 2\mathbf{j}$ and a vector $\mathbf{D}$ perpendicular to $\mathbf{i} + 2\mathbf{j}$ such that $-3\mathbf{i} + 4\mathbf{j} = \mathbf{C} + \mathbf{D}$.

(b) Draw the vectors in (a) to check that your answer is reasonable.

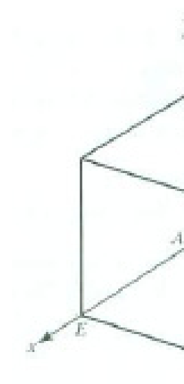
19.[R]

(a) Find a vector $\mathbf{C}$ parallel to $2\mathbf{i} - \mathbf{j}$ and a vector $\mathbf{D}$ perpendicular to $2\mathbf{i} - \mathbf{j}$ such that $3\mathbf{i} + 4\mathbf{j} = \mathbf{C} + \mathbf{D}$.

(b) Draw the vectors in (a) to check that your answer is reasonable.

20.[M] Give an example of a vector in the xy plane that is perpendicular to $3\mathbf{i}-2\mathbf{j}$.

21.[M] Give an example of a vector that is perpendicular to $5\mathbf{i}-3\mathbf{j}+4\mathbf{k}$.



Exercises 22 to 26 refer to the cube in Figure 14.2.12.

Figure 14.2.12:

22.[M] Find $\cos(\overrightarrow{AC}, \overrightarrow{BD})$, the cosine of the angle between $\overrightarrow{AC}$ and $\overrightarrow{BD}$.
23.[M] Find $\cos(\overrightarrow{AF}, \overrightarrow{BD})$, the cosine of the angle between $\overrightarrow{AF}$ and $\overrightarrow{BD}$.
24.[M] Find $\cos(\overrightarrow{AC}, \overrightarrow{AM})$, the cosine of the angle between $\overrightarrow{AC}$ and $\overrightarrow{AM}$.
25.[M] Find $\cos(\overrightarrow{MD}, \overrightarrow{MF})$, the cosine of the angle between $\overrightarrow{MD}$ and $\overrightarrow{MF}$.
26.[M] Find $\cos(\overrightarrow{EF}, \overrightarrow{BD})$, the cosine of the angle between $\overrightarrow{EF}$ and $\overrightarrow{BD}$.

27.[R] How far is the point $(1,2,3)$ from the line through the points $(1,4,2)$ and $(2,1,-4)$?

28.[M] If $\mathbf{A}\cdot\mathbf{B} = \mathbf{A}\cdot\mathbf{C}$ and $\mathbf{A}$ is not $\mathbf{0}$, must $\mathbf{B} = \mathbf{C}$?

29.[C] If $\|\mathbf{A}\| = 3$ and $\|\mathbf{B}\| = 5$,

(a) how large can $\|\mathbf{A} + B\|$ be?

(b) how small?

30.[C] By considering the dot product of the two unit vectors $\mathbf{u}_1 = \cos\theta_1\mathbf{i} + \sin\theta_1\mathbf{j}$ and $\mathbf{u}_2 = \cos\theta_2\mathbf{i} + \sin\theta_2\mathbf{j}$, prove that

$$\cos(\theta_1 - \theta_2) = \cos\theta_1\cos\theta_2 + \sin\theta_1\sin\theta_2.$$

31.[C] Consider a tetrahedron (not necessarily regular). It has six edges. Show that the line segment joining the midpoints of two opposite edges is perpendicular to the line segment joining another pair of opposite edges if any only if the remaining two edges are of the same length.

32.[C] The output of a firm that manufactures x_1 washing machines, x_2 refrigerators, x_3 dishwashers, x_4 stoves, and x_5 clothes dryers is recorded by the five-dimensional production vector $\mathbf{P} = \langle x_1, x_2, x_3, x_4, x_5 \rangle$. Similarly, the cost vector $\mathbf{C} = \langle y_1, y_2, y_3, y_4, y_5 \rangle$ records the cost of producing each item; for instance, each refrigerator costs the firm y_2 dollars.

(a) What is the economic significance of $\mathbf{P} \cdot \mathbf{C} = \langle 20, 0, 7, 9, 15 \rangle \cdot \langle 50, 70, 30, 20, 10 \rangle$?

(b) If the firm doubles the production of all items in (a), what is its new production vector?

33.[C] Let P_1 be the profit from selling a washing machine and P_2, P_3, P_4, and P_5 be defined analogously for the firm of Exercise 32. (Some of the P's may be negative.) What does it mean to the firm to have $\langle P_1, P_2, P_3, P_4, P_5 \rangle$ "perpendicular" to $\langle x_1, x_2, x_3, x_4, x_5 \rangle$?

34.[C] If a_1, a_2, b_1, b_2 are four numbers, explain why

$$|a_1 b_1 + a_2 b_2| \leq \sqrt{a_1^2 + a_2^2}\sqrt{b_1^2 + b_2^2}.$$

SHERMAN: This exercise is also #46 in 14.1. Keep only one. In Chapter Summary? Your thoughts?

35.[R] Prove that $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A}$

(a) using the geometric definition of the dot product,

(b) using the formula for the dot product in terms of components.

36.[R] Prove that $\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C}$

(a) using the geometric definition of the dot product,

(b) using the formula for the dot product in terms of components.

37.[C] Don't try to obtain the equation $\mathbf{A}\cdot(\mathbf{B}+\mathbf{C}) = \mathbf{A}\cdot\mathbf{B}+\mathbf{A}\cdot\mathbf{C}$ geometrically. If you use the geometric definition of the dot product, what does that distributive law say? Picture **B** and **C** in a horizontal plane and **A** not in that plane, as in Figure 14.2.13.

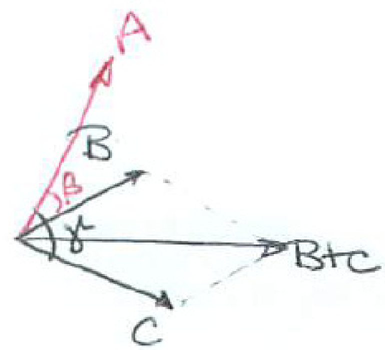


Figure 14.2.13:

It's not so obvious is it?

38.[R] Prove that $\langle a_1, a_2, a_3\rangle\cdot\langle b_1, b_2, b_3\rangle = a_1b_1 + a_2b_2 + a_3b_3$ HINT: Read the proof in the case of planar vectors on page 1138.

39.[C] Let $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$ be unit vectors such that each two are perpendicular. Let **A** be a vector.

(a) Draw a picture that shows that there are scalars x, y, and z such that $\mathbf{A} = x\mathbf{u}_1 + y\mathbf{u}_2 + z\mathbf{u}_3$.

(b) Express x as a dot product.

(c) Express $x - z$ as a dot product.

40.[M]

(a) Let **A** be a vector in the xy plane and $\mathbf{u}_1$ and $\mathbf{u}_2$ perpendicular unit vectors in that plane. If $\mathbf{A}\cdot\mathbf{u}_1 = 0$ and $\mathbf{A}\cdot\mathbf{u}_2 = 0$, must $\mathbf{A} = \mathbf{0}$?

(b) Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be nonparallel unit vectors in the xy plane. If $\mathbf{A}\cdot\mathbf{v}_1$ and $\mathbf{A}\cdot\mathbf{v}_2 = 0$, must $\mathbf{A} = \mathbf{0}$?

41.[C] A firm sells x chairs at C dollars per chair and y desks at D dollars per week. It costs the firm c dollars to make a chair and d dollars to make a desk. What is the economic interpretation of

(a) Cx?

(b) $(x\mathbf{i} + y\mathbf{j}) \cdot (C\mathbf{i} + D\mathbf{j})$?

(c) $(x\mathbf{i} + y\mathbf{j}) \cdot (c\mathbf{i} + d\mathbf{j})$?

(d) $(x\mathbf{i} + y\mathbf{j}) \cdot (C\mathbf{i} + D\mathbf{j}) > (x\mathbf{i} + y\mathbf{j}) \cdot (c\mathbf{i} + d\mathbf{j})$?

42.[C] A force $\mathbf{F}$ of 10 newtons has the direction of the vector $2\mathbf{i} + 3\mathbf{j} + \mathbf{k}$. This force pushes an object on a ramp in a straight line from the point $(3, 1, 5)$ to the point $(4, 3, 7)$, where coordinates are measured in meters. How much work does the force accomplish?

43.[C] Show that if the two diagonals of the parallelogram are perpendicular, then the four sides have the same length (forming a rhombus). HINT: Use the dot product.

44.[C] Some molecules consist of 4 atoms arranged as the vertices of a regular tetrahedron, for instance at the points labeled A, B, C, and D in Figure 14.2.14.

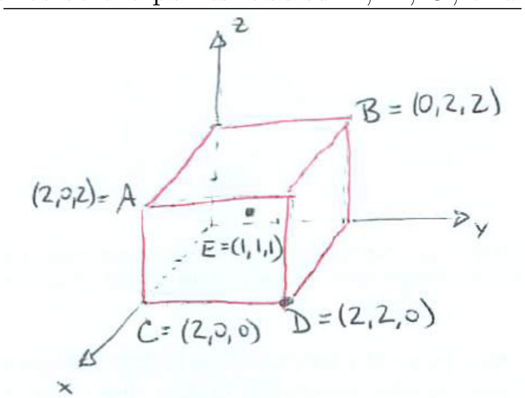


Figure 14.2.14:

(a) Show that A, B, C, and D are vertices of a regular tetrahedron. HINT: Show that the four faces are equilateral triangles.

(b) Chemists are interested in the angle $\theta = AEB$. Show that $\cos(\theta) = -1/3$.

(c) Find θ (approximately).

45.[M] The key to obtaining the expression for the dot product in terms of components is from trigonometry: the Law of Cosines. In view of this, it makes sense to see why the Law of Cosines is true. The proof is quite easy, since it consists just of two applications of the Pythagorean Theorem. Figure 14.2.15 shows a triangle with sides a, b, c, with angle θ opposite side c. (We are concerned, for the moment, in the case when θ is less than $\frac{\pi}{2}$.

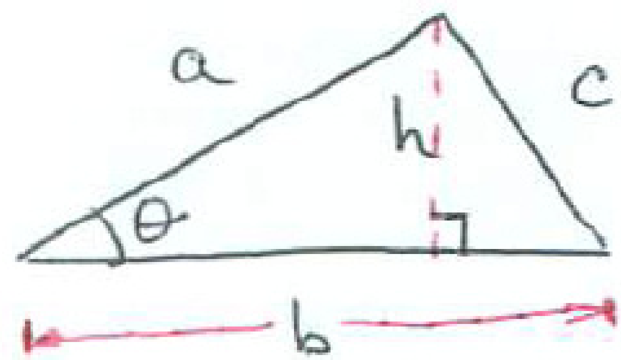


Figure 14.2.15:

(a) Show that $h^2 = a^2 - a^2\cos^2(\theta)$.

(b) Show that $h^2 = c^2 - (b - a\cos(theta))^2$.

(c) By equating the two expressions for h^2 found in (a) and (b), obtain the Law of Cosines.

46.[C] In the Exercise 45 the altitude of length h meets the side of length b. If $\theta > \pi/2$, that altitude has its base outside of side b. Prove the Law of Cosines in this case.

47.[R] What is $\text{proj}_{\mathbf{B}}\,\mathbf{A}$ if $\mathbf{A} = 2\mathbf{i} + \mathbf{j} - 3\mathbf{k}$ and $\mathbf{B} = \mathbf{i} + \mathbf{j} + \mathbf{k}$?

48.[C] How far is the point $(2, 3, 5)$ from the line through the origin and $(1, -1, 2)$?

49.[R] Express the vector $\mathbf{i} + \mathbf{j} + \mathbf{k}$ as the sum of a vector parallel to $\mathbf{i} - \mathbf{j} + 2\mathbf{k}$ and a vector perpendicular to $\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

50.[M]

Jane: I don't like the way the author found how to express **A** as the sum of a vector parallel to **B** and a vector perpendicular to **B**.

Sam: It was O.K. for me. But I had to memorize a formula.

Jane: My goal is to memorize nothing. I simply write $\mathbf{A} = x\mathbf{B} + \mathbf{C}$, when $\mathbf{C}$ is perpendicular to $\mathbf{A}$. Then *I* dot with $\mathbf{B}$, getting

$$\mathbf{A} \cdot \mathbf{B} = x\mathbf{B} \cdot \mathbf{B} + \mathbf{C} \cdot \mathbf{B}.$$

Since $\mathbf{C}$ is perpendicular to $\mathbf{B}$, $\mathbf{C} \cdot \mathbf{B} = 0$, and lo and behold, I have

$$x = \frac{\mathbf{A} \cdot \mathbf{B}}{\mathbf{B} \cdot \mathbf{B}}.$$

So the vector parallel to $\mathbf{B}$ is $\frac{\mathbf{A} \cdot \mathbf{B}}{\mathbf{B} \cdot \mathbf{B}}\mathbf{B}$.

Sam: Cool. So why did the author go through all that stuff?

Jane: Maybe they wanted to reinforce the definition of the dot product and the rule of the angle.

Sam: O.K. But how do I get the vector $\mathbf{C}$ perpendicular to $\mathbf{B}$?

Jane: Simple...

Complete Jane's reply.

14.3 The Cross Product of Two Vectors

The dot product of two vectors is a scalar. The product of two vectors we define in this section is a vector. This vector has the property that it is perpendicular to each of the given vector.

Definition of the Cross Product

Let $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ be two non-zero vectors that are not parallel. We will construct a vector $\mathbf{C}$ that is perpendicular to both $\mathbf{A}$ and $\mathbf{B}$. Of course $\mathbf{C}$ is not unique since any vector parallel to $\mathbf{C}$ is also perpendicular to $\mathbf{A}$ and $\mathbf{B}$.

Let $\mathbf{C} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. We want $\mathbf{C} \cdot \mathbf{A}$ and $\mathbf{C} \cdot \mathbf{B}$ to be 0. This gives us the equations

$$a_1x + a_2y + a_3z = 0 \tag{14.3.1}$$

$$b_1x + b_2y + b_3z = 0 \tag{14.3.2}$$

We eliminate x by subtracting b_1 times (14.3.1) from a_1 times (14.3.2), as follows.

a_1 times (14.3.2) $$a_1b_1x + a_1b_2y + a_1b_3z = 0 \tag{14.3.3}$$

b_1 times (14.3.1) $$b_1a_1x + b_1a_2y + b_1a_3z = 0 \tag{14.3.4}$$

Subtracting the bottom equation (14.3.4) from the top equation (14.3.3) gives us

$$(a_1b_2 - a_2b_1)y + (a_1b_3 - a_3b_1)z = 0 \tag{14.3.5}$$

This is like solving $2y + 3z = 0$ by letting $y = -3$ and $z = 2$.

A simple non-zero solution of (14.3.5) is

$$y = -(a_1b_3 - a_3b_1), \quad z = a_1b_2 - a_2b_1$$

To find the corresponding x, substitute the value found for y and z into (14.3.1). As Exercise 39 shows, the straightforward algebra yields

$$x = a_2b_3 - a_3b_2.$$

So the vector

$$(a_2b_2 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k} \tag{14.3.6}$$

is perpendicular to $\mathbf{A}$ and $\mathbf{B}$. It is denoted $\mathbf{A} \times \mathbf{B}$ and is called the **vector product** of $\mathbf{A}$ and $\mathbf{B}$ or the cross product of $\mathbf{A}$ and $\mathbf{B}$. This vector is defined even if $\mathbf{A}$ and $\mathbf{B}$ are parallel or if one of them is 0.

Determinants and the Cross Product

The expression (14.3.6) for the cross product is not easy to memorize. Fortunately, determinants provide a convenient memory aid.

Four numbers arranged in a square from a matrix of order 2, for instance

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$$

The determinant of this matrix is the number $a_1b_2 - a_2b_1$, denoted

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \qquad \text{or} \qquad \det\begin{pmatrix} cca_1 & a_2 \\ b_1 & b_2 \end{pmatrix}.$$

Each term in the cross product, (14.3.6), is itself the determinant of a matrix of order 2, namely

$$\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \quad \text{and} \quad \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}$$

Nine numbers arranged in a square for a matrix of order 3, for instance

$$\begin{pmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

Its determinant is defined with the aid of determinants of order 2:

$$c_1\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - c_2\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + c_3\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2. \end{vmatrix}$$

The coefficient of each c_i is plus or minus the determinant of the matrix of order 2 that remains when the row and column in which c_i appears are deleted, as shown in Figure 14.3.1 for the coefficient of c_i.

Therefore we can write (14.3.6) as a determinant of a matrix, and we have

$$\mathbf{A} \times \mathbf{B} = \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \qquad (14.3.7)$$

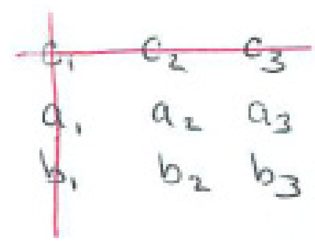


Figure 14.3.1:

DEFINITION *(Cross product (vector product).)* Let

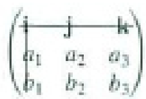
$$\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{k} + a_3\mathbf{k} \text{ and } \mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}.$$

The vector

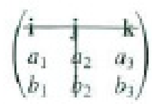
$$\begin{aligned} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} &= \mathbf{i}\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} - \mathbf{j}\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} + \mathbf{k}\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \\ &= (a_2b_3 - a_3b_2)\mathbf{i} - (a_1b_3 - a_3b_1)\mathbf{j} + (a_1b_2 - a_2b_1)\mathbf{k}. \end{aligned}$$

is called the **cross product** (or **vector product**) of **A** and **B**. It is denoted $\mathbf{A} \times \mathbf{B}$.

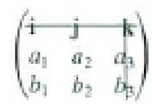
The determinant for $\mathbf{A} \times \mathbf{B}$ is expanded along its first row:



Delete the two lines through **i**. The determinant of the remaining square is the coefficient of **i** in $\mathbf{A} \times \mathbf{B}$.



Delete the two lines through **j**. The determinant of the remaining square is the coefficient of **j** in $\mathbf{A} \times \mathbf{B}$.



Delete the two lines through **k**. The determinant of the remaining square is the coefficient of **k** in $\mathbf{A} \times \mathbf{B}$.

EXAMPLE 1 Compute $\mathbf{A} \times \mathbf{B}$ if $\mathbf{A} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k}$ and $\mathbf{B} = 3\mathbf{i} + 4\mathbf{j} + \mathbf{k}$.
SOLUTION By definition,

$$\begin{aligned}\mathbf{A} \times \mathbf{B} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -1 & 3 \\ 3 & 4 & 1 \end{vmatrix} = \mathbf{i}\begin{vmatrix} -1 & 3 \\ 4 & 1 \end{vmatrix} - \mathbf{j}\begin{vmatrix} 2 & 3 \\ 3 & 1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} 2 & -1 \\ 3 & 4 \end{vmatrix} \\ &= -13\mathbf{i} + 7\mathbf{j} + 11\mathbf{k}\end{aligned}$$

◇

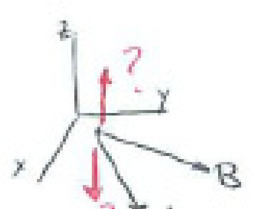
Recall: The zero vector is, by definition, perpendicular to every vector.

The cross cross product has these properties:

1. $\mathbf{A} \times \mathbf{B}$ is perpendicular to both **A** and **B**.
2. $\mathbf{A} \times \mathbf{B} = -(\mathbf{B} \times \mathbf{A})$.
3. $\mathbf{A} \times \mathbf{B} = \mathbf{0}$ if **A** and **B** are parallel or at least one of them is 0.
4. $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$.

See Exercises 27 and 28.

The first property holds because that is how we constructed the cross product. The second and third are established by straightforward computations, using (14.3.7). Exercises 16 and 17 take care of property 4.



The Direction of $\mathbf{A} \times \mathbf{B}$?

We know that $\mathbf{A} \times \mathbf{B}$ is perpendicular to $\mathbf{A}$ and $\mathbf{B}$, but there are two possible directions, as Figure 14.3.2 shows,

To find out, take a specific case and we compute $\mathbf{i} \times \mathbf{j}$:

$$\mathbf{i} \times \mathbf{j} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 0\mathbf{i} - 0\mathbf{j} + \mathbf{k} = \mathbf{k}.$$

This suggests the general situation. The direction of $\mathbf{A} \times \mathbf{B}$ is given by the right hand rule:

Left-handed people must use their right hand here.

Curl the fingers of the right hand to go from $\mathbf{A}$ and $\mathbf{B}$. The thumb points in the direction of $\mathbf{A} \times \mathbf{B}$.

EXAMPLE 2 Check that the right hand rule is correct in the case for $\mathbf{j} \times \mathbf{i}$.

SOLUTION

$$\mathbf{j} \times \mathbf{i} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 0\mathbf{i} - 0\mathbf{j} - \mathbf{k} = -\mathbf{k}.$$

In this case, $\mathbf{j} \times \mathbf{i}$, points downward, the opposite of $\mathbf{i} \times \mathbf{j}$.

The right hand rule is illustrated in Figure 14.3.4.

The thumb indeed points downward. ◇

Figure 14.3.4:

How Long is $\mathbf{A} \times \mathbf{B}$

To find a geometric meaning for $\|\mathbf{A} \times \mathbf{B}\|$ we will find $|\mathbf{A} \times \mathbf{B}|^2$ with the aid of (4). That is, we will compute $(\mathbf{A} \times \mathbf{B}) \times (\mathbf{A} \times \mathbf{B})$ and interpret the results. By (4)

Check these steps by multiplying everything out.

$$\begin{aligned} \|\mathbf{A} \times \mathbf{B}\|^2 &= (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2 \\ &= a_2^2b_3^2 + a_3^2b_3^2 + a_3^2b_1^2 + a_1^2b_2^2 + a_1^2b_2^2 + a_2^2b_1^2 - 2(a_2a_3b_2b_3 + a_1a_3b_1b_3 + a_1a_2b_1b_2) \\ &= (a_1^2 + a_2^2 + a_3^2)(b_1^2 + b_2^2 + b_3^2) - (a_1b_1 + a_2b_2 + a_3b_3)^2 \\ &= \|\mathbf{A}\|^2\|\mathbf{B}\|^2 - (\mathbf{A} \cdot \mathbf{B})^2 \\ &= \|\mathbf{A}\|^2\|\mathbf{B}\|^2 - (\|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta))^2 \\ &= \|\mathbf{A}\|^2\|\mathbf{B}\|^2(1 - \cos^2(\theta)) \\ &= \|\mathbf{A}\|^2\|\mathbf{B}\|^2\sin^2(\theta). \end{aligned}$$

θ is the angle betwe

Then

$$\|\mathbf{A} \times \mathbf{B}\| = \|\mathbf{A}\|\|\mathbf{B}\|\sin(\theta) \qquad \sin(\theta) \text{ is not negative since } 0 \le \theta \le \pi. \tag{14.3.8}$$

We then have

Let $\mathbf{A}$ and $\mathbf{B}$ be nonzero vectors and θ the angle between them. Then $\|\mathbf{A} \times \mathbf{B}\| = \|\mathbf{A}\|\|\mathbf{B}\|\sin(\theta)$.

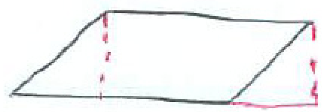


Figure 14.3.5: This figure

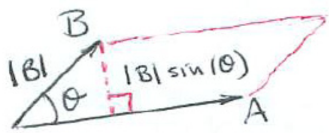


Figure 14.3.6:

In some texts the cross product is defined geometrically: It is the vector where length is the area of the parallelogram mentioned above and where direction is given by the right and rule. Then the author must obtain its formula in terms of components.

With the aid of this fact we now give a simple geometric meaning for the length of $\mathbf{A} \times \mathbf{B}$. A glance at the parallelogram spanned by $\mathbf{A}$ and $\mathbf{B}$ shows that it area is

$$\underbrace{\|\mathbf{A}\|}_{base}\ \underbrace{\|\mathbf{B}\|\sin(\theta)}_{height} = \text{area of parallelogram}$$

So now we have a simple geometric description of the length of $\mathbf{A} \times \mathbf{B}$.

The length of $\mathbf{A} \times \mathbf{B}$ is the area of the parallelogram spanned by $\mathbf{A}$ and $\mathbf{B}$.

EXAMPLE 3 Find the area of the parallelogram spanned by $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j}$ and $\mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j}$.

SOLUTION First write $\mathbf{A}$ as $a_1\mathbf{i} + a_2\mathbf{j} + 0\mathbf{k}$ and $b_\mathbf{i} + b_2\mathbf{j} + 0\mathbf{k}$. Then the area of this parallelogram is the length of $\mathbf{A} \times \mathbf{B}$. So we compute $\mathbf{A} \times \mathbf{B}$.

$$\mathbf{A} \times \mathbf{B} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & 0 \\ b_1 & b_2 & 0 \end{vmatrix} = (a_1b_2 - a_2b_1)\mathbf{k}.$$

The area is therefore $|a_1b_2 - a_2b_1|$. In other words, it is the absolute value of the determinant

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

◇ The next example is typical of the geometric applications of the cross product.

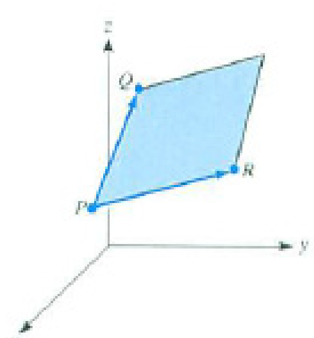


Figure 14.3.7:

EXAMPLE 4 Find a vector perpendicular to the plane determined by the three points $P = (1, 3, 2)$, $Q = (4, -1, 1)$, and $R = (3, 0, 2)$.

SOLUTION The vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ lie in a plane (see Figure 14.3.7). The vector $\mathbf{N} = \overrightarrow{PQ} \times \overrightarrow{PR}$ being perpendicular to both $\overrightarrow{PQ}$ and $\overrightarrow{PR}$, is perpendicular to the plane. Now, $\overrightarrow{PQ} = 3\mathbf{i} - 4\mathbf{j} - \mathbf{k}$ and $\overrightarrow{PR} = 2\mathbf{i} - 3\mathbf{j} + 0\mathbf{k}$.

Thus

$$\mathbf{N} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 3 & -4 & -1 \\ 2 & -3 & 0 \end{vmatrix} = -3\mathbf{i} - 2\mathbf{j} - \mathbf{k}.$$

◇

The Scalar Triple Product

The scalar $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$ is called the **scalar triple product**. It has an important geometric meaning. (The vector $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$ is also called the **vector triple product**.)

The vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ span a parallelepiped, as shown in Figure 14.3.8. The angle between $\mathbf{B}\times\mathbf{C}$ and $\mathbf{A}$ is θ (which could be greater than $\pi/2$). The area of the base of the parallelogram is $\|\mathbf{B}\times\mathbf{C}\|$. The height of the parallelepiped is $\|\mathbf{A}\||\cos(\theta)|$. Thus its volume is the absolute value of

$$\underbrace{|\mathbf{A}|\cos\theta}_{\text{height}}\ \underbrace{|\mathbf{B}\times\mathbf{C}|}_{\text{area of base}}.$$

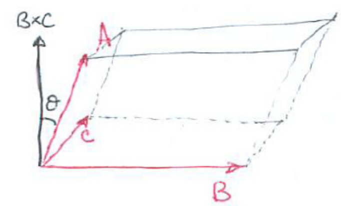


Figure 14.3.8:

This is the definition of the dot of product of $\mathbf{A}$ and $(\mathbf{B}\times\mathbf{C})$.

> $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C})$ is plus or minus the volume of the parallelepiped spanned by $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

The scalar triple product can also be expressed as a determinant. To see why, note that the dot product of $\mathbf{A}$ and $\mathbf{B}\times\mathbf{C}$ is

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = a_1\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3\end{vmatrix} + a_2(-\begin{vmatrix} b_1 & b_3 \\ c_1 & c_3\end{vmatrix}) + a_3\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2\end{vmatrix}. \quad (14.3.9)$$

Comparison of Dot Product and Vector Product

$\mathbf{A}\cdot\mathbf{B}$	$\mathbf{A}\times\mathbf{B}$
$\mathbf{B}\cdot\mathbf{B}=\mathbf{B}\cdot\mathbf{A}$	$\mathbf{A}\times\mathbf{B}=-\mathbf{B}\times\mathbf{A}$
$\lvert\mathbf{A}\cdot\mathbf{B}\rvert = \|\mathbf{A}\|\|\mathbf{B}\|\lvert\cos(\theta)\rvert$	$\|\mathbf{A}\times\mathbf{B}\| = \|\mathbf{A}\|\|\mathbf{B}\|\sin(\theta)$
$\mathbf{A}\cdot\mathbf{B}=0$ is a test for perpendicularity	$\mathbf{A}\times\mathbf{B}=\mathbf{0}$ is a test for parallel vectors
formula in components involves a_ib_i (same indices)	formula in components involves a_ib_j (unequal indices)

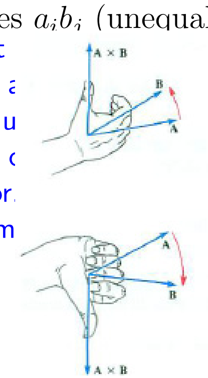


Figure 14.3.9:

Equation (14.3.9) can now be recognized the determinant of a matrix of order 3:

$$\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3\end{vmatrix}$$

So this determinant is plus or minus the volume of the parallelepiped spanned by $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

This should not be a surprise. As Example 3 showed, the determinant $\begin{vmatrix} a_1 & a_2 \\ v_1 & b_2\end{vmatrix}$ is plus or minus the area of the parallelogram spanned by the vectors $\langle a_1, a_2\rangle$ and $\langle b_1, b_2\rangle$.

Summary

We constructed a vector **C** perpendicular to vectors **A** and **B** by demanding that $\mathbf{C}\cdot\mathbf{A}=0$ and $\mathbf{C}\cdot\mathbf{B}=0$. A convenient formula for such a vector

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

It is denoted $\mathbf{A}\times\mathbf{B}$ and called the vector product or cross product of **A** and **B**. It also may be described as the vector whose length is the area of the parallelogram spanned by **A** and **B** and whose direction is given by the right-hand rule (the finger curling from **A** and **B**). These are some of its properties:

Item 5 appeared in finding the length of $\mathbf{A}\times\mathbf{B}$. It will be used in the next chapters.

1. $\mathbf{A}\times\mathbf{B} = -(\mathbf{B}\times\mathbf{A})$ (anticommunitive)
2. $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})$ *is not* usually equal to $(\mathbf{A}\times\mathbf{B})\times\mathbf{C})$ (not associative)
3. $\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = (\mathbf{C}\cdot\mathbf{A})\mathbf{B} - (\mathbf{B}\cdot\mathbf{A})\mathbf{C}$ (See Exercise 17.)
4. $(\mathbf{A}\times\mathbf{B})\cdot(\mathbf{A}\times\mathbf{B}) = (\mathbf{A}\cdot\mathbf{A})(\mathbf{B}\cdot\mathbf{B}) - (\mathbf{A}\cdot\mathbf{B})(\mathbf{A}\cdot\mathbf{B})$
5. $\mathbf{A}\cdot(\mathbf{B}\times\mathbf{C}) = \pm$ volume of parallelepiped spanned by **A**, **B**, and **C**.

EXERCISES for Section 14.3 *Key:* R–routine, M–moderate, C–challenging

SHERMAN: Move some exercises about lines, planes, etc. to Section 14.4 or to Chapter Summary for Chapter 14.

In Exercises 1 to 4 compute and sketch $\mathbf{A} \cdot \mathbf{B}$.

1.[R] $\mathbf{A} = \mathbf{k}$, $\mathbf{B} = \mathbf{j}$

2.[R] $\mathbf{A} = \mathbf{i} + \mathbf{j}$, $\mathbf{B} = \mathbf{i} - \mathbf{j}$

3.[R] $\mathbf{A} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{B} = \mathbf{i} + \mathbf{j}$

4.[R] $\mathbf{A} = \mathbf{k}$, $\mathbf{B} = \mathbf{i} + \mathbf{j}$

In Exercises 5 and 6, find $\mathbf{A} \times \mathbf{B}$ and check that it is perpendicular to both $\mathbf{A}$ and $\mathbf{B}$.

5.[R] $\mathbf{A} = 2\mathbf{i} - 3\mathbf{j} + \mathbf{k}$, $\mathbf{B} = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$

6.[R] $\mathbf{A} = \mathbf{i} - \mathbf{j}$, $\mathbf{B} = \mathbf{j} + 4\mathbf{k}$

In Exercises 7 to 10 use the cross product to find the area of each region.

7.[R] The parallelogram three of whose vertices are $(0, 0, 0)$, $(1, 5, 4)$, and $(2, -1, 3)$.

8.[R] The parallelogram three of whose vertices are $(1, 2, -1)$, $(2, 1, 4)$, and $(3, 5, 2)$.

9.[R] The triangle two of whose sides are $\mathbf{i} + \mathbf{j}$ and $3\mathbf{i} - \mathbf{j}$.

10.[R] The triangle two of whose sides are $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$ and $2\mathbf{i} - \mathbf{j} + 2\mathbf{k}$.

In Exercises 11 to 14 find the volumes of the parallelepipeds spanned by the given vectors.

11.[R] $\langle 2, 1, 3\rangle$, $\langle 3, -1, 2\rangle$, $\langle 4, 0, 3\rangle$

12.[R] $3\mathbf{i} + 4\mathbf{j} + 3\mathbf{k}$, $2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$, $\mathbf{i} - \mathbf{j} - \mathbf{k}$.

13.[R] $\overrightarrow{PQ}$, $\overrightarrow{PR}$, $\overrightarrow{PS}$, where $P = (1, 1, 1)$, $Q = (2, 1, -2)$, $R = (3, 5, 2)$, and $S = (1, -1, 2)$.

14.[R] $\overrightarrow{PQ}$, $\overrightarrow{PR}$, $\overrightarrow{PS}$, where $P = (0, 0, 0)$, $Q = (3, 3, 2)$, $R = (1, 4, -1)$, and $S = (1, 2, 3)$.

15.[R] Evaluate $\mathbf{A} \cdot (\mathbf{A} \times \mathbf{B})$.

16.[R] Prove that $\mathbf{B} \times \mathbf{A} = -(\mathbf{A} \times \mathbf{B})$ in two ways:

(a) using the algebraic definition of the cross product;

(b) using the geometric description of the cross product.

17.[R] Show that if $\mathbf{B} = c\mathbf{A}$, then $\mathbf{A} \times \mathbf{B} = \mathbf{0}$:

(a) using the algebraic definition of the cross product;

(b) using the geometric description of the cross product.

18.[M] Show that the points $(0,0,0)$, (x_1,y_1,z_1), (x_2,y_2,z_2) and (x_3,y_3,z_3) lie on a plane if and only if

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = 0.$$

19.[M]

(a) If $\mathbf{B}$ is parallel to $\mathbf{C}$, is $\mathbf{A}\times\mathbf{B}$ parallel to $\mathbf{A}\times\mathbf{C}$?

(b) If $\mathbf{B}$ is perpendicular to $\mathbf{C}$, is $\mathbf{A}\times\mathbf{B}$ perpendicular to $\mathbf{A}\times\mathbf{C}$?

20.[M] Let $\mathbf{A}$ be a nonzero vector. If $\mathbf{A}\times\mathbf{B}=\mathbf{0}$ and $\mathbf{A}\cdot\mathbf{B}=0$, must $\mathbf{B}=\mathbf{0}$?

21.[R] Show that $\mathbf{A}\times(\mathbf{A}\times\mathbf{B}) = (\mathbf{A}\cdot\mathbf{B})\mathbf{A} - (\mathbf{A}\cdot\mathbf{A})\mathbf{B}$.

22.[R] Show that $(\mathbf{A}\times\mathbf{B})\times(\mathbf{C}\times\mathbf{D}) = ((\mathbf{A}\times\mathbf{B})\cdot\mathbf{D})\mathbf{C} - ((\mathbf{A}\times\mathbf{B})\cdot\mathbf{C})\mathbf{D}$. HINT: Think of $\mathbf{A}\times\mathbf{B}$ as a single vector, $\mathbf{E}$.

23.[M]

(a) Give an example of a vector perpendicular to the vector $3\mathbf{i}-\mathbf{j}+\mathbf{k}$.

(b) Give an example of a unit vector perpendicular to the vector $3\mathbf{i}-\mathbf{j}+\mathbf{k}$.

24.[M] Let $\mathbf{u}$ be a unit vector and $\mathbf{B}$ be a vector. What happens as you keep "crossing by $\mathbf{u}$," that is, as you form the sequence $\mathbf{B}$, $\mathbf{u}\times\mathbf{B}$, $\mathbf{u}\times(\mathbf{u}\times\mathbf{B})$ and so on? (See Exercise 21)

25.[C] *(Crystallography)* A crystal is described by three vectors $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

They span a "fundamental" parallelepiped, whose copies fill out the crystal lattice. (See Figure 14.3.10.) The atoms are at the corners. In order to study the diffraction of x-rays and light through a crystal, crystallographers work with the "reciprocal lattice," as follows. Its fundamental parallelepiped is spanned by three vectors, $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$. The vector $\mathbf{k}_1$ is perpendicular to the parallelogram spanned by $\mathbf{v}_2$ and $\mathbf{v}_3$ and has a length equal to the reciprocal of the distance between that parallelogram and the opposite parallelogram of the fundamental parallelepiped. The vectors $\mathbf{k}_2$ and $\mathbf{k}_3$ are defined similarly in terms of the other four faces of the fundamental parallelepiped.

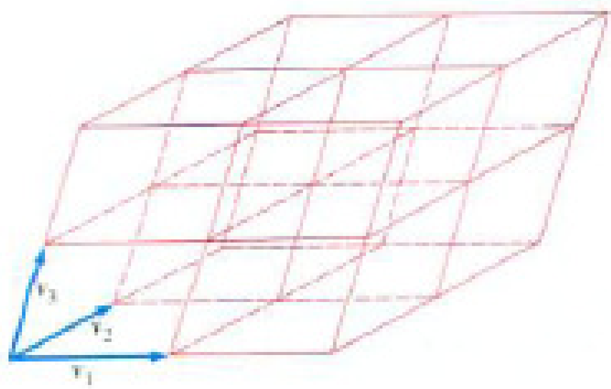


Figure 14.3.10:

(a) Show that $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$ may be chosen to be

$$\mathbf{k}_1 = \frac{\mathbf{v}_2 \times \mathbf{v}_3}{\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)}, \qquad \mathbf{k}_2 = \frac{\mathbf{v}_3 \times \mathbf{v}_1}{\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)}, \qquad \mathbf{k}_3 = \frac{\mathbf{v}_1 \times \mathbf{v}_2}{\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)}$$

(b) Show that the volume of the fundamental parallelopiped determined by $\mathbf{k}_1$, $\mathbf{k}_2$, and $\mathbf{k}_3$ is the reciprocal of the volume of the one determined by $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$.

(c) Is the reciprocal of the reciprocal lattice the original lattice? For instance, is

$$\mathbf{v}_1 = \frac{\mathbf{k}_2 \times \mathbf{k}_3}{\mathbf{k}_1 \cdot (\mathbf{k}_2 \times \mathbf{k}_3)}?$$

26.[M] Let $\mathbf{B}$ and $\mathbf{C}$ be nonzero, nonparallel vectors and $\mathbf{A}$ a vector that is perpendicular neither to $\mathbf{B}$ nor $\mathbf{C}$.

(a) Why are their scalars x and y such that

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = x\mathbf{B} + y\mathbf{C}?$$

(b) Why is $0 = x(\mathbf{A} \cdot \mathbf{B}) + y(\mathbf{A} \cdot \mathbf{C})$?

(c) Using (b), show that there is a scalar z such that

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = z[(\mathbf{A} \cdot \mathbf{C})\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\mathbf{C}].$$

(d) It would be nice if there were a simple geometric way to show that z is a constant and equals 1. Of course we could show that $z = 1$ by writing **A**, **B**, and **C** in components and grinding out a tedious calculation. But that would hardly be instructive. Can you figure out why $z = 1$ in a simpler way?

(This identity, known as Jacobi's Identity, will come in handy in Chapter 18 when dealing with electric currents and magnetic fields.)

OMIT? In this section $\mathbf{A} \times \mathbf{B}$ was defined in terms of components, and then its geometric description was obtained. This is the opposite of the way we dealt with the dot product. Exercises 27 to 29 outline a different approach to the cross product. We define $\mathbf{A} \times \mathbf{B}$ as follows. If **A** or **B** is **0** or if **A** is parallel to **B**, we define $\mathbf{A} \times \mathbf{B}$ to be **0**. Otherwise, $\mathbf{A} \times \mathbf{B}$ is the vector whose direction is given by the right-hand rule.

27.[R] Let **A** be a nonzero vector and **B** be a vector. Let $\mathbf{B}_1$ be the projection of **B** on a plane perpendicular to **A**. Let $\mathbf{B}_2$ be obtained by rotating $\mathbf{B}_1$ 90° in the direction given by the right-hand rule with thumb pointing in the same direction as **A**

(a) Show that $\mathbf{A} \times \mathbf{B} = \mathbf{A} \times \mathbf{B}_1$. (Draw a clear diagram.)

(b) Show that $\mathbf{A} \times \mathbf{B} = \|\mathbf{A}\|\mathbf{B}_2$.

28.[R] Using Exercise 27(b), show that for **A** not **0**, $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$. HINT: Draw a large, clear picture.

29.[R]

(a) From the distributive law $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$, and the fact that $\mathbf{D} \times \mathbf{E} = -\mathbf{E} \times \mathbf{D}$, deduce the distributive law $(\mathbf{B} + \mathbf{C}) \times \mathbf{A} = \mathbf{B} \times \mathbf{A} + \mathbf{C} \times \mathbf{A}$.

(b) From the distributive law $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$, deduce that $\mathbf{A} \times (\mathbf{B} + \mathbf{C} + \mathbf{D}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C} + \mathbf{A} \times \mathbf{D}$. HINT: Think of $\mathbf{B} + \mathbf{C}$ as a single vector **E**.

30.[R] Check that $-13\mathbf{i} + 7\mathbf{j} + 11\mathbf{k}$ in Example 1 is perpendicular to **A** and to **B**.

31.[R] Show, using (14.3.7), that $\mathbf{0} \times \mathbf{B} = \mathbf{0}$.

32.[R] Show, using (14.3.7), that $\mathbf{B} \times \mathbf{A} = -\mathbf{A} \times \mathbf{B}$.

33.[M] Using (14.3.7), show that if $\mathbf{B}$ is parallel to $\mathbf{A}$, then $\mathbf{A} \times \mathbf{B} = \mathbf{0}$. Suggestion: If $\mathbf{B}$ is parallel to $\mathbf{A}$, there is a scalar t such that $\mathbf{B} = t\mathbf{A}$.

34.[M] In finding $|\mathbf{A} \times \mathbf{B}|^2$ we stated that

$$a_2^2 b_3^2 + a_3^2 b_2^2 + a_3^2 b_1^2 + a_1^2 b_3^2 + a_1^2 b_2^2 + a_2^2 b_1^2 - 2(a_2 a_3 b_2 b_3 + a_1 a_3 b_1 b_3 + a_1 a_2 b_1 b_2)$$

equals

$$a_1^2 + a_2^2 + a_3^2 b_1^2 + b_2^2 + b_3^2 - (a_1 b_1 + a_2 b_2 + a_3 b_3).$$

Take nothing as faith. Check that the claim is correct.

35.[C] We showed that the direction of $\mathbf{i} \times \mathbf{j}$ is given by the right hand rule. Then we said that the right hand rule hold for any non-zero vector $\mathbf{A}$ and $\mathbf{B}$. Why is such a leap justified? HINT: Imagine moving a gradually changing pair of vectors through space, starting with $\mathbf{i}$ and $\mathbf{j}$ and ending with the pair $\mathbf{A}$ and $\mathbf{B}$.

36.[C]

(a) Thinking in terms of parallelograms, explain why $\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$ is $+$ or $-$ $\mathbf{B} \cdot (\mathbf{C} \times \mathbf{A})$.

(b) Using properties of 3 by 3 determents, decide which it's $+$ or $-$.

37.[C] In some expositions of the cross product, $\mathbf{a} \times \mathbf{b}$ is simply defined as the determinant of a matrix of order 3. If we start with this definition, use a property of determents to show that $\mathbf{a} \times \mathbf{b}$ is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. (This approach bypasses the need to consider simultaneous equations. On the other hand, it may appear unmotivated.)

38.[M]

(a) How could you use cross products to produce a vector perpendicular to $2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$? Give an example.

(b) How could you use cross product to produce two vectors perpendicular to $2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$ and to each other? Give an example.

39.[R] Use the exhibited values for y and z when solving equations (14.3.3) and (14.3.4). Substitute these values into (14.3.1) and solve for x.

40.[R] By carrying out the necessary calculations, show that $\mathbf{A} \times (\mathbf{B} + \mathbf{C}) = \mathbf{A} \times \mathbf{B} + \mathbf{A} \times \mathbf{C}$. If you wish, you may use properties of determinants.

41.[M] Let $\mathbf{A}$ and $\mathbf{B}$ be non zero, nonparallel vectors. Show that $\mathbf{A} \times (\mathbf{A} \times \mathbf{B})$ is never equal to $(\mathbf{A} \times \mathbf{A} \times \mathbf{B})$. This shows that the cross product is not associative. You **cannot omit** the parentheses in $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$.

14.4 Lines, Planes and Components

This section uses the dot product and cross product to deal with lines, planes and projections ("shadows") of a vector or a line or on a plane.

Equation of a Plane

We find an equation of the plane through the point $P_0 = (x_0, y_0, z_0)$ and perpendicular to the vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$, shown in Figure 14.4.1.

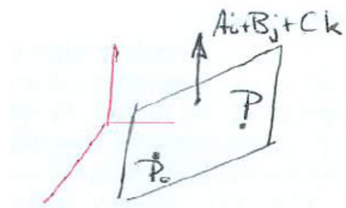


Figure 14.4.1:

Let $P = (x, y, z)$ be any point on the plane. The vector $\overrightarrow{P_0P}$ is perpendicular to $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$. (Imagine sliding it so that P_0 coincides with the tail of $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$.) Thus

$$(A\mathbf{i} + B\mathbf{j} + C\mathbf{k}) \cdot ((x - x_0)\mathbf{i} + (y - y_0)\mathbf{j} + (z - z_0)\mathbf{k}) = 0.$$

So

$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0. \qquad (14.4.1)$$

In (14.4.1) we have an equation for the plane. The vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ is called a **normal** to the plane.

EXAMPLE 1 Find an equation of the plane through $(2, -3, 4)$ and perpendicular to $\mathbf{i} + 2\mathbf{j} + 3\mathbf{k}$.
SOLUTION An equation for the plane is

$$1(x - 2) + 2(x - (-3)) + 3(z - 4) = 0$$

which simplifies to

$$x + 2y + 3z - 8 = 0$$

◇

The graph of an equation of the form $Ax + By + Cz + D = 0$, where not all of A, B, and C are 0 is a plane perpendicular to the vector $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$. To show this, first pick any point (x_0, y_0, z_0) that satisfies the equation: $Ax_0 + Bv_0 + C_0 + D = 0$. Subtracting this from the original equation gives

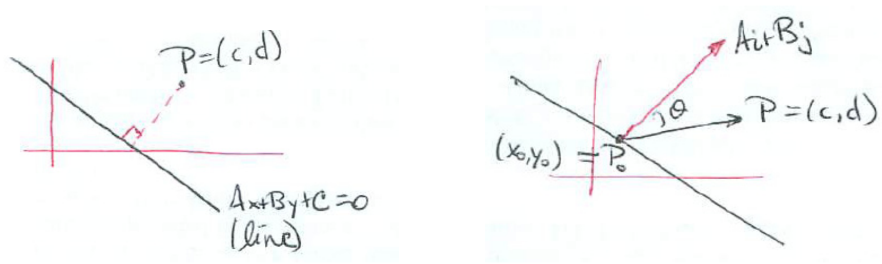
$$A(x - x_0) + B(y - y_0) + C(z - z_0) = 0,$$

which is an equation of the plane through (x_0, y_0, z_0) perpendicular to $A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$.

Similarly, we have

An equation for the line through (x_0, y_0) and perpendicular to the vector $A\mathbf{i} + B\mathbf{j}$ is $A(x - x_0) + B(y - y_0) = 0$.

Distance From a Point to the Line $Ax + By + C = 0$ or Plane $Ax + By + Cz + D = 0$



(a) (b)

Figure 14.4.2:

Let us find the distance from $P = (c, d)$ to the line whose equation is $Ax + By + C = 0$, shown in Figure 14.4.2(a).

Pick any point $P_0 = (x_0, y_0)$ on the line and place $A\mathbf{i} + B\mathbf{j}$ with its tail at P_0, as in Figure 14.4.2(b).

Let θ be the angle between $\overrightarrow{P_0P}$ and $A\mathbf{i} + B\mathbf{j}$. Then the distance from P to the line is

$\cos(\theta)$ could be negative

$$\begin{aligned}\|\overrightarrow{P_0P}\||\cos(\theta)| &= \|\overrightarrow{P_0P}\|\frac{(A\mathbf{i} + B\mathbf{j}) \cdot ((c - x_0)\mathbf{i} + (d - y_0)\mathbf{j})}{\|P_0P\|\|A\mathbf{i} + B\mathbf{j}\|} \\ &= \frac{A(c - x_0) + B(d - y_0)}{\sqrt{A^2 + B^2}} \\ &= \frac{Ac + Bd - (Ax_0 + By_0)}{\sqrt{A^2 + B^2}}.\end{aligned}$$

Since $Ax_0 + By_0 + C = 0$, we have

Distance from (c, d) to the line $Ax + By + C = 0$ is

$$\frac{|Ac + Bd + C|}{\sqrt{A^2 + B^2}}$$

In short, to find that distance simply substitute the coordinates of the point (c, d) into the expression $Ax + By + C$ and divide by $\sqrt{A^2 + B^2}$ and take its absolute value.

EXAMPLE 2 How far is the point $(1, 3)$ from the line $2x - 4y = 5$?
SOLUTION First, write the equation in the form $2x - 4y - 5 = 0$. Then the

distance is

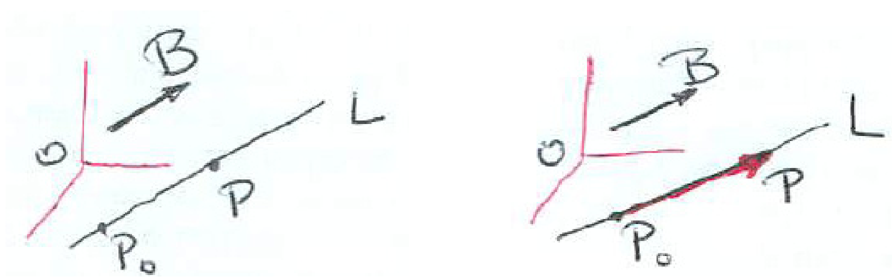
$$\frac{|2(1) - 4(3) - 5|}{\sqrt{2^2 + 4^2}} = \frac{|-15|}{\sqrt{20}} = \frac{3\sqrt{5}}{2}.$$

$\diamond$ A similar result holds for the distance from a point $P = (x_0, y_0, z_0)$ to a plane:

> The distance from (x_0, y_0, z_0) to the plane $Ax + By + C = 0$ is
> $$\frac{|Ax_0 + By_0 + C_0 + D|}{\sqrt{A^2 + B^2 + C^2}}$$

Using Vectors to Parameterize a Line

Let L be the line through the point $P_0 = (x_0, y_0, z_0)$ parallel to the vector $\mathbf{B}$, shown in Figure 14.4.3(a).



(a) (b)

Figure 14.4.3:

Let P be any point on L. Then the vector $\overrightarrow{P_0P}$ which is parallel to $\mathbf{B}$, is of the form $t\mathbf{B}$ for some scalar t. See Figure 14.4.3(b).

The $\overrightarrow{OP} = \overrightarrow{OP_0} + \overrightarrow{P_0P} = \overrightarrow{OP_0} + t\mathbf{B}$. As t varies the vector from 0 to P varies, thus parameterizing the line L.

EXAMPLE 3 The line L passes through the point $(1, 1, 2)$ and is parallel to the vector $3\mathbf{i} + 4\mathbf{j} + 5\mathbf{k}$. Use this information to parameterize the line.
SOLUTION In this case $\overrightarrow{OP_0} = \mathbf{i} + \mathbf{j} + 2\mathbf{k}$ and $B = 3\mathbf{i} + 4\mathbf{j} + 5\mathbf{k}$. Thus

$$\begin{aligned} \overrightarrow{OP} &= \mathbf{i} + \mathbf{j} + 2\mathbf{k} + f(3\mathbf{i} + 4\mathbf{j} + 5\mathbf{k}) \\ &= (3t + 1)\mathbf{i} + (4t + 1)\mathbf{j} + (5t + 2)\mathbf{k}. \end{aligned}$$

One vector equation does the work of three scalar equation.

If P is the point (x, y, z), then $\overrightarrow{OP}$ is the vector $x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Thus

$$\begin{cases} x &= 3t + 1 \\ y &= 4t + 1 \\ z &= 5t + 2. \end{cases}$$

◇

Describing the Direction of Vectors and Lines

The direction of a vector in the plane is described by a single angle, the angle it makes with the positive x-axis. The direction of a vector in space involves three angles, two of which almost determine the third.

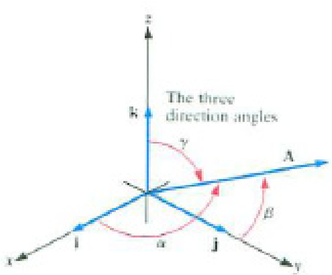


Figure 14.4.4:

DEFINITION *(Direction of a vector.)* Let **A** be a nonzero vector in space. The angle between

A and **i** is denoted α,

A and **j** is denoted β,

A and **k** is denoted γ.

The angles α, β and γ are called the **direction angles of A**. (See Figure 14.4.4.)

DEFINITION *(Direction cosines of a vector)* The **direction cosines** of a vector are the cosines of its direction angles, $\cos(\alpha)$, $\cos(\beta)$, and $\cos(\gamma)$.

EXAMPLE 4 The angle between a vector **A** and **k** is $\pi/6$. Find γ and $\cos(\gamma)$ for

1. **A**,

2. $-$**A**.

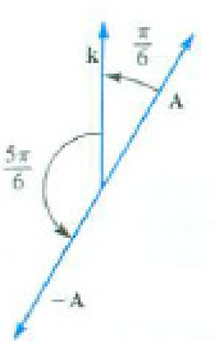


Figure 14.4.5:

SOLUTION

1. By definition, the direction angle γ for **A** is $\pi/6$. It follows that $\cos(\gamma) = \cos(\pi/6) = \sqrt{3/2}$.

2. To find γ and $\cos(\gamma)$ for $-$**A**, we draw Figure 14.4.5. For $-$**A**, $\gamma = 5\pi/6$ and $\cos(\gamma) = \cos(5\pi/6) = -\sqrt{3/2}$.

◇

As Example 4 illustrates, if the direction angles of **A** are α, β, γ, then the direction angles of $-\mathbf{A}$ are $\pi-\alpha$, $\pi-\beta$, and $\pi-\gamma$. The direction cosines of $-\mathbf{A}$ are the negatives of the direction cosines of **A**.

The three direction angles are not independent of each other, as is shown by the next theorem. Two of them determine the third up to sign.

Theorem 14.4.1. *If α, β, γ are the direction angles of the vector* **A**, *then* $\cos^2(\alpha)+\cos^2(\beta)+\cos^2(\gamma)=1$.
Proof

It is no loss of generality to assume that **A** *is a unit vector. Its component on the y-axis, for instance, is* $\cos(\beta)$, *as the right triangle OPQ in Figure 14.4.6 shows.* **A** *lies along the hypotenuse.*

Since **A** *is a unit vector,* $|\mathbf{A}|^2=1$, *and we have* $\cos^2(\alpha)+\cos^2(\beta)+\cos^2(\gamma)=1^2=1$. •

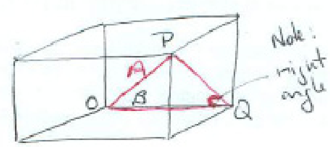


Figure 14.4.6:

EXAMPLE 5 The vector **A** makes an angle of 60° with the x and y axes. What angle does it make with the z-axis?
SOLUTION Here $\alpha=60°$ and $\beta=60°$; hence

$$\cos(\alpha)=\frac{1}{2} \quad \text{and} \quad \cos(\beta)=\frac{1}{2}.$$

Since

$$\cos^2(\alpha)+\cos^2(\beta)+\cos^2(\gamma)=1,$$

it follows that

$$\begin{aligned}(\frac{1}{2})^2+(\frac{1}{2})^2+\cos^2(\gamma) &= 1,\\ \cos^2(\gamma) &= \frac{1}{2}.\end{aligned}$$

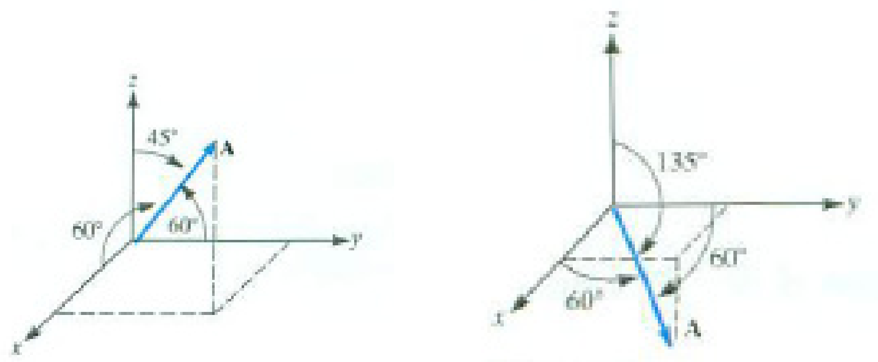
Thus

$$\cos(\gamma)=\frac{\sqrt{2}}{2} \quad \text{or} \quad \cos(\gamma)=-\frac{\sqrt{2}}{2}.$$

Hence

$$\gamma=45° \quad \text{or} \quad \gamma=135°.$$

Figures 14.4.7(a) and (b) show the two possibilities for **A**. ◇



(a) (b)

Figure 14.4.7:

(a) (b) (c)

Figure 14.4.8:

Dot Products and Flow

Let the vector **v** whose magnitude is v describe the velocity of a river, as in Figure 14.4.8(a). Place an imaginary horizontal stick of length L in the water. The amount of water crossing the stick depends on the position of the stick. If the stick is parallel to **v**, no water crosses the stick. If the stick is perpendicular to **v** water crosses it. The question then arises, "How does the angle at which we place the stick affect the amount of water that crosses in a given time?

SHERMAN: what word did you intend to go between imaginary and stick? My best guess is horizontal, but this makes no sense to me.

To answer this question, we begin by introducing a unit vector **n** perpendicular to the stick, and record its position, as in Figure 14.4.8(b). Let the angle between **n** and **v** be θ.

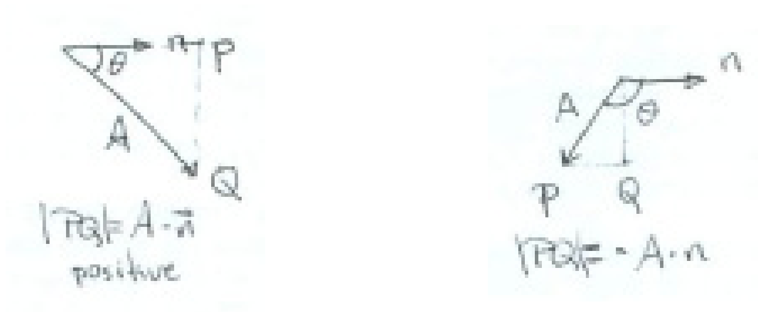
The amount of water that crosses the stick during time Δt is proportional to the area of the parallelogram in Figure 14.4.8(c). The base of the parallelogram has length $v\Delta t$ (speed times time). The height is $L\cos(\theta)$. The area of the parallelogram is therefore

$$vL\cos(\theta)$$

.

But $vL\cos(\theta)$ is equal to $\mathbf{v}\cdot\mathbf{n}$. So $\mathbf{v}\cdot\mathbf{n}$ measures the tendency of water to cross the stick.

As a check, when the stick is parallel to **v**, $\theta = \pi/2$ and $\cos(\pi/2) = 0$. Then $\mathbf{v}\cdot\mathbf{n} = 0$ and no water crosses the stick. When the stick is perpendicular to **v**, $\theta = 0$, and $\mathbf{v}\cdot\mathbf{n} = v$. For any angle $\theta < \pi/2$, $\mathbf{v}\cdot\mathbf{n} = v\cos(\theta)$ which is less than v. For any unit vector **n** and vector **A** the scalar $\mathbf{A}\cdot\mathbf{n}$ is called the



(a) (b)

Figure 14.4.9:

scalar component of A along n. It equals $\|vA\|\cos(\theta)$, where θ is the angle between **A** and **n**. It can be positive or negative, as shown in Figure 14.4.9.

EXAMPLE 6 When a stick is perpendicular to **v**, water crosses it at the rate of 100 cubic feet per second. When the stick is placed at an angle of $\pi/6$ to **v** at what rate does water cross it?

SOLUTION Figure 14.4.10 shows the position of the stick PQ.

The angle between the normal to the stick, **n**, and **v** is $\pi/2 - \pi/6 = \pi/3$. Let x be the rate at which the water crosses the stick. Since the rate of flow across the stick is proportional to $v\cos(\theta)$, where θ is the angle between the normal **n** and **v**, we have

$$\frac{100}{v\cos(0)} = \frac{x}{v\cos(\pi/3)}.$$

this tells us that

$$\frac{100}{v} = \frac{x}{(v)(1/2)},$$

have $x = 50$. The flow is half the maximum possible. ◇

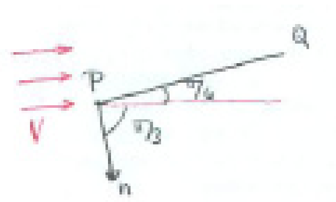


Figure 14.4.10:

Summary

We used the dot product to obtain an equation of a plane (or line in the xy plane) and to find the distance from a point to a line or plane. We also showed how to parameterize a line with the aid of a vector parallel to the line.

Direction angles and cosines of a vector were defined. Finally, we showed how the dot product describes the rate of flow across a line segment, a concept that will be needed in Chapters 17 and 18, where we deal with flows across curves and surfaces.

EXERCISES for Section 14.4 *Key:* R–routine, M–moderate, C–challenging

SHERMAN: These exercises need to be reordered. Some could move to the Chapter Summary.

In each of Exercises 1 to 4 find an equation of the line through the given point and perpendicular to the given vector.

1.[R] $(2, 3)$, $4\mathbf{i} + 5\mathbf{j}$

2.[R] $(1, 0)$, $2\mathbf{i} - \mathbf{j}$

3.[R] $(4, 5)$, $1\mathbf{i} + 3\mathbf{j}$

4.[R] $(2, -1)$, $\mathbf{i} + 3\mathbf{j}$

In each of Exercises 5 to 8 find a vector in the xy plane that is perpendicular to the given line.

5.[R] $2x - 3y + 8 = 0$

6.[R] $\pi x - \sqrt{2}y = 7$

7.[R] $y = 3x + 7$

8.[R] $2(x - 1) + 5(y + 2) = 0$

9.[M] Find an equation of the plane through $(1, 2, 3)$ that contains the line given parametrically as $\overrightarrow{OP} = 2\mathbf{i} - \mathbf{j} + 3\mathbf{k} + t(3\mathbf{i} + 2\mathbf{j} + \mathbf{k})$.

10.[M] Is the point $(21, -3, 28)$ on the line given parametrically as $\overrightarrow{OP} = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k} + t(4\mathbf{i} - \mathbf{j} + 5\mathbf{k})$?

11.[M] A line segment has projections of lengths a, b, and c on the coordinates axes. What, if anything, can be said about its length, L?

12.[C] A line segment has projections of lengths d, e, and f on the coordinates planes. What, if anything, can be said about its length, L?

13.[C] Explain why the projection of a circle is an ellipse. HINT: Set up coordinate systems in the plane of the circle and in the plane of its shadow (which might as well be taken to be the xy plane). Choose the axes for these coordinate systems to be as convenient as possible. Then express the equation of the shadow in terms of x and y by utilizing the equation of the circle.

14.[R] Find a vector perpendicular to the plane through $(2, 1, 3)$, $(4, 5, 1)$ and $(-2, 2, 3)$.

15.[R] How far is the point $(1, 2, 2)$ from the plane through $(0, 0, 0)$, $(3, 5, -2)$, and $(2, -1, 3)$?

16.[R] How far is the point $(1, 2, 3)$ from the line through $(-2, -1, 3)$, and $(4, 1, 2)$?

17.[R] Find the parametric equations of the line through $(1, 1, 2)$ and perpendicular to the plane $3x - y + z = 6$.

18.[R] How far apart are the lines whose vector equations are $2\mathbf{i}+4\mathbf{j}+\mathbf{k}+t(\mathbf{i}+\mathbf{j}+\mathbf{k})$ and $\mathbf{i}+3\mathbf{j}+2\mathbf{k}+s(2\mathbf{i}-\mathbf{j}-\mathbf{k})$?

19.[R] Find the point on the line through $(1, 2, 1)$ and $(2, -1, 3)$ that is closest to the line through $(3, 0, 3)$ and parallel to the vector $\mathbf{i}+2\mathbf{j}+5\mathbf{k}$.

20.[R]

(a) Describe how you would find an equation for the plane through points $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, and $P_3 = (x_3, y_3, z_3)$?

(b) Find an equation for the plane through $(2, 2, 1)$, $(0, 1, 5)$ and $(2, -1, 0)$.

21.[R]

(a) Describe how you would decide whether the line through $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$, is parallel to the line through $P_3 = (x_3, y_3, z_3)$ and $P_4 = (x_4, y_4, z_4)$?

(b) Is the line through $(1, 2, -3)$ and $(5, 9, 4)$ parallel to the line through $(-1, -1, 2)$ and $(1, 3, 5)$?

22.[R]

(a) Describe how you would decide whether the line through $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ is parallel to the plane $Ax + By + Cz + D = 0$?

(b) Is the line through $(1, -2, 3)$ and $(5, 3, 0)$ parallel to the plane $2x-y+z+3 = 0$?

23.[R]

(a) Describe how you would decide whether the line through P_1 and P_2 is parallel to the plane through Q_1, Q_2, and Q_3?

(b) Is the line through $(0, 0, 0)$ and $(1, 1, -1)$ parallel to the plane through $(1, 0, 1)$, $(2, 1, 0)$, and $(1, 3, 4)$?

24.[R]

(a) How would you decide whether the plane through P_1, P_2 and P_3 is parallel to the plane through Q_1, Q_2, and Q_3?

(b) Is the plane through $(1, 2, 3)$, $(4, 1, -1)$, and $(2, 0, 1)$ parallel to the plane through $(2, 3, 4)$, $(5, 2, 0)$, and $(3, 1, 2)$?

25.[M]

(a) How would you find the angle between the planes $A_1x + B_1y + C_1z + D_1 = 0$ and $A_2x + B_2y + C_2z + D_2 = 0$?

(b) Find the angle between $x - y - z - 1 = 0$ and $x + y + z + 2 = 0$.

26.[C] Assume that the planes $A_1x+B_1y+C_1z+D_1 = 0$ and $A_2x+B_2y+C_2z+D_2 = 0$ met in a line L.

(a) How would you find a vector parallel to L?

(b) How would you find a point on L?

(c) Find parametric equations for the line that is the intersection of the planes $2x - y + 3z + 4 = 0$ and $3x + 2y + 5z + 2 = 0$.

27.[C]

(a) How would you decide whether the four points $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, $P_3 = (x_3, y_3, z_3)$ and $P_4 = (x_4, y_4, z_4)$ lie in a plane?

(b) Do the points $(1, 2, 3)$, $(4, 1, -5)$, $(2, 1, 6)$, and $(3, 5, 3)$ lie in a plane?

28.[C] What is the angle between the line through $(1, 2, 1)$ and $(-1, 3, 0)$ and the plane $x + y - 2z = 0$?

29.[M]

(a) If you know the coordinates of point P and parametric equations of line L, how would you find an equation of the plane that contains P and L? (Assume P is not on L.)

(b) Find an equation for the plane through $(1, 1, 1)$ that contains the line

$$\begin{cases} x &= 2 + t \\ y &= 3 - t \\ z &= 4 + 2t. \end{cases}$$

30.[R]

(a) How many unit vectors are perpendicular to the plane $Ax+By+Cz+D=0$?

(b) How would you find one of them?

(c) Find a unit vector perpendicular to the plane $3x-2y+4z+6=0$.

31.[R]

(a) How would you go about producing a specific point on the plane $Ax+By+Cz+D=0$?

(b) Give the coordinates of a specific point that lies on the plane $3x-y+z+10=0$.

32.[R]

(a) How would you go about producing a specific point that lies on both planes $A_1x+B_1y+C_1z+D_1=0$ and $A_2x+B_2y+C_2z+D_2=0$?

(b) Find a point that lies on both planes $3x+z+2=0$ and $x-y-z+5=0$.

33.[C] The planes $A_1x+B_1y+C_1z+D_1=0$ and $A_2x+B_2y+C_2z+D_2=0$ intersect in a line L. Find the direction cosines of a vector parallel to L.

34.[R]

(a) Let $\mathbf{A}$ and $\mathbf{B}$ be vectors in space. How would you find the area of the parallelogram they span?

(b) Find the area of the parallelogram spanned by $(2,3,1)$ and $(4,-1,5)$.

35.[C] How far is the point $(2,1,3)$ from the line through $(1,5,2)$ and $(2,3,4)$?

36.[C] How far is the point P from the line through Q and R.

37.[C] How far apart are the lines given parametrically as $2\mathbf{i}+\mathbf{j}-3\mathbf{k}+t(3\mathbf{i}-5\mathbf{j}+2\mathbf{k})$ and $3\mathbf{i}+\mathbf{j}+5\mathbf{k}+s(2\mathbf{i}+6\mathbf{j}+7\mathbf{k})$? (We use different letters, s and t, for the parameters because they are independent of each other.)

38.[M]

(a) Sketch four points P, Q, R, and S, not all in one plane, such that $\overrightarrow{PQ}$ and $\overrightarrow{RS}$ are not parallel. Explain way there is a unique pair of parallel planes one of which contains P and Q and are of which contains R and S.

(b) Express a normal vector to these planes in terms of P, Q, R, and S.

39.[M] Find an equation for the plane through P_1 that is parallel to the non-parallel segments P_2P_3 and P_4P_5.

40.[C]

(a) Using properties of determinents, show that

$$\begin{vmatrix} x & y & 1 \\ a_1 & a_2 & 1 \\ b_1 & b_2 & 1 \end{vmatrix} = 0$$

is the equation of a line through the points (a_1, a_2) and (b_1, b_2).

(b) What determinant of order 4 would give an angular equation for the plane through these given points?

41.[C]

(a) Review the Folium of Descartes in Section 9.3 on page 818.

(b) Show that the part in the fourth quadrant is asymptotic to the line $x+y+1 = 0$.

42.[M] Find where the line L through $P_0 = (2, 1, 3)$ and $P_1 = (4, -2, 5)$ meets the plane whose equation is $2x + y - 4z + 5 = 0$.

43.[M]

(a) Graph the line and the parabola. Identify, graphically, the point on the parabola closest to the line.

(b) Find, analytically, the point on the parabola $y = x^2$ closest to the line $y = x - 3$.

(c) The tangent to the parabola at the point found in (b) looks as if it might be parallel to the line. Is it?

44.[C] Let f be a differential function and L a line that does not meet the graph of F. Assume that P_0 is the point as the graph that is nearest the line.

(a) Using calculus, show that the tangent there is parallel to L.

(b) Why is the result in (a) to be expected?

In Exercises 45 and 46, find the distance from the given point to the given line.

45.[R] The point $(0,0)$ to $3x+4y-10=0$

46.[R] The point $(3/2, 2/3)$ to $2x-y+5=0$

In Exercises 47 and 48 find a normal and a unit normal to the given planes.

47.[R] $2x-3y+4z+11=0$

48.[R] $z=2x-3y+4$

In Exercises 49 to 52 find the distance from the given point to the given plane.

49.[R] The point $(0,0,0)$ to the plane $2x-4y+3z+2=0$

50.[R] The point $(1,2,3)$ to the plane $x+2y-3z+5=0$.

51.[R] The point $(2,2,-1)$ to the plane that passes through $(1,4,3)$ and has a normal $2\mathbf{i}-7\mathbf{j}+2\mathbf{k}$.

52.[R] The point $(0,0,0)$ to the plane that passes through $(4,1,0)$ and is perpendicular to the vector $\mathbf{i}+\mathbf{j}+\mathbf{k}$.

53.[R] Find the direction cosines of the vector $2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$.

54.[R] Find the direction cosines of the vector from $(1,3,2)$ to $(4,-1,5)$.

55.[R] Let $P_0=(2,1,5)$ and $P_1=(3,0,4)$. Find the direction cosines and direction angles of

(a) $\overrightarrow{P_0P_1}$ and

(b) $\overrightarrow{P_1P_0}$.

56.[R] Give parametric equations for the line through $(1/2, 1/3, 1/2)$ with direction numbers 2, -5 and 8 in

(a) scalar form,

(b) vector form.

57.[R] Give parametric equations for the line through $(1, 2, 3)$ and $(4, 5, 7)$ in

(a) scalar form,

(b) vector form.

58.[R] Give symmetric equations for the line through the points $(7, -1, 5)$ and $(4, 3, 2)$.

59.[R] A vector **A** has direction angles $\alpha = 70°$ and $\beta = 80°$. Find the third direction angle γ and show the possibile angles for γ on a diagram.

60.[M] Suppose that the three direction angles of a vector are equal. What can they be? Draw the cases.

61.[R] Find the angle between the line through $(3, 2, 2)$ and $(4, 3, 1)$ and the line through $(3, 2, 2)$ and $(5, 2, 7)$.

62.[R] Find the angle between the planes $2x + 3y + 4z = 11$ and $3x - y + 2z = 13$. The angle between two planes is the angle between their normals.

63.[R] Find where the line through $(1, 2)$ and $(3, 5)$ meets the line through $(1, -1)$ and $(2, 3)$.

64.[M] Find where the line through $(1, 2, 1)$ and $(2, 1, 3)$ meets the plane that is perpendicular to the vector $2\mathbf{i} + 5\mathbf{j} + 7\mathbf{k}$ and passes through the point $(1, -2, -3)$.

65.[M] Are the three points $(1, 2, -3)$, $(1, 6, 2)$, and $(7, 14, 11)$ on a single line?

66.[R] Where does the line through $(1, 2, 4)$ and $(2, 1, -1)$ meet the plane $x + 2y + 5z = 0$?

67.[R] Give parametric equations for the line through $(1, 3, -5)$ that is perpendicular to the plane $2x - 3y + 4z = 11$.

68.[R] Give parametric equations for the line through $(1,3,4)$ that is parallel to the line through $(2,4,6)$ and $(5,3,-2)$.

69.[C] A square of a side a lies in the plane $2x+3y+2z=8$. What is the area of its projection

(a) on the xy plane?

(b) on the yz plane?

(c) on the xz plane?

70.[M] If α, β, and γ are direction angles of a vector, what is $\sin^2(\alpha)+\sin^2(\beta)+\sin^2(\gamma)$?

71.[M] Find the angle between the line through $(1,3,2)$ and $(4,1,5)$ and the plane $x-y-2z+15=0$.

72.[C] A disk of radius a is situated in the plane $x+3y+4z=5$. What is the area of its projection in the plane $2x+y-z=6$?

73.[M] What point on the line through $(1,2,5)$ and $(3,1,1)$ is closest to the point $(2,-1,5)$?

74.[C] Does the line through $(5,7,10)$ and $(3,4,5)$ meet the line through $(1,4,0)$ and $(3,6,4)$? If so, where?

> **WARNING** *(Do Not Confuse Parameters from Different Curves)* Use parametric equations but give the parameters of the lines different names, such as t and s.

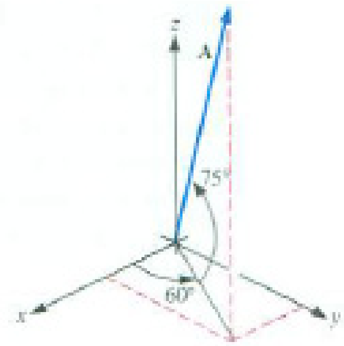
75.[C] Develop a general formula for determining the distance from the point $P_1=(x_1,y_1,z_1)$ to the line through the point $P_0=(x_0,y_0,z_0)$ and parallel to the vector $\mathbf{A}=a_1\mathbf{i}+z_2\mathbf{j}+a_3\mathbf{k}$. The formula should be expressed in terms of the vectors $\overrightarrow{P_0P_1}$ and $\mathbf{A}$.

76.[C] How far is the point $(1,2,-1)$ from the line through $(1,3,5)$ and $(2,1,-3)$?

(a) Solve by calculus, minimizing a certain function.

(b) Solve by vectors.

77.[R] Find the direction cosines of the vector **A** shown in Figure 14.4.11. HINT: First

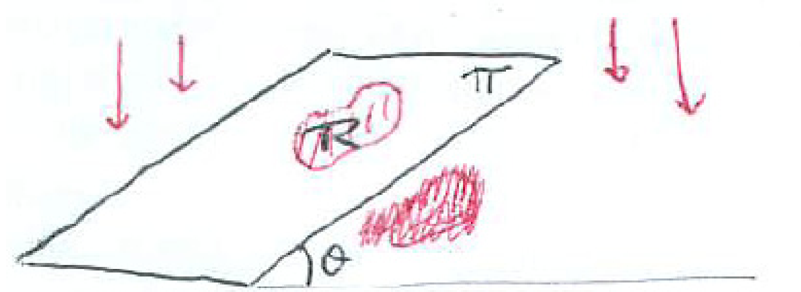


draw a large diagram.

Figure 14.4.11:

78.[C] How small can the largest of three direction angles ever be?

79.[C] A plane π is tilted at an angle θ to a horizontal plane. A convex region R in π has area A. Show that the area of its shadow ("projection") on the horizontal plane is $A\cos(\theta)$. Assume that the rays of light are perpendicular to the horizontal plane.



(See Figure 14.4.12.)

Figure 14.4.12:

80.[M]

(a) Find the point on the curve $y = \sin(x)$, $0 \le x \le \pi$, nearest the line $y = x/2+2$.

(b) Check your answer by sketching the curve to the line.

81.[M]

(a) Find the point on the curve $y = \sin x$, $0 \leq x \leq \pi$, nearest the line $y = 2x + 4$.

(b) Check your answer by drawing the curve and the line.

82.[R] Three points $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$, and $P_3 = (x_3, y_3, z_3)$ are the vertices of a triangle.

(a) What is the area of that triangle?

(b) What is the area of the projection of that triangle on the xy plane?

83.[M] How can you decide whether the line through P and Q is parallel to the plane $Ax + By + Cz + D = 0$?

84.[M] Find where the line through $(1, 1)$ and $(2, 3)$ meets the line $x+2y+3 = 0$.

85.[R] Show that the line through $(1, 1, 1)$ and $(2, 3, 4)$ is perpendicular to the plane $x_1 + 2y + 3z + 4 = 0$.

86.[C] How would you decide whether the angle and a point $P = (x_0, y_0, z_0)$ are on the same side or opposite sides of the plane $Ax + Bx + Cz + D = 0$?

87.[M]

(a) Give an example of a vector perpendicular to the plane $2x + 3y - z + 4 = 0$.

(b) Give an example of a vector parallel to that plane.

88.[C] How would you decide whether the points P and Q are on the same side, or opposite sides, of the plane $Ax + By + Cz + D = 0$?

89.[R] A plane contains the points P_0, P_1, and P_2, which do not lie on a line. Find a vector perpendicular to the plane

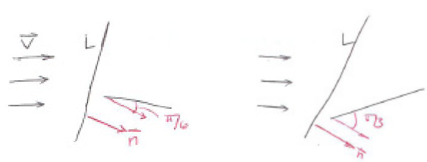
90.[C] Devise a procedure for determining whether the point $P = (x, y)$ is inside the triangle whose three vertices are $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ and $P_3 = (x_3, y_3)$.

91.[C] Devise a procedure for determining whether the point $P = (x, y, x)$ is inside the four vertices are $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$ $P_3 = (x_3, y_3, z_3)$ and $P_4 = (x_4, y_4, z_4)$.

92.[M] How far apart are the planes $Ax+By+Cz+D=0$ and $Ax+By+Cz+E=0$? Explain.

93.[R] We showed that the distance from (c,d) to the line $Ax+By+C=0$ is $\frac{|Ac+Bd+C|}{\sqrt{A^2+B^2}}$. Show, following a similar argument, that the distance from (c,d,e) to the plane $Ax+By+Cx+D=0$ is $\frac{|Ac+Bd+Ce+D|}{\sqrt{A^2+B^2+C^2}}$.

94.[M] What is the ratio of the flows across the two sticks in Figure 14.4.13(a) and (b)?



(a) (b)

Figure 14.4.13:

95.[R] Why is the angle θ shown in Figure 14.4.13 the same as the angle between $\overrightarrow{v}$ and $\widehat{\mathbf{n}}$.

96.[R] How far is the point $(1,5)$ from the line through $(4,2)$ and $(3,7)$? Hint: Draw a picture and think in terms of vectors.

97.[R] How far is the point $(1,2,-3)$ from the line through $(2,1,4)$ and $(1,5,-2)$?

98.[C] (Contributed by Melvyn Kopald Stein.) An industrial hopper is shaped as shown in Figure 14.4.14. Its top and bottom are squares of different sizes. The angle between the plane ABD and the plane BDC is 70°. The angle between the plane ABD and the plane ABC is 80°. What is the angle between plane ABC and plane BCD? NOTE: The angle is needed during the fabrication of the hopper, since the planes ABC and BCD are made from a single piece of heavy-gauge sheet metal bent along the edge BC.

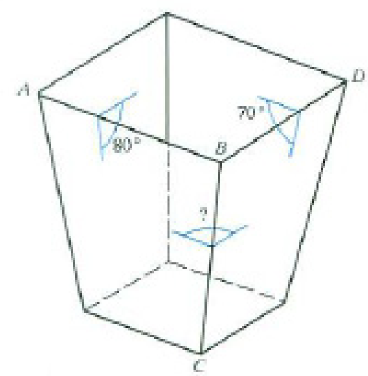


Figure 14.4.14:

99.[C]

(a) Let L_1 be the line through P_1 and Q_1 and let L_2 be the line through P_2 and Q_2. Assume that L_1 and L_2 are skew lines. How would you find the point R_1 on L_1 and point R_2 on L_2 such that $\overrightarrow{R_1R_2}$ is perpendicular to both L_1 and L_2?

(b) Find R_1 and R_2 when $P_1 = (3, 2, 1)$, $Q_1 = (1, 1, 1)$, $P_2 = (0, 2, 0)$, $R_2 = (2, 1, -1)$.

14.S Chapter Summary

DOUG/SHERMAN: Mention $\cos(\mathbf{A}, \mathbf{B})$ in text.

Because there are no limits in this chapter, it is, strictly speaking, not part of calculus. In the next chapter, which concern derivatives of functions whose inputs are scalars and whose outputs are vectors, we return to calculus.

The following table summarizes the basic concepts of vectors in space.

For plane vectors, disregard the third component.

Assume $\mathbf{A} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$, $\mathbf{B} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, and $\mathbf{C} = c_1\mathbf{i} + c_2\mathbf{j} + c_3\mathbf{k}$.

Symbol	Name	Geometric Descriptions	Algebraic Formula
$\mathbf{A}$	Vector	Direction and magnitude (Figure)	$a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ or $\langle a_1, a_2, a_3\rangle$
$\|\mathbf{A}\|$	Length (norm, magnitude)	Length of $\mathbf{A}$	$\sqrt{a_1^2 + a_2^2 + a_3^2}$
$-\mathbf{A}$	Negative, or opposite, of $\mathbf{A}$	Figure	$-a_1\mathbf{i} - a_2\mathbf{j} - a_3\mathbf{k}$ or $\langle -a_1, -a_2, -a$
$\mathbf{A} + \mathbf{B}$	Sum of $\mathbf{A}$ and $\mathbf{B}$	Figure	$(a_1 + b_1)\mathbf{i} + (a_2 + b_2)\mathbf{j} + (a_3 + b_3)\mathbf{k}$ or $\langle a_1 + b_1,$
$\mathbf{A} - \mathbf{B}$	Difference of $\mathbf{A}$ and $\mathbf{B}$	Figure	$(a_1 - b_1)\mathbf{i} + (a_2 - b_2)\mathbf{j} + (a_3 - b_3)\mathbf{k}$ or $\langle a_1 - b_1,$
$c\mathbf{A}$	Scalar multiple of $\mathbf{A}$	Figure	$ca_1\mathbf{i} + ca_2\mathbf{j} + ca_3\mathbf{k}$ or $\langle ca_1, ca_2, ca$
$\mathbf{A} \cdot \mathbf{B}$	Dot, or scalar, product	$\|\mathbf{A}\|\|\mathbf{B}\|\cos(\theta)$	$a_1b_1 + a_2b_2 + a_3b_3$
$\mathbf{A} \times \mathbf{B}$	Cross, or vector, product	Magnitude: area of parallelogram spanned by $\mathbf{A}$ and $\mathbf{B}$, $\|\mathbf{A}\|\|\mathbf{B}\|\sin(\theta)$ Direction: perpendicular to $\mathbf{A}$ and $\mathbf{B}$, direction by right-hand rule	$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b+3 \end{vmatrix}$
$\text{proj}_{\mathbf{B}}\,\mathbf{A}$	(Vector) Projection of $\mathbf{A}$ on $\mathbf{B}$	Figure	$(\mathbf{A} \cdot \mathbf{u})\mathbf{u}$, where $u = \mathbf{B}/\|\mathbf{B}\|$
$\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$	Scalar triple product	$\pm$ volume of parallelepiped spanned by $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$	$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b+3 \\ c_1 & c_2 & c+3 \end{vmatrix}$
$\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$	Vector triple product		

Table 14.S.1:

Some Common Applications and Definitions

$\mathbf{A}\cdot\mathbf{B}=0$	$\mathbf{A}$ is perpendicular to $\mathbf{B}$ (assuming neither $\mathbf{A}$ nor $\mathbf{B}$
$\mathbf{A}\cdot\langle x-x_0, y-y_0, z-z_0\rangle = 0$	plane through (x_0,y_0,z_0) perpendicular to $\mathbf{A}$
$\frac{\lvert D\rvert}{\sqrt{A^2+B^2+C^2}}$	distance from the plane $Ax+By+Cz+D=0$ to
$\frac{\lvert Ax_1+By_1+Cz_1+D\rvert}{\sqrt{A^2+B^2+C^2}}$	distance from the plane $Ax+By+Cz+D=0$ to th
$\frac{\mathbf{A}\cdot\mathbf{B}}{\lVert\mathbf{A}\rVert\lVert\mathbf{B}\rVert}=\cos(\theta)$	θ is the angle between $\mathbf{A}$ and $\mathbf{B}$, $0<\theta<\pi$

When the angles between a vector $\mathbf{A}$ and $\mathbf{i}$, $\mathbf{j}$, vk are respectively α, β, and γ, the numbers $\cos(\alpha)$, $\cos(\beta)$, and $\cos(\gamma)$ are called the **direction cosines** of $\mathbf{A}$. They are linked by the equation $\cos(\alpha)^2+\cos(\beta)^2+\cos(\gamma)^2=1$.

The line through $P_0=(x_0,y_0,z_0)$ parallel to $\mathbf{A}=a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ is given parametrically as

$$\begin{cases} x &= x_0+a_1t \\ y &= y_0+a_2t \\ z &= z_0+a_3t, \end{cases}$$

or vectorially as

$$\overrightarrow{OP}=\overrightarrow{OP_0}+t\mathbf{A}.$$

Assuming none of a_1, a_2, and a_3 are zero.

Also, the line has the description in the symmetric form

$$\frac{x-x_0}{a_1}=\frac{y-y_0}{a_2}=\frac{z-z_0}{a_3}.$$

EXERCISES for 14.S *Key:* R–routine, M–moderate, C–challenging

1.[R] Find a vector perpendicular to the plane determined by the points $(1,2,1)$, $(2,1,-3)$, and $(0,1,5)$.

2.[R] Find a vector perpendicular to the plane determined by the points $(1,3,-1)$, $(2,1,1)$, and $(1,3,4)$.

3.[R] Find a vector that is perpendicular to the line through the points $(3,6,1)$ and $(2,7,2)$ and also to the line through the points $(2,1,4)$ and $(1,-2,3)$.

4.[R] Find a vector perpendicular to the line through $(1,2,1)$ and $(4,1,0)$ and also to the line through $(3,5,2)$ and $(2,6,-3)$.

5.[C] Figure 14.S.1 shows a tetrahedron $OABC$ with three edges of the indicated lengths.

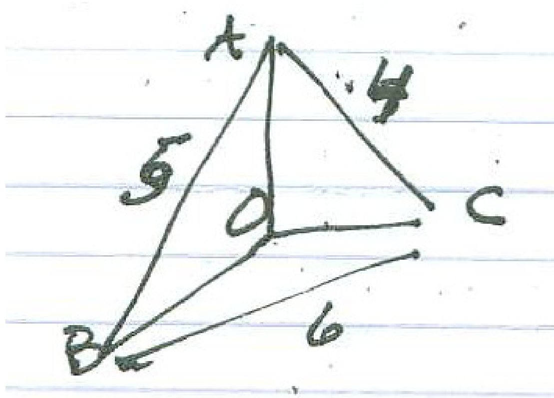


Figure 14.S.1:

(a) Find the coordinates of A, B, and C.

(b) Find the volume of the tetrahedron.

(c) Find the area of triangle ABC.

(d) Find the distance from O to the plane in which triangle ABC lies.

(e) Find the cosine of angle ABC.

Calculus is Everywhere # 16

Space Flight: The Gravitational Slingshot

For vector-algebra chapter

In a "slingshot" or "gravitational assist" a spacecraft picks up speed as it passes near a planet and exploits the planet's gravity. For instance, New Horizons, launched on January 19, 2006, enjoys a gravitational assist as it passed by Jupiter, February 27, 2007 on its long journey to Pluto. With the aid of that slingshot the speed of the spacecraft increased from 47,000 to 50,000 miles per hour (mph). As a result, it will arrive near Pluto in 2015, instead of 2018.

Before we see how this technique works, let's look at a simple situation on earth that illustrates the idea. Later we will replace the truck with a planet's gravitational field.

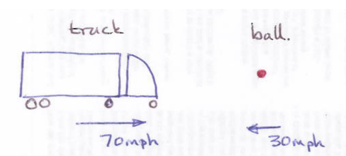


Figure C.16.1:

A playful lad throws a perfectly elastic tiny ball at 30 mph directly at a truck approaching him at 70 miles per hour, as shown in Figure C.16.1.

The truck driver sees the ball coming toward her at $70 + 30 = 100$ mph. The balls hits the windshield and, because the ball is perfectly elastic, the driver sees it bounce off at 100 mph in the opposite direction.

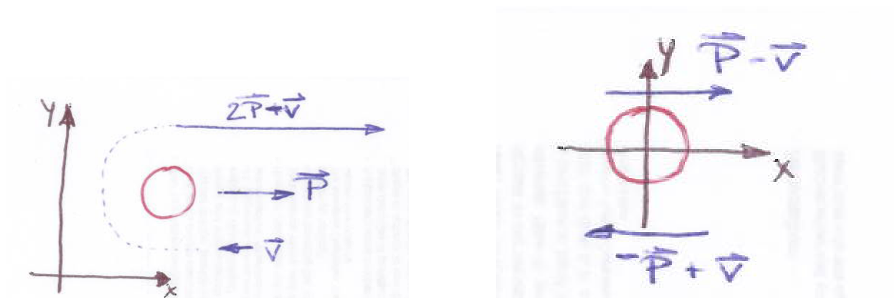
However, because the truck is moving in the same direction as the ball, the ball is moving through the air at $100 + 70 = 170$ mph as it returns to the boy. The ball has gained 140 mph, twice the speed of the truck.

Now, instead of picturing a truck, think of a planet whose velocity relative to the solar system is represented by the vector $\mathbf{P}$. A spacecraft, moving in the opposite direction with the velocity $\mathbf{v}$ relative to the solar system comes close to the planet.

An observer on the planet sees the spacecraft approaching with velocity $-vP + \mathbf{v}$. The spacecraft swings around the planet as gravity controls its orbit and sends it off in the opposite direction. Whatever speed it gained as it arrived, it loses as it exits. Its velocity vector when it exits is $-(-vP + \mathbf{v}) = \mathbf{P} - \mathbf{v}$, as viewed by the observer on the planet. Since the planet is moving through the solar system with velocity vector $\mathbf{P}$, the spacecraft is now moving through the solar system with velocity $\mathbf{P}+(\mathbf{P}-\mathbf{v}) = 2\mathbf{P}-\mathbf{v}$. See Figure C.16.2.

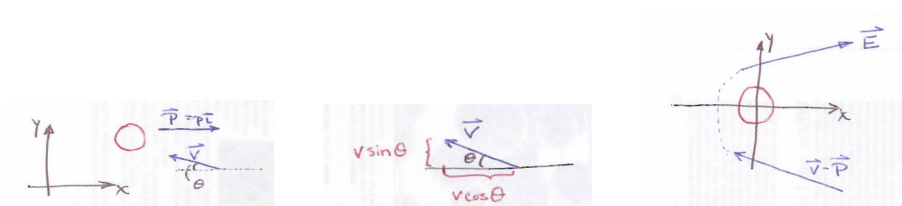
If $\mathbf{P} = 70\mathbf{i}$ and $\mathbf{v} = -30\mathbf{i}$, we have the vector $2(70\mathbf{i}) - (-30\mathbf{i}) = 170\mathbf{i}$, the case of the ball and truck.

But the direction of the spacecraft as it arrives may not be exactly opposite the direction of the planet. To treat the more general case, assume that $\mathbf{P} = p\mathbf{i}$, where p is positive and $\mathbf{v}$ makes an angle θ, $0 \leq \theta \leq \pi/2$, with $-\mathbf{i}$, as shown in Figure C.16.3(a). Let $v = |\mathbf{v}|$ be the speed of the spacecraft relative to the solar system. We will assume that the spacecraft's speed (relative to the planet) as it exits is the same as its speed relative to the planet on its arrival. (Figure C.16.3(b)) shows the arrival and exit vectors. Note that $\mathbf{E}$ and $\mathbf{v} - \mathbf{P}$



(a) (b)

Figure C.16.2: (a) The velocity vector relative to the solar system. (b) The velocity vector relative to the planet.



(a) (b) (c)

Figure C.16.3:

have the same y-components, but the x-component of $\mathbf{E}$ is the negative of the x-component of $\mathbf{v} - \mathbf{P}$.

Figure C.16.3(c) shows the arrival vector relative to the solar system. So, $\mathbf{v} = -w\cos(\theta)\mathbf{i} + v\sin(\theta)\mathbf{j}$.

Relative to the planet we have

$$\begin{array}{lrcl} \text{Arrival Vector:} & \mathbf{v} - \mathbf{P} & = & -p\mathbf{i} + (-v\cos(\theta)\mathbf{i} + v\sin(\theta)\mathbf{j}) \\ \text{Exit Vector:} & \mathbf{E} & = & p\mathbf{i} + v\cos(theta)\mathbf{i} + v\sin(\theta)\mathbf{j} \end{array}$$

The exit vector relative to the solar system, $\mathbf{E}$, is therefore

$$\mathbf{E} = (2p + v\cos(\theta))\mathbf{i} + v\sin(theta)\mathbf{j}.$$

The magnitude of $\mathbf{E}$ is

$$\sqrt{(2p + v\cos(\theta))^2 + (v\sin(theta))^2} = \sqrt{v^2 + 2pv\cos(\theta) + 4p^2}.$$

When $\theta = 0$, we have the case of the truck and ball or the planet and spacecraft in Figure C.16.2. Then $\cos(\theta) = 1$ and $|\mathbf{E}| = \sqrt{v^2 + 2pv + 4p^2} = v + 2p$, in agreement with our earlier observations.

The scientists controlling a slingshot carry out much more extensive calculations, which take into consideration the masses of the spacecraft and the planet, and involve an integration while the spacecraft is near the planet. Incidentally, the diameter of Jupiter is 86,000 miles.

"Near" in the case of the slingshot around Jupiter means 1.4 million miles. If the spacecraft gets too close, the atmosphere slows down or destroys the craft.

The gravity assist was proposed by Michael Minovitch in 1963 when he was still a graduate student at UCLA. Before then it was felt that to send a spacecraft to the outer solar system and beyond would require launch vehicles with nuclear reactors to achieve the necessary thrust.

Calculus is Everywhere # 17
How to Find Planets around Stars

Astronomers have discovered that other stars than the sun have planets circling them. How do they do this, given that the planets are too small to be seen? It turns out that they combine some vector calculus with observations of the star. Let us see what they do.

Imagine a star S and a planet P in orbit around S. To describe the situation, we are tempted to choose a coordinate system attached to the star. In that case the star would appear motionless, hence having no acceleration. However, the planet exerts a gravitational force F on the star and the equation force = mass × acceleration would be violated. After introducing the appropriate mathematical tools, we will choose a proper coordinate system.

Let $\mathbf{X}$ be the position vector of the planet P and $\mathbf{Y}$ be the position vector of the star S, relative to our inertial system. Let M be the mass of the sun and m the mass of planet P. Let $\mathbf{r} = \mathbf{X} - \mathbf{Y}$ be the vector from the star to the planet, as shown in Figure C.17.1.

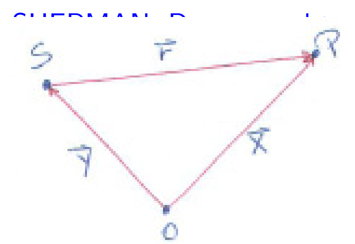


Figure C.17.1:

The gravitational pull of the star on the planet is proportional to the product between them:

$$\mathbf{F} = \frac{-GmM\mathbf{r}}{r^3}.$$

Here G is a universal constant, that depends on the units used to measure mass, length, time, and force. Equating the force with mass times acceleration, we have

$$\begin{aligned} M\mathbf{X}'' &= \tfrac{-GmM\mathbf{r}}{r^3}. \\ \text{Thus} \qquad \mathbf{X}'' &= \tfrac{-Gm\mathbf{r}}{r^3}. \end{aligned}$$

Similarly, by calculating the force that the planet exerts on the star, we have

$$\mathbf{Y}'' = \frac{Gm\mathbf{r}}{r^3}.$$

The center of gravity of the system consisting of the planet and the star, which we will denote C (see Figure C.17.2), is given by

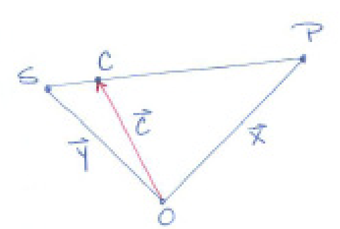


Figure C.17.2:

$$\mathbf{C} = \frac{M\mathbf{Y} + m\mathbf{X}}{M + m}.$$

The center of gravity is much closer to the start than to the planet. In the case of our sun and Earth, the center of gravity is a mere 300 miles from the center of the sun.

The acceleration of the center of gravity is

$$\mathbf{C}'' = \frac{M\mathbf{Y}'' + m\mathbf{X}''}{M+m} = \frac{1}{M+m}\left(M\left(\frac{Gm\mathbf{r}}{r^3}\right) + m\left(\frac{-Gm\mathbf{r}}{r^3}\right)\right) = \mathbf{0}.$$

Because the center of gravity has **0**-acceleration, it is moving at a constant velocity relative to the coordinate system we started with. Therefore a coordinate system rigidly attached to the center of gravity may also serve as an inertial system in which the laws of physics still hold.

We now describe the position of the star and planet to this new coordinate system. Star S has the vector $\mathbf{x}$ from $\mathbf{C}$ to it and planet P has the vector $\mathbf{y}$ from $\mathbf{C}$ to it, as shown in Figure C.17.3. Note that $\mathbf{r} = \mathbf{x} - \mathbf{y}$.

Figure C.17.3:

To obtain a relation between $\mathbf{x}$ and $\mathbf{y}$, we first express each in terms of $\mathbf{r}$. We have

$$\mathbf{y} = \mathbf{Y} - \vec{OC} = \mathbf{Y} - \frac{M\mathbf{Y} - m\mathbf{X}}{M+m} = \frac{m}{M+m}\mathbf{Y} + \frac{m}{M+m}\mathbf{X}.$$

Letting $k = m/M$, a very small quantity, we have

$$\mathbf{y} = \frac{k}{1+k}(\mathbf{Y} - \mathbf{X}) = \frac{-k}{1+k}\mathbf{r}. \qquad \text{(C.17.1)}$$

Since $\mathbf{r} = \mathbf{x} - \mathbf{y}$, it follows that $\mathbf{x} = \mathbf{r} + \mathbf{y}$, hence

$$\mathbf{x} = \mathbf{r} + \left(\frac{-k}{1+k}\right)\mathbf{r} = \frac{1}{1+k}\mathbf{r}. \qquad \text{(C.17.2)}$$

Combining (C.17.1) and (C.17.2) shows that

$$\boxed{\mathbf{y} = -k\mathbf{x}.} \qquad \text{(C.17.3)}$$

SHERMAN: First use of "second inertial system;" what is the first?

Equation (C.17.3) tells us a good deal about the relation between the orbits of the star and planet in terms of the second inertial system:

1. The star and planet remain on opposite sides of C on a straight line through C.

2. The star is always much closer to C than the planet is.

3. The orbit of the star is similar in shape to the orbit of the planet, but smaller and reflected through C.

4. If the orbit of the star is periodic so is the orbit of the planet, and both have the same period.

Equation (C.17.3) is the key to the discover of planets around stars. The astronomers look for a star that "wobbles" a bit. That wobble is the sign that the star is in orbit around the center of gravity of it and some planet. Moveover, the time it takes for the planet to orbit the star is simply the time it takes for the star to oscillate back and forth once.

The reference cited below shows that the star and the planet sweep out elliptical orbits in the second coordinate system (the one relative to C).

Astronomers have found over two hundred stars with planets, some with several planets. A registry of these **exoplanets** is maintained at `http://exoplanets.org/`.

Reference: Robert Osserman, *Kepler's Laws, Newton's Laws, and the Search for New Planets*, Am. Math. Monthly **108** (2001), pp. 813–820.

EXERCISES 1.[R] The mass of the sun is about 330,000 times that of Earth.

The closest Earth gets to the sun is about 91,341,000 miles, and the farthest from it is about 94,448,000 miles. What is the closest the center of the sun gets to the center of gravity of the sun-Earth system? What is the farthest it gets from it? HINT: It lies within the sun itself.

2.[M] Find the condition that must be satisfied if the center of gravity of a sun-planet system will lie outside the sun.

SHERMAN: See `http://en.wikipedia.org/wiki/Center_of_mass`, particularly the animations at the end of the section on "Barycenter in astrophysics and astronomy".

Chapter 15

Derivatives and Integrals of Vector Functions

In Section 9.3 we studied parametric curves in the plane. With the aid of calculus we saw how to compute arc length, speed and curvature. For instance, we defined curvature as the rate at which a certain angle changes as a function of arc length.

In this chapter we examine curves in the plane or in space. Of particular interest will be velocity and acceleration. For a particle moving along a straight line, say, the x-axis, these were simply the derivatives dx/dt and d^2x/dt^2. For a particle moving in space, velocity and acceleration involve both magnitude and direction. How should we calculate them?

How can we define curvature for a curve that does not lie in a plane? While arc length still makes sense, there is no angle to differentiate with respect to arc length.

While we could answer these questions using the cumbersome component notation for parameterization $\langle x(t), y(t)\rangle$ or $\langle x(t), y(t), x(t)\rangle$, we will emphasize the efficient vector notation, where a vector-valued concept is denoted by one letter. We will resort to the component notation to carry out computations or a proof. This point of view becomes increasingly important in the final chapters, particularly in Chapter 18.

15.1 The Derivative of a Vector Function: Velocity and Acceleration

In the case of motion on a horizontal line the derivative of position with respect to time is sufficient to describe the motion of the particle. If the derivative is positive, the particle is moving to the right. If the derivative is negative, the particle is moving to the left. The speed is simply the absolute value of the derivative. But the study of motion in the plane or in space depends on the concept of the derivative of a vector function.

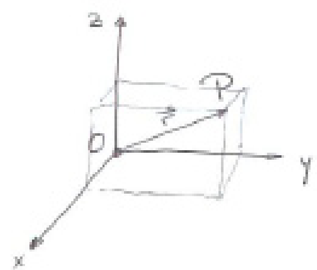


Figure 15.1.1:

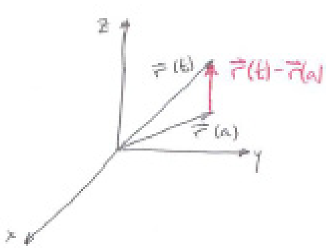


Figure 15.1.2:

In this section we introduce the calculus of a vector function and apply it to motion along a curve in a plane or in space.

Assume that a curve in the plane is parameterized as $\langle x(t), y(t)\rangle$ or, in space, by $\langle x(t), y(t), z(t)\rangle$. Let $P = P(t)$ be the point corresponding to t, which we may think of as "time," though it can be any parameter, such as arc length.

We introduce the **position vector**, $\mathbf{r} = \mathbf{r}(t)$, whose tail is at the origin O and whose tip is at P. Then $\mathbf{r} = \overrightarrow{OP}$, as shown in Figure 15.1.1

We will assume that $\mathbf{r}(t)$ is continuous, in that each of its components is continuous. The limit of $\mathbf{r}(t)$ as t approaches a we define as the vector

$$\langle \lim_{t\to a} x(t), \lim_{t\to a} y(t), \lim_{t\to a} z(t)\rangle.$$

We denote this as $\lim_{t\to a}\mathbf{r}(t)$. Figure 15.1.2 shows this geometrically. As t approaches a, the vector $\mathbf{r}(t) - \mathbf{r}(a)$ gets shorter and shorter as it approaches the zero vector $\mathbf{0}$.

We will say that $\mathbf{r}(t)$ is differentiable at $t = a$ if its components are differentiable at $t = a$. Then the derivative of $\mathbf{r}(t)$ is defined as the vector.

$$\langle x'(a), y'(a), z'(a)\rangle.$$

In vector notation,

$$\mathbf{r}'(a) = \lim_{t\to a}\frac{\mathbf{r}(t)-\mathbf{r}(a)}{t-a} \quad \text{or} \quad \mathbf{r}'(a) = \lim_{\Delta t\to 0}\frac{\mathbf{r}(a+\Delta t)-\mathbf{r}(a)}{\Delta t}.$$

and, if $\Delta\mathbf{r} = \mathbf{r}(r+\Delta t) - \mathbf{r}(r)$, $\mathbf{r}'(\alpha) = \lim_{\Delta t\to 0}\frac{\Delta \mathbf{r}}{\Delta t}$. When t is near a (or Δt is near 0) the vector in the numerator will be short. However, it is divided by $t-a$ (or Δt), which is small, so the quotient could be a vector of any size.

Some Derivative Formulas

In order to exploit the efficient vector notation when computing, we state some of the useful identities:

If $\mathbf{r}$ and $\mathbf{s}$ are differentiable vector functions, and f is a differentiable scalar, then

$(\mathbf{r}+\mathbf{s})'$	$=$	$\mathbf{r}'+\mathbf{s}'$	
$(\mathbf{r}\times\mathbf{s})'$	$=$	$\mathbf{r}'\times\mathbf{s}+\mathbf{r}\times\mathbf{s}'$	differentiate a cross product
$(\mathbf{r}\cdot\mathbf{s})'$	$=$	$(\mathbf{r}'\cdot\mathbf{s})+(\mathbf{r}\cdot\mathbf{s}')$	differentiate a dot product
$(f\mathbf{r})'$	$=$	$f'\mathbf{r}+f\mathbf{r}'$	product rule (f is a scalar function)
$(\mathbf{r}(f(t)))'$	$=$	$\mathbf{r}(f(t))f'(t)$	Chain Rule.

The proofs are straightforward calculations. We prove the formula for $(\mathbf{r}\cdot\mathbf{s})'$ in both the component notation and vector notation. For convenience, assume $\mathbf{r}(t)$ and $\mathbf{s}(t)$ are vectors in the xy plane: $\mathbf{r}(t) = \langle x(t), y(t)\rangle$ and $\mathbf{s}(t) = \langle u(t), v(t)\rangle$.

Proof in components language:

$$\begin{aligned}(\mathbf{r}\cdot\mathbf{s})' &= (x(t)u(t) + y(t)v(t))' = x'u + xu' + y'v + yv' \\ &= (x'u + y'v) + (u'x + v'y) = \mathbf{r}'\cdot\mathbf{s} + \mathbf{r}\cdot\mathbf{s}'.\end{aligned}$$

Now, the same proof, but *in vector language*:

$$\begin{aligned}(\mathbf{r}\cdot\mathbf{s})' &= \lim_{\Delta t\to 0}\frac{\mathbf{r}(a+\Delta t)\cdot\mathbf{s}(a+\Delta t) - \mathbf{r}(a)\cdot\mathbf{s}(a)}{\Delta t} \\ &= \lim_{\Delta t\to 0}\frac{(\mathbf{r}(a)+\Delta\mathbf{r})\cdot(\mathbf{s}(a)+\Delta\mathbf{s}) - \mathbf{r}(a)\cdot\mathbf{s}(a)}{\Delta t} \\ &= \lim_{\Delta t\to 0}\frac{\mathbf{r}(a)\cdot\mathbf{s}(a) + \Delta\mathbf{r}\cdot\mathbf{s}(a) + \mathbf{r}(a)\Delta\mathbf{s} + \Delta\mathbf{r}\cdot\Delta\mathbf{s} - \mathbf{r}(a)\cdot\mathbf{s}(a)}{\Delta t} \\ &= \lim_{\Delta t\to 0}\frac{\Delta\mathbf{r}}{\Delta t}\cdot\mathbf{s}(a) + \mathbf{r}(a)\frac{\Delta\mathbf{s}}{\Delta t} + \Delta\mathbf{r}\cdot\frac{\Delta\mathbf{s}}{\Delta t} \\ &= \mathbf{r}'(a)\cdot\mathbf{s}'(a) + \mathbf{r}(a)\cdot\mathbf{s}'(a) + \mathbf{0}\cdot\mathbf{s}'(a) \\ &= \mathbf{r}'(a)\cdot\mathbf{s}(a) + \mathbf{r}(a)\cdot\mathbf{s}'(a).\end{aligned}$$

This is almost the same as the proof for the derivative of the product in Section 4.3. •

EXAMPLE 1 At the time t, a particle has the position vector $\mathbf{r}(t) = 3\cos(2\pi t)\mathbf{i} + 3\sin(2\pi t)\mathbf{j} + 5t\mathbf{k}$. Describe its path.

SOLUTION At time t the particle is at the point

$$\begin{cases} x &= 3\cos(2\pi t) \\ y &= 3\sin(2\pi t) \\ z &= 5t.\end{cases}$$

Notice that $x^2 + y^2 = (3\cos(2\pi t))^2 + (3\sin(2\pi t))^2 = 9$. Thus the point is always above or below the circle

$$x^2 + y^2 = 9.$$

Moreover, as t increases, $z = 5t$ increases.

The path is thus the spiral spring sketched in Figure 15.1.3. When t increases by 1, the angle $2\pi t$ increases by 2π, and the particle goes once around the spiral. This type of corkscrew path is called a **helix**. You see it in the spiral on the cardboard tube inside a roll of kitchen paper towels and in a DNA molecule. ◇

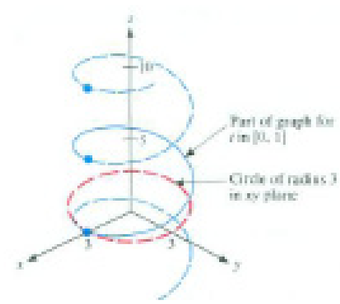


Figure 15.1.3:

A biology building on the University of California at Davis campus contains a large exact-to-scale model of the DNA molecule, 18 inches in diameter and 48 feet long. For additional information, please visit http://biosci.ucdavis.edu/sculpture.html.

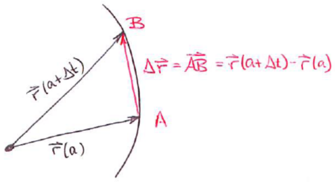


Figure 15.1.4:

The Meaning of r′ and r″

The vector $\mathbf{r}'(a)$ is the limit of

$$\frac{\mathbf{r}(a+\Delta t)-\mathbf{r}(a)}{\Delta t}$$

as $\Delta t \to 0$. The numerator $\mathbf{r}(a+\Delta t)-\mathbf{r}(a) = \Delta\mathbf{r}$ is shown in Figure 15.1.4.

Since $\Delta\mathbf{r}$ coincides with a chord, it points almost along the tangent line at the head of $\mathbf{r}(a)$ when Δt is small. Dividing a vector by a scalar (in this case, by Δt) produces a parallel vector. The corresponding position vector is $\mathbf{r}(t)$, so the vector

$$\frac{\mathbf{r}(a+\Delta t)-\mathbf{r}(a)}{\Delta t}$$

approximates a vector tangent to the curve at a. We conclude that

$$\mathbf{r}'(a) = \lim_{\Delta t\to 0}\frac{\mathbf{r}(t+\Delta t)-\mathbf{r}(a)}{\Delta t}$$

is a vector tangent to the curve at $\mathbf{r}(a)$. That is the geometric meaning of the derivative $\mathbf{r}'$:

> $\mathbf{r}'$ is tangent to the curve.

To see what $\mathbf{r}'$ means when we interpret the parameter t as time, we compute the length of $\mathbf{r}'(t)$.

Since $\mathbf{r}'(t) = \langle x'(t), y'(t), z'(t)\rangle$, its length is

$$\sqrt{(x'(t)^2)+(y'(t)^2)+(z'(t)^2)}.$$

As we saw in Section 9.3, this is the speed of the moving particle.

> The length of $\mathbf{r}'(t)$, $\|\mathbf{r}'(t)\|$, is the speed.

Since $\mathbf{r}'(t)$ points in the direction of motion and its length is the speed, we call $\mathbf{r}'(t)$ the **velocity vector**. Note that **velocity** is a vector, while speed is a scalar. That is a big distinction. The velocity carries much more information than speed: it also tells the direction of the motion.

> The velocity $\mathbf{r}'(t)$ is also denoted $\mathbf{v}$ or $\mathbf{v}(t)$. The speed is $\|\mathbf{v}\|$, denoted v or $v(t)$.

The **acceleration vector**, $\mathbf{a}(t)$ is the derivative of the velocity vector:

The acceleration is $\mathbf{a}(t) = \mathbf{v}'(t) = \frac{d\mathbf{v}}{dt} = \mathbf{r}''(t) = \frac{d^2\mathbf{r}}{dt^2}$.

EXAMPLE 2 Let $\mathbf{r}(t) = \langle t, t^3 \rangle$.

(a) Draw, and label, $\mathbf{r}$, $\mathbf{v}$, and $\mathbf{a}$ at $t = 1$.

(b) Draw $\mathbf{v}(1.1)$.

SOLUTION

(a) $\mathbf{r}(t) = \langle t, t^3 \rangle$, $\mathbf{v}(t) = \langle 1, 3t^2 \rangle$ and $\mathbf{a} = \langle 0, 6t \rangle$. So $\mathbf{r}(1) = \langle 1, 1 \rangle$, $\mathbf{v}(1) = \langle 1, 3 \rangle$ and $a(1) = \langle 0, 6 \rangle$. We show these in Figure 15.1.5.

(b) Before we compute $\mathbf{v}(1.1)$, let us predict how it may change from $\mathbf{v}(1)$. We think of the acceleration vector as representing a force. Since it's almost in the direction of $\mathbf{v}(1)$, we would expect it to be speeding up the moving particle. That is, $\mathbf{v}(1,1)$ should be longer than $\mathbf{v}(1)$.

Also, it would tend to rotate the velocity vector counterclockwise. So the direction of $\mathbf{v}(1,1)$ should be a bit counterclockwise from that of $\mathbf{v}(1)$. To check, we compute $\mathbf{v}(1,1) = \langle 1, 3(1.1)^2 \rangle = \langle 1, 3.63 \rangle$.

It is longer than $\mathbf{v}(1) = \langle 1, 3 \rangle$ since $\sqrt{1 + (3.63)^2}$ is larger than $\sqrt{1 + 3^2}$. Figure 15.1.6 shows that it is turned a bit counterclockwise, as expected. Its tail is placed at

$$\mathbf{r}(1.1) = \langle 1.1, 1.331 \rangle = 1.1\mathbf{i} + 1.331\mathbf{j}.$$

◇

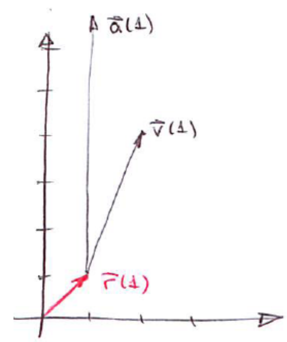


Figure 15.1.5:

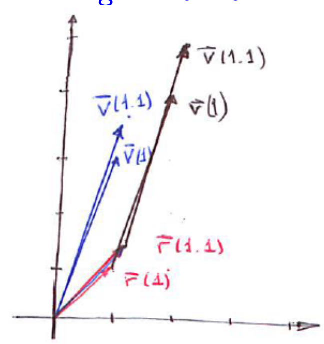


Figure 15.1.6:

EXAMPLE 3 Find the speed at time t of the particle described in Example 1.

SOLUTION

$$\begin{aligned} \text{Speed} = \|\mathbf{r}'(t)\| &= \sqrt{(-6\pi \sin 2\pi t)^2 + (6\pi \cos 2\pi t)^2 + 5^2} \\ &= \sqrt{36\pi^2(\sin^2 2\pi t + \cos^2 2\pi t) + 25} = \sqrt{36\pi^2 + 25}. \end{aligned}$$

The particle travels at a constant speed along it helical path. In t units of time it travels the distance $\sqrt{36\pi^2 + 25}t$.

Note that the velocity vector *is not constant*; its *direction always changes*. However, its length remains constant, and so the speed is constant. ◇

EXAMPLE 4 Sketch the path of a particle whose position vector at time $t \geq 0$ is $\mathbf{r}(t) = \cos(t^2)\mathbf{i} + \sin(t^2)\mathbf{j}$. Find its speed at time t.

SOLUTION Note that

$$\|\mathbf{r}(t)\| = \sqrt{\cos^2(t^2) + \sin^2(t^2)} = 1.$$

So the path of the particle is on the circle of radius 1 and center $(0,0)$. The speed of the particle is

$$\begin{aligned}\|\mathbf{v}(t)\| = \|\mathbf{r}'(t)\| &= \|-2t\sin(t^2)\mathbf{i} + 2t\cos(t^2)\mathbf{j}\| \\ &= \sqrt{(-2t\sin(t^2))^2 + (2t\cos(t^2))^2} \\ &= |2t|\sqrt{\sin^2(t^2) + \cos^2(t^2)} = 2t.\end{aligned}$$

The particle travels faster and faster around a circle of radius 1. ◇

EXAMPLE 5 If the acceleration vector is always perpendicular to the velocity vector, show that the speed is constant.

SOLUTION The speed is $\|\mathbf{v}\|$. Rather than writing this in terms of components and showing that its derivative is zero, let's use a trick that will be useful later.

We will show that the square of the speed, $\|\mathbf{v}\|^2$, is constant by showing that its derivative, with respect to time, is zero. Since $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$, we have

$$\frac{d}{dt}\left(\|\mathbf{v}\|^2\right) = \frac{d}{dt}(\mathbf{v} \cdot \mathbf{v}) \quad = \quad \mathbf{v}' \cdot \mathbf{v} + \mathbf{v} \cdot \mathbf{v}' = 2\mathbf{v} \cdot \mathbf{v}' = 2\mathbf{v} \cdot \mathbf{a}.$$

Since $\mathbf{a}$ is perpendicular to $\mathbf{v}$: $\mathbf{v} \cdot \mathbf{a} = 0$.

Thus $\mathbf{v} \cdot \mathbf{v}$ is constant. This implies that the speed is constant. ◇

The force of a magnetic field on a moving electron is perpendicular to the velocity vector that describes the motion of the electron. Since the acceleration vector is parallel to the vector representing the force, the speed of the electron remains constant (unless affected by other forces). Its direction, however, changes.

The calculation in Example 5 implies that if $\mathbf{r}(t)$ is always perpendicular to $\mathbf{r}'(t)$, then the length of $\mathbf{r}(t)$ is constant. The converse of this is also true:

> If the length of $\mathbf{r}(t)$ is constant, then its derivative $\mathbf{r}'(t)$ is perpendicular to $\mathbf{r}(t)$.

This is not surprising. When $\mathbf{r}(t)$ lies on a sphere of radius c (centered at the origin), the length of $\mathbf{r}(t)$ always has the same value, c. A tangent to the curve at the typical point P is tangent to the sphere. The tangent vector at P is perpendicular to the radius to P, as indicated in Figure 15.1.7(a), and the result follows.

(a) (b) (c)

Figure 15.1.7:

EXAMPLE 6 Is the particle shown in Figure 15.1.7(b) speeding up or slowing down? Is its direction turning clockwise or counterclockwise?

SOLUTION Think of $\mathbf{a}$ as the sum of two vectors, one parallel to $\mathbf{v}$, and the other perpendicular to $\mathbf{v}$, as shown in Figure 15.1.7(c). Since $\mathbf{b}$ is in the same direction as $\mathbf{v}$, the particle is speeding up. The direction of $\mathbf{c}$ indicates that the direction is shifting counterclockwise. ◇

Summary

Instead of parameterizing a curve by displaying the varying components $(x(t), y(t))$ or $(x(t), y(t), z(t))$, we introduced the position vector $\overrightarrow{OP} = \mathbf{r}(t)$. If $\mathbf{r}(t)$ describes the position of a moving particle at "time" t, then $\mathbf{r}'(t)$ is the velocity of the particle and $\|\mathbf{r}'(t)\|$ is its speed. The acceleration $\mathbf{a}(t)$ is the second derivative of $\mathbf{r}(t)$: $\mathbf{a} = \mathbf{r}''$. One may think of it as being proportional to the force operating in the particle.

Also, we showed that if $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are perpendicular, then the length of $\mathbf{r}(t)$, $\|\mathbf{r}(t)\|$, is constant. The converse holds: If $\mathbf{r}(t)$ has constant length, then $\mathbf{r}'(t)$ is perpendicular to $\mathbf{r}(t)$, and $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$.

EXERCISES for Section 15.1 *Key:* R–routine, M–moderate, C–challenging

1.[R] At time t a particle has the position vector $\mathbf{r}(t) = t\mathbf{i} + r^2\mathbf{j}$.

(a) Compute and draw $\mathbf{r}(1)$, $\mathbf{r}(2)$ and $\mathbf{r}(3)$.

(b) Show that the path is a parabola.

2.[R] At time t a particle has the position vector $\mathbf{r}(t) = (2t+1)\mathbf{i} + 4t\mathbf{j}$.

(a) Compute and draw $\mathbf{r}(0)$, $\mathbf{r}(1)$ and $\mathbf{r}(2)$.

(b) Show that the path is a straight line.

3.[R] Let $\mathbf{r}(t) = 2t\mathbf{i} + t^2\mathbf{j}$.

(a) Compute and draw $\mathbf{r}(1.1)$, $\mathbf{r}(1)$ and their difference $\Delta\mathbf{r} = \mathbf{r}(1.1) - \mathbf{r}(1)$.

(b) Compute and draw $\Delta\mathbf{r}/0.1$, where $\Delta\mathbf{r}$ is defined in part (a).

(c) Compute and draw $\mathbf{r}'(1)$. NOTE: Use one set of axes for all of the graphs.

4.[R] Let $\mathbf{r}(t) = 3t\mathbf{i} + t^2\mathbf{j}$.

(a) Compute and draw $\Delta\mathbf{r} = \mathbf{r}(2.01) - \mathbf{r}(2)$.

(b) Compute and draw $\Delta\mathbf{r}/0.01$.

(c) Compute and draw $\mathbf{r}'(2)$. NOTE: Use one set of axes for all the graphs.

5.[R] At time t the position vector of a thrown ball is $\mathbf{r}(t) = 32t\mathbf{i} - 16t^2\mathbf{j}$.

(a) Draw $\mathbf{r}(1)$ and $\mathbf{r}(2)$.

(b) Sketch the path.

(c) Compute and draw $\mathbf{v}(0)$, $\mathbf{v}(1)$, and $\mathbf{v}(2)$. In each case place the tail of the vector at the head of the corresponding position vector.

6.[R] At the time $t \geq 0$ a particle is at the point $x = 2t$, $y = 4t^2$.

(a) What is the position vector $\mathbf{r}(t)$ at time t?

(b) Sketch the path.

(c) How fast is the particle moving when $t = 1$?

(d) Draw $\mathbf{v}(1)$ with its tail at the head of $\mathbf{r}(1)$.

7.[R] Let $\mathbf{r}(t)$ describe the path of a particle moving in the xy plane. If $\mathbf{r}(1) = 2.3\mathbf{i} + 4.1\mathbf{j}$ and $\mathbf{r}(1.2) = 2.31\mathbf{i} + 4.05\mathbf{j}$, estimate

(a) how much does the position of the particle change during the time interval $[1, 1.2]$.

(b) the slope of the tangent vector to the path at $\mathbf{r}(1)$.

(c) the velocity vector $\mathbf{r}'(1)$.

(d) the speed of the particle at time $t = 1$.

8.[R] Let $\mathbf{r}(t)$ describe the path of a particle moving in space. If $\mathbf{r}(2) = 1.7\mathbf{i} + 3.6\mathbf{j} + 8\mathbf{k}$ and $\mathbf{r}(2.01) = 1.73\mathbf{i} + 3.59\mathbf{j} + 8.02\mathbf{k}$, estimate

(a) how far the particle moves during the time interval $[2, 2.01]$.

(b) the velocity vector $\mathbf{r}'(2)$.

(c) the speed of the particle at time $t = 1$.

In Exercises 9 and 12 compute the velocity vectors and speeds for the given paths.

9.[R] $\mathbf{r}(t) = \cos 3t\mathbf{i} + \sin 3t\mathbf{j} + 6t\mathbf{k}$.

10.[R] $\mathbf{r}(t) = 3\cos 5t\mathbf{i} + 2\sin 5t\mathbf{j} + t^2\mathbf{k}$.

11.[R] $\mathbf{r}(t) = \ln(1 + t^2)\mathbf{i} + e^{3t}\mathbf{j} + \frac{\tan t}{1+2t}\mathbf{k}$.

12.[R] $\mathbf{r}(t) = \sec^2 3t\mathbf{i} + \sqrt{1 + t^2}\mathbf{j}$.

13.[R] At time t the position vector of a particle is

$$\mathbf{r}(t) = 2\cos(4\pi t)\mathbf{i} + 2\sin(4\pi t)\mathbf{j} + t\mathbf{k}.$$

(a) Sketch its path.

(b) Find its speed.

(c) Find a unit tangent vector to the path at time t.

In each of Exercises 14 to 21 the figure shows a velocity vector and an acceleration vector. Decide whether (a) the particle is speeding up, slowing down, or neither, (b) the velocity vector is turning clockwise, counter-clockwise, or neither, at the moment.

14.[R] Figure 15.1.8(a)

15.[R] Figure 15.1.8(b)

16.[R] Figure 15.1.8(c)

17.[R] Figure 15.1.8(d)

18.[R] Figure 15.1.9(a)

19.[R] Figure 15.1.9(b)

20.[R] Figure 15.1.9(c)

21.[R] Figure 15.1.9(d)

(a) (b) (c) (d)

Figure 15.1.8:

(a) (b) (c) (d)

Figure 15.1.9:

22.[R] At time t a particle is at $(4t, 16t^2)$.

(a) Show that the particle moves on the curve $y = x^2$.

(b) Draw $\mathbf{r}(t)$ and $\mathbf{v}(t)$ for $t = 0, 1/4, 1/2$.

(c) What happens to $\|\mathbf{v}(t)\|$ and the direction of $\mathbf{v}(t)$ for large t?

23.[R] At time $t \geq 1$ a particle is at the point $(x, y) = (t, t^{-1})$.

(a) Draw the path of the particle.

(b) Draw $\mathbf{r}(1)$, $\mathbf{r}(2)$, and $\mathbf{r}(3)$.

(c) Draw $\mathbf{v}(1)$, $\mathbf{v}(2)$ and $\mathbf{v}(3)$.

(d) As times go on, what happens to dx/dt, dy/dt, $\|\mathbf{v}\|$, and $\mathbf{v}$?

24.[R] At time t a particle is at $(2\cos(t^2), \sin(t^2))$.

(a) Show that it moves on an ellipse.

(b) Compute $\mathbf{v}(t)$.

(c) How does $\|\mathbf{v}(t)\|$ behave for large t? What does this say about the particle?

25.[R] An electron travels at constant speed clockwise in a circle of radius 100 feet 200 times a second. At time $t = 0$ it is at $(100, 0)$.

(a) Compute $\mathbf{r}(t)$ and $\mathbf{v}(t)$.

(b) Draw $\mathbf{r}(0)$, $\mathbf{r}(1/800)$, $\mathbf{v}(0)$, $\mathbf{v}(1/800)$.

(c) How do $\|\mathbf{r}(t)\|$ and $\|\mathbf{v}(t)\|$ behave as time goes on?

26.[R] A ball is thrown up at an initial speed of 200 feet per second and at an angle of 50° from the horizontal. If we disregard air resistance, then at time t it is at $(100t, 100\sqrt{3}t - 16t^2)$, as long as it is in flight. Compute and draw $\mathbf{r}(t)$ and $\mathbf{v}(t)$ (a) when $t = 0$, (b) when the ball reaches its maximum height, and (c) when the ball strikes the ground.

27.[R] A particle moves in a circular orbit of radius a. At time t its position vector is

$$\mathbf{r}(t) = a\cos(2\pi t)\mathbf{i} + a\sin(2\pi t)\mathbf{j}.$$

(a) Draw its position vector when $t = 0$ and when $t = \frac{1}{4}$.

(b) Draw its velocity when $t = 0$ and when $t = \frac{1}{4}$.

(c) Show that its velocity vector is always perpendicular to its position vector.

28.[R] Use a computer or graphing calculator to graph $\mathbf{r} = \mathbf{r}(t) = (2\cos(t) + \cos(3t))\mathbf{i} + (3\sin(t) + \sin(3t))\mathbf{j}$; $0 \leq t \leq 2\pi$.

29.[R] If $\mathbf{r}(t)$ is the position vector, $\mathbf{v}$ the velocity vector, and $\mathbf{a}$ the acceleration vector, show that $\frac{d}{dt}(\mathbf{r} \times \mathbf{v}) = \mathbf{r} \times \mathbf{a}$.

30.[M] Let $\mathbf{r}(t) = t^2\mathbf{i} + t^3\mathbf{j}$.

(a) Sketch the vector $\Delta\mathbf{r} = \mathbf{r}(1.1) - \mathbf{r}(1)$.

(b) Sketch the vector $\Delta\mathbf{r}/\Delta t$, where $\Delta\mathbf{r}$ is given in (a) and $\Delta t = 0.1$.

(c) Sketch $\mathbf{r}'(1)$.

(d) Find $\|\Delta\mathbf{r}/\Delta t - \mathbf{r}'(1)\|$, where $\Delta\mathbf{r}$ is given in (a) and $\Delta t - 0.1$.

31.[M] Instead of time t, use the arc length s along the path as a parameter, $\mathbf{r} = \mathbf{r}(s)$.

(a) Show that $d\mathbf{r}/ds$ is a unit vector.

(b) Sketch $\Delta\mathbf{r}$ and the arc of length Δs. Why is it reasonable that $\|\Delta\mathbf{r}/\Delta s\|$ is near 1 when Δs is small?

32.[M] A particle at time $t = 0$ is at the point (x_0, y_0, z_0). It moves on the line through that point in the direction of the unit vector $\mathbf{u} = \cos(\alpha)\mathbf{i}+\cos(\beta)\mathbf{j}+\cos(\gamma)\mathbf{k}$. It travels at the constant speed of 3 feet per second.

(a) Give a formula for its position vector $\mathbf{r} = \mathbf{r}(t)$.

(b) Find its velocity vector $\mathbf{v} = \mathbf{r}'(t)$.

33.[M] A rock is thrown up at an angle θ from the horizontal and at a speed v_0.

(a) Show that

$$\mathbf{r}(t) = (v_0\cos(\theta))t\mathbf{i} + ((v_0\sin(\theta))t - 16t^2))\mathbf{j}.$$

NOTE: At time $t = 0$, the rock is at $(90, 0)$; the x-axis is horizontal. Time is in seconds and distance is in feet.

(b) Show that the horizontal distance that the rock travels by the time it returns to its initial height is the same whether the angle is θ or its complement $(\pi/2) - \theta$.

(c) What value of θ maximizes the horizontal distance traveled?

(This is similar to Exercise 25 in Section 9.3, but this version uses vector ideas.)

34.[M]

(a) Solve Example 5 by writing the speed as $\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}$ and differentiating.

(b) Which way do you prefer? The vector method in Example 5 or the component method in (a)?

35.[C] At time t the position vector of a particle is

$$\mathbf{r}(t) = t\cos(2\pi t)\mathbf{i} + t\sin(2\pi t)\mathbf{j} + t\mathbf{k}.$$

Sketch the path of the particle.

36.[C] A spaceship outside any gravitational field is on the path $\mathbf{r}(t) = t^2\mathbf{i} + 3t\mathbf{j} + 4t^3\mathbf{k}$. At time $t = 1$ it shuts off it rockets and coasts along the tangent line to the curve at that point.

(a) Where is it at time $t > 1$?

(b) Does it pass through the point $(9, 15, 50)$?

(c) If not, how close does it get to that point? (At what time?)

37.[C] A particle traveling on the curve $\mathbf{r}(t) = \ln(t)\mathbf{i} + \cos(3t)\mathbf{j}$, $t \geq 1$, leaves the curve when $t = 2$ and travels through space along the tangent to the curve at $\mathbf{r}(2)$. Where is it when $t = 3$?

38.[C] Drawing a picture of $\mathbf{r}(t)$, $\mathbf{r}(t + \Delta t)$, and $\mathbf{r}(t + \Delta t)$, explain why $\left\|\frac{\Delta\mathbf{r}}{\Delta t}\right\|$ is an estimate of the speed of a particle moving on the curve $\mathbf{r}(t)$.

39.[C] The moment a ball is dropped straight down from a tall tree, you shoot an arrow directly at it. Assume that there is no air resistance. Show that the arrow will hit the ball. (Assume that the ball does not hit the ground first.)

(a) Solve with the aid of the formulas in Exercise 33.

(b) Solve with a maximum of intuition and a minimum of computation.

40.[C]

(a) At time t a particle has the position vector $\mathbf{r}(t)$. Show that for small Δt the area swept out by the position vector is approximately $\frac{1}{2}\|\mathbf{r}(t) \times \mathbf{v}(t)\|\Delta t$. (See Figure 15.1.10.) HINT: $\mathbf{v}(t)$ is approximated by $\Delta\mathbf{r}/\Delta t$.

(b) Assume that the curve in (a) is parameterized over the time interval $[a, b]$. Show that the arc length swept out is $\frac{1}{2}\int_a^b \|\mathbf{r} \times \mathbf{v}\|\ dt$.

(c) Must the curve in (a) and (b) lie in a plane for the formula in (b) to hold?

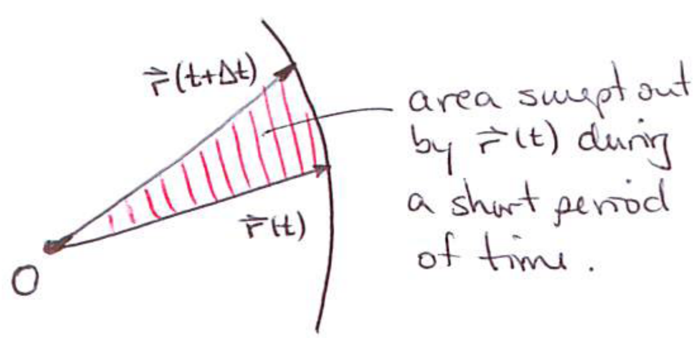


Figure 15.1.10:

Skill Drill

In Exercises 41 to 47 $\mathbf{v}(t)$ is the velocity vector at time t for a moving particle and $\mathbf{r}(0)$ is the particle's position at time $t = 0$. Find $\mathbf{v}(t)$, the position vector of the particle at time t. (These review the integration techniques of Chapter 8.)

41.[C] $\mathbf{v}(t) = \sin^2(3t)\mathbf{i} + \frac{t}{3t^2+1}\mathbf{j}$; $\mathbf{r}(0) = \mathbf{j}$

42.[C] $\mathbf{v}(t) = \frac{t}{t^2+t+1}\mathbf{i} + \tan^{-1}(3t)\mathbf{j}$; $\mathbf{r}(0 = \mathbf{i} + \mathbf{j}$

43.[C] $\mathbf{v}(t) = \frac{t^3}{t^4+1}\mathbf{i} + \ln(t+1)\mathbf{j}$; $\mathbf{r}(0) = \mathbf{0}$

44.[C] $\mathbf{v}(t) = e^{2t}\sin(3t)\mathbf{i} + \frac{t^3}{3t+2}\mathbf{j}$; $\mathbf{r}(0) = \mathbf{i} + 3\mathbf{j}$

45.[C] $\mathbf{v}(t) = \frac{t}{(t+1)(t+2)(t+3)}\mathbf{i} + \frac{t^2}{(t+2)^3}\mathbf{j}$; $\mathbf{r}(0) = \mathbf{i} - \mathbf{j}$

46.[C] $\mathbf{v}(t) = \frac{(\ln(t+1))^3}{t+1}\mathbf{i} + \frac{1}{\sqrt{1-4t^2}}\mathbf{j} + \sec^2(3t)\mathbf{k}$; $\mathbf{r}(0) = \mathbf{i} + \mathbf{j} + \mathbf{k}$

47.[C] $\mathbf{v}(t) = t^3e^{-t}\mathbf{i} + (1+t)(2+t)\mathbf{j}$; $\mathbf{r}(0) = 2\mathbf{i} - \mathbf{j}$

15.2 Curvature and Components of Acceleration

In Section 9.3 we defined the curvature of a plane curve as the absolute value of the derivative $d\phi/ds$, where ϕ is the angle the tangent makes with the x-axis and s is the arc length. This definition does not work for a curve that does not lie in a plane. Why not? In this section we use vectors to define the curvature of a curve, whether it lies in a plane or not. Curvature is then used to analyze the acceleration vector.

Definition of Curvature

A particle whose position vector at the time t is $\mathbf{r}(t)$ has velocity $\mathbf{v}(t)$. When $\mathbf{v}(t)$ is not the zero vector, the unit vector in the direction of $\mathbf{v}(t)$ is

$$\mathbf{T}(t) = \frac{\mathbf{v}(t)}{\|\mathbf{v}\|} = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}.$$

To save writing so many "t"s' we will just say

$$\mathbf{T} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\mathbf{r}'}{\|\mathbf{r}'\|} \qquad \text{(assuming } \mathbf{v} \neq \mathbf{0}\text{)}$$

All that $\mathbf{T}$ does is record the direction of motion.

As the particle moves along the curve the direction of $\mathbf{T}$ changes most rapidly where the curve is curviest. This suggests the following definition of curvature for any curve in the plane or in space:

Let s denote the length of arc of a curve, measured from a fixed starting point. Then, curvature, κ, is the length of $\frac{d\mathbf{T}}{ds}$, $\kappa = \left\|\frac{d\mathbf{T}}{ds}\right\|$.

We first check that this definition of curvature agrees with the definition for curvature given in the case of a plane curve in Section 9.3. We carry out this check in Example 1.

EXAMPLE 1 Show that the definition of curvature as $\|d\mathbf{T}/ds\|$ agrees with the definition $|d\phi/ds|$ given earlier for plane curve.

SOLUTION As Figure 15.2.1 shows, ϕ is the angle that $\mathbf{T}$ makes with the x-axis. Since $\mathbf{T}$ is a unit vector, $\mathbf{T} = \cos(\phi)\mathbf{i} + \sin(\phi)\mathbf{j}$. Thus

$$\begin{aligned}\kappa = \left\|\frac{d\mathbf{T}}{ds}\right\| &= \left\|\frac{d(\cos\phi\mathbf{i} + \sin\phi\mathbf{j}}{ds}\right\| = \left\|\frac{d(\cos\phi\mathbf{i} + \sin\phi\mathbf{j})}{d\phi}\frac{d\phi}{ds}\right\| \\ &= \left|(-\sin\phi\mathbf{i} + \cos\phi\mathbf{j})\frac{d\phi}{ds}\right| = \|(-\sin\phi\mathbf{i} + \cos\phi\mathbf{j})\| \left|\frac{d\phi}{ds}\right| = \left|\frac{d\phi}{ds}\right|\end{aligned}$$

so that $\left\|\frac{d\mathbf{T}}{ds}\right\| = \left|\frac{d\phi}{ds}\right|$. ◇

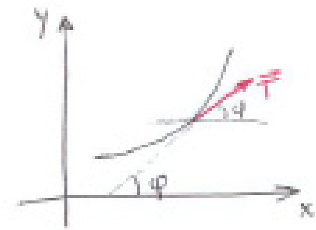


Figure 15.2.1:

Define the **radius of curvature**, r, as the **reciprocal** of κ as in Section 9.6.

EXAMPLE 2 Compute the curvature of the helical path

$$\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + 3t\mathbf{k}.$$

SOLUTION To find $\mathbf{T}$ we first must compute $\mathbf{v} = -\sin(t)\mathbf{i} + \cos(t)\mathbf{j} + 3\mathbf{k}$ and $\|\mathbf{v}\| = \sqrt{(-\sin(t))^2 + (\cos(t))^2 + 3^2} = \sqrt{10}$. Thus $\mathbf{T} = \frac{1}{\sqrt{10}}(-\sin(t)\mathbf{i} + \cos(t)\mathbf{j} + 3\mathbf{k})$.

Using the fact that speed is both $v = \|\mathbf{v}\|$ and the rate of change of arc length, the curvature equals

$$\left\|\frac{d\mathbf{T}}{ds}\right| = \frac{\left\|\frac{d\mathbf{T}}{dt}\right\|}{\left|\frac{ds}{dt}\right|} = \frac{\left\|\frac{d\mathbf{T}}{dt}\right\|}{v} = \frac{\left\|\frac{1}{\sqrt{10}}(-\cos(t)\mathbf{i} - \sin(t)\mathbf{j})\right\|}{\sqrt{10}}$$

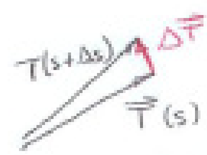


Figure 15.2.2:

The curvature is 1/10 and the radius of curvature is 10. For a helix the curvature and radius of curvature are both constant. For this particular helix, $\kappa = 1/10$ and $r = 1/\kappa = 10$. ◇

The Unit Normal N

Since $\mathbf{T}(t)$ is a constant length, $d\mathbf{T}/ds$ is perpendicular to $\mathbf{T}$. By considering small Δs and $\Delta\mathbf{T}$, as in Figure 15.2.2, we see that $d\mathbf{T}/ds$ points in the direction in which $\mathbf{T}$ is turning. Since the length of $d\mathbf{T}/ds$ is the curvature κ, we may write

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N}$$

where $\mathbf{N}$ is a unit normal called the "principal normal." That κ is positive reminds us that $d\mathbf{T}/ds$ and $\mathbf{N}$ point in the same direction. The vectors $\mathbf{T}$ and $\mathbf{N}$, if placed with their tails at a point P on the curve, determine a plane. The part of the curve near P stays close to that plane. (See Figure 15.2.3.)

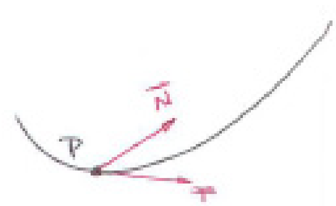


Figure 15.2.3:

The Acceleration Vector and T and N

The acceleration vector, $\mathbf{a}$, is defined as the second derivative of the position vector, $\mathbf{r}$. We will show that $\mathbf{a}$ is parallel to the plane determined by $\mathbf{T}$ and $\mathbf{N}$. That is, $\mathbf{a}$ can be written in the form $c_1\mathbf{T} + c_2\mathbf{N}$, where c_1 and c_2 are scalars, which we will express in terms of the motion of the particle.

Since $\mathbf{a} = \frac{d\mathbf{v}}{dt}$, we begin by computing $\mathbf{v}$ in terms of $\mathbf{T}$ and $\mathbf{N}$. This is easy: by the definition of $\mathbf{T}$, $\mathbf{v} = v\mathbf{T}$, where $v = \|\mathbf{v}\|$, the speed. $\mathbf{N}$ is not needed to express the velocity vector $\mathbf{v}$.

Thus

$$\begin{aligned} \mathbf{a} = \frac{d\mathbf{v}}{dt} &= \frac{d(v\mathbf{T})}{dt} \\ &= \frac{dv}{dt}\mathbf{T} + v\frac{d\mathbf{T}}{dt} && \text{product rule} \\ &= \frac{d^2s}{dt^2}\mathbf{T} + v\frac{d\mathbf{T}}{ds}\frac{ds}{dt} && v = \frac{ds}{dt} \text{ and chain rule} \\ &= \frac{d^2s}{dt^2}\mathbf{T} + v^2\frac{d\mathbf{T}}{ds} \end{aligned}$$

Thus, replacing $\frac{d\mathbf{T}}{ds}$ with $\kappa\mathbf{N}$, we find

Acceleration in terms of T, N, and curvature (κ)

$$\mathbf{a} = \frac{d^2s}{dt^2}\mathbf{T} + v^2\kappa\mathbf{N} \qquad (15.2.1)$$

$r = 1/\kappa$ ($\kappa \neq 0$)

If κ is not 0, then we have

Acceleration in terms of T, N, and radius of curvature (r)

$$\mathbf{a} = \frac{d^2s}{dt^2}\mathbf{T} + \frac{v^2}{r}\mathbf{N}. \qquad (15.2.2)$$

Tangential component of acceleration: $\mathbf{a}\cdot\mathbf{T} = \frac{d^2s}{dt^2}$

Normal component of acceleration: $\mathbf{a}\cdot\mathbf{N} = \frac{v^2}{r}$

The tangential component of acceleration, $\frac{d^2s}{dt^2}$, is positive if the particle is speeding up and is negative if the particle is slowing down. The normal component of acceleration, v^2/r, is always positive.

Figure 15.2.4 indicates how $\mathbf{a}$ may look relative to $\mathbf{T}$ and $\mathbf{N}$. In both cases $\mathbf{T}$ turns in the direction indicated by $\mathbf{N}$. In Figure 15.2.4 that means that $\mathbf{T}$ is turning counterclockwise.

Computing Curvature, κ

We can compute the curvature directly from its definition. There is also a shorter formula for κ. To develop this formula, we compute

$$\mathbf{T}\times\mathbf{a} = \mathbf{T}\times\left(\frac{d^2s}{dt^2}\mathbf{T} + v^2\kappa\mathbf{N}\right). \qquad (15.2.3)$$

(a) (b)

Figure 15.2.4: The tangential and normal components of acceleration: (a) $d^2s/dt^2 > 0$ and (b) $d^2s/dt^2 < 0$.

We do this for two reasons. First, $\mathbf{T} \times \mathbf{T} = 0$. Second $\|\mathbf{T} \times \mathbf{N}\| = 1$, since $\mathbf{T}$ and $\mathbf{N}$ span a unit square. By (15.2.3), then, we have

$$\mathbf{T} \times \mathbf{a} = \kappa v^2(\mathbf{T} \times \mathbf{N}).$$

Thus

$$\|\mathbf{T} \times \mathbf{a}\| = \kappa v^2.$$

Recalling that $\mathbf{T} = \mathbf{v}/v$, we have

$$\frac{\|\mathbf{v} \times \mathbf{a}\|}{v} = \kappa v^2$$

and finally

Curvature in terms of speed, velocity and acceleration

$$\kappa = \frac{\|\mathbf{v} \times \mathbf{a}\|}{v^3}. \tag{15.2.4}$$

We illustrate (15.2.4) by applying it to the helical path of Example 2.

EXAMPLE 3 Compute the curvature of the helical path $\mathbf{r}(t) = \cos(t)\mathbf{i} + \sin(t)\mathbf{j} + 3t\mathbf{k}$ using (15.2.4).

SOLUTION We compute $\mathbf{v}$, v and $\mathbf{a}$. First, $\mathbf{v} = -\sin t\mathbf{i} + \cos t\mathbf{j} + 3\mathbf{k}$, so $v = \sqrt{(-\sin(t))^2 + (\cos(t))^2 + 3^2} = \sqrt{10}$. Then

$$\mathbf{a} = \frac{d\mathbf{v}}{dt} = -\cos(t)\mathbf{i} - \sin(t)\mathbf{j}.$$

Next we compute $\mathbf{v} \times \mathbf{a}$:

$$\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin(t) & \cos(t) & 3 \\ -\cos(t) & -\sin(t) & 0 \end{pmatrix} = 3\sin(t)\mathbf{i} - 3\cos(t)\mathbf{j} + (\sin^2(t) + \cos^2(t))\mathbf{k}$$
$$= 3\sin(t)\mathbf{i} - 3\cos(t)\mathbf{j} + \mathbf{k}.$$

Finally,

$$k = \frac{|\mathbf{v} \times \mathbf{a}|}{v^3} = \frac{|3\sin(t)\mathbf{i} - 3\cos(t)\mathbf{j} + \mathbf{k}|}{(\sqrt{10})^3} = \frac{\sqrt{(3\sin(t))^2 + (-3\cos(t))^2 + 1^2}}{\sqrt{10}^3} = \frac{\sqrt{10}}{(\sqrt{10})^3} = \frac{1}{10}.$$

◇

Though curvature is defined as a derivative with respect to arc length s, we seldom use that definition in computations.

First of all, we seldom can obtain a formula for the arc length. Second, if the curve is described in terms of a parameter t, such as time or angle, then we can use the chain rule to express $d\mathbf{T}/ds$ as the directly calculable

$$\frac{\frac{d\mathbf{T}}{dt}}{\frac{ds}{dt}}.$$

The Meaning of the Components of a

If no force acts as a moving particle, it would move in a line at a constant speed. But if there is a force $\mathbf{F}$, then, according to Newton, it is related to the acceleration vector $\mathbf{a}$ by the equation $\mathbf{F} = m\mathbf{a}$ (when the mass m is constant.) So we can think of $\mathbf{a}$ as a representative of the force $\mathbf{F}$.

If $\mathbf{F}$ is parallel to $\mathbf{T}$, the particle moves in a line with an acceleration $dv/dt = d^2s/dt^2$. So we would expect $\mathbf{a}$ to equal $d^2s/dt^2\mathbf{T}$.

If $\mathbf{F}$ is perpendicular to $\mathbf{T}$, it would not change the speed, but it would push the particle away from a straight path, as shown in Figure 15.2.5

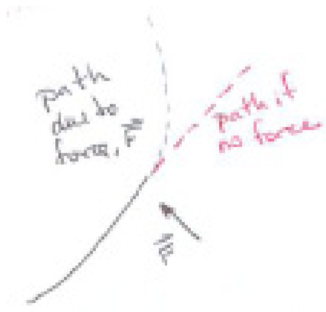


Figure 15.2.5:

If you spin a pail of water at the end of a rope (or a discus at the end of your arm) you can feel this force. It is proportional to the square of the speed and inversely proportional to the length of the rope. No wonder driving a car around a sharp curve too fast can cause it to skid: the friction of the tire against road cannot supply the necessary force (whose magnitude is the speed squared divided by the radius of the turn) to prevent skidding.

The Third Unit Vector, B

The vector $\mathbf{T} \times \mathbf{N}$ has length 1 and is perpendicular to both $\mathbf{T}$ and $\mathbf{N}$. We may think of it as a normal to the plane through a given point P on the curve parallel to $\mathbf{T}$ and $\mathbf{N}$. The unit vector $\mathbf{T} \times \mathbf{N}$ is denoted $\mathbf{B}$ and is called the **binormal**. It is shown in Figure 15.2.6

The plane that contains the point P and the two vectors $\mathbf{T}$ and $\mathbf{N}$ is called the osculating plane. See also Section 9.6.

Figure 15.2.6:

The three vectors, $\mathbf{T}$, $\mathbf{N}$, and $\mathbf{B}$, may change direction as the point P moves along the curve. However, they remain a rigid frame, where $\mathbf{T}$ indicates the direction of motion, $\mathbf{N}$ the direction of turning, and $\mathbf{B}$ the tilt of the plane through $\mathbf{T}$ and $\mathbf{N}$ if their tails are at P.

Summary

We defined the curvature of a curve in space (or in the xy-plane) using vectors. This definition agrees with the definition of curvature for curves in the xy-plane given in Section 9.6. The curvature, or its reciprocal, the radius of the curvature, appears in the normal component of the acceleration vector.

The section concluded with the definition of the binormal, $\mathbf{B} = \mathbf{T} \times \mathbf{N}$, which records the tilt of the plane determined by $\mathbf{T}$ and $\mathbf{N}$.

Several different formulas for the curvature were found in this section. More formulas for κ are found in Exercises 21, 22, and 23. In reality when given explicit formulas for a curve, it is often easiest to use a computer algebra system such as Mathematica or Maple to find the curvature and the vectors $\mathbf{T}$, $\mathbf{N}$, and $\mathbf{B}$.

EXERCISES for Section 15.2

Key: R–routine, M–moderate, C–challenging

In Exercises 1 to 4, $\mathbf{v}$ denotes the velocity and $\mathbf{a}$ denotes the acceleration. Evaluate the indicated scalar component.

1.[R] What is $\mathbf{v} \cdot \mathbf{T}$?

2.[R] What is $\mathbf{a} \cdot \mathbf{T}$?

3.[R] What is $\mathbf{v} \cdot \mathbf{N}$?

4.[R] What is $\mathbf{a} \cdot \mathbf{N}$?

5.[R]

(a) Why is $\mathbf{T} \times \mathbf{N}$ a unit vector?

(b) Why is $\mathbf{N}$ perpendicular to $\mathbf{T}$?

In Exercises 6 and 7, $\mathbf{v}$ and $\mathbf{a}$ are given at a certain instant. In each case, find the (i) curvature, (ii) radius of curvature, and (iii) d^2s/dt^2.

6.[R] $\mathbf{v} = 2\mathbf{i} + 3\mathbf{j} + 4\mathbf{k}$, $\mathbf{a} = \mathbf{i} - \mathbf{j} + \mathbf{k}$

7.[R] $\mathbf{v} = \mathbf{i} + \mathbf{j} + \mathbf{k}$, $\mathbf{a} = -\mathbf{i} + \mathbf{j} + \mathbf{k}$

In Exercises 8 and 9 compute the curvature using the formula $\kappa = |d\mathbf{T}/dt|/v$.

8.[R] $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$

9.[R] $\mathbf{r}(t) = 3\cos(2t)\mathbf{i} + 3\sin(2t)\mathbf{j} + 4t$

In Exercises 10 and 11, compute the curvature using the speed, velocity, and acceleration, that is, with the formula $\kappa = \frac{\|\mathbf{v}\times\mathbf{a}\|}{v^3}$.

10.[R] $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$

11.[R] $\mathbf{r}(t) = 3\cos(2t)\mathbf{i} + 3\sin(2t)\mathbf{j} + 4t\mathbf{k}$

12.[R] We showed that $d|v|/dt = \mathbf{v} \cdot \mathbf{a}/|v|$, using vectors. To emphasize the value of the vector approach, derive the same result starting with the fact that

$$\|\mathbf{v}\| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}.$$

13.[R] Let a and b be constants. A particle moves in a helical path described by

$$\mathbf{r}(t) = 3\cos(at)\mathbf{i} + 3\sin(at)\mathbf{j} + bt\mathbf{k}$$

(a) Compute its curvature.

(b) As $b \to \infty$ what happens to the curvature?

(c) Why is the answer to (b) reasonable?

(d) As $a \to \infty$, what happens to the curvature?

(e) Why is the answer to (d) reasonable?

14.[M] Show that the formula $\frac{\|(\mathbf{v}\times\mathbf{a})\|}{v^3}$ in the case $\mathbf{r}(x) = x\mathbf{i} + f(x)\mathbf{j} + 0\mathbf{k}$, gives the formula in Section 9.6 for curvature of the curve $y = f(x)$.

15.[M] Show that $\frac{d\mathbf{r}}{ds}$ is a unit vector,

(a) by drawing $\mathbf{r}(s + \Delta s)$ and $\mathbf{r}(s)$ and considering $\frac{\mathbf{r}(s+\Delta s)-\mathbf{r}(s)}{\Delta s}$

(b) by writing it as $(d\mathbf{r}/dt)/(ds/dt)$.

16.[M] Express the area of the parallelogram spanned by $\mathbf{v}$ and $\mathbf{a}$ in terms of the curvature and speed.

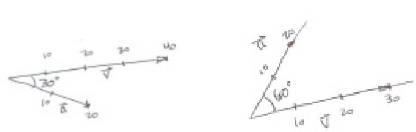
17.[M] If a particle reaches a maximum speed at time t_0, must d^2s/dt^2 be 0 at t_0? Must $d^2\mathbf{r}/dt^2$ be 0 at t_0? Assume the time interval is $(-\infty, \infty)$.

18.[M] If $\mathbf{r}(t)$ is the position vector, is $d^2\mathbf{r}/dt^2$ parallel to $d^2\mathbf{r}/ds^2$, where s denotes arc length?

In Exercises 19 and 20 the figures show the velocity and acceleration vectors at a point P and a curve. Use that information to find (i) v, (ii) d^2s/dt^2, (iii) κv^2. Then (iv) find r, the radius of the curvature, (v) draw the osculating circle, and (vi) using the osculating circle, draw an approximation of a short piece of the path near P.

19.[M] Figure 15.2.7(a)

20.[M] Figure 15.2.7(b)



(a) (b)

Figure 15.2.7:

21.[M]

Jane: After doing Exercises 19 and 20, I have a simpler way to get a formula for curvature. Just look at the right triangle whose hypotenuse has length $\|\mathbf{a}\|$ and its component along $\mathbf{v}$. By trigonometry,

$$\kappa v^2 = \|\mathbf{a}\| |\sin(\mathbf{a}, \mathbf{v})|. \tag{15.2.5}$$

All that's left is getting $\sin(\mathbf{a}, \mathbf{v})$ out and $\cos(\mathbf{a}, \mathbf{v})$ in because we know how to express $\cos(\mathbf{a}, \mathbf{v})$ in terms of a dot product. Squaring (15.2.5) gives

$$\kappa^2 v^4 = \|\mathbf{a}\|^2 \left(1 - \cos^2(\mathbf{a}, \mathbf{v})\right).$$

If you use the fact that

$$\cos(\mathbf{a}, \mathbf{v}) = \frac{\mathbf{a} \cdot \mathbf{v}}{\|\mathbf{a}\|(v)}.$$

and a little algebra, you get

$$\kappa^2 = \frac{(\mathbf{v} \cdot \mathbf{v})(\mathbf{a} \cdot \mathbf{a}) - (\mathbf{a} \cdot \mathbf{v})^2}{v^6}$$

My way is simpler than using the cross product. I guess the authors don't understand trigonometry.

(a) Fill in the missing steps.

(b) Check that Jane's formula agrees with (15.2.4).

22.[M]

Sam: You used trigonometry. I can do it with just the Pythagorean Theorem. Look at that triangle with hypotenuse $|\mathbf{a}|$. Its two legs have lengths $|d^2s/dt^2|$ and κv^2. So

$$\|\mathbf{a}\|^2 = \left(\frac{d^2 s}{dt^2}\right)^2 + (\kappa v^2)^2.$$

Solve this for κ. JANE But you have to express everything in vectors. We're in the chapter on vectors.

Sam: O.K. First $\|\mathbf{a}\|^2 = \mathbf{a} \cdot \mathbf{a}$ and $v^2 = \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$.

Jane: But d^2s/dt^2?

Sam: That's dv/dt. So I just differentiate both sides of $v^2 = \mathbf{v}\cdot\mathbf{v}$, getting $2v\frac{dv}{dt} = 2\mathbf{v}\cdot\mathbf{a}$. So $dv/dt = (\mathbf{v}\cdot\mathbf{a})/v$. So $(dv/dt)^2 = (\mathbf{v}\cdot\mathbf{a})^2/v^2$. So

$$\mathbf{a}\cdot\mathbf{a} = \frac{(v\cdot\mathbf{a})^2}{v^2} + \kappa^2(v^2)^2.$$

Then solve for κ^2.

I get the same result that you got in Exercise 21. It seems quite straightforward. The authors should have used my formula.

Find the formula for curvature that is obtained from this line of reasoning.

23.[M] This equation provides yet another way to find a formula for curvature. Consider the right triangle whose hypotenuse is $|\mathbf{a}|$ and whose legs are parallel to $\mathbf{T}$ and $\mathbf{N}$. Show that

$$\kappa^2 v^4 = (\ddot{x})^2 + (\ddot{y})^2 + (\ddot{z})^2 - (\ddot{s})^2.$$

NOTE: Two dots over a variable denotes the second derivative of the variable with respect to t.

24.[M] Assume that you are prone to car sickness on curvy roads such as Highway 1 north of San Francisco or the small state highways in southern Ohio. Which matters more to you, $|d\mathbf{T}/ds|$ where s is arc length or $|d\mathbf{T}/dt|$ where t is time? Explain the difference in the two quantities.

25.[M] Let $\mathbf{r} = \mathbf{r}(s)$, where s is arc length. Show that the curvature is $\kappa = \left\|\frac{d^2\mathbf{r}}{ds^2}\right\|$.

26.[M] Consider curves situated on the surface of a sphere $\mathcal{S}$ of radius a. (Recall that a sphere is the surface of a ball.)

(a) Show that there are curves on $\mathcal{S}$ that have very large curvature.

(b) Exhibit a curve whose curvature is as small as $1/a$.

(c) Show that there are no curves with curvature smaller than $1/a$. HINT: See Exercise 25 and start with $\mathbf{r}\cdot\mathbf{r} = a^2$.

The Frenet formulas concern the derivatives of $\mathbf{T}$, $\mathbf{B}$, and $\mathbf{N}$ with respect to arc length s:

$$\frac{d\mathbf{T}}{ds} = \kappa\mathbf{N}, \quad \frac{d\mathbf{B}}{ds} = \tau\mathbf{N}, \quad \frac{d\mathbf{N}}{ds} = -\kappa\mathbf{T} + \tau\mathbf{B}.$$

Here κ is curvature and τ is "torsion," the measure of the tendency of the plane through $\mathbf{T}$ and $\mathbf{N}$ to turn.

Exercises 27 and 28 develop the formulas for $d\mathbf{B}/ds$ and $d\mathbf{N}/ds$.

27.[M] This exercise develops $d\mathbf{B}/ds$.

(a) Why is $d\mathbf{B}/ds$ perpendicular to $\mathbf{B}$?

(b) Why are there scalars p and q such that $\frac{d\mathbf{B}}{ds} = p\mathbf{T} + q\mathbf{N}$?

(c) Using the fact that $\mathbf{B}$ and $\mathbf{T}$ are always perpendicular show that

$$(p\mathbf{T} + q\mathbf{N}) - \mathbf{T} = \mathbf{0}.$$

(d) From (c) show that $p = 0$. Thus $d\mathbf{B}/ds = q\mathbf{N}$. The scalar function q is usually denoted τ ("tau"). Thus $d\mathbf{B}/ds = \tau\mathbf{N}$.

28.[M] This exercise develops $d\mathbf{N}/ds$.

(a) Why are there scalars c and d such that $\frac{d\mathbf{N}}{ds} = c\mathbf{T} + d\mathbf{B}$?

(b) Using the fact that $\mathbf{B}$ and $\mathbf{N}$ are always perpendicular, show that $\tau\mathbf{N} \cdot \mathbf{N} + \mathbf{B} \cdot (c\mathbf{T} + d\mathbf{B}) = 0$.

(c) From (b) show that $d = -\tau$.

(d) Similarly, starting with $\mathbf{T} \cdot \mathbf{N} = 0$, show that $c = -\kappa$. Thus $\frac{d\mathbf{N}}{ds} = -\kappa\mathbf{T} - \tau\mathbf{N}$.

29.[M] You are swinging a pail of water at the end of a rope. You slowly increase the amount of rope until the radius of the circle the pail sweeps out doubles. Does the force of your pull remain the same? Increase? Decrease? Explain.

30.[M] In Example 1 we used calculus to show that for a plane curve $|d\mathbf{T}/ds| = |d\phi/ds|$, when ϕ is the angle that $\mathbf{T}$ makes with the x-axis. This suggests that for small values of Δs, $|\Delta\phi| = |\phi(s + \Delta s) - \phi(s)|$ is a good approximate of $|\Delta\mathbf{T}| = |\mathbf{T}(s + \Delta s) - \mathbf{T}(s)|$.

(a) Draw $\mathbf{T}(s + \Delta s)$ and $\mathbf{T}(s)$ with their tails at the origin.

(b) Using the diagram in (a), show why you would expect $|\Delta\mathbf{T}|$ and $|\Delta\phi|$ to be close to each other in the sense that $|\Delta\mathbf{T}/\Delta\phi|$ would be near 1.

31.[C] Show that a curve that has a constant curvature $\kappa = 0$ is a line. HINT: Start with the definition, $\kappa = |d\mathbf{T}/ds|$. NOTE: Don't say, "Oh, it's a curve with infinite radius of curvature. So it must be a line".

32.[C] Express $d\mathbf{T}/ds$ in terms of the curvature and $\mathbf{N}$.

33.[C]

Sarah: I don't like the way the authors got the formula for curvature. I'm sure they didn't need to drag in the components of the acceleration vector. It's not elegant.

Sam: They're trying to save space. Calculus books are too long.

Sarah: My way is neat and short: just calculate

$$\left\|\frac{d\mathbf{T}}{ds}\right\| = \frac{\|\frac{d\mathbf{T}}{dt}\|}{\|\mathbf{v}\|}.$$

To begin I write $\mathbf{T}$ as $\mathbf{v}/\|\mathbf{v}\|$. Then I just differentiate the quotient $\mathbf{v}/\|\mathbf{v}\|$. Along the way I'll need $d\|\mathbf{v}\|/dt$, but I get that by differentiating $\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$. That will give me

$$\frac{d\mathbf{T}}{ds} = \frac{v^2\mathbf{a} - (\mathbf{a}\cdot\mathbf{v})\mathbf{v}}{|v|^3} \tag{15.2.6}$$

Sam: That's a nice formula but its not got the cross product.

Sarah: If you like cross products, then use (15.2.6) to find $(d\mathbf{T})/ds \cdot d\mathbf{T}/ds$ and call on that identity that appeared when getting the length of the cross product $\|\mathbf{A}\times\mathbf{B}\|$ (page 1155). I'll let you fill in the steps. I don't want to deprive you of a little fun.

Fill in all the missing steps.

15.3 Line Integrals and Conservative Vector Fields

In Section 6.2, we defined the integral of a function $f(x)$ over an interval $[a, b]$ as the limit of sums of the form $\sum_{i=1}^{n} f(c_i)\Delta x_i$. Now we use similar definitions for integrals over curves. In the next section we apply these concepts to work, fluid flow, and geometry.

The Integral with Respect to Arc Length s

Let $\mathbf{r}(t)$ be the position vector corresponding to a parameter value t in $[a, b]$. Assume that $\mathbf{r}(t)$ sweeps out a curve C with a continuous unit tangent vector $\mathbf{T}(t)$. Let f be a scalar-valued function defined at least on C. We will define the integral of f over C with respect to arc length.

Partition $[a, b]$ by $t_0 = a$, t_1, $\dots$, $t_n = b$ and let $\mathbf{r}(t_0) = \overrightarrow{OP_0}$, $\mathbf{r}(t_1) = \overrightarrow{OP_1}$, $\mathbf{r}(t_n) = \overrightarrow{OP_n}$ be the corresponding position vector as shown in Figure 15.3.1. The points P_0, P_1, $\dots$, P_n break the curve into n shorter curves of lengths Δs_1, Δs_2, $\dots$, Δs_n. Then form the Riemann sum

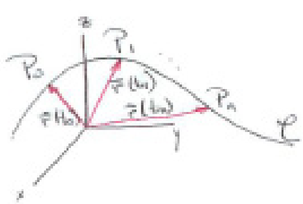


Figure 15.3.1:

$$\sum_{i=1}^{n} f(P_i)\Delta s_i \tag{15.3.1}$$

The limit of sums of the form (15.3.1) as all the lengths Δs_i are chosen smaller and smaller is denoted $\int_C f(P)ds$. That is,

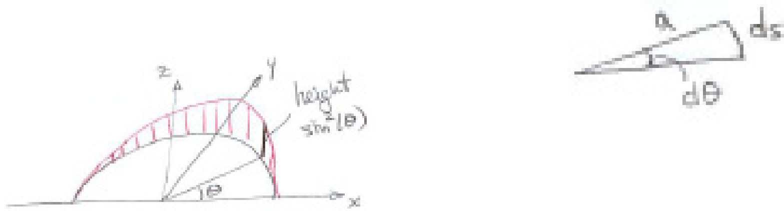
$$\int_C f(P)\ ds = \lim_{\Delta s \to 0} \sum_{i=1}^{n} f(P_i)\Delta s_i.$$

This limit does not depend on the particular parameterization. In particular it does not depend on the direction in which the curve is swept out. For computational purposes we have

$$\int_C f(P)\ ds = \int_a^b f(P)\left|\frac{ds}{dt}\right| dt$$

EXAMPLE 1 A fence is built as a semi-circle of radius a with center at the origin. At the point on the circle of angle θ, its height is $\sin^2(\theta)$. What is the area of one side of the fence?

SOLUTION Let $f(P)$ be the height of the fence at $P = (r, \theta)$ in polar coordinates. (See Figure 15.3.2(a). Then $f(r, \theta) = \sin^2(\theta)$. Let θ be the



(a) (b)

Figure 15.3.2:

parameter, θ in $[0, \pi]$. Let $s = a\theta$ be the arc length subtended by the angle θ, as in Figure 15.3.2(b). Then $ds = ad\theta$ and we have

$$\text{Area} = \int\limits_C \sin^2(\theta)\ ds \quad = \quad \int\limits_0^{\pi} \sin^2(\theta) a\ d\theta = 2\int\limits_0^{\pi/2} \sin^2(\theta) a\ d\theta = 2a\frac{\pi}{4} = \frac{\pi a}{2}.$$

◇

The Integrals with Respect to x, y, or z

The integral with respect to arc length is so similar to the integral over an interval that it presents little novelty. However, the integrals with respect to x, y, or z are quite different.

As before, we start with a parameterized curve C and a scalar function f defined at least on C. We divide the interval $[a, b]$ into n sections by $t_0 = a, t_1, \cdots, t_n = b$. For convenience, take the sections to be of equal lengths.

Let $\mathbf{r}(t_i) = \langle x(t_i), y(t_i), z(t_i)\rangle$.

Instead of considering the arc lengths Δs_i of each short interval *we consider instead the change in the x coordinate, $x_{i+1} - x_i = \Delta x_i$. This change can be positive or negative.* We then make the following definition.

The integral of f over the curve C with respect to x is the limit of sums of the form

$$\sum_{i=1}^{n} f(x(t_i), y(t_i), z(t_i))\Delta x_i$$

as n approaches infinity. It is denoted

$$\int_C f\ dx \quad \text{or} \quad \int_C f(x, y, z)\ dx \quad \text{or} \quad \int_C f(P)\ dx$$

For computational purposes, when C is parameterized, $\int_C f\ dx$ is expressed as $\int_a^b f(x(t), y(t), z(t))\frac{dx}{dt}\ dt$.

In contrast to an integral with respect to arc length, *the value of* $\int_C f(P)\ dx$ *depends on the orientation in which the curve is swept out.* If we reverse the orientation, the expression $x_{i+1} - x_i$ changes sign. For instance, if x is an increasing function of the parameter in one parameterization, then $\Delta x_i = x_{i+1} - x_i$ is positive; but in the reverse orientation x is a decreasing function of the parameter, so $\Delta x_i = x_{i+1} - x_i$ is negative.

If $-C$ denotes the curve C swept out in the opposite orientation, then

$$\int_{-C} f(P)\ dx = -\int_C f(P)\ dx$$

In any case, it is necessary to pay attention to the orientation of C.

If $\mathbf{r}(a) = \mathbf{r}(b)$, that is, the finish is the same as the start, the curve is called **closed**. If the curve does not intersect itself except, perhaps, at its endpoints, we call the curve **simple**. These ideas are independent. A curve can be neither closed nor simple, closed but not simple, simple but not closed, or both simple and closed. (See Figure 15.3.3.)

closed curve

simple curve

(a) (b) (c) (d)

Figure 15.3.3: ARTIST: Please replace **G** with **r**, throughout.

When C is a closed curve we will usually use the notation $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for a line integral.

EXAMPLE 2 A smooth closed convex curve C is situated in the first quadrant, as shown in Figure 15.3.4. Find $\oint_C y\, dx$ if the curve is oriented counterclockwise.

SOLUTION Let A and B be the contact points of the two vertical tangents to C. Break C into a lower curve C_1 from A to B and an upper curve C_2 from B to A, both swept out counterclockwise.

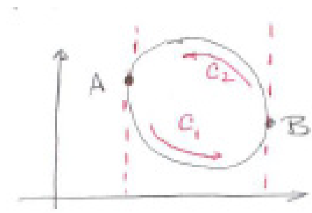
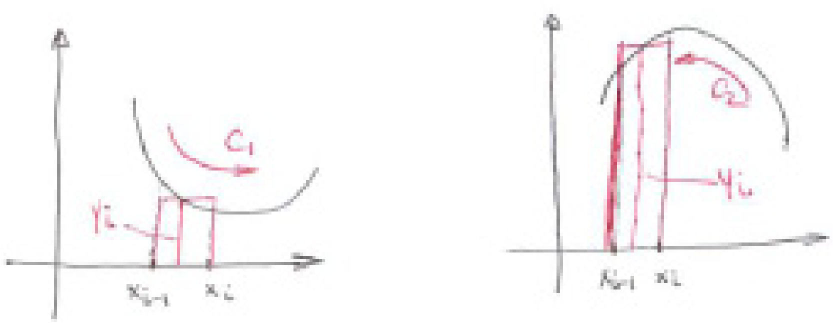


Figure 15.3.4:

First we interpret $\int_{C_1} y\, dx$. On C_1, $\Delta x_i = x_{i+1} - x_i$ is positive. Let y_i be the y-coordinate of some point on C_1 above $[x_i, x_{i+1}]$. Then $(x_{i+1} - x_i)y_i$ approximates the area under C_1 and above $[x_i, x_{i+1}]$, as shown in Figure 15.3.5(a)



(a) (b)

Figure 15.3.5:

We may think of "$y\, dx$" as the local approximation to the area under C_1. Thus

$$\int_{C_1} y\, dx = \text{area below } C_1 \text{ and above the } x\text{-axis.}$$

On C_2, x is a decreasing function of the parameter. Now $\Delta x_i = x_{i+1} - x_i$ is negative, as Figure 15.3.5(b) suggests.

Again let y_i be the y-coordinate of a point on the curve C_2 above the interval whose ends are x_i and x_{i+1}. Now $(x_{i+1} - x_i)y_i$ is the negative of an approximation of the area below C_2 and above the x-axis. We conclude that

$$\int_{C_1} y\, dx = \text{negative of the area below } C_2 \text{ and above the } x\text{-axis.}$$

Since $\int_C y\ dx = \int_{C_1} y\ dx + \int_{C_2} y\ dx$, it follows that when C is oriented counterclockwise

$$\int\limits_C y\ dx = \text{negative of the area inside } C.$$

◇

The integrals of f over a curve C with respect to y or z are defined similarly and denoted

$$\int\limits_C f(x,y,z)\ dy \quad \text{and} \quad \int\limits_C f(x,y,z)\ dz.$$

The integrals with respect to x, y, or z are called **line integrals**. It is more natural to call them curve integrals, but by tradition, they are known as line integrals.

The most general line integral is the sum of the three types,

$$\int\limits_C \left(P(x,y,z)\ dx + Q(x,y,z)\ dy + R(x,y,z)\ dz\right). \tag{15.3.2}$$

The "integrand" for this line integral, $Pdx + Qdy + Rdz = \mathbf{F} \cdot d\mathbf{r}$, is sometimes referred to as a **differential form**. This language will be encountered again in Chapter 18.

Of course, it we are dealing only with curves in the xy-plane, then the most general line integral would be

$$\int\limits_C \left(P(x,y,z)\ dx + Q(x,y,z)\ dy\right). \tag{15.3.3}$$

Both (15.3.2) and (15.3.3) are easily expressed in the compact language of vectors.

A **vector field** assigns a vector to each point in some region of space (or the plane). The use of the term "field" instead of "function" is in deference to physicists and engineers, who speak of "magnetic field" and "electric field," both of which are examples of vector fields.

By comparison, a function that assigns a scalar (real number) to each point in a region in space (or the plane) is called a **scalar field**. The function that assigns the temperature at a point in space is a scalar function; so is the function that describes the density at a point.

The typical vector field in space is $\mathbf{F}(x,y,z) = P(x,y,z)\mathbf{i} + Q(x,y,z)\mathbf{j} + R(x,y,z)\mathbf{k}$ where $P(x,y,z)$, $Q(x,y,z)$, and $R(x,y,z)$ are scalar fields. A vector field $\mathbf{F}$ in the plane is described by two scalar fields: $\mathbf{F}(x,y) = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$.

To take advantage of vector notation, we write the formal vector $d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$. Then (15.3.2) becomes simply

$$\int_C \mathbf{F}(x, y, z) \cdot d\mathbf{r} \qquad \text{or} \qquad \int_C \mathbf{F}(\mathbf{r}) \cdot d\mathbf{r}.$$

Or, even simpler yet,

$$\int_C \mathbf{F} \cdot d\mathbf{r}.$$

In this setting, the dot product $\mathbf{F} \cdot d\mathbf{r}$ is called a **differential term**.

For computational purposes, this may be written as

$$\int_a^b \mathbf{F}(x(t), y(t), z(t)) \cdot \frac{d\mathbf{r}}{dt}\, dt$$

or simply

$$\int_a^b \mathbf{F} \cdot \frac{d\mathbf{r}}{dt} dt.$$

Line integrals in the plane, such as (15.3.3), would be expressed in exactly the same way.

Another standard notation uses the unit vector $\mathbf{T} = \frac{d\mathbf{r}}{ds}$. Writing $d\mathbf{r}$ as $\mathbf{T}ds$ we rewrite $\int_C \mathbf{F} \cdot d\mathbf{r}$ as $\int_C \mathbf{F} \cdot \mathbf{T}\, ds$. (Why is $\frac{d\mathbf{r}}{ds} = \mathbf{T}$?)

The integrand depends on the orientation of the curve because switching the orientation changes $\mathbf{T}$ to $-\mathbf{T}$.

The next example shows that different paths with the same initial point and terminal point may yield different integrals.

EXAMPLE 3 Let C_1 be the path from $(1, 0)$ to $(0, 1)$ along the unit circle with center at the origin. Let C_2 be the path that starts at $(1, 0)$, goes to $(1, 1)$ on the line $x = 1$, and then to $(0, 1)$ on the line $y = 1$. Compute $\int_{C_1} xy\, dx$ and $\int_{C_2} xy\, dx$.

SOLUTION Figure 15.3.6 shows the two paths C_1 and C_2, together with two more curves, C_3 and C_4, that also will be used.

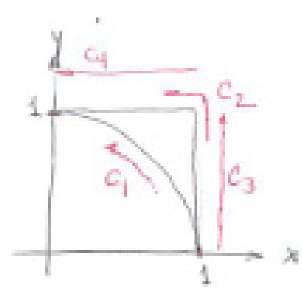


Figure 15.3.6:

To compute $\int_{C_1} xy\, dx$, we parameterize the circle by angle θ in $[0, \pi/2]$. Thus $x = \cos(\theta)$ and $y = \sin(\theta)$ and $dx = \frac{dx}{d\theta} d\theta = -\sin(\theta)\, d\theta$

$$\begin{aligned} \int_{C_1} xy\, dx &= \int_0^{\pi/2} (\cos(\theta))(\sin(\theta))(-\sin(\theta))\, d\theta \\ &= -\int_0^{\pi/2} \sin^2(\theta)\cos(\theta)\, d\theta = -\left.\frac{\sin^3(\theta)}{3}\right|_0^{\pi/2} = -\frac{1}{3}. \end{aligned}$$

We used $u = \sin(\theta)$ to find an antiderivative of $\sin^2(\theta)\cos(\theta)$.

To calculate $\int_{C_2} xy\ dx$ we break C_2 into two straight paths: C_3 from $(1,0)$ to $(1,1)$ and C_4 from $(1,1)$ to $(0,1)$. (See Figure 15.3.6.)

On C_3, $x = 1$ and $dx = 0$. Thus $\int_{C_3} xy\ dx = 0$.

On C_4, $y = 1$ and x begins at 1 and ends at 0. A standard parameterization of C_4 is $x = 1-t$, $y = 1$ for $0 \le t \le 1$. Then

$$\int_{C_4} xy\ dx = \int_0^1 (1-t)(1)(-dt) = \int_0^1 (t-1) = \left.\frac{t^2}{2} - t\right|_0^1 = -\frac{1}{2}.$$

On C_4 we could have used the parameter x itself, which starts at 1 and goes down to 0. In that case we would have $\int_{C_4} xy\ dx = \int_1^0 x\ dx = \left.\frac{x^2}{2}\right|_1^0 = -\frac{1}{2}$.

Thus on the path C_2 made up of C_3 followed by C_4 we have $\int_{C_2} xy\ dx = 0 + (-1/2) = -1/2$.

The line integrals $\int_{C_1} xy\ dx$ and $\int_{C_2} xy\ dx$ are not equal even though they start at the same point $(1,0)$ and end at the same point $(0,1)$ and have the same integrand. ◇

As Example 3 shows, $\int_C xy\ dx$ is not determined by the end points of the curve C. This raises a question: Which line integrals $\int_C (P\ dx + Q\ dy + R\ dz)$ depend only on the end points of C?

EXAMPLE 4 Compute $\int_C \frac{x\ dx + y\ dy}{x^2+y^2}$ on the two paths, C_1 and C_2 in Example 3.

SOLUTION On the circular path C_1 we use θ as a parameter and have

$$\int_{C_1} \frac{x\ dx + y\ dy}{x^2+y^2} = \int_0^{\pi/2} \frac{(\cos(\theta))(-\sin(\theta)\ d\theta) + \sin(\theta)(\cos(\theta))\ d\theta}{\cos^2(\theta) + \sin^2(\theta)} = \int_0^{\pi/2} \frac{0}{1} d\theta = 0.$$

Next we compute the integral on the linear path from $(1,0)$ to $(1,1)$ to $(0,1)$. The path from $(1,0)$ to $(1,1)$ is C_3. There $x = 1$, so $dx = 0$. Therefore, using y itself as the parameter, we find that

$$\begin{aligned}\int_{C_3} \frac{x\ dx + y\ dy}{x^2+y^2} &= \int_{C_3} \frac{1\cdot 0 + y\ dy}{1+y^2} = \int_{C_3} \frac{y}{1+y^2}\ dy \\ &= \int_0^1 \frac{y\ dy}{1+y^2} = \left.\frac{\ln(1+y^2)}{2}\right|_0^1 = \frac{\ln 2}{2}.\end{aligned}$$

On the path C_4, from $(1,1)$ to $(0,1)$, we use x as the parameter starting at

$x = 1$ and $y = 1$, so $dy = 0$ and we have

$$\int\limits_{C_4} \frac{x\ dx + y\ dy}{x^2+y^2} = \int\limits_1^0 \frac{x\ dx}{x^2+1} \quad = \quad \left.\frac{\ln(x^2+1)}{2}\right|_1^0 = -\frac{\ln 2}{2}.$$

Thus $\int_{C_2} \frac{x\ dx + y\ dy}{x^2+y^2} = -\frac{\ln 2}{2} + \frac{\ln 2}{2} = 0$. This is the same value as the integral over the circular arc C_1. ◇

In Section 18.1 we will show that $\int_C \frac{x\ dx+y\ dy}{x^2+y^2}$ depends only on the end points of C. That is, if C_1 and C_2 are any two curves from point A to point B then

$$\int\limits_{C_1} \frac{x\ dx + y\ dy}{x^2+y^2} = \int\limits_{C_2} \frac{x\ dx + y\ dy}{x^2+y^2}.$$

However, as Example 3 shows, $\int_C xy\ dx$ does depend on the particular curve C joining two points.

An expression of the form $P(x,y,z)\ dx + Q(x,y,z)\ dy + R(x,y,z)\ dz$ is called **conservative** if its line integrals depend only on the endpoints of the curves over which the integration takes place. For instance,

$$\frac{x}{x^2+y^2}\ dx + \frac{y}{x^2+y^2}\ dy + 0\ dz$$

is conservative. Better yet, in Section 18.6 we will develop a simple criterion for determining whether a form $P\ dx + Q\ dy + R\ dz$ is conservative. In applications, such as gravity, conservative expressions are much easier to work with.

Summary

We defined four integrals for curves in space (three for curves in the xy plane). The first, $\int_C (f(P) ds$, is the limit of sums of the form $\sum_{i=1}^n f(P_i)\Delta s_i$, which is an integral defined in Chapter 6. The other three integrals $\int_C f(P)\ dx$, $\int_C f(P)\ dy$, $\int_C f(P)\ dz$ are quite different. For instance, the first is the limit of sums of the form $\sum_{i=1}^n f(P_i)\Delta x_i$, where x is the x-coordinate of a point on the curve. Putting these three together we have the general line integral

$$\int\limits_C (P(x,y,z)\ dx + Q(x,y,z)\ dy + R(x,y,z)\ dz) = \int\limits_C (P\ dx + Q\ dy + R\ dz).$$

Introducing the vector field $\mathbf{F}(x,y,z) = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, we developed three compact notations for a line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$, $\int_C \mathbf{F}\cdot \mathbf{r}'dt$, and $\int_C \mathbf{F}\cdot \mathbf{T}\ ds$.

EXERCISES for Section 15.3 *Key:* R–routine, M–moderate, C–challenging

1.[R] Following the approach in Example 2, show that if C were oriented clockwise, then $\oint_C y\ dx$ would equal the area inside C.

2.[R] Let C in Example 2 be oriented counterclockwise. Show why $\oint_C x\ dy$ equals the area inside C.

3.[R] Show that the area within a convex curve C is $\frac{1}{2}\oint_C (x\ dy - y\ dx)$ if C is oriented counterclockwise.

4.[R] (See Example 3.) Compute $\int_C xy\ dx$ on the path that goes from $(1, 0)$ directly to $(0, 0)$, and then directly on to $(0, 1)$.

5.[R] If $\mathbf{F}(P)$ is perpendicular to the curve C at every point P on C, what is $\int_C \mathbf{F} \cdot d\mathbf{r}$?

6.[R] If $\mathbf{F}(P)$ equals $\mathbf{T}(P)$ for P on the curve C, what is $\int_C \mathbf{F} \cdot d\mathbf{r}$?

7.[R] Let a and b be positive numbers. Let C be the curve bounding the rectangle with vertices $(0, 0)$, $(a, 0)$, (a, b), and $(0, b)$, where a and b are positive numbers. By calculating $\oint_C x\ dy$ with C oriented counterclockwise, confirm the result of Example 2. That is, check that the line integral over the closed curve C equals the area of the rectangle.

8.[R] Let a and b be positive numbers. Let C be the curve bounding the triangle with vertices $(0, 0)$, $(a, 0)$, and $(0, b)$. By calculating $\oint_C y\ dx$ with C oriented clockwise, show that the integral equals the area of the triangle.

9.[R] Let C be the curve bounding the circle of radius a with center at the origin. By calculating $\int_C x\ dy$ counterclockwise, check that the integral equals the area of the circle.

10.[R] Let $\mathbf{F}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = \mathbf{r}$. Let C be any curve starting at (x_0, y_0, z_0) and ending at (x_1, y_1, z_1). Calculate $\int_C \mathbf{F} \cdot d\mathbf{r}$ by rewriting it as $\int_a^b (\mathbf{F} \cdot \mathbf{r}')\ dt$. Note that $\int_C \mathbf{F} \cdot d\mathbf{r}$ depends only on the endpoints of C.

In Exercises 11 to 14, sketch the curve described by the given parameterization and label its start and finish.

11.[R] $\mathbf{r}(t) = t\mathbf{i} + t^2\mathbf{j}$, t in $[0, 1]$.

12.[R] $\mathbf{r}(t) = (1 - t)\mathbf{i} + (1 - t)^2\mathbf{j}$, t in $[0, 1]$.

13.[R] $\mathbf{r}(t) = (2t + 1)\mathbf{i} + 3t\mathbf{j}$, t in $[0, 2]$.

14.[R] $\mathbf{r}(t) = 4\cos t\mathbf{i} + 5\sin t\mathbf{j}$, t in $[0, 1]$.

In Exercises 15 to 18, parameterize the given curve with the indicated orientation.

15.[R] Figure 15.3.7(a)

16.[R] Figure 15.3.7(b)

17.[R] Figure 15.3.7(c)

18.[R] Figure 15.3.7(d)

(a) (b) (c) (d)

Figure 15.3.7:

In Exercises 19 to 22, evaluate the given line integrals.

19.[R] $\int_C xy\ dx$, where C is the straight line from $(1,1)$ to $(3,3)$.

20.[R] $\int_C x^2\ dy$, where C is the straight line from $(2,0)$ to $(2,5)$.

21.[R] $\int_C x^2\ dy$, where C is the straight line from $(3,2)$ to $(7,2)$.

22.[R] $\int_C (xy\ dx + x^2\ dy)$, where C is the straight line from $(1,0)$ to $(0,1)$.

In Exercises 23 to 26 evaluate with minimum effort. C is a counterclockwise curve bounding a region of area 5.

23.[R] $\oint_C 3y\ dx$

24.[R] $\oint_C (2y\ dx + 6x\ dy)$

25.[R] $\oint_C [2x\ dx + (x+y)\ dy]$

26.[R] $\oint_C [(x+2y+3)\ dx + (2x-3y+4)\ dy]$

In Exercises 10, the value of the line integral depends only on the endpoints, not on the particular path that joins them. Exercises 27 and 28 are examples where the path matters.

27.[R] Evaluate $\int_C (xy\ dx + 7\ dy)$ on

(a) the straight path from $(1,1)$ to $(2,4)$;

(b) the path from $(1,1)$ to $(2,4)$ that lies on the parabola $y = x^2$.

28.[R] Evaluate $\int_C x\ dy$ on

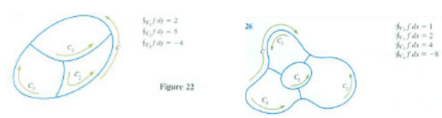
(a) the straight path from $(0,0)$ to $(\pi/2,1)$;

(b) the path from $(0,0)$ to $(\pi/2,1)$ that lies on the curve $y = \sin(x)$.

In Exercises 29 and 30, the values of certain line integrals are given for curves oriented as shown. Use this information to find $\int_C f\ dy$. HINT: Pay attention to the orientations. carefully.)

29.[R] Figure 15.3.8(a)

30.[R] Figure 15.3.8(b)



(a) (b)

Figure 15.3.8:

31.[M] Let the closed curve C bound the region $\mathcal{R}$, which is broken into regions $\mathcal{R}_i$, $i \leq i \leq n$, and each $\mathcal{R}_i$ is bounded by its own C_i. Let $\mathbf{F}$ be a vector field in $\mathcal{R}$. If all the $n+1$ curves are swept out counterclockwise, show that $\oint_C \mathbf{F}\cdot d\mathbf{r} = \sum_{i=1}^{n} \oint_{C_i} \mathbf{F}\cdot d\mathbf{r}$.

32.[M] Show that $\int_C \frac{x\ dx}{s^2+y^2}$ is not conservative by calculating $\int_C \frac{x\ dx}{s^2+y^2}$ on two paths joining $(1,0)$ to $(1,1)$ for which the integrals are not equal.

33.[M] Let k be a constant. Show that $\oint_C k\ dy = 0$.

34.[C] Let $\mathbf{r} = \mathbf{r}(t)$ describe a curve C in the plane or in space. What is the geometric interpretation of

$$\frac{1}{2}\int_C \|\mathbf{r} \times \mathbf{T}\|\ ds?$$

NOTE: See also CIE 18.

35.[C] If t represents time and $\mathbf{r}(t)$ describes a curve C, what is the meaning of $\int_C \mathbf{T}\cdot d\mathbf{r}$? HINT: Draw a picture of a small section of the curve.

15.4 Four Applications of Line Integrals

In the previous section we defined line integrals and showed that $\oint_C y\,dx$ and $\oint_C x\,dy$ in the plane are related to the area of the region bounded by the closed curve C. In this section we show how line integrals occur in the study of work, fluid flow, and in the angle subtended by a planar curve.

Each application will be developed by following the same basic idea as we used when defining definite integrals: divide the domain into smaller pieces, approximate the quantity on each piece, add the contribution form each piece, and take a limit as the pieces get smaller and smaller.

Work Along a Curve

Consider a force $\mathbf{F}$ that remains constant (in direction and magnitude) and pushes a particle in a straight line from A to B. Let $\mathbf{R} = \overrightarrow{AB}$. The work accomplished by $\mathbf{F}$ is defined as $\mathbf{F} \cdot \mathbf{R}$:

$$\text{Work} = \mathbf{F} \cdot \mathbf{R}.$$

This is the product of the scalar component of $\mathbf{F}$ in the direction of $\mathbf{R}$ and the distance the particle moves. (See Figure 15.4.1)

But what if the force $\mathbf{F}$ varies and pushes the particle along a curve that is not straight? (See Figure 15.4.2(a).)

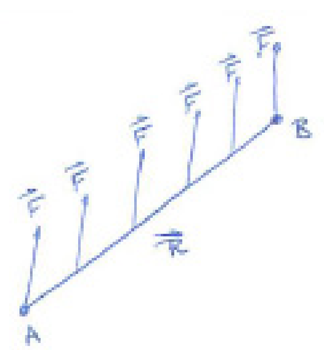


Figure 15.4.1:

(a) (b)

Figure 15.4.2:

Let's follow the process used in Section 15.3 that led to line integrals. Assume the curve, C, is parameterized by $\mathbf{r}(t)$ for t in $[a, b]$. Partition $[a, b]$ by $t_0 = a$, t_1, ..., $t_n = b$ and let $\mathbf{r}(t_0) = \overrightarrow{OP_0}$, $\mathbf{r}(t_1) = \overrightarrow{OP_1}$, ..., $\mathbf{r}(t_n) = \overrightarrow{OP_n}$, be the corresponding position vectors. (See Figure 15.4.2(b).) The points P_0, P_1, ..., P_n break the curve into n shorter curves. The work done by $\mathbf{F}$ along C between P_i and P_{i+1} is approximately $\mathbf{F}(\mathbf{r}(t_i)) \cdot \Delta\mathbf{r}_i$ where $\Delta\mathbf{r}_i = \overrightarrow{P_iP_{i+1}}$. The total work done by $\mathbf{F}$ along C is approximated by

$$\sum_{i=1}^{n} \mathbf{F}(\mathbf{r}(t_i)) \cdot \Delta\mathbf{r}_i. \qquad (15.4.1)$$

Taking the limit as $\max_i \|\Delta\mathbf{r}_i\|$ approaches 0, we conclude

$$\text{Work done by } \mathbf{F} \text{ along } C \text{ is } \int_C \mathbf{F}\cdot\ d\mathbf{r}. \qquad (15.4.2)$$

If $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$, where P and Q are functions of x and y, and $d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j}$, then

$$\text{Work done by } \mathbf{F} = P\mathbf{i} + Q\mathbf{j} \text{ along } C \text{ is } \int_C (P\ dx + Q\ dy).$$

Physicists and engineers commonly use (15.4.2) as a starting point when expressing work.

The vector notation $\mathbf{F}\cdot d\mathbf{r}$ is far more suggestive than the scalar notation $P\ dx + Q\ dy$. It reminds us that "work is the dot product of force and displacement." That implies that only the component of the force in the direction of motion accomplishes work.

EXAMPLE 1 How much work is accomplished by the force $\mathbf{F}(x,y) = xy\mathbf{i} + y\mathbf{j}$ in pushing a particle from $(0,0)$ to $(3,9)$ along the parabola $y = x^2$?

SOLUTION

Figure 15.4.3 shows the path of the particle. Call this path C. Then

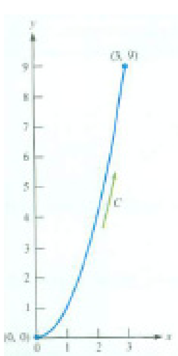


Figure 15.4.3:

$$\text{Work} \;=\; \int_C \mathbf{F}\cdot d\mathbf{r} = \int_C (xy\mathbf{i} + y\mathbf{j})\cdot(dx\mathbf{i} + dy\mathbf{j}) = \int_C (xy\ dx + y\ dy).$$

To evaluate this line integral, let us use x as the parameter, with x in $[0,3]$. Then $y = x^2$ and $dy = 2x\ dx$, so

$$\int_C (xy\ dx + y\ dy) \;=\; \int_0^3 \left(x\cdot x^2\ dx + x^2(2x\ dx)\right) = \int_0^3 3x^3\ dx = \frac{243}{4}.$$

◇

Circulation of a Fluid

Consider a fluid (liquid or gas) flowing on a portion of the xy plane. Let its density and velocity at the point P be given by $\sigma(P)$ and $\mathbf{v}(P)$, respectively. The product

$$\mathbf{F}(P) = \sigma(P)\mathbf{v}(P)$$

represents the rate and direction of the flow of the fluid at P. Now put an imaginary closed wire loop C on the fluid as in Figure 15.4.4 or Figure 15.4.5 and keep it fixed. In Figure 15.4.4, C surrounds a whirlpool and there is a tendency for fluid to flow along C rather than across it. The opposite case is shown in Figure 15.4.5, where most of the fluid flow is *across* C rather than parallel to it. The component of $\mathbf{F}$ parallel to the tangent vector determines the tendency of the fluid to flow along C. Now, $\mathbf{F}\cdot d\mathbf{r}$ represents flow in the *direction* of $d\mathbf{r}$, a small section of the curve C. Thus

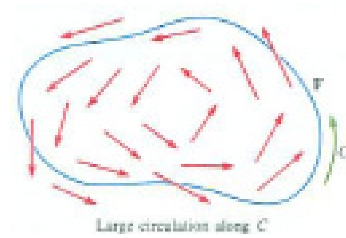


Figure 15.4.4:

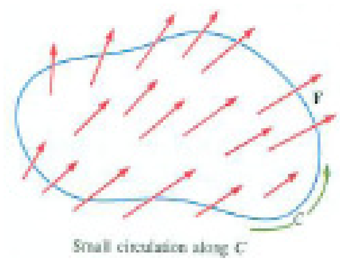


Figure 15.4.5:

$$\oint_C \mathbf{F}\cdot d\mathbf{r}$$

represents the tendency of the fluid to flow along C. If C is counterclockwise and $\oint_C \mathbf{F}\cdot d\mathbf{r}$ is positive, the flow of $\mathbf{F}$ would be counterclockwise as well. If $\oint_C \mathbf{F}\cdot d\mathbf{r}$ is negative, the flow would tend to be clockwise. The line integral $\oint_C \mathbf{F}\cdot d\mathbf{r}$ is called the **circulation** of $\mathbf{F}$ along C.

circulation

Note that the very same integral, $\oint_C \mathbf{F}\cdot d\mathbf{r}$, occurs in the study of work and in the study of fluids.

EXAMPLE 2 Find the circulation of the planar flow $\mathbf{F}(x,y) = xy\mathbf{i} + y\mathbf{j}$ around closed curve that follows $y = x^2$ from $(0,0)$ to $(3,9)$, then horizontally to $(0,9)$ and straight down to $(0,0)$.

SOLUTION The closed curve C comes in three parts: $C = C_1 + C_2 + C_3$ where C_1 is $y = x^2$ for $0 \le x \le 3$, $-C_2$ is $y = 9$, $0 \le x \le 3$, and $-C_3$ is $x = 0$, $0 \le y \le 9$.

The circulation is

$$\begin{aligned}\oint_C \mathbf{F}\cdot d\mathbf{r} &= \oint_{C_1} \mathbf{F}\cdot d\mathbf{r} + \oint_{C_2} \mathbf{F}\cdot d\mathbf{r} + \oint_{C_3} \mathbf{F}\cdot d\mathbf{r} \\ &= \oint_{C_1} \mathbf{F}\cdot d\mathbf{r} - \oint_{-C_2} \mathbf{F}\cdot d\mathbf{r} - \oint_{-C_3} \mathbf{F}\cdot d\mathbf{r}.\end{aligned}$$

Notice that we work with $-C_2$ and $-C_3$ because they are easier to parameterize than C_2 and C_3.

By Example 1, $\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \frac{243}{4}$. And, by direct calculations:

$$\oint_{-C_2} \mathbf{F} \cdot d\mathbf{r} = \int_0^3 \langle 9x, 9\rangle \cdot \langle dx, 0\rangle = \int_0^3 9x\ dx = \frac{81}{2}$$

and

$$\oint_{-C_3} \mathbf{F} \cdot d\mathbf{r} = \int_0^9 \langle 0, y\rangle \cdot < 0, dy\rangle = \int_0^9 y\ dy = \frac{81}{2}.$$

The circulation of $\mathbf{F}$ around C is $\oint_C \mathbf{F} \cdot d\mathbf{r} = \frac{243}{4} - \frac{81}{2} - \frac{81}{2} = \frac{-81}{4}$. ◇

Loss or Gain of a Fluid (Flux)

Imagine again that we place an imaginary wire loop C on the surface of a stream.

We raise the question: At what rate is fluid escaping or entering the region R whose boundary is C?

If the fluid tends to escape, then it is thinning out in R, becoming less dense at some points. If the fluid tends to accumulate, it is becoming denser at some points. (Think of this ideal fluid as resembling a gas rather than a liquid; gases can vary widely in density while liquids tend to have constant density.)

Since the fluid is escaping or entering R only along its boundary, it suffices to consider the total loss or gain across C. Where $\mathbf{v}$, the fluid velocity, is tangent to C, fluid neither enters nor leaves. Where $\mathbf{v}$ is not tangent to C, fluid is either entering or leaving across C, as indicated in

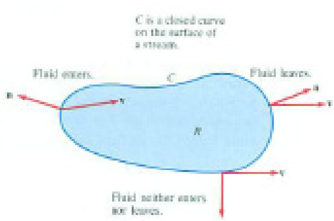


Figure 15.4.6:

The vector $\mathbf{n}$ is a unit vector perpendicular to the curve C and pointing away from the region it bounds. It is called the **exterior normal** or **outward normal**. Recall that $\mathbf{F} = \sigma\mathbf{v}$, the product of density and velocity, so $\mathbf{F}$ and $\mathbf{v}$ have the same direction.

To find the *total* loss or gain of fluid past C, let us look at a very short section of C, which we will vew as a vector $d\mathbf{r}$. How much fluid crosses $d\mathbf{r}$ in a short interval of time Δt.

During time Δt the fluid moves a distance $\|\mathbf{v}\|\Delta t$ across $d\mathbf{r}$. The fluid that crosses $d\mathbf{r}$ during the time Δt forms approximately the parallelogram shown in Figure 15.4.7.

The area of the parallelogram is the product of its height h and its base $\|d\mathbf{r}\|$. That is,

$$\text{Area of parallelogram} = \|\operatorname{proj}_{\mathbf{n}}(\mathbf{v}\Delta t)\|\|d\mathbf{r}\| = (\mathbf{v}\Delta t) \cdot \mathbf{n}\|d\mathbf{r}\| \quad (15.4.3)$$

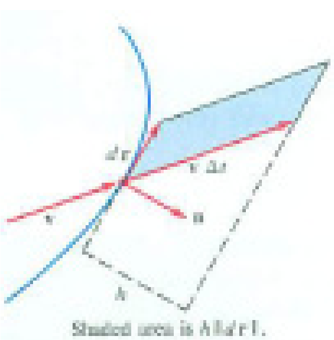


Figure 15.4.7:

Since the density of the fluid is σ,

$$\text{Mass in parallelogram} = \sigma(\mathbf{v}\Delta t)\cdot\mathbf{n}\|d\mathbf{r}\| = (\sigma\mathbf{v})\cdot\mathbf{n}\|d\mathbf{r}\|\Delta t = \mathbf{F}\cdot\mathbf{n}\|d\mathbf{r}\|\Delta t$$

Thus the rate at which fluid crosses $d\mathbf{r}$ per unit time is approximately

$$\frac{\mathbf{F}\cdot\mathbf{n}\|d\mathbf{r}\|\Delta t}{\Delta t} = \mathbf{F}\cdot\mathbf{n}\|d\mathbf{r}\|. \tag{15.4.4}$$

Since $d\mathbf{r}$ approximates a short piece of the curve, its length $\|d\mathbf{r}\|$ approximates the length ds of a short piece of the curve. Therefore, the rate at which the fluid crosses a short part of C, of length ds, is approximately

$$\mathbf{F}\cdot\mathbf{n}\; ds.$$

Hence the line integral

$$\oint_C \mathbf{F}\cdot\mathbf{n}\; ds$$

represents the rate of net loss or gain of fluid inside R. It it is positive, fluid tends to *leave* R, and the mass of fluid in R decreases. If it is negative, fluid tends to *enter* R, and the mass of fluid in R increases. In short,

Net loss or gain of fluid inside the region bounded by C is $\oint_R \mathbf{F}\cdot\mathbf{n}\; ds$.

The quantity $\oint_C \mathbf{F}\cdot\mathbf{n}\; ds$ is called the **flux** of $\mathbf{F}$ across C. So flux is "the integral of the normal component of $\mathbf{F}$." Circulation, $\oint_C \mathbf{F}\cdot d\mathbf{r}$, on the other hand, can be written as $\oint_C \mathbf{F}\cdot(\mathbf{T}\; ds)$, where $\mathbf{T}$ is the unit tangent vector in the direction of C. ($\mathbf{T}\; ds$ and $d\mathbf{r}$ have the same direction and same length ds, so they may be used interchangeably.) Hence

"Flux" comes from the Latin *fluxus* (flow), from which we also get "influx," "reflux" and "fluctuate," but, oddly, not "flow," which comes from the Latin *pluere* (to rain).

The circulation across C is $\oint_R \mathbf{F}\cdot\mathbf{T}\; ds$.

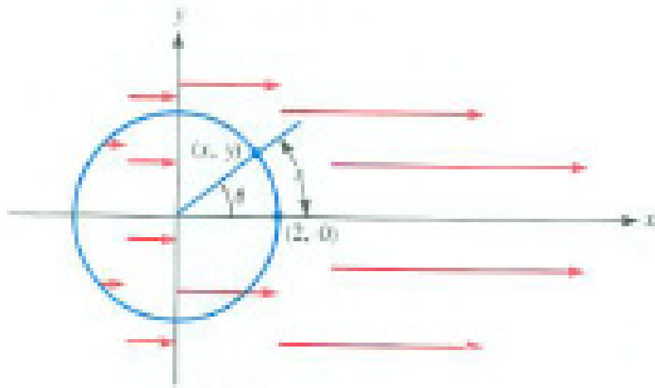


Figure 15.4.8:

Flux is the "integral of the normal component of **F**." Circulation is "the integral of the tangential component of **F**."

Stop! Before doing any calculations, what is your answer?

EXAMPLE 3 Let $\mathbf{F} = (2 + x)\mathbf{i}$ describe the flow of a fluid in the xy plane. Does the amount of fluid within the circle C of radius 2 and center $(0, 0)$ tend to increase or decrease?

SOLUTION Figure 15.4.8 shows the circle and a few of the vectors of **F**, calculated by the formula $\mathbf{F}(x, y) = (2 + x)\mathbf{i}$. Since the flow increases as we move to the right, there appears to be more fluid leaving the disk than entering. We expect the flux $\int_C \mathbf{F} \cdot \mathbf{n}\, ds$ to be positive. To compute $\int_C \mathbf{F} \cdot \mathbf{n}\, ds$, introduce angle θ as the parameter. Then

$$x = 2\cos(\theta), \qquad y = 2\sin(\theta) \qquad \text{for } 0 \leq \theta \leq 2\pi.$$

Since the circle has radius 2, $s = 2\theta$ and therefore

$$ds = 2d\theta.$$

The unit normal is parallel to the radius vector $x\mathbf{i} + y\mathbf{j}$. Therefore,

$$\mathbf{n} = \frac{x\mathbf{i} + y\mathbf{j}}{\|x\mathbf{i} + y\mathbf{j}\|} = \frac{2\cos(\theta)\mathbf{i} + 2\sin(\theta)\mathbf{j}}{2} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j},$$

which leads to the following calculation for the flux:

$$
\begin{aligned}
\text{Flux} &= \int_C \mathbf{F}\cdot\mathbf{n}\,ds = \int_0^{2\pi} \underbrace{[(2+x)\mathbf{i}\cdot\mathbf{n}]}_{\mathbf{F}\cdot\mathbf{n}}\,\underbrace{2d\theta}_{ds} \\
&= \int_0^{2\pi} (2+2\cos(\theta))\mathbf{i}\cdot(\cos(\theta)\mathbf{i}+\sin(\theta)\mathbf{j})2\;d\theta = \int_0^{2\pi}(4\cos(\theta)+4\cos^2(\theta))d\theta \\
&= \int_0^{2\pi}(4\cos(\theta)+2+2\cos(2\theta))d\theta = \left.(4\sin(\theta)+2\theta+\sin(2\theta))\right|_0^{2\pi} = 4\pi.
\end{aligned}
$$

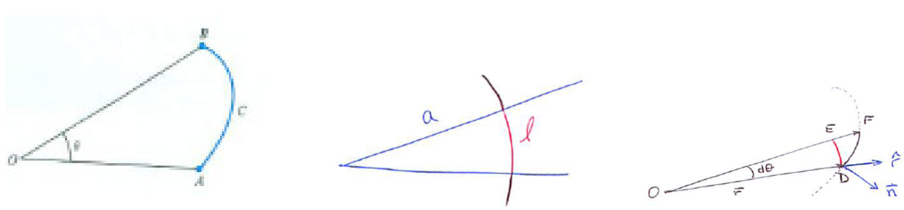
You can evaluate these definite integrals in your head. Why is $\int_0^{2\pi}\cos\theta d\theta = 0$ and $\int_0^{2\pi}\cos^2\theta d\theta = \frac{\pi}{2}$? HINT: $\int_0^{2\pi}\cos^2(\theta)\;d\theta = \int_0^{2\pi}\sin^2(\theta)\;d\theta$.

As expected, the flux is positive since there is a net flow out of the disk. ◇

The Angle Subtended by a Curve

Our fourth illustration of a line integral concerns the angle subtended at a point O by a curve C in the plane. (We assume that each ray from O meets C in at most one point.) We include this example as background for "the solid angle subtended by a surface," an important concept in Chapter 18.

The curve C in Figure 15.4.9(a) subtends an angle θ at the point O. We will show that θ can be expressed as a line integral of a suitable function. Of course, we do not need such an integral to find θ. Just knowing the points A, O, and B is enough. What is important is that θ can be expressed as a line integral. It is this idea that generalizes from a curve to a surface.



(a) (b) (c)

Figure 15.4.9:

First, recall the definition of radian measure of an angle whose vertex is at O, as in Figure 15.4.9(b). One draws a circle of any radius, say, a, with center at O. The angle intercepts an arc of length ℓ on the circle. The ratio ℓ/a is the radian measure of the angle.

To express θ in Figure 15.4.9(a) as an integral over the curve C we develop the "local estimate," $d\theta$, of the radians subtended by a short part of the curve, of length ds, as shown in Figure 15.4.9(c). Here, $\overline{DF}$ is part of the curve, and $\overline{DE}$ is part of the circle of radius r. Treating them as being almost straight, we have

$$DE \approx DF\ \cos(\widehat{\mathbf{r}}, \mathbf{n}) = DF\ \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{\|\widehat{\mathbf{r}}\|\|\mathbf{n}\|} = DF\ \widehat{\mathbf{r}} \cdot \mathbf{n} \approx \widehat{\mathbf{r}} \cdot \mathbf{n}\ ds.$$

Thus

$$d\theta = \frac{DE}{r} \approx \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r}\ ds.$$

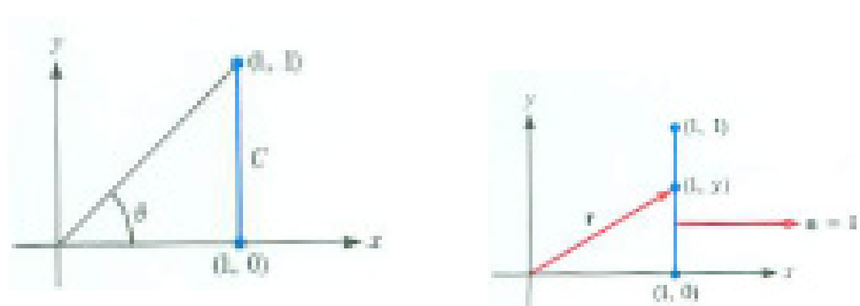
From this local estimate we conclude that

$$\text{The angle } \theta \text{ subtended by arc } C \text{ is } \int_C \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r}\ ds. \qquad (15.4.5)$$

Therefore, *the angle subtended by C is the integral with respect to arclength of the normal component of the vector function* $\widehat{\mathbf{r}}/\|\mathbf{r}\|$. In short, it is the flux of the vector field $\widehat{\mathbf{r}}/r$ (in the plane).

EXAMPLE 4 Verify (15.4.5) for the angle subtended at the origin by the line segment that joins $(1,0)$ and $(1,1)$.

SOLUTION The subtended angle θ is shown in Figure 15.4.10(a); obviously $\theta = \pi/4$.



(a) (b)

Figure 15.4.10:

Now let us evaluate the integral in (15.4.5) in this instance. Figure 15.4.10(b) shows that $\mathbf{n} = \mathbf{i}$ and $\mathbf{r} = \mathbf{i} + y\mathbf{j}$. Using $s = y$,

$$\begin{aligned}\theta &= \int_C \frac{\mathbf{n}\cdot\widehat{\mathbf{r}}}{\|\mathbf{r}\|}\,ds = \int_C \frac{\mathbf{i}\cdot\left(\frac{\mathbf{i}+y\mathbf{j}}{\sqrt{1+y^2}}\right)}{\sqrt{1+y^2}}\,ds = \int_C \frac{1}{1+y^2}\,ds \\ &= \int_0^1 \frac{1}{1+y^2}\,dy = \tan^{-1}(y)\Big|_0^1 = \frac{\pi}{4}.\end{aligned}$$

This agrees with our observation. ◇

Summary

Application	Work	Circulation	Flux	Angle Subtended
Integral	$\int_C \mathbf{F}\cdot\mathbf{T}\,ds = \int_C \mathbf{F}\cdot d\mathbf{r}$	$\oint_C \mathbf{F}\cdot\mathbf{T}\,ds$	$\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$	$\int_C \frac{\widehat{\mathbf{r}}\cdot\mathbf{n}}{r}\,ds$
Description	integral of tangential component of force $\mathbf{F}$ along C	integral of tangential component of flow $\mathbf{F}$ around closed curve C	integral of normal component of flow $\mathbf{F}$ along closed curve C	integral of normal component of $\widehat{\mathbf{r}}/r$ along C
Common Notations	$\int_C (P\ dx + Q\ dy)$ if $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$	$\int_C (Pdx+Qdy+Rdz)$ if $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$	$\oint_C (-Q\ dx+P\ dy)$ if $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$ (and C is oriented counterclockwise)	

EXERCISES for Section 15.4 *Key:* R–routine, M–moderate, C–challenging

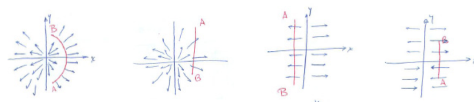
In Exercises 1 to 4 decide whether the work accomplished by the indicated vector field in moving a particle along the curve from A to B is positive, negative, or zero.

1.[R] Figure 15.4.11(a)

2.[R] Figure 15.4.11(b)

3.[R] Figure 15.4.11(c)

4.[R] Figure 15.4.11(d)



(a) (b) (c) (d)

Figure 15.4.11:

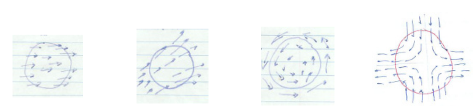
In Exercises 5 to 8 decide whether fluid is tending to leave, or enter or neither.

5.[R] Figure 15.4.12(a)

6.[R] Figure 15.4.12(b)

7.[R] Figure 15.4.12(c)

8.[R] Figure 15.4.12(d)



(a) (b) (c) (d)

Figure 15.4.12:

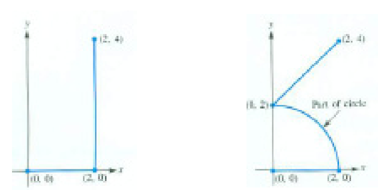
In Exercises 9 to 12 compute the work accomplished by the force $\mathbf{F} = x^2y\mathbf{i} + y\mathbf{j}$ along the given curve.

9.[R] From $(0,0)$ to $(2,4)$ along the parabola $y = x^2$.

10.[R] From $(0,0)$ to $(2,4)$ along the line $y = 2x$.

11.[R] From $(0,0)$ to $(2,4)$ along the path in Figure 15.4.13(a).

12.[R] From $(0,0)$ to $(2,4)$ along the path in Figure 15.4.13(b).



(a) (b)

Figure 15.4.13:

13.[R] Verify (15.4.5) for the angle subtended at the origin by the line segment that joins $(2,0)$ to $(2,3)$.

14.[R] Verify (15.4.5) for the angle subtended at the origin by the line segment that joins $(1,0)$ to $(0,1)$.

15.[R] Find the work done by the force $-3\mathbf{j}$ in moving a particle from $(0,3)$ to $(3,0)$ along

(a) The circle of radius 3 with center at the origin.

(b) The straight path from $(0,3)$ to $(3,0)$.

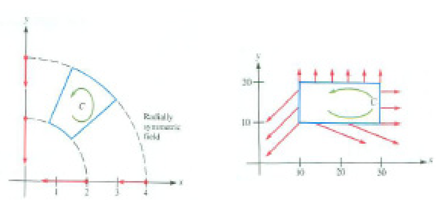
(c) The answers to (a) and (b) are the same. Will they by the same for all curves from $(0,3)$ to $(3,0)$?

16.[R] Figure 15.4.14(a) shows some representative vectors for the vector field $\mathbf{F}$ and curve C. Use this information to estimate

(a) the circulation of $\mathbf{F}$ along the boundary curve C and

(b) the flux of $\mathbf{F}$ across C.

(Since you have no formula for $\mathbf{F}$, there is a range of "correct" answers.)



(a) (b)

Figure 15.4.14:

17.[M] Repeat Exercise 16 for the vector field represented in Figure 15.4.14(b).

18.[M] The gravitational force $\mathbf{F}$ of the earth, which is located at the origin $(0,0)$ of a rectangular coordinate system, on a particle at the point (x,y) is

$$\frac{-x\mathbf{i}}{(\sqrt{x^2+y^2})^3}+\frac{-y\mathbf{j}}{(\sqrt{x^2+y^2})^3}=\frac{-\mathbf{r}}{\|\mathbf{r}\|^3}=\frac{-\widehat{\mathbf{r}}}{r^2}.$$

Compute the total work done by $\mathbf{F}$ if the particles goes from $(2,0)$ to $(0,1)$ along

(a) the portion of the ellipse $x = 2\cos(t)$, $y = \sin(t)$ in the first quadrant;

(b) the line parameterized as $x = 2 - 2t$, $y = t$.

19.[M]

(a) Let $W(b)$ be the work done by the force in Exercise 18 in moving a particle along the straight line from $(1, 0)$ to $(b, 0)$.

(b) What is $\lim_{b\to\infty} W(b)$?

20.[M] Let the vector field describing a fluid flow have at the point (x, y) the value $(x + 1)^2\mathbf{i} + y\mathbf{j}$. Let C be the unit circle described parametrically as $x = \cos(t)$, $y = \sin(t)$, for t in $[0, 2\pi]$.

(a) Draw $\mathbf{F}$ at eight convenient, equally spaced points on the circle.

(b) Is fluid tending to leave or enter the region bounded by C; that is, is the net outward flow positive or negative? NOTE: Answer on the basis of your diagram in (a).

(c) Compute the net outward flow with the aid of a line integral.

21.[M] Like Exercise 20 where $\mathbf{F}(x, y) = (2 - x)\mathbf{i} + y\mathbf{j}$ and C is the square with vertices $(0, 0)$, $(1, 0)$, $(1, 1)$, and $(0, 1)$.

22.[M] Let $\mathbf{F}(x, y) = \sigma\mathbf{v}$, the fluid flow, and C be a closed curve in the xy plane. If $\oint_C \mathbf{F} \cdot d\mathbf{r}$ is positive and C is counterclockwise, does the motion along C tend to be clockwise or counterclockwise?

23.[M] Let $\mathbf{F}(x, y) = \sigma\mathbf{v}$, the fluid flow, and C be a closed curve in the xy plane. If $\oint_C \mathbf{F} \cdot \mathbf{n}\ ds$ is positive, is fluid tending to leave the region bounded by C or to enter it?

24.[M] Let C be a closed convex curve that encloses the point O. Let $\mathbf{r}$ be the position vector $\overrightarrow{OP}$ for points P on the curve. Determine the value of $\oint_C (\widehat{\mathbf{r}} \cdot \mathbf{n})/r\ ds$, where $\mathbf{n}$ is the outward unit normal to C.

25.[M] Let C be a closed convex curve. Let O be a point not on C and not in the region C bounds. Let $\mathbf{r}$ be the position vector $\overrightarrow{OP}$ for points P on the curve. Determine the value of $\oint_C (\widehat{\mathbf{r}} \cdot \mathbf{n})/r\ ds$, where $\mathbf{n}$ denotes the external unit normal to

C. HINT: Draw a picture and pay attention to the angle between $\mathbf{n}$ and $\mathbf{r}$.

26.[C] Write up in your own words and diagrams why $\int_C \mathbf{F}\cdot d\mathbf{r}$ represents the work done by force $\mathbf{F}$ along the curve C.

27.[C] Write up in your own words and diagrams why $\int_C \mathbf{F}\cdot\mathbf{n}\ ds$ represents the net loss of fluid across C if $\mathbf{F}$ is the fluid flow and $\mathbf{n}$ is a unit external normal to C. Include the definition of $\mathbf{F}$.

28.[C] Explain why $\int_C \mathbf{F}\cdot d\mathbf{r}$ represents the tendency of a fluid to move along C, if $\mathbf{F}$ is the fluid flow.

29.[C] Explain why $\int_C (\widehat{\mathbf{r}}\cdot\mathbf{n})/r\ ds$ represents the angle subtended by a curve C at the origin. (Assume that each ray from the origin meets C at most once.)

30.[C] Let C be a curve in space and C^* its projection on the xy plane. Assume that distinct points of C project onto distinct points of C^*. The line integral $\int_C 1\ ds$ equals the arc length of C. What integral over C equals the arc length of C^*?

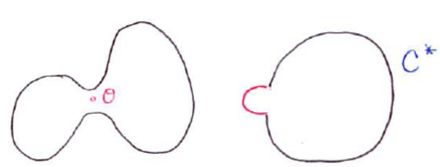
31.[C] Sam, Jane, and Sarah are debating a delicate issue.

Sam: Let C be the circle in the xy-plane whose polar equation is $r = 2\cos(\theta)$. It is a unit circle that passes through the origin O. Let $\mathbf{F}$ be the inverse first power central field $\widehat{\mathbf{r}}/r$. What is the flux of $\mathbf{F}$ across C?

Jane: The field blows up at O, so the flux is an improper integral.

Sam: Yes, but if I move C rigidly just a tiny bit so O is inside it, the flux is 2π. So I say the flux across C is 2π.

Sarah: I say it's π. Just draw a figure 8 made of two copies of C joined smoothly to form one curve, as in Figure 15.4.15(a).



(a) (b)

Figure 15.4.15:

The flux across the curve is 2π. Each half must have flux π. Since each half looks like C, the flux across C must be π.

Settle the issue by

(a) Evaluating the integral $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$ by the Fundamental Theorem of Calculus.

(b) Considering the flux across the curve C^* obtained from C by replacing the small part of C near O by a semicircle C, as in Figure 15.4.15(b).

(c) By considering the angle the curve C subtends at O.

32.[C] Let $\mathbf{F}(P) = \sigma(P)\mathbf{v}(P)$ represent the flow of a fluid as described in the discussion of circulation and flux. Let C by a closed curve that bounds the region R. Let $Q(t)$ be the total mass of the fluid in R at time t. Express dQ/dt in terms of a line integral.

Skill Drill

33.[R] Differentiate for practice.

(a) $\frac{1}{2a} \ln \left| ax^2 + c \right|$

(b) $\frac{2(3ax-2b)}{15a^2} \sqrt{(ax+b)^3}$

(c) $\frac{1}{a} \sin(ax) - \frac{1}{3a} \sin^3(ax)$

(d) $\frac{1}{a} \tan(ax) - x$

(e) $\frac{1}{a^2} \cos(ax) + \frac{x}{a} \sin(ax)$

(f) $x \arctan(ax) - \frac{1}{2a} \ln(1 + a^2x^2)$

15.S Chapter Summary

This chapter concerns the derivatives of vector functions and integrals over curves.

Let $\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$ be the position vector from the origin to a point on a curve. We defined its derivative, $\mathbf{r}'(t)$, in terms of the derivatives of the components. But, we could just as well define it without mentioning components:

$$\mathbf{r}'(t) = \lim_{\Delta t\to 0} \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t)}{\Delta t} = \lim_{\Delta t\to 0} \frac{\Delta \mathbf{r}}{\Delta t}. \qquad (15.S.1)$$

This definition reveals the underlying geometry, as Figure 15.S.1 shows. For small Δt, the direction of $\Delta \mathbf{r}$ is almost along the tangent. The length of $\Delta \mathbf{r}$ is almost the same as the scalar length Δs along the curve. Thus, $\Delta\mathbf{r}/\Delta t$ is a vector pointing almost in the direction of motion and with a magnitude approximating the instantaneous speed.

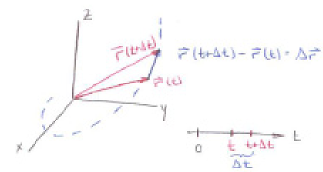


Figure 15.S.1:

The limit in (15.S.1) is called the derivative of the function $\mathbf{r}(t)$. If we think of t as time, then $\mathbf{r}'$ is called the velocity vector, denoted $\mathbf{v}$. The derivative of $\mathbf{v}$ is the acceleration vector: $\mathbf{v}' = \mathbf{a}$.

The vector $\mathbf{T} = \mathbf{r}'/\|\mathbf{r}'\|$ is a unit tangent vector. The magnitude of its derivative with respect to arclength, s, is the curvature, κ, of the path, as suggested by Figure 15.S.2. Keep in mind that the curve may not lie in a plane. Nevertheless, this figure resembles Figure 15.2.6 in Section 15.2.

(a) (b)

Figure 15.S.2:

It was shown that curvature equals $\|\mathbf{v}\times\mathbf{a}\|/\|\mathbf{v}\|^3$.

Checking that this follows from the definition of curvature provides a good review of the derivative of a vector function.

The vector $d\mathbf{T}/ds$ is perpendicular to $\mathbf{T}$. (Why?) The unit vector $\mathbf{N} = \frac{d\mathbf{T}/ds}{\|d\mathbf{T}/ds\|}$ is called the principal normal to the curve at the given point. The vector $\mathbf{T}\times\mathbf{N} = \mathbf{B}$ is the third unit vector forming a frame that moves along the curve, with $\mathbf{T}$ and $\mathbf{N}$ indicating the plane in which the curve locally "almost lies."

The acceleration vector $\mathbf{a}$, even for space curves, can be expressed relative to $\mathbf{T}$ and $\mathbf{N}$ ($\mathbf{B}$ is not involved):

$$\mathbf{a} = \frac{d^2s}{dt^2}\mathbf{T} + \frac{v^2}{r}\mathbf{N}$$

where $r = 1/\kappa$ is the radius of curvature. The first coefficient is to be expected. The second is more complicated, indicating the force needed to keep the particle in the path is proportional to the square of the velocity and inversely proportional to the radius of curvature.

This chapter then introduced four types of integrals involving a curve C:

$$\int\limits_C f(P)\ ds, \qquad \int\limits_C f(P)\ dx, \qquad \int\limits_C f(P)\ dy, \quad \text{and} \quad \int\limits_C f(P)\ dz,$$

whose definitions resemble those in Chapter 6 for definite integrals. In the last three the orientation of the curve matters: switching the direction in which the curve is swept out changes the sign of dx, dy, and dz, and thus puts a minus sign in front of the integral.

In particular, for a closed curve taken counterclockwise $\oint_C y\ dx$ is the negative of the area enclosed by the curve. (Why?) On the other hand, $\int_C x\ dy$ taken counterclockwise is the area enclosed.

The most general integral considered is

$$\int\limits_C \left(P(x,y,z)\ dx + Q(x,y,z)\ dy + R(x,y,z)\ dz\right).$$

The integrand in this form is called a differential form. For $\mathbf{F} = \langle P, Q, R\rangle$, this can be written much more compactly as $\int_C \mathbf{F}\cdot\ d\mathbf{r}$. However, in proofs or computations one must often return to the longer differential form.

If $\int_C \mathbf{F}\cdot d\mathbf{r}$ depends only on the ends of C, $\mathbf{F}$ is called a conservative vector field, a concept that will be important in Chapter 18.

Line integrals were applied to work, circulation, flux, and the angle subtended by a curve (the last in preparation for the "solid angle" subtended by a surface).

EXERCISES for 15.S *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 6, evaluate $\int_C \mathbf{F}\cdot d\mathbf{r}$ for the given vector field $\mathbf{F}$ and given curve C.

1.[R] $\mathbf{F}(x,y) = 2x\mathbf{i}$ and C is a semicircle, $\mathbf{r}(\theta) = 3\cos\theta\mathbf{i} + 3\sin\theta\mathbf{j}$, $0 \le \theta \le \pi$.

2.[R] $\mathbf{F}(x,y) = x^2\mathbf{i} + 2xy\mathbf{j}$ and C is a line segment, $\mathbf{r}(t) = 2t^2\mathbf{i} + 3t^2\mathbf{j}$, $1 \le t \le 2$.

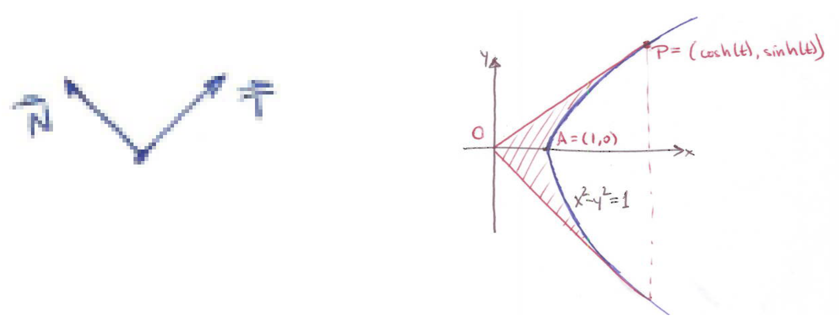
3.[R] $\mathbf{F}(x,y,z) = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ and C is a helix, $\mathbf{r}(t) = \cos t\mathbf{i}+\sin t\mathbf{j}+3t\mathbf{k}$, $0 \le t \le 4\pi$.

4.[R] $\mathbf{F}(x,y,z) = x^2\mathbf{i} + xy\mathbf{j} + 3\mathbf{k}$ and C is a line segment, $\mathbf{r}(t) = 2t\mathbf{i} + (3t+1)\mathbf{j} + t\mathbf{k}$, $1 \le t \le 2$.

5.[R] $\mathbf{F}(\mathbf{r}) = \widehat{\mathbf{r}}/\|\mathbf{r}\|^2$ and C is a line, $\mathbf{r}(t) = 2t\mathbf{i} + 3t\mathbf{j} + 4t\mathbf{k}$, $0 \le t \le 2$.

6.[R] $\mathbf{F}(\mathbf{r}) = \mathbf{r}$ and C is the circle, $\mathbf{r}(t) = \cos\theta\mathbf{i} + \sin\theta\mathbf{j} + 2\mathbf{k}$, $0 \le \theta \le 2\pi$.

7.[R] Figure 15.S.3(a) shows **T** and **N** for one point P on a curve C. The curve is not shown. Sketch what a short part of C may look like.



(a) (b)

Figure 15.S.3:

8.[M]

(a) Express the area under the hyperbola $x^2 - y^2 = 1$ and above the interval $[1, \cosh(t)]$ as a line integral.

(b) Evaluate the line integral found in (a).

(c) What is the area of the shaded region in Figure 15.S.3(b)?

NOTE: See also Exercises 64 in Section 6.5 and 77 in Section 8.6.

The CIE at the end of Chapter 3 developed the reflection properties of parabolas and ellipses. Exercises 9 and 10 show how vectors provide a much shorter way to obtain those results.

9.[C] A parabola consists of the points P equidistant from a fixed point F and fixed line L, as in Figure 15.S.4.

Let O be some point on L and let $\mathbf{u}$ be a unit vector perpendicular to L aimed toward P. Let $\mathbf{r} = \overrightarrow{OP}$ and $\mathbf{F} = \overrightarrow{OF}$. (We assume the curve is parameterized in such a manner that there is a well-defined tangent vector, $\mathbf{r}'$.)

(a) Show that $\|\mathbf{r} - \mathbf{F}\| = \mathbf{r} \cdot \mathbf{u}$.

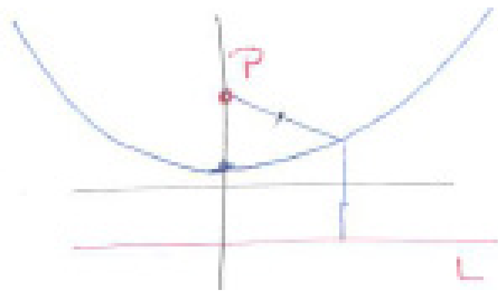


Figure 15.S.4:

(b) From (a) deduce that

$$\frac{\mathbf{r}-\mathbf{F}}{\|\mathbf{r}-\mathbf{F}\|}\cdot\mathbf{r}'=\mathbf{r}'\cdot\mathbf{u}.$$

(c) From (b) deduce that

$$\|\mathbf{r}'\|\cos(\mathbf{r}',\mathbf{r}-\mathbf{F})=\|\mathbf{r}\|\cos(\mathbf{r}',\mathbf{u}).$$

(d) From (c) deduce the reflection principle of a parabola.

This proof, which starts with the geometric definition of a parabola rather than the equation $y=x^2$, appears in Harley Flanders', "The Optical Properties of the Conics," American Mathematics Monthly, 1968, p. 399.

10.[C] This exercise develops the reflection property of an ellipse. Start with its geometric definition as the locus of points such that the sum of its distances from two fixed points is constant. Let $\mathbf{p}$ and $\mathbf{q}$ be the position vectors of the two fixed points and $\mathbf{r}$ the position vector for a typical point P on the ellipse, which is parameterized so we may speak of $\mathbf{r}'$, a tangent vector.

(a) Differentiate both sides of $\|\mathbf{r}-\mathbf{p}\|+\|\mathbf{r}-\mathbf{q}\|=c$, a constant.

(b) Let $\mathbf{u}_1$ be the unit vector in the direction of $\|\mathbf{r}-\mathbf{p}\|$ and $\mathbf{u}_2$ be the unit vector in the direction of $\|\mathbf{r}-\mathbf{q}\|$. Show that $\mathbf{u}_1\cdot\mathbf{r}'+\mathbf{u}_2\cdot\mathbf{r}'=\mathbf{0}$.

(c) Show that $\mathbf{u}_1+\mathbf{u}_2$ is normal to the curve at P.

(d) Show that $\mathbf{u}_1$ and $\mathbf{u}_2$ make equal angles with $\mathbf{u}_1+\mathbf{u}_2$.

(e) From (d) deduce the reflection property of an ellipse.

Skill Drill

In Exercises 11 to 13, $\mathbf{a}(t)$ is the acceleration vector at time t for a moving particle and $\mathbf{r}(t_0)$ and $\mathbf{a}(t_0)$ are the particle's position and acceleration at time $t = t_0$. Find the velocity and position vectors, $\mathbf{v}(t)$ and $\mathbf{r}(t)$, of the particle at time t. (These review the integration techniques of Chapter 8.)

11.[C] $\mathbf{a}(t) = t108\,(\ln(t))^2\,\mathbf{i} + \ln(1+t^2)\mathbf{j} + t\arctan(t)\mathbf{k}$; $\mathbf{r}(1) = 19\mathbf{i} - \mathbf{j} + 4(\pi - 2 - \ln(2))\mathbf{k}$, $\mathbf{v}(1) = 27\mathbf{i} - 2\mathbf{j} + 6(\pi - 2)\mathbf{k}$

12.[C] $\mathbf{a}(t) = \dfrac{\tan(t)+\sin(t)}{\sec(t)}\mathbf{i} + \dfrac{t^4}{t^2+4}\mathbf{j} + \dfrac{2t-4}{t^2+2t+1}\mathbf{k}$; $\mathbf{r}(0) = \frac{1}{4}\mathbf{i} + \mathbf{j} - 8\mathbf{k}$, $\mathbf{v}(0) = \frac{-3}{2}\mathbf{i} + 4\mathbf{j} + 6\mathbf{k}$

13.[C] $\mathbf{a}(t) = (t^2+4t+5)^{-1}\mathbf{i} + t^2\cos(t)\mathbf{j} + \dfrac{1}{t^2+4}\mathbf{k}$; $\mathbf{r}(0) = 6\mathbf{j}$, $\mathbf{v}(0) = \arctan(2)\mathbf{i}$

Calculus is Everywhere # 18
Newton's Law Implies Kepler's Three Laws

After hundreds of pages of computation based on observations by the astronomer Tycho Brahe (1546—1601) in the last 30 years of the sixteenth century, plus lengthy detours and lucky guesses, Kepler (1571–1630) arrived at these three laws of planetary motion:

Kepler's Three Laws

1. Every planet travels around the sun in an elliptical orbit such that the sun is situated at one focus (discovered in 1605, published in 1609).
2. The velocity of a planet varies in such a way that the line joining the planet to the sun sweeps out equal areas in equal times (discovered 1602, published 1609).
3. The square of the time required by a planet for one revolution around the sun is proportional to the cube of its mean distance from the sun (discovered 1618, published 1619).

The work of Kepler shattered the crystal spheres which for 2,000 years had carried the planets. Before him astronomers admitted only circular motion and motion compounded of circular motions. Copernicus (1473–1543), for instance, used five circles to describe the motion of Mars.

The ellipse got a cold reception.

The ellipse was not welcomed. In 1605 Kepler complained to a skeptical astronomer:

> You have disparaged my oval orbit .... If you are enraged because I cannot take away oval flight how much more you should be enraged by the motions assigned by the ancients, which I did take away .... You disdain my oval, a single cart of dung, while you endure the whole stable. (If indeed my oval is a cart of dung.)

But the astronomical tables that Kepler based on his theories, and published in 1627, proved to be more accurate than any other, and the ellipse gradually gained acceptance.

The three laws stood as mysteries alongside a related question: If there are no crystal spheres, what propels the planets? Bullialdus (1605–1694), a French astronomer and mathematician, suggested in 1645:

The inverse square law was conjectured.

> The force with which the sun seizes or pulls the planets, a physical force which serves as hands for it, is sent out in straight lines into all the world's space . . . ; since it is physical it is decreased in greater space; . . . the ratio of this distance is the same as that for light, namely as the reciprocal of the square of the distance.

In 1666, Hooke (1635–1703), more of an experimental scientist than a mathematician, wondered:

> why the planets should move about the sun . . . being not included in any solid orbs . . . nor tied to it . . . by any visible strings . . . . I cannot imagine any other likely cause besides these two: The first may be from an unequal density of the medium . . . ; if we suppose that part of the medium, which is farthest from the centre, or sun, to be more dense outward, than that which is more near, it will follow, that the direct motion will be always deflected inwards, by the easier yielding of the inwards . . . .
>
> But the second cause of inflecting a direct motion into a curve may be from an attractive property of the body placed in the center; whereby it continually endeavours to attract or draw it to itself. For if such a principle be supposed all the phenomena of the planets seem possible to be explained by the common principle of mechanic motions. . . . By this hypothesis, the phenomena of the comets as well as of the planets may be solved.

In 1675, Hooke, in an announcement to the Royal Society, went further:

> All celestial bodies have an attraction towards their own centres, whereby they attract not only their own parts but also other celestial bodies that are within the sphere of their activity . . . . All bodies that are put into direct simple motion will so continue to move forward in a single line till they are, by some other effectual powers, deflected and bent into a motion describing a circle, ellipse, or some other more compound curve . . . . These attractive powers are much more powerful in operating by how much the nearer the body wrought upon is to their own centers . . . . It is a notion which if fully prosecuted as it ought to be, will mightily assist the astronomer to reduce all the celestial motions to a certain rule . . . .

Hooke pressed Newton to work on the problem.

Trying to interest Newton in the question, Hooke wrote on November 24, 1679: "I shall take it as a great favor if . . . you will let me know your thoughts of that of compounding the celestial motion of planets of a direct motion by the tangent and an attractive motion toward the central body." But four days later Newton replied:

> My affection to philosophy [science] being worn out, so that I am almost as little concerned about it as one tradesman used to be about another man's trade or a countryman about learning. I must acknowledge myself averse from spending that time in writing about it which I think I can spend otherwise more to my own content and the good of others ....

In a letter to Newton on January 17, 1680, Hooke returned to the problem of planetary motion:

> It now remains to know the properties of a curved line (not circular ...) made by a central attractive power which makes the velocities of descent from the tangent line or equal straight motion at all distances in a duplicate proportion to the distances reciprocally taken. I doubt not that by your excellent method you will easily find out what that curve must be, and its properties, and suggest a physical reason for this proportion.

Hooke succeeded in drawing Newton back to science, as Newton himself admitted in his *Principia*, published in 1687: "I am beholden to him only for the diversion he gave me from the other studies to think on these things and for his dogmaticalness in writing as if he had found the motion in the ellipse, which inclined me to try it."

It seems that Newton then obtained a proof — perhaps containing a mistake (the history is not clear) — that if the motion is elliptical, the force varies as the inverse square. In 1684, at the request of the astronomer Halley, Newton provided a correct proof. With Halley's encouragement, Newton spent the next year and a half writing the *Principia*.

Halley, of Halley's comet, paid for publication of the *Principia*.

In the *Principia*, which develops the science of mechanics and applies it to celestial motions, Newton begins with two laws:

1. Every body continues in its state of rest, or of uniform motion in a straight line, unless it is compelled to change this state by forces impressed upon it.

2. The change of momentum is proportional to the motive force impressed, and is made in the direction of the straight line in which that force is impressed.

To state these in the language of vectors, let $\mathbf{v}$ be the velocity of the body, $\mathbf{F}$ the impressed force, and m the mass of the body. The first law asserts that $\mathbf{v}$ is constant if $\mathbf{F}$ is $\mathbf{0}$. **Momentum** is defined as $m\mathbf{v}$; the second law asserts that

$$\mathbf{F} = \frac{d}{dt}(m\mathbf{v}).$$

If m is constant, this reduces to

$$\mathbf{F} = m\mathbf{a},$$

where $\mathbf{a}$ is the acceleration vector.

Newton assumed a universal **law of gravitation**. Any particle P exerts an attractive force on any other particle Q, and the direction of the force is from Q toward P. Then *assuming* that the orbit of a planet moving about the sun (both treated as points) is an ellipse, he *deduced* that this force is inversely proportional to the square of the distance between the particles P and Q.

Nowhere in the *Principia* does he deduce from the inverse-square law of gravity that the planets' orbits are ellipses. (However, there are general theorems in *Principia* on the basis of which this deduction could have been made.) In the *Principia* he showed that Kepler's second law (concerning areas) was equivalent to the assumption that the force acting on a planet is directed toward the sun. Finally, he deduced Kepler's third law.

Newton's universal law of gravitation asserts that any particle, of mass M, exerts a force on any other particle, of mass m, and that the magnitude of this force is proportional to the product of the two masses, mM, inversely proportional to the square of the distance between them, and is directed toward the particle with the larger mass. (Here, we assume $M > m$.)

We will assume that the sun is fixed at O.

Assume that the sun has mass M and is located at point O and that the planet has mass m and is located at point P. (See Figure C.18.1.) Let $\mathbf{r} = \vec{OP}$ and $r = \|\mathbf{r}\|$. Then the sun exerts a force $\mathbf{F}$ on the planet given by the formula

$$\mathbf{F} = -\frac{GmM}{r^3}\mathbf{r}, \tag{C.18.1}$$

where G is a universal constant. It is convenient to introduce the unit vector $\mathbf{u} = \mathbf{r}/r$, which points in the direction of $\mathbf{r}$. Then (C.18.1) reads

$$\mathbf{F} = -\frac{GmM}{r^2}\mathbf{u}.$$

Now, $\mathbf{F} = m\mathbf{a}$, where $\mathbf{a}$ is the acceleration vector of the planet. Thus

$$m\mathbf{a} = -\frac{GmM}{r^2}\mathbf{u},$$

from which it follows that

$$\mathbf{a} = -\frac{q\mathbf{u}}{r^2}, \tag{C.18.2}$$

where $q = GM$ is independent of the planet.

The vectors $\mathbf{u}$, $\mathbf{r}$, and $\mathbf{a}$ are indicated in Figure C.18.1.

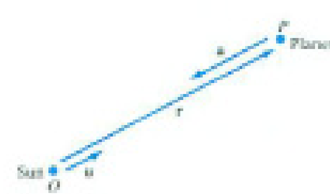


Figure C.18.1:

The following exercises show how to obtain Kepler's three laws from the single law of Newton, $\mathbf{a} = -q\mathbf{u}/r^2$.

EXERCISES

Exercises 1 to 3 obtain Kepler's "area" law.

1.[R] Let $\mathbf{r}(t)$ be the position vector of a given planet at time t. Let $\Delta\mathbf{r} = \mathbf{r}(t+\Delta t) - \mathbf{r}(t)$. Show that for small Δt,

$$\frac{1}{2}\left\|\mathbf{r} \times \Delta\mathbf{r}\right\|$$

approximates the area swept out by the position vector during the small interval of time Δt. HINT: Draw a picture.

Figure C.18.2:

2.[R] From Exercise 1 deduce that $\frac{1}{2}\left\|\mathbf{r} \times \frac{d\mathbf{r}}{dt}\right\|$ is the rate at which the position vector $\mathbf{r}$ sweeps out area. (See Figure C.18.2.)

Let $\mathbf{v} = d\mathbf{r}/dt$. The vector $\mathbf{r} \times \mathbf{v}$ will play a central role in the argument leading to Kepler's area law. (See also Exercise 34 in Section 15.3.)

3.[R] With the aid of (C.18.2), show that the vector $\mathbf{r} \times \mathbf{v}$ is constant, independent of time.

Since $\mathbf{r} \times \mathbf{v}$ is constant, $\frac{1}{2}\left\|\mathbf{r} \times \mathbf{v}\right\|$ is constant. In view of Exercise 2, it follows that the radius vector of a given planet sweeps out area at a constant rate. **To put it another way, the radius vector sweeps out equal areas in equal times. This is Kepler's second law.**

Introduce an xyz-coordinate system such that the unit vector $\mathbf{k}$, which points in the direction of the positive z axis, has the same direction as the constant vector $\mathbf{r} \times \mathbf{v}$. Thus there is a positive constant h such that

$$\mathbf{r} \times \mathbf{v} = h\mathbf{k}. \qquad \text{(C.18.3)}$$

Exercises 4 to 13 obtain Kepler's "ellipse" law.

4.[R] Show that h in (C.18.3) is twice the rate at which the position vector of the planet sweeps out area.

5.[R] Show that the planet remains in the plane perpendicular to $\mathbf{k}$ that passes through the sun.

By Exercise 5, the orbit of the planet is planar. We may assume that the orbit lies in the xy plane; for convenience, locate the origin of the xy coordinates at the sun. Also introduce polar coordinates in this plane, with the pole at the sun and the polar axis along the positive x axis, as in Figure C.18.3.

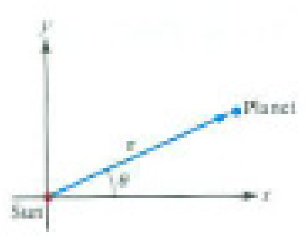


Figure C.18.3:

6.[R]

(a) Show that during the time interval $[t_0, t]$ the position vector of the planet sweeps out the area

$$\frac{1}{2}\int_{t_0}^{t} r^2 \frac{d\theta}{dt} dt.$$

(b) From (a) deduce that the radius vector sweeps out area at the rate $\frac{1}{2}r^2\frac{d\theta}{dt}$.

Henceforth use the dot notation for differentiation with respect to time. Thus $\dot{\mathbf{r}} = \mathbf{v}$, $\dot{\mathbf{v}} = \mathbf{a}$, and $\dot{\theta} = \frac{d\theta}{dt}$.

7.[R] Show that $\mathbf{r} \times \mathbf{v} = r^2\dot{\theta}\mathbf{k}$.

8.[R] Show that $\dot{\mathbf{u}} = \frac{d\mathbf{u}}{d\theta}\dot{\theta}$ and is perpendicular to $\mathbf{u}$. Recall that $\mathbf{u}$ is defined as $\mathbf{r}/\|\mathbf{r}\|$.

9.[R] Recalling that $\mathbf{r} = r\mathbf{u}$, show that $h\mathbf{k} = r^2(\mathbf{u} \times \dot{\mathbf{u}})$.

10.[R] Using (C.18.2) and Exercise 9, show that $\mathbf{a} \times h\mathbf{k} = q\dot{\mathbf{u}}$. HINT: What is the vector identity for $\mathbf{A} \times (\mathbf{B} \times \mathbf{C})$?

11.[R] Deduce from Exercise 10 that $\mathbf{v} \times h\mathbf{k}$ and $q\mathbf{u}$ differ by a constant vector.

By Exercise 11, there is a constant vector $\mathbf{C}$ such that

$$\mathbf{v} \times h\mathbf{k} = q\mathbf{u} + \mathbf{C}. \tag{C.18.4}$$

Then the angle between $\mathbf{r}$ and $\mathbf{C}$ is the angle θ of polar coordinates.

The next exercise requires the vector identity $(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot (\mathbf{B} \times \mathbf{C})$, which is valid for any three vectors $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$.

12.[R]

(a) Show that $(\mathbf{r} \times \mathbf{v}) \cdot h\mathbf{k} = h^2$.

(b) Show that $\mathbf{r} \cdot (\mathbf{v} \times h\mathbf{k}) = rq + \mathbf{r} \cdot \mathbf{C}$.

(c) Combining (a) and (b), deduce that $h^2 = rq + rc\cos(\theta)$, where $c = \|\mathbf{C}\|$

It follows from Exercise 12 that the polar equation for the orbit of the planet is given by

$$r(\theta) = \frac{h^2}{q + c\cos(\theta)}. \tag{C.18.5}$$

13.[R] By expressing (C.18.5) in rectangular coordinates, show that it describes a conic section.

Since the orbit of a planet is bounded and is also a conic section, it must be an ellipse. This establishes Kepler's first law.

Kepler's third law asserts that the square of the time required for a planet to complete one orbit is proportional to the cube of its mean distance from the sun.

First the term **mean distance** must be defined. For Kepler this meant the average of the shortest distance and the longest distance from the planet to the sun in its orbit. Let us compute this average for the ellipse of semimajor axis a and semiminor axes b, shown in Figure C.18.4. The sun is at the focus F, which is also the pole of the polar coordinate system we are using. The line through the two foci contains the polar axis.

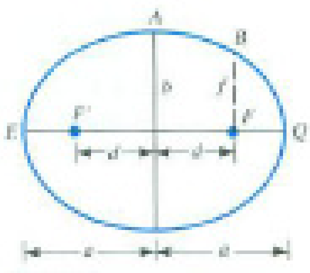


Figure C.18.4:

Recall that an ellipse is the set of points P such that the sum of the distances from P to the two foci F and F' is constant, $2a$. The shortest distance from the planet to the sun is $\overline{FQ} = a - d$ and the longest distance is $\overline{EF} = a + d$. Thus Kepler's mean distance is

$$\frac{(a-d)+(a+d)}{2} = a.$$

Now let T be the time required by the given planet to complete one orbit. Kepler's third law asserts that T^2 is proportional to a^3. Exercises 14 to 18 establish this law by showing that T^2/a^3 is the same for all planets.

14.[R] Using the fact that the area of the ellipse in Figure C.18.4 is πab, show that $Th/2 = \pi ab$, hence that

$$T = \frac{2\pi ab}{h}. \qquad \text{(C.18.6)}$$

The rest of the argument depends only on (C.18.5) and (C.18.6) and the "fixed sum of two distances" property of an ellipse.

15.[R] Using (C.18.5), show that f in Figure C.18.4 equals h^2/q.

16.[R] Show that $b^2 = af$, as follows:

(a) From the fact that $\bar{F'}A + \bar{F}A = 2a$, deduce that $a^2 = b^2 + d^2$.

(b) From the fact that $\bar{F'}B + \bar{F}B = 2a$, deduce that $d^2 = a^2 - af$.

(c) From (a) and (b), deduce that $b^2 = af$.

17.[R] From Exercises 15 and 16, deduce that $b^2 = ah^2/q$.

18.[R] Combining (C.18.6) and Exercise 17, show that

$$\frac{T^2}{a^3} = \frac{4\pi^2}{q}.$$

Since $4\pi^2/q$ is a constant, the same for all points, Kepler's third law is established.

For Further Reading

1. M. Caspar, *Kepler*, Hellman, New York, 1959.
2. J. Kepler, "Copernican astronomy," *Great Books of the Western World*, Encyclopedia Britannica, Chicago, 1952.
3. A. Koestler, *The Watershed*, Doubleday, New York, 1960.
4. I. Newton, *Principia*, University of California, Berkeley, 1947.
5. V. Frederick Rickey, "Isaac Newton: Man, Myth, and Mathematician," *College Mathematics Journal*, **18** (1987), 362—388.
6. J. L. Russell, "Kepler's Laws of Planetary Motion 1609–1666," *British Journal for the History of Science*, **2** (1964), 1—24.
7. S. Stein, "Exactly how did Newton deal with his planets?", *Math. Intell.*, **18** (1996), 6—11.

Calculus is Everywhere # 19

The Suspension Bridge and the Hanging Cable

In a suspension bridge the roadway hangs from a cable, as shown in Figure C.19.1. We will use calculus to find the shape of the cable. To begin, we assume that the weight of any section of the roadway is proportional to its length. That is, there is a constant k such that x feet of the roadway weighs kx pounds. We will assume that the cable itself is weightless. That is justified for it weighs little in comparison to the roadway.

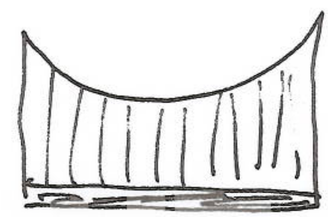
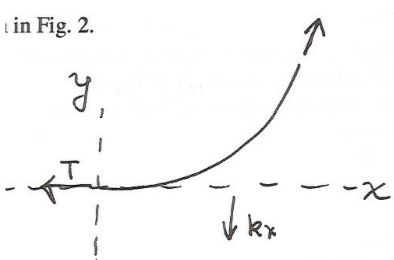
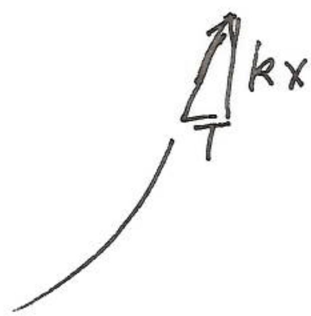


Figure C.19.1:

We introduce an xy-coordinate system with origin at the lowest point of the cable, and consider a typical section of the cable, which goes from $(0,0)$ to (x,y), as shown in Figure C.19.2(a). Three forces act on this section. The



(a)



(b)

Figure C.19.2:

force at $(0,0)$ is horizontal and pulls the cable to the left. Call its magnitude T. Gravity pulls the cable down with the force kx (the weight of the roadway beneath the cable). At the top of the section, at (x,y) the cable above it pulls the cable to the right and upward, along the tangent line to the cable.

The section does not move, neither up nor down, neither to the left nor to the right. That means the horizontal part of the force at (x,y) must have magnitude T and the vertical part of the force must have magnitude kx, as shown in Figure C.19.2(b). (Think of one person pulling horizontally at (x,y) and another pulling vertically to duplicate the effect of the part of the cable above (x,y) that is pulling on the section.)

Since the force at the point (x,y) is directed along the tangent line there, we have

$$\frac{dy}{dx} = \frac{kx}{T}. \tag{C.19.1}$$

Therefore,

$$y = \frac{kx^2}{T} + C.$$

for some constant C. Since $(0, 0)$ is on the curve, $C = 0$, and the cable has the equation

$$y = \frac{kx^2}{tT}.$$

The cable forms a parabola.

But what if, instead, we have the cable but no roadway? That is the case with a laundry line or a telephone wire or a hanging chain. In this case the downward force is due to the weight of the cable. If s feet of cable weighs ks pounds, reasoning almost identical to the case of the suspension bridge leads to the equation

$$\frac{dy}{dx} = \frac{ks}{T}. \tag{C.19.2}$$

Since

$$s = \int_0^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

we have the equation

$$\frac{dy}{dx} = \frac{k \int_0^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx}{T}. \tag{C.19.3}$$

We get rid of the integral by differentiating both sides of (C.19.3), and using part of the fundamental theorem of calculus, obtaining

$$\frac{d^2y}{dx^2} = \frac{k}{T}\sqrt{1 + \left(\frac{dy}{dx}\right)^2}. \tag{C.19.4}$$

This equation is solved in a differential equations course, where it is shown that

$$y = \frac{k}{T}\left(e^{\frac{kx}{T}} + e^{\frac{-kx}{T}}\right) - 2\frac{k}{T}. \tag{C.19.5}$$

This curve is called a **catenary**, after the Latin "catena," meaning "chain." (Hence the word "concatenation," referring to a chain of events.) It may look like a parabola, but it isn't. The 630-foot tall Gateway Arch in St. Louis, completed October 28, 1965, is the most famous catenary.

EXERCISES

1.[M] Check that the solution to

$$\frac{d^2y}{dx^2} = \frac{k}{T}\sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

that passes through $(0, 0)$ is

$$y = \frac{k}{T}\left(e^{\frac{kx}{T}} + e^{\frac{-kx}{T}}\right) - 2\frac{k}{T}. \qquad \text{(C.19.6)}$$

Calculus is Everywhere # 20

The Path of the Rear Wheel of a Scooter

When the front wheel of a scooter follows a certain path, what is the path of the rear wheel? This question could be phrased in terms of a bicycle or car, but the scooter is more convenient for carrying out real-life experiments.

In 13 we considered the special case when the front wheel moves in a straight line, as may occur when parking a car. Now, using vectors, we will look at the case when the front wheel sweeps out a circular path.

The Basic Equation

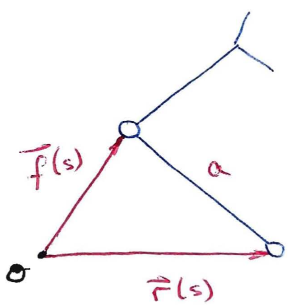


Figure C.20.1:

Figure C.20.1 shows the geometry at any instant. Let s denote the arc length of the path swept out by the rear wheel as measured from its starting point. Let a be the length of the wheel base, that is, the distance between the front and rear axels. The vector $\mathbf{r}(s)$ records the position of the rear wheel and $\mathbf{f}(s)$ records the position of the front wheel. Because the rear wheel is parallel to $\mathbf{f}(s) - \mathbf{r}(s)$, the vector $\mathbf{r}'(s)$ points directly toward the front wheel or directly away from it. Note that $\mathbf{r}'(s)$ is a unit vector.

Thus

$$\mathbf{f}(s) = \mathbf{r}(s) + a\mathbf{r}'(s) \tag{C.20.1}$$

or

$$\mathbf{f}(s) = \mathbf{r}(s) - a\mathbf{r}'(s). \tag{C.20.2}$$

In short, we will write $\mathbf{f}(s) = \mathbf{r}(s) \pm a\mathbf{r}'(s)$.

Assume that the front wheel moving, say, counterclockwise traces out a circular path with center O and radius c. Because

$$\mathbf{f}(s) \cdot \mathbf{f}(s) = c^2,$$

we have

$$(\mathbf{r}(s) \pm a\mathbf{r}'(s)) \cdot (\mathbf{r}(s) \pm a\mathbf{r}'(s)) = c^2.$$

By distributivity of the dot product,

$$\mathbf{r}(s) \cdot \mathbf{r}(s) + a^2\mathbf{r}'(s) \cdot \mathbf{r}'(s) \pm 2a\mathbf{r}(s) \cdot \mathbf{r}'(s) = c^2. \tag{C.20.3}$$

Letting $r(s) = \|\mathbf{r}(s)\|$, we may rewrite (C.20.3) as

$$(r(s))^2 + a^2 \pm 2a\mathbf{r}(s) \cdot \mathbf{r}'(s) = c^2. \tag{C.20.4}$$

Differentiating $\mathbf{r}(s) \cdot \mathbf{r}(s) = r(s)^2$ to obtain, $\mathbf{r}(s) \cdot \mathbf{r}'(s) = r(s)r'(s)$, which changes (C.20.3) to an equation involving the scalar function $r(s)$. For simplicity, we write $r(s)$ as r and $r'(s)$ as r', obtaining

$$r^2 + a^2 \pm 2arr' = c^2. \tag{C.20.5}$$

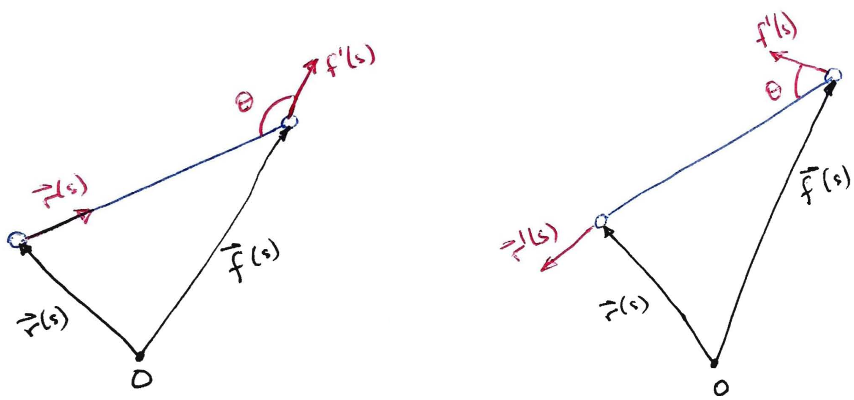
This is the basic equation we will use to analyze the path of the rear wheel of a scooter.

The Direction of r′

Before going further we examine when $\mathbf{r}'$ points towards the front wheel and when it points away from the front wheel.

The movement of the back wheel is determined by the projection of $\mathbf{f}'$ on the line of the scooter. That projection is the same as $\mathbf{r}'$.

Thus, when the angle θ between the front wheel and the line of the scooter is obtuse, as in Figure C.20.2(a), $\mathbf{r}'$ points towards the front wheel. When θ is acute, the scooter backs up and $\mathbf{r}'$ points away from the front wheel, as shown in Figure C.20.2(b).



(a) (b)

Figure C.20.2: The direction of $\mathbf{r}'$ depends on the angle θ between the front wheel and the line of the scooter. (a) θ is obtuse, (b) θ is acute.

When the direction of $\mathbf{r}'$ abruptly shifts from pointing towards the front wheel to pointing away from the front wheel, the path of the rear wheel also abruptly changes, as shown in Figure C.20.3.

The path of the rear wheel is continuous but the unit tangent vector $\mathbf{r}'$ is not defined at the point where its direction suddenly shifts. The path is said to contain a "cusp" and the point at which $\mathbf{r}'(s)$ shifts direction by the angle π is the "vertex" of the cusp.

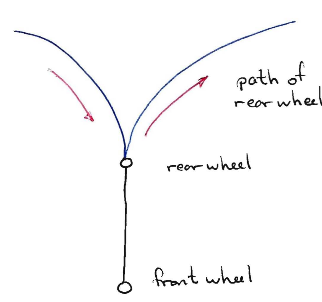


Figure C.20.3:

The Path of the Rear Wheel for a Short Scooter

Assume that the wheel base, a, is less than the radius of the circle, c, that, initially, θ is obtuse, and that r^2 is less than $c^2 - a^2$. Thus, $c^2 - a^2 - r^2$ is positive. (Exercise 1 shows the special significance of $c^2 - a^2$.)

We rewrite the equation $c^2 = a^2 + r^2 + 2rr'a$ in the form

$$\frac{-2rr'}{c^2 - a^2 - r^2} = \frac{-1}{a}. \tag{C.20.6}$$

Integration of both sides of (C.20.6) with respect to arc length s shows that there is a constant k such that

$$\ln(c^2 - a^2 - r^2) = \frac{-s}{a} + k,$$

hence

$$c^2 - a^2 - r^2 = e^k e^{-s/a}. \tag{C.20.7}$$

Equation (C.20.7) tells us that r^2 increases but remains less than $c^2 - a^2$, and approaches $c^2 - a^2$ as s increases. Thus the rear wheel traces a spiral path that gets arbitrarily close to the circle of radius $\sqrt{c^2 - a^2}$ and center O, as in Figure C.20.4.

The Path of the Rear Wheel for a Long Scooter

Assume that the wheel base is longer than the radius of the circle on which the front wheel moves, that is, $a > c$. Assume also that initially the scooter is moving forward, so we again have the equation

$$c^2 = a^2 + r^2 + 2rr'a. \tag{C.20.8}$$

The initial position is indicated in Figure C.20.5(a).

Now $c^2 - a^2 - r^2$ is negative, and we have

$$\frac{2rr'}{a^2 + r^2 - c^2} = \frac{-1}{a},$$

where the denominator on the left-hand side is positive. Thus there is a constant k such that

$$a^2 + r^2 - c^2 = e^k e^{-s/a}. \tag{C.20.9}$$

If s gets arbitrarily large, (C.20.9) implies that r^2 approaches $c^2 - a^2$. *But, $c^2 - a^2$ is negative, so this cannot happen.* Our assumption that (C.20.8) holds for all s must be wrong. Instead, there must be a cusp and the governing equation switches to

$$c^2 = a^2 + r^2 - 2arr'.$$

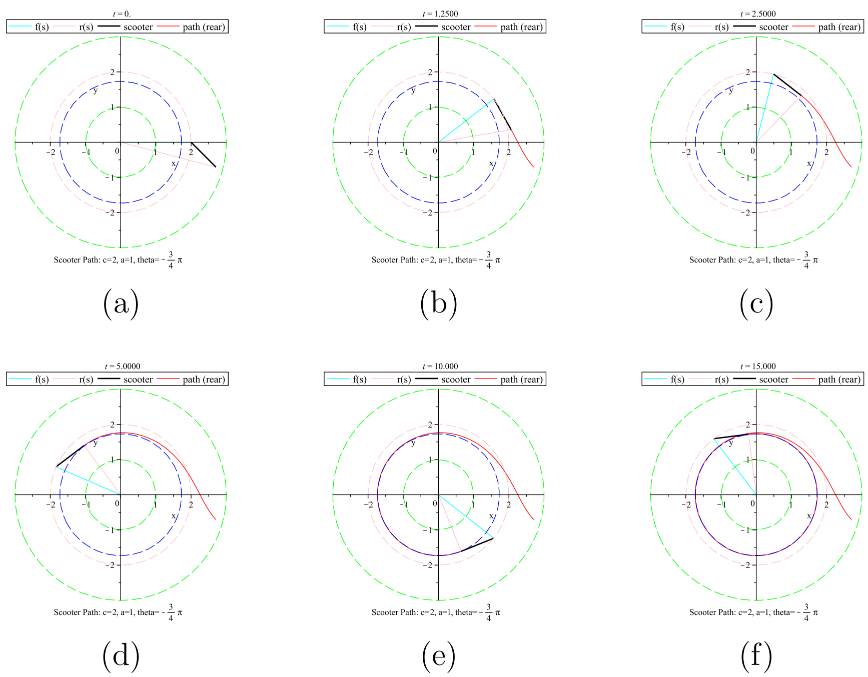


Figure C.20.4: The path of the rear wheel of a scooter with length $a = 1$, whose front wheel moves counter-clockwise around the circle with radius $c = 2$ from the point $(2, 0)$ with the line of the scooter at an angle $\theta = -3\pi/4$ with the front wheel. The snapshots are taken when (a) $s = 0$, (b) $s = 1.25$, (c) $s = 2.50$, (d) $s = 5.0$, (e) $s = 10.0$, and (f) $s = 15.0$. Because this is a short scooter ($a < c$), the rear wheel approaches the circle with radius $r = \sqrt{c^2 - a^2} = \sqrt{3}$. (Recall that s is the arclength of the rear wheel's path.)

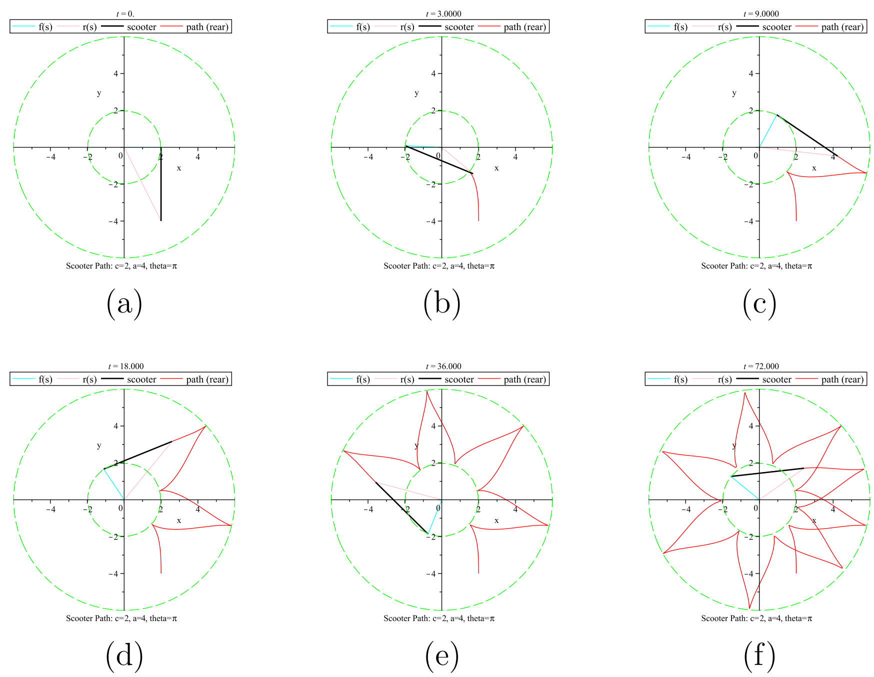


Figure C.20.5: The path of the rear wheel of a scooter with length $a = 4$, whose front wheel moves counter-clockwise around the circle with radius $c = 2$ from the point $(2, 0)$ with the line of the scooter at an angle $\theta = pi$ with the front wheel. The snapshots are taken when (a) $s = 0$, (b) $s = 3$, (c) $s = 9$, (d) $s = 18$, (e) $s = 36$, and (f) $s = 72$. Because this is a long scooter ($a > c$), the rear wheel travels along path that has cusps whenever $r = c + a$ and $f = |c - a|$. (Recall that s is the arclength of the rear wheel's path.)

This leads to the equation

$$a^2 + r^2 - c^2 = e^k e^{s/a}. \qquad \text{(C.20.10)}$$

Equation (C.20.10) implies that as s increases r becomes arbitrarily large. However, r can never exceed $c + a$. So, another cusp must form.

It can be shown that the cusps occur when $r = a - c$ (assuming $a > c$) and $r = a + c$. At the vertex of a cusp, $\mathbf{r}'$ is not defined; it changes direction by π.

Figure C.20.5(b) shows the shape of the path of the rear wheel for a long scooter, $a > c$. (For $a > 2c$, that path remains outside the circle.)

EXERCISES

1.[R]

(a) Assume a and c are positive numbers with $c > a$ and that the front wheel moves on a circle of radius c. Show that when the front wheel moves along a circle of radius c the rear wheel could remain on a concentric circle of radius $b = \sqrt{(}c^2 - a^2)$.

(b) Draw the triangle whose sides are a, b, and c and explain why the result in (a) is plausible.

2.[M] We assumed in the case of the short scooter that initially $r^2 < c^2 - a^2$. Examine the case in which initially $r^2 > c^2 - a^2$. Again, assume that initially the scooter is not backing up.

3.[M] We assumed in the case of the short scooter that initially $r^2 < c^2 - a^2$ and that the scooter is not backing up. Investigate what happens when we assume that initially $r^2 < c^2 - a^2$ and the scooter is backing up.

(a) Draw such an initial position.

(b) Predict what will happen.

(c) Carry out the mathematics.

4.[R] It is a belief among many bicyclists that the rear tire wears out more slowly than the front tire. Decide whether their belief is justified. (Assume both tires support the same weight.)

5.[M] Show that if the path of the front wheel is a circle and a cusp forms in the path of the rear wheel, the scooter at that moment lies on a line through the center of the circle.

6.[M] In the case of the long scooter, $a > c$, do cusps always form, whatever the initial value of r and θ?

7.[C] Extend the analysis of the scooter to the case when $a = c$.

8.[C] Assume that the path of the front wheel is a straight line. For convenience, choose that line as the x-axis. Write $\mathbf{r}(s)$ as $x(s)\mathbf{i} + y(s)\mathbf{j}$.

(a) Show that $y(s) + y'(s)a = 0$.

(b) Deduce that there is a constant k such that $y(s) = ke^{-s/a}$. Thus the distance from the rear wheel to the x-axis "decays" exponentially.

Chapter 16

Partial Derivatives

The use of contour lines to help understand a function whose domain is part of the plane goes back to the year 1774. A group of surveyors had collected a large number of the elevations of points on Mount Schiehalli in Scotland. They were doing this in order to estimate its mass and by its gravitational attraction, the mass of the earth. They asked the mathematician Charles Hutton for help in using the data entered as a map. Hutton saw that if he connected points on the map that showed the same elevation, the resulting curves — contour lines — suggested the shape of the mountain.

Reference: Bill Bryson, *A Short History of Nearly Everything*, Broadway Books, New York, 2003, p. 57.

16.1 Picturing a Function of Several Variables

The graph of $y = f(x)$, a function of just one variable, x, is a curve in the xy-plane. The graph of a function of two variables, $z = f(x, y)$ is a surface in space. It consists of the points (x, y, z) for which $z = f(x, y)$. For instance, if $z = 2x + 3y$, the graph is the plane $2x + 3y - z = 0$.

A vector field in the xy-plane is a vector-valued function of x and y. We pictured it by drawing a few vectors with their tails placed at the arguments.

This section describes some of the ways of picturing a scalar-valued functions of two or three variables.

Contour Lines

This is similar to what we did for vector fields.

For a function, $z = f(x, y)$, the simplest method is to attach at some point (x, y) the value of the function, $z = f(x, y)$. For instance, if $z = xy$, Figure 16.1.1 shows this method. This conveys a sense of the function. Its

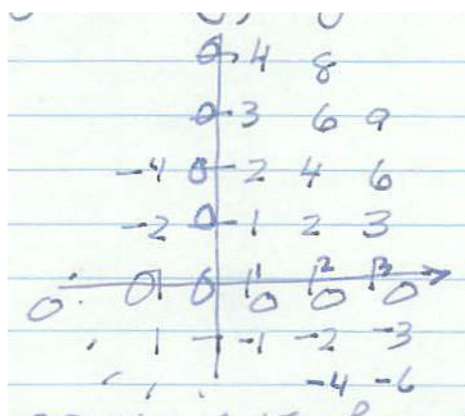


Figure 16.1.1:

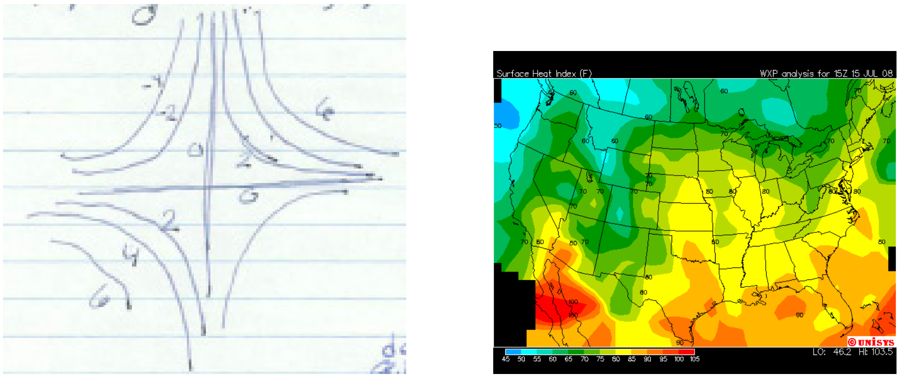
values are positive in the first and third quadrants, negative in the second and fourth. For (x, y) far from the origin near the lines $y = x$ or $y = -x$ the values are large.

Rather than attach the values at points, we could indicate all the points where the function has a specific fixed value. In other words we could graph, for a constant c, all the points where $f(x, y) = c$. Such as graph is called a **contour** or **level curve**.

contours and level curves

For the function $z = xy$, the contours are hyperbolas $xy = c$. In Figure 16.1.2(a) the contours corresponding to $c = 2,\ 4,\ 6,\ 0,\ -2,\ -4,\ -6$ are shown.

Many newspapers publish a daily map showing the temperature throughout the nation with the aid of contour lines. Figure 16.1.2(b) is an example.



(a) (b)

Figure 16.1.2:

At a glance you can see where it is hot or cold and in what direction to travel to warm up or cool off.

Traces

SHERMAN: xrcs Katrina wind / pressure?

Another way to get some idea of what the surface $z = f(x, y)$ looks like is to sketch the intersection of various planes with the surface. These intersections (or cross sections) are called **traces**.

For instance, Figure 16.1.3 exhibits the notion of a trace by a plane parallel to the xy-coordinate plane, namely, the plane $z = k$ This trace is an exact copy of the contour $f(x, y) = k$, as shown in Figure 16.1.3.

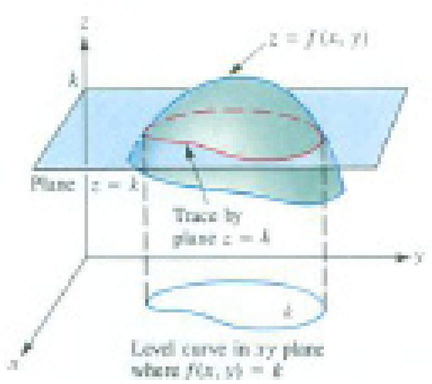


Figure 16.1.3:

Doug: maybe $z = x^2 - y^2$ is better? SHERMAN: I'm OK with xy, it's just a matter of perspective.

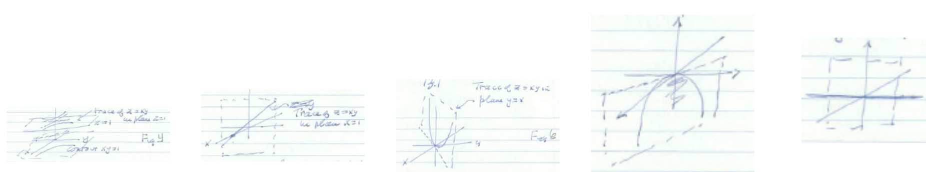
EXAMPLE 1 Sketch the traces of the surface $z = xy$ with the planes

1. $z = 1$,

2. $x = 1$,
3. $y = x$,
4. $y = -x$,
5. $x = 0$.

SOLUTION

1. The trace with the plane $z = 1$ is shown in Figure 16.1.4. For points (x, y, z) on this trace $xy = 1$. The trace is a hyperbola. In fact, it is just the contour line $xy = 1$ in the xy plane raised by one unit as in Figure 16.1.4(a)
2. The trace in the plane $x = 1$ satisfies the equation $z = 1 \cdot y = y$. It is a straight line, shown in Figure 16.1.4(b)
3. The trace in the plane $y = x$ satisfies the equation $z = x^2$. It is the parabola shown in Figure 16.1.4(c).
4. The trace in the plane $y = -x$ satisfies the equation $z = x(-x) = -x^2$. It is an "upside-down" parabola, shown in Figure 16.1.4(d).
5. The intersection with the coordinate plane $x = 0$ satisfies the equation $z = 0 \cdot y = 0$. It is the y-axis, shown in Figure 16.1.4(e).

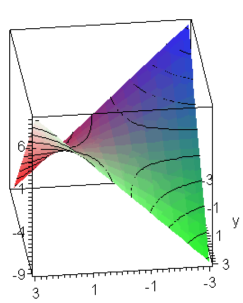


(a) (b) (c) (d) (e)

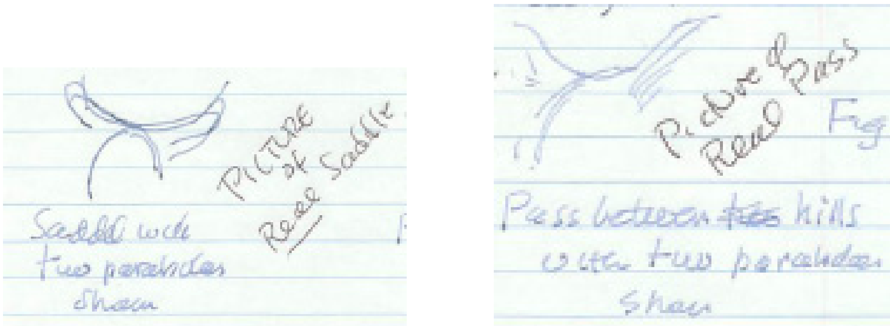
Figure 16.1.4:

So the surface can be viewed as made up of lines, or of parabolas or of hyperbolas.

The surface $z = xy$ is shown in Figure 16.1.5 with some of the traces drawn on it. ◇



The surface $z = xy$ looks like a saddle or the pass between two hills, as shown in Figure 16.1.6.



(a) (b)

Figure 16.1.6:

Functions of Three Variables

The graph of $y = f(x)$ consists of certain points in the xy plane. The graph of $z = f(x, y)$ consists of certain points in the xyz space. But what if we have a function of three variables, $u = f(x, y, x)$? (The volume V of a box of sides x, y, z is given by the equation $V = xyz$; this is an example of the function of three variables.) We cannot graph the set of points (x, y, z, u) where $u = f(x, y, z, u)$ since we live in space of only three dimensions. What we could do is pick a constant k and draw the "level surfaces," the set of points where $f(x, y, z) = k$. Varying k may give an idea of this function's behavior, just as varying the k of $f(x, y) = k$ yields information about the behavior of a function of two variables.

For example, let $T = f(x, y, z)$ be the temperature (Fahrenheit) at the point (x, y, z). Then the level surface

$$68 = f(x, y, z)$$

consists of all points where the temperature is 68°.

EXAMPLE 2 Describe the level surfaces of the function $u = x^2 + y^2 + z^2$.
SOLUTION For each k we examine the equation $u = x^2 + y^2 + z^2$. If k is negative, there are no points in the "level surface." If $k = 0$, there is only one point, the origin $(0, 0, 0)$. If $k = 1$, the equation $1 = x^2 + y^2 + z^2$, which describes a sphere of radius 1 center $(0, 0, 0)$. If k is positive, the level surface $f(x, y, z) = k$ is a sphere of radius $\sqrt{k}$, center $(0, 0, 0)$. See Figure 16.1.7 ⋄

Summary

We introduced the idea of a function of two variables $z = f(P)$ is in some region in the xy plane. The graph of $z = f(P)$ is usually a surface. But it is

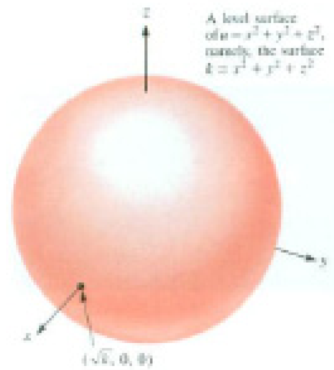


Figure 16.1.7:

often more useful to sketch a few of its level curves than to sketch that surface. Each level curve is the projection of a trace of the surface in a plane of the form $z = k$. Note that at all points (x, y) on a level curve the function have the same value. In other words, the function f is constant on a level curve.

In particular, we used level curves to analyze the function $z = xy$ whose graph is a saddle.

For functions of three variables $u = (x, y, z)$, we defined level surfaces. When considered on a level surface, $k = f(x, y, z)$ such a function is constant, with value k.

EXERCISES for Section 16.1 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 10, graph the given function.

1.[R] $f(x, y) = y$
2.[R] $f(x, y) = x + 1$
3.[R] $f(x, y) = 3$
4.[R] $f(x, y) = -2$
5.[R] $f(x, y) = x^2$
6.[R] $f(x, y) = y^2$
7.[R] $f(x, y) = x + y + 1$
8.[R] $f(x, y) = 2x - y + 1$
9.[R] $f(x, y) = x^2 + 2y^2$
10.[M] $f(x, y) = \sqrt{x^2 + y^2}$

In Exercises 11 to 14 draw for the given functions the level curves corresponding to the values -1, 0, 1, and 2 (if they are not empty).

11.[R] $f(x, y) = x + y$
12.[R] $f(x, y) = x + 2y$
13.[R] $f(x, y) = x^2 + 2y^2$
14.[R] $f(x, y) = x^2 - 2y^2$

In Exercises 15 to 18 draw the level curves for the given functions that pass through the given points.

15.[R] $f(x, y) = x^2 + y^2$ through $(1, 1)$ Hint: First compute $f(1, 1)$.
16.[R] $f(x, y) = x^2 + 3y^2$ through $(1, 2)$
17.[R] $f(x, y) = x^2 - y^2$ through $(3, 2)$
18.[R] $f(x, y) = x^2 - y^2$ through $(2, 3)$

19.[R]

(a) Draw the level curves for the functions $f(x, y) = x^2 + y^2$ corresponding to the values $k = 0, 1, \ldots, 9$.

(b) By inspection of the curves in (a), decide where the functions changing most rapidly. Explain why you think so.

20.[R] Let $f(P)$ be the average daily solar radiation at the point P (measured in langleys). The level curves corresponding to 350, 400, 450, and 500 langley are shown

in Figure 16.1.8.

Figure 16.1.8:

(a) What can be said about the ratio between the maximum and minimum solar radiation at points in the United States?

(b) Why are there rather sharp bends in the level curves in two areas?

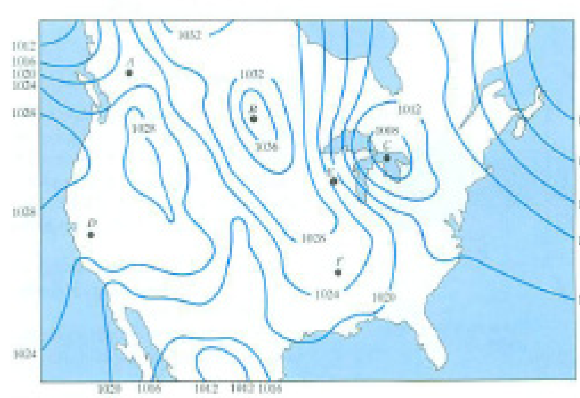
21.[R] Let $u = g(x, y, z)$ be a function of three variables. Describe the level surface $g(x, y, z) = 1$ if $g(x, y, z)$ is

(a) $x + y + z$

(b) $x^2 + y^2 + z^2$

(c) $x^2 + y^2 - z^2$

(d) $x^2 - y^2 - z^2$ Hint: For (c) and (d) are examples of quadric surfaces.



22.[R]

Figure 16.1.9:

The daily weather map shows the barometric pressure function by a few well-chosen level curves (called *isobars*), as in Figure 16.1.9. In this case, the function is 'pressure at (x, y)."

(a) Where is the lowest pressure?

(b) Where is the highest pressure?

(c) Where do you think the wind at ground level is the fastest? Why?

23.[R] A map of August, 26, 2005 showing isobars and wind vectors, day of Katrina and some questions.

24.[R] Questions about the map in Figure 16.1.2(b).

25.[M]

(a) Sketch the surface $z = x^2 + y^2$.

(b) Show that all the traces by planes parallel to the xz plane are parabolas.

(c) Show that the parabolas in (b) are all congruent. (So the surface is made up of identical parabolas.)

(d) What kind of curve is a trace in a plane parallel to the xy plane?

26.[M] Consider the surface $z = x^2 + 4y^2$. What type of curve is produced by a trace by a plane parallel to

(a) the xy plane,

(b) the xz palne,

(c) the yz plane.

27.[C]

(a) Is the parabola $y = x^2$ congruent to the parabola $y = 4x^2$?

(b) Is the parabola $y = x^2$ similar to the parabola $y = 4x^2$? (One figure is similar to another if one is simply the other magnified by the same factor in all direction.)

16.2 The Many Derivatives of $f(x, y)$.

The notions of limit, continuity and derivative carry over with similar definitions from functions $f(x)$ of one variable to functions of several variables $f(x, y)$. However, the derivatives of functions of several variable involves some new ideas.

SHERMAN: There are more new ideas for limits than derivatives. In fact, partial

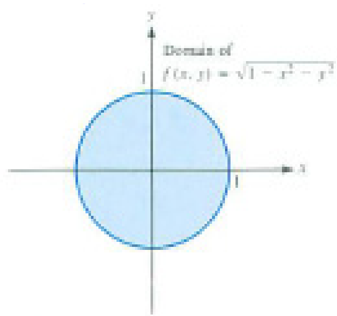


Figure 16.2.1:

Limits and Continuity of $f(x, y)$

The **domain** of function $f(x, y)$ is the set of points where it is defined. The domain of $f(x, y) = x + y$ is the entire xy plane. The domain of $f(x, y) = \sqrt{1 - x^2 - y^2}$ is much smaller. In order for the square root of $1 - x^2 - y^2$ to be defined, $1 - x^2 - y^2$ must not be negative. In other words, we must have $x^2 + y^2 \leq 1$. The domain is the disk bounded by the circle $x^2 + y^2 = 1$, shown in Figure 16.2.1.

A point P_0 is on the **boundary** of a set if every disk centered at P_0, no matter how small, contains points in the set and points not in the set. (See Figure 16.2.3.) The boundary of the circle $x^2 + y^2 \leq 1$ is the circle $x^2 + y^2 = 1$. The domain of $f(x, y) = \sqrt{1 - x^2 - y^2}$ includes every point on its boundary.

The domain of $f(x, y) = 1/\sqrt{1 - x^2 - y^2}$ is even smaller. Now we must not let $1 - x^2 - y^2$ be 0 or negative. The domain of $1/\sqrt{1 - x^2 - y^2}$ consists of the points (x, y) such that $x^2 + y^2 < 1$. It is the disk in Figure 16.2.1 *without* its boundary.

The function $f(x, y) = 1/(y - x)$ is defined everywhere except on the line $y - x = 0$. Its domain is the xy plane from which the line $y = x$ is removed. (See Figure 16.2.2.)

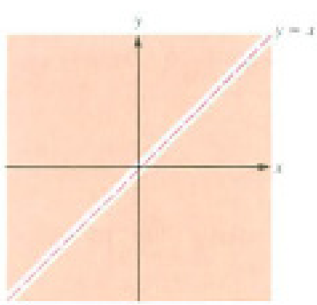
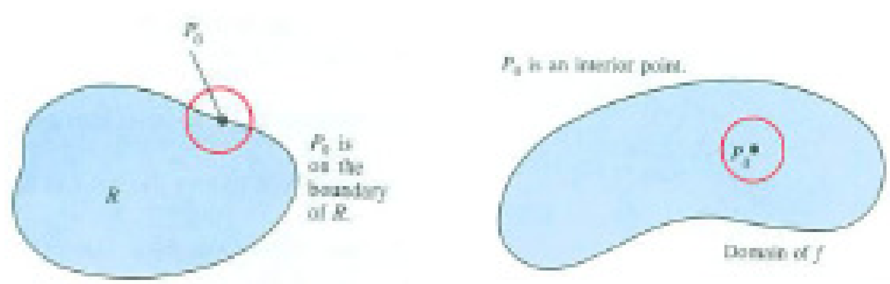


Figure 16.2.2:

The domain of a function of interest to us will either be the entire xy plane or some region bordered by curves or lines, or perhaps such a region with a few points omitted. Let P_0 be a point in the domain of a function f. If there is



(a) (b)

Figure 16.2.3:

a disk with center P_0 that lies within the domain of f, we call P_0 an **interior** point of the domain. (See Figure 16.2.3(b).) When P_0 is an interior point of the domain of f, we know that $f(P)$ is defined for all points P sufficiently near P_0. Every point P_0 in the domain not on its boundary is an interior point. A set R is called **open** if each point P of R is an interior point of R. The entire xy plane is open. So is any disk without its circumference. More generally, the set of points inside some closed curve but not on it forms an open set.

The definition of the limit of $f(x, y)$ as (x, y) approaches $P_0 = (a, b)$ will not come as a surprise.

> **DEFINITION** *(Limit of $f(x, y)$ at $P_0 = (a, b)$)* Let f be a function defined at least at every point in some disk with center P_0, except perhaps at P_0. If there is a number L such that $f(P)$ approaches L whenever P approaches P_0 we call L the **limit of** $f(P)$ **as** P **approaches** P_0. We write
>
> $$\lim_{P \to P_0} f(P) = L$$
>
> or
>
> $$f(P) \to L \qquad \text{as} \qquad P \to P_0.$$
>
> We also write
>
> $$\lim_{(x,y) \to (a,b)} f(x, y) = L.$$

For most of the functions of interest the limit will always exist throughout its domain. However, even a formula that is easily defined may not have a limit at some points.

EXAMPLE 1 Let $f(x, y) = \frac{x^2 - y^2}{x^2 + y^2}$. Determine whether $\lim_{P \to (0,0)} f(P)$ exists.

SOLUTION The function is not defined at $(0, 0)$. When (x, y) is near $(0, 0)$, both the numerator and denominator of $(x^2 - y^2)/(x^2 + y^2)$ are small numbers. There are, as in Chapter 2, two influences. The numerator is pushing the quotient towards 0 while the denominator is influencing the quotient to be large. We must be careful.

We try a few inputs near $(0, 0)$. For instance, $(0.01, 0)$ is near $(0, 0)$ and

$$f(0.01, 0) = \frac{(0.01)^2 - 0^2}{(0.01)^2} + 0^2 = 1$$

Also, $(0, 0.01)$ is near $(0, 0)$ and

$$f(0, 0.01) = \frac{0^2 - (0.01)^2}{0^2 + (0.01)^2} = -1$$

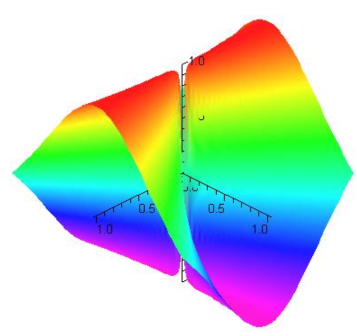


Figure 16.2.4:

More generally, for $x \neq 0$,

$$f(x,0) = 1;$$

while, for $y \neq 0$,

$$f(0,y) = -1$$

Since x can be as near 0 as we please and y can be as near 0 as we please, it is *not* the case that $\lim_{P \to (0,0)} f(P)$ exists. Figure 16.2.4 shows the graph of $z = \frac{x^2-y^2}{x^2+y^2}$. ◇

Continuity of $f(x,y)$ at $P_0 = (a,b)$

With only slight changes, the definition of continuity for $f(x)$ in Section 2.4 easily generalizes to the definition of continuity for $f(x,y)$.

> **DEFINITION** *(Continuity of $f(x,y)$ at $P_0 = (a,b)$).* Assume that $f(P)$ is defined throughout some disk with center P_0. Then f is **continuous** at P_0 if $\lim_{P \to P_0} f(P) = f(P_0)$.
>
> This means
>
> 1. $f(P_0)$ is defined (that is, P_0 is in the domain of f),
> 2. $\lim_{P \to P_0} f(P)$ exists, and
> 3. $\lim_{P \to P_0} f(P) = f(P_0)$.
>
> Continuity at a point on the boundary of the domain can be defined similarly. A function $f(P)$ is **continuous** if it is continuous at every point in its domain.

EXAMPLE 2 Determine whether $f(x,y) = \frac{x^2-y^2}{x^2+y^2}$ is continuous at $(1,1)$.
SOLUTION This is the function explored in Example 1. First, $f(1,1)$ is defined. (It equals 0.) Second, $\lim_{(x,y)\to(1,1)} \frac{x^2-y^2}{x^2+y^2}$. (It is $\frac{0}{2} = 0$.) Third, $\lim_{(x,y)\to(1,1)} f(x,y) = f(1,1)$.

Hence, $f(x,y)$ is continuous at $(1,1)$. ◇

In fact, the function of Example 2 is continuous at every point (x,y) in its domain. We do not need to worry about the behavior of $f(x,y)$ when (x,y) is near $(0,0)$ because $(0,0)$ is *not* in the domain. Since $f(x,y)$ is continuous at every point *in its domain*, it is a continuous function.

The Two Partial Derivatives of $f(x, y)$

Let (a, b) be a point on the domain of $f(x, y)$. The trace on the surface $z = f(x, y)$ by a plane through (a, b) and parallel to the z-axis is a curve, as shown in Figure 16.2.5.

If f is well behaved at the point $P = (a, b, f(a, b))$ the trace has a slope. This slope depends on the plane through (a, b). In this section we consider only the two planes parallel to the coordinate planes $y = 0$ and $x = 0$. In the next section we treat the general cases.

Consider the function $f(x, y) = x^2y^3$. If we hold y constant and differentiate with respect to x, we obtain $d(x^2y^3)/dx = 2xy^3$. This derivative is called the "partial derivative" of x^2y^3 with respect to x. We could hold x fixed instead and find the derivative of x^2y^3 with respect to y, that is, $d(x^2y^3)/dy = 3x^2y^2$. This derivative is called the "partial derivative" of x^2y^3 with respect to y. This example introduces the general idea of partial derivative. First we define them. Then we will see what they mean in terms of slope and rate of change.

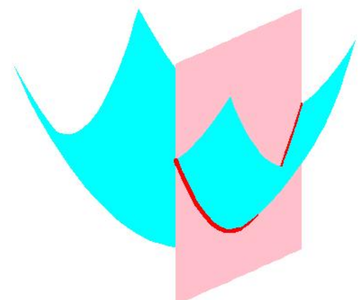


Figure 16.2.5:

DEFINITION *(Partial derivatives.)* Assume that the domain of $f(x, y)$ includes the region within some disk with center (a, b). If

$$\lim_{\Delta x \to 0} \frac{f(a + \Delta x, b) - f(a, b)}{\Delta x}$$

exists, this limit is called the **partial derivative of f with respect to x** at (a, b). Similarly, if

$$\lim_{\Delta x \to 0} \frac{f(a, b + \Delta y) - f(a, b)}{\Delta y}$$

exists, it is called the **partial derivative of f with respect to y** at $9a, b)$.

The following notations are used for the partial derivatives of $z = f(x, y)$ with respect to x:

Notations for partial derivatives.

$$\frac{\partial z}{\partial x}, \frac{\partial f}{\partial y}, f_x, f_1, \quad \text{or} \quad z_x.$$

And the following are used for partial derivative of $z = f(x, y)$ with respect to y:

$$\frac{\partial z}{\partial y}, \frac{\partial f}{\partial y}, f_y, f_2, \quad \text{or} \quad z_y.$$

Since physicists and engineers use the subscript notation in study of vectors, they prefer to use

$$\frac{\partial f}{\partial x} \quad \text{and} \quad \frac{\partial f}{\partial y}$$

to denote the two partial derivatives. The symbol $\partial\partial f/\partial x$ may be viewed as "the rate at which the function $f(x, y)$ changes when x varies and y is kept fixed." The symbol $\partial f/\partial y$ records "the rate at which the function $f(x, y)$ changes when y varies and x is kept fixed."

The value of $\partial f/\partial x$ at (a, b) is denoted

$$\frac{\partial f}{\partial x}(a, b) \quad \text{or} \quad \left.\frac{\partial f}{\partial x}\right|_{(a,b)}.$$

In the middle of a sentence, we will write it as $f_x(a, b)$ or $\partial f/\partial x(a, b)$.

EXAMPLE 3 If $f(x, y) = \sin(x^2 y)$, find

1. $\partial f/\partial x$,
2. $\partial f/\partial y$, and
3. $\partial f/\partial y$ at $(1, \pi/4)$.

SOLUTION

1. To find $\frac{\partial}{\partial x}(\sin x^2 y)$, differentiate with respect to x, keeping y constant:

$$\begin{aligned} \tfrac{\partial}{\partial x}(\sin x^2 y) &= \cos(x^2 y)\tfrac{\partial}{\partial x}(x^2 y) && \text{chain rule} \\ &= \cos(x^2 y)(2xy) && y \text{ is constant} \\ &= 2xy\cos(x^2 y). \end{aligned}$$

2. To find $\frac{\partial}{\partial y}(\sin x^2 y)$, differentiate with respect to y, keeping x constant:

$$\begin{aligned} \tfrac{\partial}{\partial y}(\sin x^2 y) &= \cos(x^2 y)\tfrac{\partial}{\partial y}(x^2 y) && \text{chain rule} \\ &= \cos(x^2 y)(x^2) && x \text{ is constant} \\ &= x^2\cos(x^2 y). \end{aligned}$$

3. By (b)

$$\tfrac{\partial f}{\partial y}(1, \pi/4) = x^2\cos(x^2 y)\big|_{(1,\pi/4)} = 1^2\cos\left(1^2\tfrac{\pi}{4}\right) = \tfrac{\sqrt{2}}{2}.$$

◇

As Example 3 shows, since partial derivatives are really ordinary derivatives, the procedures for computing derivatives of a function $f(x)$ of a single variable carry over to functions of two variables.

Higher-Order Partial Derivatives

Just as there are derivatives of derivatives, so are there partial derivatives of partial derivatives. For instance, if

$$z = 2x + 5x^4y^7,$$

then

$$\frac{\partial z}{\partial x} = 2 + 20x^3y^7 \qquad \text{and} \qquad \frac{\partial z}{\partial y} = 35x^4y^6.$$

We may go on and compute the partial derivatives of $\partial z/\partial x$ and $\partial z/\partial y$:

$$\begin{aligned} \tfrac{\partial}{\partial x}\left(\tfrac{\partial z}{\partial x}\right) &= 60x^2y^7 & \tfrac{\partial}{\partial y}\left(\tfrac{\partial z}{\partial y}\right) &= 140x^3y^6 \\ \tfrac{\partial}{\partial x}\left(\tfrac{\partial z}{\partial y}\right) &= 140x^3y^6 & \tfrac{\partial}{\partial y}\left(\tfrac{\partial z}{\partial y}\right) &= 210x^4y^5. \end{aligned}$$

There are four partial derivatives of the second order:

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

These are usually denoted, in the same order, as

$$\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial y \partial x}, \frac{\partial^2 z}{\partial y^2}, \frac{\partial^2 z}{\partial x \partial y}.$$

To compute $\partial^2 z/\partial x\partial y$, you first differentiate with respect to y, then with respect to x. To compute $\partial^2 z/\partial y\partial x$, you first differentiate with respect to x, then with respect to y. In both cases, "differentiate from right to left in the order that the variables occur."

The partial derivative $\frac{\partial f}{\partial x}$ is also denoted f_x and $\frac{\partial f}{\partial y}$ is denoted f_y. The second partial derivative $\frac{\partial^2 f}{\partial x\partial y} = \frac{\partial(f_y)}{\partial x} = (f_y)_x$ is denoted f_{yx}. In this case you differentiate from left to right, "first f_y, then $(f_y)_x$." In short, $f_{yx} = (f_y)_x$, $f_{yy} = (f_y)_y$, and $f_{xy} = (f_x)_y$. In both notations the mixed partial is computed in the order that resembles its definition (with the parentheses removed), Thus

The subscript notation, f_{yx}, is generally preferred in the midst of other text.

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \qquad \text{and} \qquad f_{xy} = (f_x)_y$$

are the two different mixed second partial derivatives of f.

Equality of the mixed partials

In the computations just done, the two mixed partials z_{xy} and z_{yx} are equal. For the functions commonly encountered, the two mixed partials are equal. (For a proof, see Appendix K.)

SHERMAN: V had an appedix on interchanging limits. How will we deal with this in VI?

Exercise 27 presents a function for which the two mixed particles are not equal. Such a special case mathematicians call "pathological", though the function does not view itself as sick.

EXAMPLE 4 Compute $\frac{\partial^2 z}{\partial x^2} = f_{xx}$, $\frac{\partial^2 z}{\partial y \partial x} = f_{xy}$, and $\frac{\partial^2 z}{\partial x \partial y} = z_{yx}$ for $z = y\cos(xy)$.

SOLUTION First compute

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}(-y^2\sin(xy)) = -y^3\cos(xy).$$

Then

$$\frac{\partial^2 z}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial y}(-y^2\cos(xy)) = -2y\sin(xy) - xy^2\cos(xy).$$

Finally,

$$\begin{aligned}
\frac{\partial^2 z}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial y}\right) = \frac{\partial}{\partial x}(-yx\sin(xy) + \cos(xy)) \\
&= -y\frac{\partial}{\partial x}(x\sin(xy) + \frac{\partial}{\partial x}(\cos(xy)) = -y(xy\cos(xy) + \sin(xy)) - y\sin(xy) \\
&= -xy^2\cos(xy) - y\sin(xy) - y\sin(xy) = -2y\sin(xy) - xy^2\cos(xy).
\end{aligned}$$

Notice that while the work required to compute the mixed partials is very different, the two derivatives are, as expected, are equal. ◇

Functions of More Than Two Variables

A quantity may depend on more than two variables. For instance, the volume of a box depends on three variables: the length l, width w, and height h, $V = lwh$. The "chill factor" depends on the temperature, humidity, and wind velocity. The temperature T at any point in the atmosphere is a function of the three space coordinates, x, y, and z: $T = f(x, y, z)$.

To differentiate, hold all variables constant except one.

The notions and notations of partial derivatives carry over to functions of more than two variables. If $u = f(x, y, z, t)$, there are four first-order partial derivatives. For instance, the partial derivative of u with respect to x, holding y, z, and t fixed, is denoted

$$\frac{\partial u}{\partial x}, \ \frac{\partial f}{\partial x}, \ u_x, \quad \text{etc.}$$

Insert CIE on the Vibrating String.

Higher-ordered partial derivatives are defined and denoted similarly. Many basic problems in chemistry and physics, such as vibrating string are examined in terms of equations involving partial derivatives (known as PDEs).

Summary

We defined limit, continuity, and derivatives for functions of several variables. These notions are all closely related to the one variable versions.

A key difference is that a partial derivative with respect to one variable, say x, is found by treating all other variables as constants and applying the standard differentiation rules with respect to x. Higher-order partial derivatives are also defined much like higher-order derivatives. An important property of higher-order partial derivatives is that the order in which the partial derivatives are applied can be important, but not for the functions usually met in applications.

EXERCISES for Section 16.2 *Key:* R–routine, M–moderate, C–challenging

SHERMAN: Move some of these to Chapter Summary. Emphasis is on partial derivatives.

In Exercises 1 to 8 evaluate the limits, if they exist.

1.[R] $\lim_{(x,y)\to(2,3)} \frac{x+y}{x^2+y^2}$

2.[R] $\lim_{(x,y)\to(1,1)} \frac{x^2}{x^2+y^2}$

3.[R] $\lim_{(x,y)\to(0,0)} \frac{x^2}{x^2+y^2}$

4.[R] $\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$

5.[R] $\lim_{(x,y)\to(2,3)} x^y$

6.[R] $\lim_{(x,y)\to(0,0)} (x^2)^y$

7.[R] $\lim_{(x,y)\to(0,0)} (1+xy)^{1/(xy)}$

8.[R] $\lim_{(x,y)\to(0,0)} (1+x)^{1/y}$

In Exercises 9 to 14, (a) describe the domain of the given functions and (b) state whether the functions are continuous.

9.[R] $f(x,y) = 1/(x+y)$

10.[R] $f(x,y) = 1/(x^2+2y^2)$

11.[R] $f(x,y) = 1/(9-x^2-y^2)$

12.[R] $f(x,y) = \sqrt{x^2+y^2-25}$

13.[R] $f(x,y) = \sqrt{16-x^2-y^2}$

14.[R] $f(x,y) = \sqrt{49-x^2-y^2}$

In Exercises 15 to 20, find the boundary of the given region R.

15.[R] R consists of all points (x,y) such that $x^2+y^2 \le 1$.

16.[R] R consists of all points (x,y) such that $x^2+y^2 < 1$.

17.[R] R consists of all points (x,y) such that $1/(x^2+y^2)$ is defined.

18.[R] R consists of all points (x,y) such that $1/(x+y)$ is defined.

19.[R] R consists of all points (x,y) such that $y < x^2$.

20.[R] R consists of all points (x,y) such that $y \le x$.

In Exercises 21 to 24 concern the precise definition of $\lim_{(x,y)\to P_0} f(x,y)$.

21.[R] Let $f(x,y) = x+y$.

(a) Show that if $P=(x,y)$ lies within a distance 0.01 of $(1,2)$, then $|x-1| < 0.01$ and $|y-2| < 0.01$. (See Figure 16.2.6).

(b) Show that if $|x - 1| < 0.01$ and $|y - 2| < 0.01$, then $|f(x, y) - 3| < 0.02$.

(c) Find a number $\delta > 0$ such that if $P = (x, y)$ is in the disk of center $(1, 2)$ and radius δ, then $|f(x, y) - 3| < 0.001$.

(d) Show that for any positive number ϵ, no matter how small, there is a positive number δ such that when $P = (x, y)$ is in the disk of radius δ and center $(1, 2)$, then $|f(x, y) - 3| < \epsilon$. (Give δ as a function of ϵ.)

(e) What may we conclude on the basis of (d)?

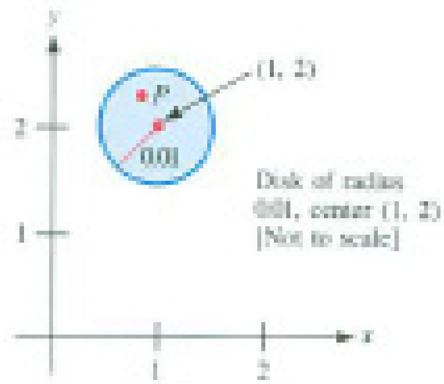


Figure 16.2.6:

22.[R] Let $f(x, y) = 2x + 3y$.

(a) Find a disk with center $(1, 1)$ such that whenever P is in that disk, $|f(P) - 5| < 0.01$

(b) Let ϵ be any positive number. Show that there is a disk with center $(1, 1)$ such that whenever P us in that disk, $|f(P) - 5| < \epsilon$. (Give δ as a function of ϵ.)

(c) What may we conclude on the basis of (b)?

23.[R] Let $f(x, y) = s^2y/(x^4 + 2y^2)$.

(a) What is the domain of f?

(b) Fill in this table:

(x, y)	$(0.01, 0.01)$	$(0.01, 0.02)$	$(0.001, 0.003)$
$f(x, y)$			

(c) On the basis of (b), do you think $\lim_{P \to (0,0)} f(P)$ exists? If so, what is its value?

(d) Fill in this table:

(x,y)	$(0.5, 0.25)$	$(0.1, 0.01)$	$(0.001, 0.000001)$
$f(x,y)$			

(e) On the basis of (d), do you think $\lim_{P\to(0,0)} f(P)$ exists? If so, what is its value?

(f) Does $\lim_{P\to(0,0)} f(P)$ exist? If so, what is it? Explain.

24.[R] Let $f(x,y) = 5x^2y/(2x^4+3y^2)$.

(a) What is the domain of f?

(b) As P approaches $(0,0)$ on the line $y = 2x$, what happens to $f(P)$?

(c) As P approaches $(0,0)$ on the line $y = 3x$, what happens to $f(P)$?

(d) As P approaches $(0,0)$ on the parabola $y = x^2$, what happens to $f(P)$?

(e) Does $\lim_{P\to(0,0)} f(P)$ exist? If so, what is it? Explain.

25.[R] Show that for any polynomial $P(x,y)$, P_{yx} equals P_{xy}. Suggestion: It is enough to show it for an arbitrary monomial ax^my^n, where a is constant and m and n are non-negative integers. The case where m or n is 0 should be treated separately.

26.[M] Let $T(x,y,z) = 1/\sqrt{x^2+y^2+z^2}$, if (x,y,z) is not the origin $(0,0,0)$. Show that

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0$$

This equation arises in the theory of heat as we will show in Section 16.4.

27.[C] This exercise presents a function $f(x,y)$ such that its two mixed partial derivatives at $(0,0)$ are not equal.

(a) Let $g(x,y) = \frac{x^2-y^2}{x^2+y^2}$ for (x,y) not $(0,0)$. Show that $\lim_{k\to 0}(\lim_{h\to 0} g(h,k)) = -1$ but $\lim_{h\to 0}(\lim_{k\to 0} g(h,k)) = 1$.

(b) Let $f(x,y) = xyg(x,y)$ if (x,y) is not $(0,0)$ and $f(0,0) = 0$. Show that $f(x,y) = 0$ if x or y is 0.

(c) Show that $f_{xy}(0,0) = \lim_{k\to 0} \frac{f_x(0,k)-f_x(0,0)}{k}$.

(d) Show that $f_{xy}(0,0) = \lim_{k\to 0} \left(\lim_{h\to 0} \frac{f(h,k)-f(j,k)-f(h,0)+f(0,0)}{hk}\right)$.

(e) Show that $f_{xy}(0,0) = -1$.

(f) Similarly, show that $f_{xy}(0,0) = 1$.

(g) Show that in polar coordinates the value of f at the point (r,θ) is $r^2 \sin(4\theta)/4$.

16.3 Change and the Chain Rule

For a function of one variable, $f(x)$, the change in the value of the function as the input changes from a to $a+\Delta x$ is approximately $f'(a)\Delta x$. In this section we estimate the change in $f(x,y)$ as (x,y) moves from (a,b) to $(a+\Delta x, b+\Delta y)$.

That type of estimate is the key to obtaining the chain rule for functions of several variables. We will find that the chain rule involves the *sum* of terms that resemble the product $\frac{dy}{du} \cdot \frac{du}{dx}$ that appear in the chain rule for a function of one variable.

Estimating the Change of Δf

Let $z = f(x,y)$ be a function of two variables with continuous partial derivatives at least throughout a disk centered at the point (a,b). We will express $\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b)$ in terms of f_x and f_y. This change is shown in Figure 16.3.1. We can view this change as obtained in two steps. First, the

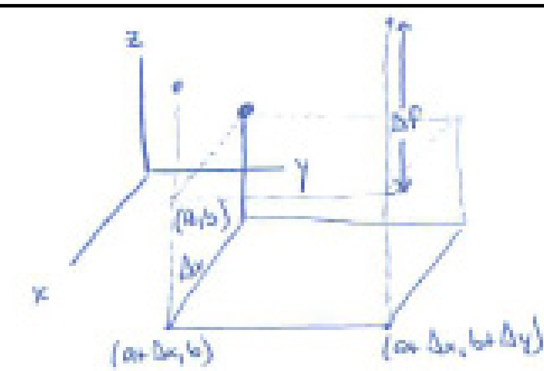


Figure 16.3.1:

change as x goes from a to $a+\Delta x$, that is, $f(a+\Delta x, b) - f(a,b)$. Second, the change from $f(a+\Delta x, b)$ to $f(a+\Delta x, b+\Delta y)$, as y changes from b to $b+\Delta y$.

In short,

$$\Delta f = (f(a+\Delta x, b) - f(a,b)) + (f(a+\Delta x, b+\Delta y) - f(a+\Delta, b)). \quad (16.3.1)$$

(16.3.1) is clear algebraically because the two $f(a+\Delta x, b)$ terms cancel.

By the mean-value theorem, there is a number c_1 between a and $a+\Delta x$ such that

$$f(a+\Delta x, b) - f(a,b) = \frac{\partial f}{\partial x}(c_1, b)\Delta x \quad (16.3.2)$$

Similarly, applying the mean-value theorem to the second bracket expression as (16.3.2), we see that there is a number c_3 between b and $b+\Delta y$ such that

$$f(a+\Delta x, b+\Delta y) - f(a+\Delta x, b) = \frac{\partial f}{\partial y}(a+\Delta x, c_2)\Delta y. \qquad (16.3.3)$$

Combining (16.3.1), (16.3.2) and (16.3.3) we obtain

$$\Delta f = \frac{\partial f}{\partial x}(c_1, b)\Delta x + \frac{\partial f}{\partial y}(z+\Delta x, c_2)\Delta y. \qquad (16.3.4)$$

When both Δx and Δy are small, the points (c_1, b) and $(a+\Delta x, c_2)$ are near the point (a, b). If we assume that the partial derivatives f_x are continuous at (a, b), then we may conclude that

$$\frac{\partial f}{\partial x}(c_1, b) = \frac{\partial f}{\partial x}(a, b) + \epsilon_1 \quad \text{and} \quad \frac{\partial f}{\partial y}(a+\Delta x, c_2) = \frac{\partial f}{\partial y}(a, b) + \epsilon_2, \quad (16.3.5)$$

where both ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0.

Combining (16.3.4) and (16.3.5) gives the key to estimating the change in the function f. We state this important result as a theorem.

Theorem 16.3.1. *Let f have continuous partial derivatives f_x and f_y for all points within some disk with center at the point (a, b). Then Δf, which is the change $f(a+\Delta x, b+\Delta y) - f(a, b)$, can be written*

$$\Delta f = \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y, \qquad (16.3.6)$$

where ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0. (Both ϵ_1 and ϵ_2 are functions of the four variables a, b, Δx and Δy.)

This equation is the core of this section.

The term $f_x(a, b)\Delta x$ estimates the change due to the change in the x-coordinate, while $f_y(a, b)\Delta y$ estimates the change due to the change in the y-coordinate.

We will call $f(x, y)$ **differentiable** at (a, b) if (16.3.6) holds. In particular if the partial derivatives f_x and f_y exist in a disk around (a, b) and are continuous at (a, b), then f is differentiable at (a, b).

Since ϵ_1 and ϵ_2 in (16.3.6) both approach 0 as Δx and Δy approach 0,

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y, \qquad (16.3.7)$$

The approximation (16.3.7) gives us a way to estimate Δf when Δx and Δy are small.

EXAMPLE 1 Estimate $(2.1)^2(0.95)^3$.

SOLUTION Let $f(x, y) = x^2y^3$. We wish to estimate $f(2.1, 0.95)$. We know that $f(2, 1)$ equals $2^21^3 = 4$. We use (16.3.7) to estimate $\Delta f = f(2.1, 0.95) - f(2, 1)$. We have

$$\frac{\partial(x^2y^3)}{\partial x} = 2xy^3 \quad \text{and} \quad \frac{\partial(x^2y^3)}{\partial y} = 3x^2y^2.$$

Then

$$\frac{\partial f}{\partial x}(2, 1) = 4 \quad \text{and} \quad \frac{\partial f}{\partial y}(2, 1) = 12.$$

Since $\Delta x = 0.1$ and $\Delta y = -0.05$, we have

$$\Delta f \;=\; 4(0.1) + 12(-0.05) = 0.4 - 0.6 = -0.2.$$

The exact value is 3.78102375.

DOUG: You may want to do more with "approximation". SHERMAN: What do you mean by this?

Thus $(2, 1)^2(0.95)^3$ is approximately $4 + (-0.2) = 3.8$. ◇

The Chain Rule

We begin with two special cases of the chain rule for functions of more than one variable. Afterward we will state the chain rule for functions of any number of variables.

The first theorem considers the case when $z = f(xy)$ and x and y are functions of just one variable t. The second theorem is more general, where x and y may be functions of two variables, t and u.

Theorem. *Chain Rule – Special Case #1 Let $z = f(x, y)$ have continuous partial derivatives f_x and f_y, and let $x = x(t)$ and $y = y(t)$ be differentiable functions of t. Then z is a differentiable function of t and*

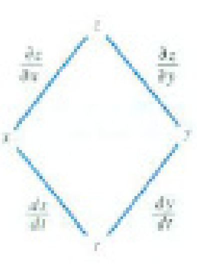
$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}. \tag{16.3.8}$$

Proof

By definition,

$$\frac{dz}{dt} = \lim_{\Delta t \to 0} \frac{\Delta z}{\Delta t}.$$

Now, Δt induces changes Δx and Δy in x and y, respectively. According to Theorem 16.3.1,



$$\Delta z = \frac{\partial f}{\partial x}(x, y)\Delta x + \frac{\partial f}{\partial y}(x, y)\Delta y + \epsilon_1\Delta x + \epsilon_2\Delta y,$$

where $\epsilon_1 \to 0$ and $\epsilon_2 \to 0$ as Δx and Δy approach 0. (Keep in mind that x and y are fixed.) Thus

$$\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}(x,y)\frac{\Delta x}{\Delta t} + \frac{\partial f}{\partial y}(x,y)\frac{\Delta y}{\Delta t} + \epsilon_1\frac{\Delta x}{\Delta t} + \epsilon_2\frac{\Delta y}{\Delta t}.$$

and

$$\frac{dz}{dt} = \lim_{\Delta t \to 0}\frac{\Delta z}{\Delta t} = \frac{\partial f}{\partial x}(x,y)\frac{dx}{dt} + \frac{\partial f}{\partial y}(x,y)\frac{dy}{dt} + 0\frac{dx}{dt} + 0\frac{dy}{dt}.$$

This proves the theorem. •

MEMORY AID: Each path produces one summand. And, each leg in each path produces one factor in that summand.

The two summands on the right-hand sides of (16.3.8) remind us of the chain rule for functions of one variable. Why is there a "+" in (16.3.8)? The "+" first appears in (16.3.4) and you can trace it back to Figure 16.3.1.

The diagram in Figure 16.3.2 helps in using this special case of the chain rule . There are two paths from the top variable z down to the bottom variable t. Label each edge with the appropriate partial derivative (or derivative). For each path there is a summand in the chain rule. The left-hand path (see Figure 16.3.3) gives us the summand

$$\frac{\partial z}{\partial x}\frac{dx}{dt}.$$

The right-hand path (see Figure 16.3.4) gives us the summand

$$\frac{\partial z}{\partial y}\frac{dy}{dt}.$$

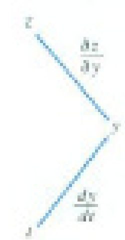


Figure 16.3.4:

Then dz/dt is the sum of those two summands.

EXAMPLE 2 Let $z = x^2y^3$, $x = 3t^2$, and $y = t/3$. Find dz/dt when $t = 1$.
SOLUTION In order to apply the special case of the chain rule, compute z_x, z_y, dx/dt, and dy/dt:

$$\frac{\partial z}{\partial x} = 2xy^3 \qquad \frac{\partial z}{\partial y} = 3x^2y^2$$

$$\frac{dx}{dt} = 6t \qquad \frac{dy}{dt} = \frac{1}{3}.$$

By the special case of the chain rule,

$$\frac{dz}{dt} = 2xy^3 \cdot 6t + 3x^2y^2 \cdot \frac{1}{3}.$$

In particular, when $t = 1$, x is 3 and y is $\frac{1}{3}$. Therefore, when $t = 1$,

$$\frac{dz}{dt} = 2 \cdot 3 \left(\frac{1}{3}\right)^3 6 \cdot 1 + 3 \cdot 3^2 \left(\frac{1}{3}\right)^2 \frac{1}{3} = \frac{36}{27} + \frac{27}{27} = \frac{7}{3}.$$

◇

In Example 2, the derivative dz/dt can be found without using the theorem. To do this, express z explicitly in terms of t:

$$z = x^2y^3 = (3t^2)^2 \left(\frac{t}{3}\right)^3 = \frac{t^7}{3}.$$

Then

$$\frac{dz}{dt} = \frac{7t^6}{3}.$$

When $t = 1$, this gives

$$\frac{dz}{dt} = \frac{7}{3},$$

in agreement with the first computation.

EXAMPLE 3 The temperature at the points (x, y) on a window is $T(x, y)$. A bug wandering on the window is at the point $(x(t), y(t))$ at time t. How fast does the bug observe that the temperature of the glass changes as he crawls about?

SOLUTION The bug is asking us to find dT/dt. The chain rule (16.3.8) tells us that

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}.$$

The bug can influence this rate by crawling faster or slower. He may want to know the direction he should choose in order to cool off as quickly as possible. But we will not be able to tell him how to do this until the next section, Section 16.4. ◇

The proof of the next chain rule is almost identical to the proof of Theorem 16.3. (See Exercise 24.)

Theorem. *Chain Rule – Special Case #2 Let $z = f(x, y)$ have continuous partial derivatives, f_x and f_t. Let $x = x(t, u)$ and $y = (t, u)$ have continuous partial derivatives*

$$\frac{\partial x}{\partial t}, \qquad \frac{\partial x}{\partial u}, \qquad \frac{\partial y}{\partial t}, \qquad \frac{\partial y}{\partial u}.$$

Then

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} \qquad \textit{and} \qquad \frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}.$$

The variables are listed in Figure 16.3.5.

To find z_t, draw all the paths from z down to t. Label the edges by the appropriate partial derivative, as shown in Figure 16.3.6.

Each path from the top variable down to the bottom variable contributes a summand in the chain rule. The only difference between Figure 16.3.2 and Figure 16.3.6 is that ordinary derivatives dx/dt and dy/dt appear in Figure 16.3.2, while partial derivatives x_t and y_t appear in Figure 16.3.6.

In the first special case of the chain rule there are two middle variables and one bottom variable. In the second chain rule there are two middle variables and two bottom variables. The chain rule holds for any number of middle variables and any number of bottom variable. For instance, there may be three middle variables and, say, four bottom variables. In that case there are three summands for each of four partial derivatives.

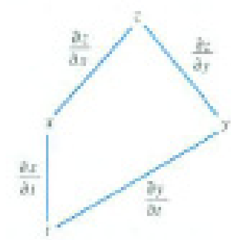


Figure 16.3.6:

In the next example there is only one middle variable and two bottom variables.

EXAMPLE 4 Let $z = f(u)$ be a function of a single variable. Let $u = 2x + 3y$. Then z is a composite function of x and y. Show that

$$2\frac{\partial z}{\partial y} = 3\frac{\partial z}{\partial x}. \tag{16.3.9}$$

SOLUTION We will evaluate both z_x and z_y by the chain rule and then check whether (16.3.9) is true.

To find z_x we consider all paths from z down to x. There is only one middle variable so there is only one path. Since $u = 2x + 3y$, $u_x = 2$. Thus

$$\frac{\partial z}{\partial x} = \frac{dz}{du}\frac{\partial u}{\partial x} = \frac{dz}{du} \cdot 2 = 2\frac{dz}{du}$$

(Note that one derivative is ordinary, while the other is a partial derivative.)

Next we find z_y. Again, there is only one summand. Since $u = 2x + 3y$, $u_y = 3$. Thus

$$\frac{\partial z}{\partial y} = \frac{dz}{du}\frac{\partial u}{\partial y} = \frac{dz}{du} \cdot 3 = 3\frac{dz}{du}.$$

Thus $z_x = 2dz/du$ and $z_y = 3dz/du$. Substitute these into the equation

$$2\frac{\partial z}{\partial y} = 3\frac{\partial z}{\partial x}$$

to see whether we obtain a true equation:

$$2\left(3\frac{dz}{du}\right) = 3\left(2\frac{dz}{du}\right). \tag{16.3.10}$$

Since (16.3.10) is true, we have verified (16.3.9). ◇

An Important Use of the Chain Rule

There is a fundamental difference between Example 2 and Example 4. In the first example, we were dealing with explicitly given functions. We did not really need to use the chain rule to find the derivative, dz/dt. As remarked after the example, we could have shown that $z = t^7/3$ and easily found that $dz/dt = 7t^6/3$. But in Example 4, we were dealing with a general type of function formed in a certain way: We showed that (16.3.9) holds for *every* differentiable function $f(u)$. No matter what $f(u)$ we choose, we know that $2z_y = 3z_x$.

Example 4 shows why the chain rule is important. It enables us to make *general statements* about the partial derivatives of an infinite number of functions, all of which are formed the same way. The next example illustrates this use again.

The wave equation also appears in the study of sound or light.

D'Alembert in 1746 obtained the partial differential equation for a vibrating string:

$$\frac{\partial^2 y}{\partial t^2} = k^2 \frac{\partial^2 y}{\partial x^2}. \tag{16.3.11}$$

(See Figure C.21.3 in the CIE about the Wave in a Rope.) This "wave equation" created a great deal of excitement, especially since d'Alembert showed that any differentiable function of the form

$$g(x + kt) + h(x - kt)$$

is a solution.

Before we show that d'Alembert is right, we note that it is enough to check it for $g(x+kt)$. If you replace k by $-k$ in it, you will also have a solution since replacing k by $-k$ in (16.3.11) doesn't change the equation.

EXAMPLE 5 Show that any function $y = g(x + kt)$ satisfies the partial differential equation (16.3.11).

SOLUTION In order to find the partial derivatives y_{xx} and y_{tt} we express $y = g(x + kt)$ as a composition of functions:

$$y = g(u) \qquad \text{where} \qquad u = x + kt.$$

Note that g is a function of just one variable. Figure 16.3.7 lists the variables.

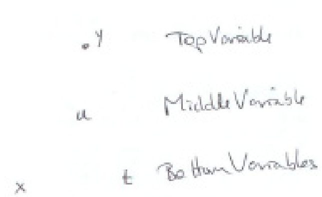


Figure 16.3.7:

We will compute y_{xx} and y_{tt} in terms of derivatives of g and then check whether (16.3.11) holds. We first compute y_{xx}. First of all,

Recall that $u = x + kt$.

$$\frac{\partial y}{\partial x} = \frac{dy}{du}\frac{\partial u}{\partial x} = \frac{dy}{du} \cdot 1 = \frac{dy}{du}. \qquad (16.3.12)$$

(There is only one path from y down to x. See Figure 16.3.7.) In (16.3.12) dy/du is viewed as a function of x and t; that is, u is replaced by $x + kt$. Next,

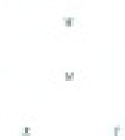
$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial y}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{dy}{du}\right).$$

Now, dz/du, viewed as a function of x and t, may be expressed as a composite function. Letting $w = dy/du$, we have

$$w = f(u), \qquad \text{where} \qquad u = x + kt.$$

Therefore

$$\begin{aligned} \frac{\partial^2 y}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial y}{\partial x}\right) = \frac{\partial w}{\partial x} \\ &= \frac{dw}{du} \cdot \frac{\partial u}{\partial x} \qquad \text{(only one path down to } x\text{)} \\ &= \frac{d}{du}\left(\frac{dy}{du}\right)\frac{\partial u}{\partial x} = \frac{d^2 y}{du^2} \cdot 1; \end{aligned}$$



hence

$$\frac{\partial^2 y}{\partial x^2} = \frac{d^2 y}{du^2}. \qquad (16.3.13)$$

Then we also express y_{tt} in terms of d^2y/du^2, as follows. First of all,

Figure 16.3.8:

$$\frac{\partial y}{\partial t} = \frac{dy}{du}\frac{\partial u}{\partial t} = \frac{dy}{du} \cdot k = k\frac{dy}{du}.$$

(See Figure 16.3.9.)

Then

$$\begin{aligned} \frac{\partial^2 y}{\partial t^2} &= \frac{\partial}{\partial t}\left(\frac{\partial y}{\partial t}\right) = \frac{\partial}{\partial t}\left(k\frac{dy}{du}\right) \\ &= k\frac{d}{dy}\left(\frac{dy}{du}\right) \cdot \frac{\partial u}{\partial t} \qquad \text{(only one path down to } t\text{)} \\ &= k\frac{d^2 y}{du^2} \cdot k; \end{aligned}$$

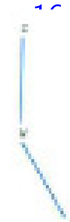


Figure 16.3.9:

hence

$$\frac{\partial^2 y}{\partial t^2} = k^2\frac{d^2 z}{du^2} \qquad (16.3.14)$$

Comparing (16.3.13) and (16.3.14) shows that

$$\frac{\partial^2 z}{\partial t^2} = k^2\frac{d^2 z}{dx^2}$$

◇

Summary

The section opened by showing that under suitable assumptions on $f(x, y)$

$$\Delta f = \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b) + \epsilon_1 \Delta x + \epsilon_2 \Delta y, \tag{16.3.15}$$

where ϵ_1 and ϵ_2 approach 0 as Δx and Δy approach 0. This gave us a way to estimate Δf, namely

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y$$

"The change is due to both the change in x and the change in y." (16.3.15) generalizes to any number of variables and also is the basis for the various chain rules for partial derivatives. This is the general case:

If z is a function of $x_1, x_2, \ldots x_m$ and each x_i is a function of $t_1, t_2 \ldots t_n$, then there are n partial derivatives of $\partial z/\partial t_j$. Each is a sum of m products of the form $(\partial z/\partial x_i)(\partial x_i/\partial t_j)$. To do the bookkeeping, first make a roster as shown in Figure 16.3.10. To compute $\partial z/\partial t_j$, list all paths from z down to t_j, as shown in Figure 16.3.11. Each path that starts at z and goes down to t_j "contributes" a product. You do not have to be a great mathematician to apply the chain rule. However, you must do careful bookkeeping. First, display the top, middle, and bottom variables. Second, keep in mind that the number of middle variables determines the number of summands.

Figure 16.3.10:

Some advice

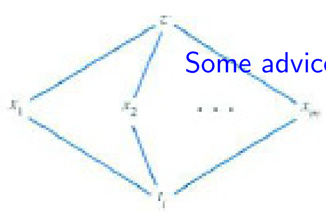


Figure 16.3.11:

EXERCISES for Section 16.3 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4 verify the chain rule (Special Case #1, on page 1294) by computing dz/dt two ways: (a) with the chain rule, (b) without the chain rule (by writing z as a function of t).

1.[R] $z = x^2y^3$, $x = t^2$, $y = t^3$

2.[R] $z = xe^y$, $x = t$, $y = 1 + 3t$

3.[R] $z = \cos(xy^2)$, $x = e^{2t}$, $y = \sec(3t)$

4.[R] $z = \ln(x + 3y)$, $x = t^2$, $y = \tan(3t)$.

In Exercises 5 and 6 verify the chain rule (Special Case #2, on page 1296) by computing dz/dt two ways: (a) with the chain rule, (b) without the chain rule (by writing z as a function of t and u).

5.[R] $z = x^2y$, $x - 3t + 4u$, $y = 5t - u$

6.[R] $z = \sin(x + 3y)$, $x = \sqrt{t/u}$, $y = \sqrt{t} + \sqrt{u}$

7.[R] Assume that $z = f(x_1, x_2, x_3, x_4, x_5)$ and that each x_i is a function of t_1, t_2, t_3.

(a) List all variables, showing top, middle, and bottom variables.

(b) Draw the paths involved in expressing $\partial z/\partial t_3$ in terms of the chain rule.

(c) Express $\partial z/\partial t_3$ in terms of the sum of products of partial derivatives.

(d) When computing $\partial z/\partial t_2$, which variables are constant?

(e) When computing $\partial z/\partial t_3$, which variables are constant?

8.[R] If $z = f(g(t_1, t_2, t_3), h(t_1, t_2, t_3))$

(a) How many middle variables are there?

(b) How many bottom variables?

(c) What does the chain rule say about $\partial z/\partial t_3$? (Include a diagram showing the paths.)

9.[R] Find dz/dt if $z_x = 4$, $x_y = 3$, $dx/dt = 4$, and $dy/dt = 1$.

10.[R] Find dz/dt if $z_x = 3$, $z_y = 2$, $dx/dt = 4$, and $dy/dt = -3$.

11.[R] Let $z = f(x, y)$, $x = u + v$, and $y = u - v$.

(a) Show that $(z_x)^2 - (z_y)^2 = (z_u)(z_v)$. (Include diagrams.)

(b) Verify (a) when $f(x, y) = x^2 + 2y^3$.

12.[R] Let $z = f(x, y)$, $x = u^2 - v^2$, and $y = v^2 - u^2$.

(a) Show that

$$u\frac{\partial z}{\partial v} + v\frac{\partial z}{\partial u} = 0.$$

(Include diagrams.)

(b) Verify (a) when $f(xy) = \sin(x + 2y)$.

13.[R] Let $z = f(t - u, -t + u)$.

(a) Show that $\frac{\partial z}{\partial t} + \frac{\partial z}{\partial u} = 0$ (Include diagrams.)

(b) Verify (a) when $f(x, y) = x^2 y$

14.[R] Let $w = f(x - y, y - z, z - x)$.

(a) Show that $\frac{\partial w}{\partial x} + \frac{\partial w}{\partial y} + \frac{\partial w}{\partial z} = 0$. (Include diagrams.)

(b) Verify (a) in the case $f(s, t, u) = s^2 + t^2 - u$.

15.[R] Let $z = f(u, v)$ where $u = ax + by$, $v = cx + dy$, and a, b, c, d are constants. Show that

(a) $\dfrac{\partial^2 z}{\partial x^2} = a^2\dfrac{\partial^2 f}{\partial u^2} + 2ac\dfrac{\partial^2 f}{\partial u \partial v} + c^2\dfrac{\partial^2 f}{\partial v^2}$

(b) $\dfrac{\partial^2 z}{\partial y^2} = b^2\dfrac{\partial^2 f}{\partial u^2} + 2bd\dfrac{\partial^2 f}{\partial u \partial v} + d^2\dfrac{\partial^2 f}{\partial v^2}$

(c) $\dfrac{\partial^2 z}{\partial x \partial y} = ab\dfrac{\partial^2 f}{\partial u^2} + (ad + bc)\dfrac{\partial^2 f}{\partial u \partial v} + cd\dfrac{\partial^2 f}{\partial v^2}$.

16.[R] Let a, b, and c be given constants and consider the partial differential equation

$$a\frac{\partial^2 z}{\partial x^2} + b\frac{\partial^2 z}{\partial x \partial y} + c\frac{\partial^2 z}{\partial y^2} = 0$$

Assume a solution of the form $z = f(y + mx)$, where m is a constant. Show that for this function to be a solution, $am^2 + bm + c$ must be 0.

17.[R]

(a) Show that any function of the form $z = f(x + y)$ is a solution of the partial differential equation

$$\frac{\partial^2 z}{\partial x^2} - 2\frac{\partial^2 z}{\partial x \partial y} + \frac{\partial^2 z}{\partial y^2} = 0.$$

(b) Verify (a) for $z = (x + y)^3$.

18.[R] Let $u(x, t)$ be the temperature at point x along a rod at time t. The function u satisfies the one-dimensional heat equation for a constant k:

$$\frac{\partial u}{\partial t} = k\frac{\partial^2 u}{\partial x^2}.$$

(a) Show that $u(x, t) = e^{kt}g(x)$ satisfies the heat equation if $g(x)$ is any function such that $g''(x) = g(x)$.

(b) Show that if $g(x) = 3e^{-x} + 4e^{x}$, then $g''(x) = g(x)$.

19.[R]

(a) Show that any function of the form $z = f(x + y) + e^{y} f(x - y)$ is a solution of the partial differential equation

$$\frac{\partial^2 z}{\partial x^2} - \frac{\partial^2 z}{\partial y^2} - \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} = 0.$$

(b) Check (a) for $z = (x + y)^2 + e^{y} \sin(x - y)$.

20.[R] Let $z = f(x, y)$ denote the temperature at the point (x, y) in the first quadrant. If polar coordinates are used, then we would write $z = f(r, \theta)$.

(a) Express z_r in terms of z_x and x_y. HINT: What is the relation between rectangular coordinates (x, y) and polar coordinates (r, θ)?

(b) Express z_θ in terms of z_x and z_y.

(c) Show that

$$\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = \left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2.$$

21.[R] Let $u = f(r)$ and $r = (x^2 + y^2 + z^2)^{1/2}$. Show that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = \frac{d^2 u}{dr^2} + \frac{2}{r}\frac{du}{dr}.$$

22.[R] At what rate is the volume of a rectangular box changing when its width is 3 feet and increasing at the rate of 2 feet per second, its length is 8 feet and decreasing at the rate of 5 feet per second, and its height is 4 feet and increasing at the the rate of 2 feet per second?

23.[R] The temperature T at (x, y, z) in space is $f(x, y, z)$. An astronaut is traveling in such a way that his x and y coordinates increase at the rate of 4 miles per second and his z coordinate decreases at the rate of 3 miles per second. Compute the rate dT/dt at which the temperature changes at a point where

$$\frac{\partial T}{\partial x} = 4, \quad \frac{\partial T}{\partial y} = 7, \quad \text{and} \quad \frac{\partial T}{\partial z} = 9.$$

24.[M] We proved Special Case #1 of the chain rule (page 1294), when there are two are two middle variables and one bottom variable. Prove Special Case #2 of the chain rule (page 1296), where there are two middle variables and two bottom variables.

25.[M] To prove the general chain rule when there are three middle variables, we need an analog of Theorem 16.3.1 concerning Δf when f is a function of three variables.

(a) Let $y = f(x, y, x)$ be a function of three variables. Show that

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

(b) Using (a) show that

$$\Delta f = \frac{\partial f}{\partial x}(x,y,z)\Delta x + \frac{\partial f}{\partial y}(x,y,z)\Delta y + \frac{\partial f}{\partial z}(x,y,z)\Delta z + \epsilon_1\Delta x + \epsilon_2\Delta y + \epsilon_3\Delta z,$$

where ϵ_1, ϵ_2, $\epsilon_3 \to 0$ as Δx, Δy, $\Delta z \to 0$.

(c) Obtain the general chain rule in the case of three middle variables and any number of bottom variables.

26.[M] Let $z = f(x,y)$, where $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Show that

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

27.[M] Let $u = f(x,y)$, where $x = r\cos(\theta)$ and $y = r\sin(\theta)$. Verify the following equation, which appears in electromagnetic theory,

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 u}{\partial \theta^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

28.[M] Let u be a function of x and y, where x and y are both functions of s and t. Show that

$$\frac{\partial^2 u}{\partial s^2} = \frac{\partial^2 u}{\partial x^2}\left(\frac{\partial x}{\partial s}\right)^2 + 2\frac{\partial^2 u}{\partial x\partial y}\frac{\partial x}{\partial s}\frac{\partial y}{\partial x} + \frac{\partial^2 u}{\partial y^2}\left(\frac{\partial y}{\partial x}\right)^2 + \frac{\partial u}{\partial x}\frac{\partial^2 x}{\partial s^2} + \frac{\partial u}{\partial y}\frac{\partial^2 y}{\partial s^2}.$$

29.[C] Let (r,θ) be polar coordinates for the point (x,y) given in rectangular coordinates.

(a) From the relation $r = \sqrt{x^2+y^2}$, show that $\partial r/\partial x = \cos(\theta)$.

(b) From the relation $r = x/\cos\theta$, show that $\partial r/\partial x = 1/\cos(\theta)$.

(c) Explain why (a) and (b) are not contradictory.

30.[C] In developing (16.3.6), we used the path that started at (x,y), went to $(x+\Delta x, y)$, and ended at $(x+\Delta x, y+\Delta y)$. Could we have used the path from (x,y), through $(x, y+\Delta y)$, to $(x+\Delta x, y+\Delta y)$ instead? If "no", explain why. If "yes," write out the argument, using the path.

In Exercises 31 to 34 concern homogeneous functions. A function $f(x, y)$ is homogeneous of degree r if $f(kx, ky) = k^r f(x, y)$ for all $k > 0$.

31.[R] Verify that each of the following functions is homogeneous of degree 1 and also verify that each satisfies the conclusion of Euler's theorem (with $r = 1$):

$$f(x, y) = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}.$$

(a) $f(x, y) = 3x + 4y$

(b) $f(x, y) = x^3 y^{-2}$

(c) $f(x, y) = xe^{x/y}$

32.[M] Show that each of the following functions is homogeneous, and find the degree r.

(a) $f(x, y) = x^2(\ln x - \ln y)$

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

(c) $f(x, y) = \sin\left(\frac{y}{x}\right)$

33.[C] (See Exercise 31.) Show that if f is homogeneous of degree r, then $xf_x + yf_y = rf$. This is the general form of Euler's theorem.

34.[C] (See Exercise 33.) Verify Euler's theorem for each of the functions in Exercise 32.

35.[C] (See Exercise 32.) Show that if f is homogeneous of degree r, then $\partial f/\partial x$ is homogeneous of degree $r - 1$.

16.4 Directional Derivatives and the Gradient

In this section we generalize the notion of a partial derivative to that of a directional derivative. Then we introduce a vector, called "the gradient," to provide a short formula for the directional derivative. The gradient will have other uses later in this chapter and in Chapter 18.

Directional Derivatives

If $z = f(x, y)$, the partial derivative $\partial f/\partial x$ tells us how rapidly z changes as we move the input point (x, y) in a direction parallel to the x-axis. Similarly, f_y tells how fast z changes as we move parallel to the y-axis. But we can ask, "How rapidly does z change when we move the input point (x, y) in any fixed direction in the xy plane?" The answer is given by the directional derivative.

It is important to remember that $\|\mathbf{u}\| = 1$.

Consider a function $z = f(x, y)$, let's say the temperature at (x, y). Let (a, b) be a point and let $\mathbf{u}$ be a unit vector in the xy plane. Draw a line through (a, b) and parallel to $\mathbf{u}$. Call it the t-axis and let its positive part point in the direction of $\mathbf{u}$. Place the 0 of the t-axis at (a, b). (See Figure 16.4.1.) Each value of t determines a point (x, y) on the t-axis and thus a value of z. Along the t-axis, z can therefore be viewed as a function of t, $z = g(t)$. The derivative dg/dt, evaluated at $t = 0$, is called the **directional derivative** of $z = f(x, y)$ at (a, b) in the direction $\mathbf{u}$. It is denoted $D_{\mathbf{u}}f$. The directional derivative is the slope of the tangent line to the curve $z = g(t)$ at $t = 0$. (See Figure 16.4.1(c).)

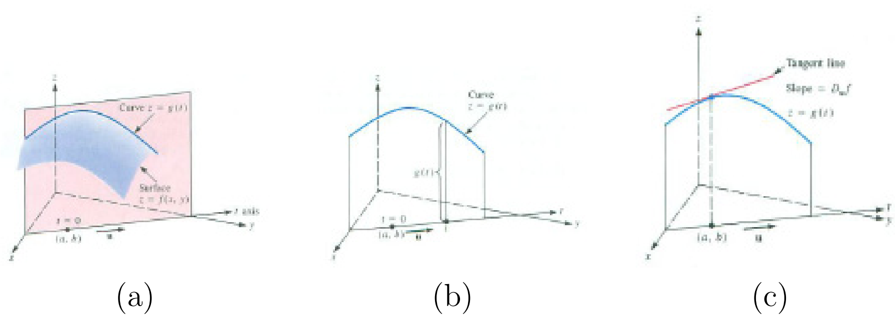


Figure 16.4.1: ARTIST: Improved figures are needed here.

When $\mathbf{u} = \mathbf{i}$, we obtain the directional derivative $D_{\mathbf{u}}f$, which is simply fx. When $\mathbf{u} = \mathbf{j}$, we obtain $D_{\mathbf{j}}f$, which is f_y.

The directional derivative generalizes the two partial derivatives f_x and f_y. After all, we can ask for the rate of change of $z = f(x, y)$ in any direction in the xy plane, not just the directions indicated by the vectors $\mathbf{i}$ and $\mathbf{j}$.

The following theorem shows how to compute a directional derivative.

Theorem. *(Directional Derivatives) If $f(x,y)$ has continuous partial derivatives f_x and f_y, then the directional derivative of f at (a,b) in the direction of $\mathbf{u} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j}$ where θ is the angle between $\mathbf{u}$ and $\mathbf{i}$ is*

$$\frac{\partial f}{\partial x}(a,b)\cos(\theta) + \frac{\partial f}{\partial y}(a,b)\sin(\theta). \tag{16.4.1}$$

Proof

The directional derivative of f at (a,b) in the direction $\mathbf{u}$ is the derivative of the function

$$g(t) = f(a + t\cos(\theta), b + t\sin(\theta))$$

when $t = 0$. (See Figure 16.3.2 and Figure 16.3.3.)

Now, g is a composite function

$$g(t) = f(x,y) \qquad \text{where} \qquad \begin{cases} x &= a + t\cos(\theta) \\ y &= b + t\sin(\theta). \end{cases}$$

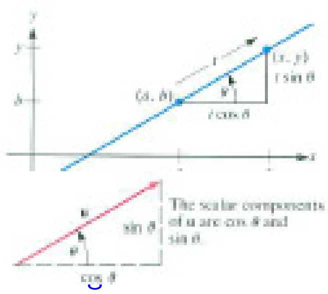


Figure 16.4.3:

The chain rule tells us that

$$g'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

Moreover,

$$\frac{dx}{dt} = \cos(\theta) \qquad \text{and} \qquad \frac{dy}{dt} = \sin(\theta).$$

Thus

$$g'(0) = \frac{\partial f}{\partial x}(a,b)\cos\theta + \frac{\partial f}{\partial y}(a,b)\sin\theta,$$

and the theorem is proved. •

Check (16.4.1) when $\theta = 0$

When $\theta = 0$, that is, $\mathbf{u} = \mathbf{i}$, (16.4.1) becomes

$$\frac{\partial f}{\partial x}(a,b)\cos(0) + \frac{\partial f}{\partial y}(a,b)\sin(0) \;=\; \frac{\partial f}{\partial x}(a,b)(1) + \frac{\partial f}{\partial y}(a,b)(0) = \frac{\partial f}{\partial x}(a,b).$$

Check (16.4.1) when $\theta = \pi$

When $\theta = \pi$, that is, $\mathbf{u} = -\mathbf{i}$, (16.4.1) becomes

$$\frac{\partial f}{\partial x}(a,b)\cos(\pi) + \frac{\partial f}{\partial y}(a,b)\sin(\pi) \;=\; \frac{\partial f}{\partial x}(a,b)(-1) + \frac{\partial f}{\partial y}(a,b)(0) = -\frac{\partial f}{\partial x}(a,b).$$

Check (16.4.1) when $\theta = \frac{\pi}{2}$

(This makes sense: If the temperature increases as you walk east, then it decreases when you walk west.)

When $\theta = \frac{\pi}{2}$, that is, $\mathbf{u} = \mathbf{j}$, (16.4.1) asserts that the directional derivative is

$$\frac{\partial f}{\partial x}(a,b)\cos(\frac{\pi}{2}) + \frac{\partial f}{\partial y}(a,b)\sin(\frac{\pi}{2}) \;=\; \frac{\partial f}{\partial x}(a,b)(0) + \frac{\partial f}{\partial y}(a,b)(1) = \frac{\partial f}{\partial y}(a,b).$$

which also is expected.

EXAMPLE 1 Compute the derivative of $f(x, y) = x^2y^3$ at $(1, 2)$ in the direction given by the angle $\pi/3$. (That is, $\mathbf{u} = \cos(\pi/3)\mathbf{i} + \sin(\pi/3)\mathbf{j}$.) Interpret the results if f describes a temperature distribution.

SOLUTION First of all,

$$\frac{\partial f}{\partial x} = 2xy^3 \quad \text{and} \quad \frac{\partial f}{\partial y} = 3x^2y^2.$$

Hence

$$\frac{\partial f}{\partial x}(1, 2) = 16 \quad \text{and} \quad \frac{\partial f}{\partial y}(1, 2) = 12.$$

Second,

$$\cos\left(\frac{\pi}{3}\right) = \frac{1}{2} \quad \text{and} \quad \sin\left(\frac{\pi}{3}\right) = \frac{\sqrt{3}}{2}.$$

Thus the derivative of f in the direction given by $\theta = \pi/3$ is

$$16\left(\frac{1}{2}\right) + 12\left(\frac{\sqrt{3}}{2}\right) = 8 + 6\sqrt{3} \approx 18.3923.$$

If x^2y^3 is the temperature in degrees at the point (x, y), where x and y are measured in centimeters, then the rate at which the temperature changes at $(1, 2)$ in the direction given by $\theta = \pi/3$, is approximately 18.4 degrees per centimeter. ◇

The Gradient

Equation (16.4.1) resembles the formula for the dot product. To exploit this similarity, it is useful to introduce the vector whose scalar components are $f_x(a, b)$ and $f_y(a, b)$.

DEFINITION *(The gradient of $f(x, y)$.)* The vector

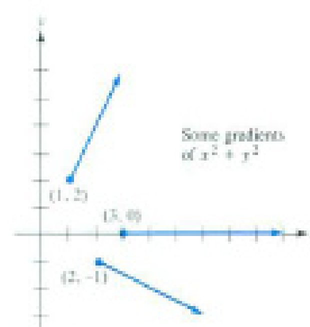
$$\frac{\partial f}{\partial x}(a, b)\mathbf{i} + \frac{\partial f}{\partial y}(a, b)\mathbf{j}$$

is the **gradient** of f at (a, b) and is denoted ∇f. (It is also called "del f," because of the upside-down delta ∇.)

The del symbol is in boldface to emphasize that the gradient of f is a vector. For instance, let $f(x, y) = x^2 + y^2$. We compute and draw ∇f at a few points, listed in the following table:

Figure 16.4.4 shows ∇f, with the tail of ∇f placed at the point where ∇f is computed.

In vector notation, Theorem 16.4 reads as follows:



(x, y)	$\frac{\partial f}{\partial x} = 2x$	$\frac{\partial f}{\partial y} = 2y$	∇f
$(1, 2)$	2	4	$2\mathbf{i} + 4\mathbf{j}$
$(3, 0)$	6	0	$6\mathbf{i}$
$(2, -1)$	4	-2	$4\mathbf{i} - 2\mathbf{j}$

Table 16.4.1:

Theorem. *Directional Derivative - Rephrased If $z = f(x, y)$ has continuous partial derivatives f_x and f_y, then at (a, b)*

$$D_{\mathbf{u}} f = \nabla f(a, b) \cdot \mathbf{u} = (f_x(a, b)\mathbf{i} + f_y(a, b)\mathbf{j}) \cdot \mathbf{u}.$$

The gradient is introduced not merely to simplify the computation of directional derivatives. Its importance is made clear in the next theorem.

A Different View of the Gradient

The gradient vector provides two important pieces of geometric information about a function. The gradient vector, $\nabla f(a, b)$, always points in the direction in which the function increases most rapidly from the point (a, b). In the same way, the negative of the gradient vector, $-\nabla f(a, b)$, always points in the direction in which the function decreases most rapidly from the point (a, b). And, the length of the gradient vector, $\|\nabla f(a, b)\|$, is the largest directional derivative of f at (a, b).

The meaning of $\|\nabla f\|$ and the direction of ∇f

Theorem. *Significance of ∇f Let $z = f(x, y)$ have continuous partial derivatives f_x and f_y. Let (a, b) be a point in the plane where ∇f is not $\mathbf{0}$. Then the length of ∇f at (a, b) is the largest directional derivative of f at (a, b). The direction of ∇f is the direction in which the directional derivative at (a, b) has its largest value.*

Proof

By the definition of the directional derivative, if $\mathbf{u}$ is a unit vector, then, at (a, b),

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}.$$

By the definition of the dot product

$$\nabla f \cdot \mathbf{u} = \|\nabla f\| \|\mathbf{u}\| \cos(\alpha),$$

where α is the angle between ∇f and $\mathbf{u}$, as shown in Figure 16.4.5. Since $|\mathbf{u}| = 1$,

$$D_{\mathbf{u}} f = \|\nabla f\| \cos(\alpha). \tag{16.4.2}$$

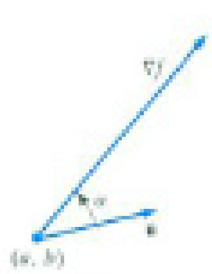


Figure 16.4.5:

The largest value of $\cos(\alpha)$ for $0 \le \alpha \le \pi$, occurs when $\cos(\alpha) = 1$; that is, when $\alpha = 0$. Thus, by (16.4.2), the largest directional derivative of $f(x, y)$ at (a, b) occurs when the direction is that of ∇f at (a, b). For that choice of $\mathbf{u}$, $D_{\mathbf{u}} f = \|\nabla f\|$. This proves the theorem. •

What does this theorem tell a bug wandering around on a flat piece of metal? If it is at the point (a, b) and wishes to get warmer as quickly as possible, it should compute the gradient of the temperature function and then go in the direction indicated by that gradient.

EXAMPLE 2 What is the largest direction derivative of $f(x, y) = x^2y^3$ at $(2, 3)$? In what direction does this maximum directional derivative occur?
SOLUTION At the point (x, y),

$$\nabla f = 2xy^3\mathbf{i} + 3x^2y^2\mathbf{j}.$$

Thus at $(2, 3)$,

$$\nabla f = 108\mathbf{i} + 108\mathbf{j},$$

which is sketched in Figure 16.4.6 (not to scale). Note that its angle θ is $\pi/4$. The maximal directional derivative of x^2y^3 at $(2, 3)$ is $\|\nabla f\| = 108\sqrt{2} \approx 152.735$. This is achieved at the angle $\theta = \pi/4$, relative to the x-axis, that is, for

$$\mathbf{u} = \cos\left(\frac{\pi}{4}\right)\mathbf{i} + \sin\left(\frac{\pi}{4}\right)\mathbf{j} = \frac{\sqrt{2}}{2}\mathbf{i} + \frac{\sqrt{2}}{2}\mathbf{j}.$$

◇

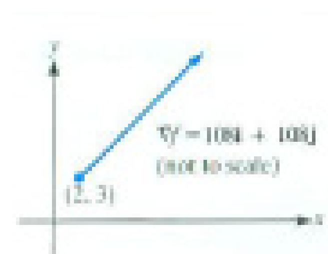
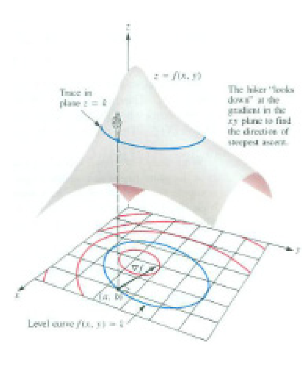


Figure 16.4.6: Direction of fastest decrease is $-\nabla f$

Incidentally, if $f(x, y)$ denotes the temperature at (x, y), the gradient ∇f helps indicate the direction in which heat flows. It tends to flow "toward the coldest," which boils down to the mathematical assertion, "Heat tends to flow in the direction of $-\nabla f$."

The gradient and directional derivative have been interpreted in terms of a temperature distribution in the plane and a wandering bug. It is also instructive to interpret these concepts in terms of a hiker on the surface of a mountain.

Consider a mountain above the xy plane. The elevation of the point on the surface above the point $x, y)$ will be denoted by $f(x, y)$. The directional derivative $D_{\mathbf{u}} f$ indicates the rate at which altitude changes per unit change in *horizontal* distance in the direction of $\mathbf{u}$. The gradient ∇f at (a, b) points in the compass direction the hiker should choose to climb in the direction of the steepest ascent. The length of ∇f tells the hiker the steepest slope available. (See Figure 16.4.7.)



Generalization to $f(x, y, z)$

The notions of directional derivative and gradient can be generalized with little effort to functions of three (or more) variables. It is easiest to interpret the directional derivative of $f(x, y, z)$ in a particular direction in space as indicating the rate of change of the function in that direction in space. A useful interpretation is how fast the temperature changes in a given direction.

Let $\mathbf{u}$ be a unit vector in space, with direction angles α, β, and γ. Then $\mathbf{u} = \cos\alpha\mathbf{i} + \cos\beta\mathbf{j} + \cos\gamma\mathbf{k}$. We now define the derivative of $f(x, y, z)$ in the direction $\mathbf{u}$.

DEFINITION *(Directional Derivative of $f(x, y, z)$.)* The **directional derivative** of f at (a, b, c) in the direction of the unit vector $\mathbf{u} = \cos(\alpha)\mathbf{i} + \cos(\beta)\mathbf{j} + \cos(\gamma)\mathbf{k}$ is $g'(0)$, where g is defined by

$$g(t) = f(a + t\cos(\alpha), b + t\cos(\beta), c + t\cos(\gamma)).$$

It is denoted $D_{\mathbf{u}}f$.

Note that t is the measure of length along the line through (a, b, c) with direction angles α, β, and γ. Therefore $D_{\mathbf{u}}f$ is just a derivative along the t-axis.

The proof of the following theorem for a function of three variables is like those given earlier in this section for functions of two variables.

Theorem. *Directional Derivative of $f(x, y, z)$ If $f(x, y, z)$ has continuous partial derivatives f_x, f_y, and f_z, then the directional derivative of f at (a, b, c) in the direction of the unit vector $\mathbf{u} = \cos(\alpha)\mathbf{i} + \cos(\beta)\mathbf{j} + \cos(\gamma)\mathbf{k}$ is*

$$\frac{\partial f}{\partial x}(a, b, c)\cos(\alpha) + \frac{\partial f}{\partial y}(a, b, c)\cos(\beta) + \frac{\partial f}{\partial z}(a, b, c)\cos(\gamma).$$

DEFINITION *(The gradient of $f(x, y, z)$.)* The vector

$$\frac{\partial f}{\partial x}(a, b, c)\mathbf{i} + \frac{\partial f}{\partial y}(a, b, c)\mathbf{j} + \frac{\partial f}{\partial z}(a, b, c)\mathbf{k}$$

is the **gradient** of f at (a, b, c) and is denoted ∇f.

This theorem thus asserts that

the derivative of $f(x, y, z)$ in the direction of the unit vector $\mathbf{u}$ equals the dot product of $\mathbf{u}$ and the gradient of f:

$$D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}.$$

Just as in the case of a function of two variables, ∇f evaluated at (a, b, c), points in the direction $\mathbf{u}$ that produces the largest directional derivative at (a, b, c). Moreover $\|\nabla f\|$ is that largest directional derivative. Just as in the two variable case, the key steps in the proof of this theorem are writing $\nabla f \cdot \mathbf{u} = \|\nabla f\|\|\mathbf{u}\| \cos(\nabla f, \mathbf{u})$ and recalling that $\mathbf{u}$ is a unit vector.

EXAMPLE 3 The temperature at the point (x, y, z) in a solid piece of metal is given by the formula $f(x, y, z) = c^{2x+y+3z}$ degrees. In what direction at the point $(0, 0, 0)$ does the temperature increase most rapidly?

SOLUTION First compute

$$\frac{\partial f}{\partial x} = 2e^{2x+y+3z}, \quad \frac{\partial f}{\partial y} = e^{2x+y+3z}, \quad \frac{\partial f}{\partial z} = 3e^{2x+y+3z}.$$

Then form the gradient vector:

$$\nabla f = 2e^{2x+y+3z}\mathbf{i} + e^{2x+y+3z}\mathbf{j} + 3e^{2x+y+3z}\mathbf{k}.$$

At $(0, 0, 0)$,

$$\nabla f = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}.$$

Consequently, the direction of most rapid increase in temperature is that given by the vector $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$. The rate of increase is then

$$\|2\mathbf{i} + \mathbf{j} + 3\mathbf{k}\| = \sqrt{14} \text{ degrees per unit length.}$$

If the line through $(0, 0, 0)$ parallel to $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$ is given a coordinate system so that it becomes the t-axis, with $t = 0$ at the origin and the positive part in the direction of $2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$, the $df/dt = \sqrt{14}$ at 0. ◇

The gradient was denoted Δ by Hamilton in 1846. By 1870 it was denoted ∇, an upside-down delta, and therefore called "atled." In 1871 Maxwell wrote, "The quantity ∇P is a vector. I venture, with much diffidence, to call it the *slope* of P." The name "slope" is no longer used, having been replaced by "gradient." "Gradient" goes back to the word "grade," the slope of a road or surface. The name "del" first appeared in print in 1901, in *Vector Analysis, A text-book for the use of students of mathematics and physics founded upon the lectures of J. Willard Gibbs*, by E.B. Wilson.

Summary

We defined the derivative of $f(x, y)$ at (a, b) in the direction of the unit vector $\mathbf{u}$ in the xy plane and the derivative of $f(x, y, z)$ at (a, b, c) in the direction of the unit vector $\mathbf{u}$ in space. Then we introduced the gradient vector ∇f in terms of its components and obtained the formula

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u}.$$

By examining this formula we saw that the length and direction of ∇f at a given point are significant:

- ∇f points in the direction $\mathbf{u}$ that maximizes $D_{\mathbf{u}} f$ at the given point
- $\|\nabla f\|$ is the maximum directional derivative of f at the given point.

EXERCISES for Section 16.4 *Key:* R–routine, M–moderate, C–challenging

As usual, we assume that all functions mentioned have continuous partial derivatives. In Exercises 1 and 2 compute the directional derivatives of x^4y^5 at $(1, 1)$ in the indicated directions.

1.[R] (a) $\mathbf{i}$, (b) $-\mathbf{i}$, (c) $\cos(\pi/4)\mathbf{i} + \sin(\pi/4)\mathbf{j}$

2.[R] (a) $\mathbf{j}$, (b) $-\mathbf{j}$, (c) $\cos(\pi/3)\mathbf{i} + \sin(\pi/3)\mathbf{j}$

In Exercises 3 and 4 compute the directional derivatives of x^2y^3 in the directions of the given vectors.

3.[R] (a) $\mathbf{j}$, (b) $\mathbf{k}$, (c) $-\mathbf{i}$

4.[R] (a) $\mathbf{i}+\mathbf{j}+\mathbf{k}$, (b) $2\mathbf{i}-\mathbf{j}+2\mathbf{k}$, (c) $\mathbf{i}+\mathbf{k}$ NOTE: These are not unit vectors. First construct a unit vector with the same direction.

5.[R] Assume that, at the point $(2, 3), \partial f/\partial x = 4$ and $\partial f/\partial y = 5$.

(a) Draw ∇f at $(2, 3)$.

(b) What is the maximal directional derivative of f at $(2, 3)$?

(c) For which $\mathbf{u}$ is $D_{\mathbf{u}}f$ at $(2, 3)$ maximal? (Write $\mathbf{u}$ in the form $x\mathbf{i} + y\mathbf{j}$.)

6.[R] Assume that, at the point $(1, 1), \partial f/\partial x = 3$ and $\partial f/\partial y = -3$.

(a) Draw ∇f at $(1, 1)$.

(b) What is the maximal directional derivative of f at $(1, 1)$?

(c) For which $\mathbf{u}$ is $D_{\mathbf{u}}f$ at $(1, 1)$ maximal? (Write $\mathbf{u}$ in the form $x\mathbf{i} + y\mathbf{j}$.)

In Exercises 7 and 8 compute and draw ∇f at the indicated points for the given functions.

7.[R] $f(x, y) = x^2y$ at (a) $(2, 5)$, (b) $(3, 1)$

8.[R] $f(x, y) = 1/\sqrt{x^2 + y^2}$ at (a) $(1, 2)$, (b) $(3, 0)$

9.[R] If the maximal directional derivative of f at (a, b) is 5, what is the minimal directional derivative there? Explain.

10.[R] For a given function $f(x, y)$ at a given point (a, b) is there always a direction in which the directional derivative is 0? Explain.

11.[R] If $(\partial f/\partial x)(a, b) = 2$ and $(\partial f/\partial t)(a, b) = 3$, in what direction should a directional derivative at (a, b) be computed in order that it be

(a) 0?

(b) as large as possible?

(c) as small as possible?

12.[R] If, at the point (a, b, c), $\partial f/\partial x = 2$, $\partial f/\partial y = 3$, $\partial f/\partial z = 4$, what is the largest directional derivative of f at (a, b, c)?

13.[R] Assume that $f(1, 2) = 2$ and $f(0.99, 2.01) = 1.98$.

(a) Which directional derivatives $D_{\mathbf{u}}f$ at $(1, 2)$ can be estimated with this information? (Give **u**.)

(b) Estimate the directional derivatives in (a).

14.[R] Assume that $f(1, 1, 1) = 3$ and $f(1.1, 1.2, 1.1) = 3.1$.

(a) Which directional derivatives $D_{\mathbf{u}}f$ at $(1, 1, 1)$ can be estimated with this information? (Give **u**.)

(b) Estimate the directional derivatives in (a).

15.[R] When a bug crawls east, it discovers that the temperature increases at the rate of 0.02° per centimeter. When it crawls north, the temperature decreases at the rate of −0.03° per centimeter.

(a) If the bug crawls south, at what rate does the temperature change?

(b) If the bug crawls 30° north of east, at what rate does the temperature change?

(c) If the bug is happy with its temperature, in what direction should it crawl to try to keep the temperature the same?

16.[R] A bird is very sensitive to the temperature. It notices that when it flies in the direction **i**, the temperature increases at the rate of 0.03° per centimeter.

When it flies in the direction **j**, the temperature decreases at the rate of $0.02°$ per centimeter. When it flies in the direction **k** the temperature increases at the rate of $0.05°$ per centimeter. It decides to fly off in the direction of the vector $(2, 5, 1)$. Will it be getting warmer or colder?

17.[R] Assume that $f(1, 2) = 3$ and that the directional derivative of f at $(1, 2)$ in the direction of the (nonunit) vector $\mathbf{i} + \mathbf{j}$ is 0.7. Use this information to estimate $f(1.1, 2.1)$.

18.[R] Assume that $f(1, 1, 2) = 4$ and that the directional derivative of f at $(1, 1, 2)$ in the direction of the vector from $(1, 1, 2)$ to $(1.01, 1.02, 1.99)$ is 3. Use this information to estimate $f(0.99, 0.98, 2.01)$.

In Exercises 19 to 24 find the directional derivative of the function in the given direction and the maximum directional derivative.

19.[R] xyz^2 at $(1, 0, 1)$; $\mathbf{i} + \mathbf{j} + \mathbf{k}$
20.[R] x^3yz at $(2, 1, -1)$; $2\mathbf{i} - \mathbf{k}$
21.[R] $e^{xy\sin(z)}$ at $(1, 1, \pi/4)$; $\mathbf{i} + \mathbf{j} + +3\mathbf{k}$
22.[R] $\arctan(\sqrt{x^2 + y + z})$ at $(1, 1, 1)$; $-\mathbf{i}$
23.[R] $\ln(1 + xyz)$ at $(2, 3, 1)$; $-\mathbf{i} + \mathbf{j}$
24.[R] $x^x y e^{z^2}$ at $(1, 1, 0)$; $\mathbf{i} - \mathbf{j} + \mathbf{k}$

25.[R] Let $f(x, y, z) = 2x + 3y + z$.

(a) Compute ∇f at $(0, 0, 0)$ and at $(1, 1, 1)$.

(b) Draw ∇f for the two points in (a), in each case putting its tail at the point.

26.[R] Let $f(x, y, z) = x^2 + y^2 + z^2$.

(a) Compute ∇f at $(2, 0, 0)$, $(0, 2, 0)$ and $(0, 0, 2)$.

(b) Draw ∇f for the three points in (a), in each case putting its tail at the point.

27.[M] Assume that ∇f at (a, b) is not $\mathbf{0}$. Show that there are two unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$, such that the directional derivatives of f at (a, b) in the direction of $\mathbf{u}_1$ and $\mathbf{u}_2$ are 0.

28.[M] Assume that ∇f at (a, b, c) is not $\mathbf{0}$. How many unit vectors $\mathbf{u}$ are there such that $D_{\mathbf{u}} f = 0$? Explain.

29.[R] Let $T(x,y,z)$ be the temperature at the point (x,y,z). Assume that ∇T at $(1,1,1)$ is $2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$.

(a) Find $D_{\mathbf{u}}T$ at $(1,1,1)$ if $\mathbf{u}$ is in the direction of the vector $\mathbf{i}-\mathbf{j}+2\mathbf{k}$.

(b) Estimate the change in temperature as you move from the point $(1,1,1)$ a distance 0.2 in the direction of the vector $\mathbf{i}-\mathbf{j}+2\mathbf{k}$.

(c) Find three unit vectors $\mathbf{u}$ such that $D_{\mathbf{u}}T=0$ at $(1,1,1)$.

30.[R] A bug at the point $(1,2)$ is very sensitive to the temperature and observes that if it moves in the direction $\mathbf{i}$ the temperature increases at the rate of 2° per centimeter. If it moves in the direction $\mathbf{j}$, the temperature decreases at the rate of 2° per centimeter. In what direction should it move if it wants

(a) to warm up most rapidly?

(b) to cool off most rapidly?

(c) to change the temperature as little as possible?

31.[R] Let $f(x,y)=1/\sqrt{x^2+y^2}$; the function f is defined everywhere except at $(0,0)$. Let $\mathbf{r}=\langle x,y\rangle$.

(a) Show that $\nabla f=-r/\|\mathbf{r}\|^3$.

(b) Show that $\|\nabla f\|=-1/\|\mathbf{r}\|^2$.

32.[R] Let $f(x,y,z)=1/\sqrt{x^2+y^2+z^2}$, which is defined everywhere except at $(0,0,0)$. (This function is related to the potential in a gravitational field due to a point-mass.) Let $\mathbf{r}=x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$. Express ∇f in terms of r.

33.[R] Let $f(x,y)=x^2+y^2$. Prove that if (a,b) is an arbitrary point on the curve $x^2+y^2=9$, then ∇f computed at (a,b) is perpendicular to the tangent line to that curve at (a,b).

34.[R] Let $f(x,y,z)$ equal temperature at (x,y,z). Let $P=(a,b,c)$ and Q be a point very near (a,b,c). Show that $\nabla f\cdot\overrightarrow{PQ}$ is a good estimate of the change in temperature from point P to point Q.

SHERMAN: Exercises 28 and 35 are similar, but different. Should (27 and) 28 be moved later, and classified as M? Or, one moved to the Chapter Summary?

35.[R]

(a) If $(\partial f/\partial x)(a,b,c) = 2$, $(\partial f/\partial y)(a,b,c) = 3$ and $(\partial f/\partial z)(a,b,c) = 1$, find three different unit vectors $\mathbf{u}$ such that $D_{\mathbf{u}}f$ at (a,b,c) is 0.

(b) How many unit vectors $\mathbf{u}$ are there such that $D_{\mathbf{u}}f$ at (a,b,c) is 0?

36.[C] Let $f(x,y) = xy$.

(a) Draw the level curve $xy = 4$ carefully.

(b) Compute ∇f at three convenient points on that level curve and draw it with its tail at the point where it is evaluated.

(c) What angle does ∇f seem to make with the curve at the point where it is evaluated?

(d) Prove that the angle is what you think it is.

37.[M] Let (x,y) be the temperature at (x,y). Assume that ∇f at $(1,1)$ is $2\mathbf{i}+3\mathbf{j}$. A bug is crawling northwest at the rate of 3 centimeters per second. Let $g(t)$ be the temperature at the point where the bug is at time t seconds. Then dg/dt is the rate at which temperature changes on the bug's journey (degrees per second.) Find dg/dt when the bug is at $(1,1)$.

38.[R] If $f(P)$ is the electric potential at the point P, then the electric field $\mathbf{E}$ at P is given by $-1/c^2\nabla f$. Calculate $\mathbf{E}$ if $f(x,y) = \sin(\alpha x)\cos(\beta y)$, where α and β are constants.

SHERMAN: There is a typo for this exercise in V. What coefficient do you want? I assumed $-1/c^2$.

39.[R] The equality $\partial^2 f/\partial x\partial y = \partial^2 f/\partial y\partial x$ can be written as $D_{\mathbf{i}}(D_{\mathbf{j}}f) = D_{\mathbf{j}}(D_{\mathbf{i}}f)$. Show for any two unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ that $D_{\mathbf{u}_2}(D_{\mathbf{u}_1}f) = D_{\mathbf{u}_1}(D_{\mathbf{u}_2}f)$. (Assume that all partial derivatives of f of all orders are continuous.)

40.[C] Without the aid of vectors, prove that the maximum value of

$$g(\theta) = \partial f/\partial x(a,b)\cos(\theta) + \partial f/\partial y(a,b)\sin(\theta)$$

is $\sqrt{(\partial f/\partial x(a,b))^2 + (\partial f/\partial y(a,b))^2}$. NOTE: This is the first part of the theorem about the significance of the gradient, on page 1310.

41.[R] Figure 16.4.8 shows two level curves of a function $f(x,y)$ near the point $(1,2)$, namely $f(x,y) = 3$ and $f(x,y) = 3.02$. Use the diagram to estimate

(a) $D_{\mathbf{i}}f$ at $(1,2)$,

(b) $D_{\mathbf{j}}f$ at $(1,2)$,

(c) Draw ∇f at $(1, 2)$.

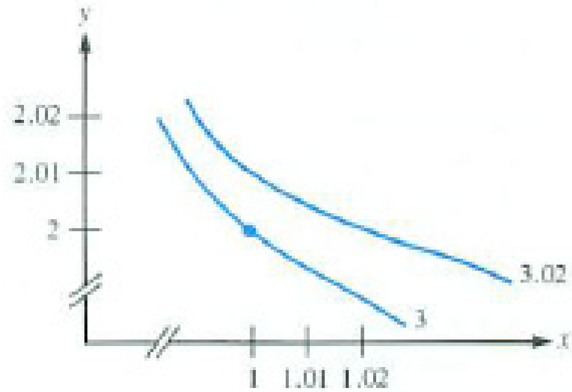


Figure 16.4.8:

42.[C] Why is a unit vector **u** in the xy-plane described by a single angle θ, but a unit vector in space is described by three angles?

43.[M] Let f and g be two vector functions defined throughout the xy-plane. Assume they have the same gradient, $\nabla f = \nabla g$. Must $f = g$? Is there any relation between f and g?

16.5 Normals and Tangent Planes

In this section we first find how to obtain a normal vector to a curve given implicitly, as a level curve $f(x, y) = k$. Then we find how to obtain a normal to a surface given implicitly, as a level surface $f(x, y, z) = k$. With the aid of this vector we define the tangent plane to a surface at a given point on the surface.

Normals to a Curve in the xy Plane

We saw in Section 14.4 how to find a normal vector to a curve when the curve is given parametrically, $\mathbf{r} = \mathbf{G}(t)$. Now we will see how to find a normal when the curve is given implicitly, as a level curve $f(x, y) = k$. Throughout this section we assume that the various functions are "well behaved." In particular, curves have continuous tangent vectors and functions have continuous partial derivatives.

Theorem. *The gradient ∇f at (a, b) is a normal to the level curve of f passing through (a, b).*

Proof
Let $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ be a parameterization of the level curve of f that passes through the point (a, b). On this curve, $f(x, y)$ is a constant and has the value $f(a, b)$. Let $\mathbf{G}'(t_0)$ be the tangent vector to the curve at (a, b) and let the gradient of f at (a, b) be $\nabla f(a, b) = f_x(a, b,)\mathbf{i} - f_y(a, b)\mathbf{j}$. We wish to show that

$$\nabla f \cdot \mathbf{G}'(t_0) = 0;$$

that is,

$$\frac{\partial f}{\partial x}(a, b)\frac{dx}{dt}(t_0) + \frac{\partial f}{\partial y}(a, b)\frac{dy}{dt}(t_0) = 0. \qquad (16.5.1)$$

The left side of (16.5.1) has the form of a chain rule. To make use of this fact, introduce the function $u(t)$ defined as

$$u(t) = f(x(t), y(t)).$$

Note that $u(t)$ is the value of f at a point on the level curve that passes through (a, b). Hence $u(t) = f(a, b)$. What is more important is that $u(t)$ is a constant function. Therefore, $du/dt = 0$.

Now, $u = f(x, y)$, where x and y are functions of t. The chain rule asserts that

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

Since $du/dt = 0$, (16.5.1) follows. Hence ∇f, evaluated at (a, b), is a normal to the level curve of f that passes through (a, b). •

What does this theorem say about the daily weather map that shows the barometric pressure? A level curve, or contour, shows the points where the pressure has a prescribed value. The gradient ∇f at anyplace on such a curve points in the direction in which the pressure increases most rapidly. So $-\nabla f$ points where the pressure is decreasing most rapidly. Since the wind tends to go from high pressure to low pressure, we can think of $-\nabla f$ as representing the wind.

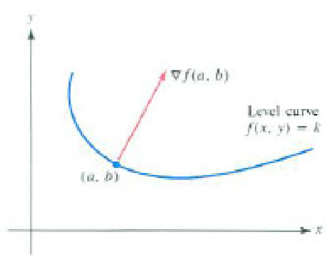


Figure 16.5.1:

Figure 16.5.1 shows a typical level curve and gradient. The gradient is perpendicular to the level curve. Moreover, as we saw in Section 16.4, the gradient points in the direction in which the function increases most rapidly.

EXAMPLE 1 Find and draw a normal vector to the hyperbola $xy = 6$ at the point $(2, 3)$.

SOLUTION Let $f(x, y) = xy$. Then $f_x = y$ and $f_y = x$. Hence,

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

In particular

$$\nabla f(2, 3) = 3\mathbf{i} + 2\mathbf{j}.$$

This gradient and level curve $xy = 6$ are shown in Figure 16.5.2. ◇

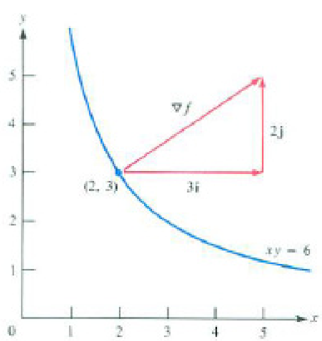


Figure 16.5.2:

EXAMPLE 2 Find an equation of the tangent line to the ellipse $x^2 + 2y^2 = 7$ at the point $(2, 1)$.

SOLUTION As we saw in Section 14.4, we may write the equation of a line in the plane if we know a point on the line and a vector normal to the line. We know that $(2, 1)$ lies on the line. We use a gradient to produce a normal.

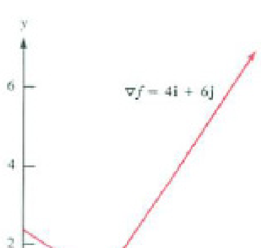
The ellipse $x^2 + 3y^2 = 7$ is a level curve of the function $f(x, y) = x^2 + 3y^2$. Since $f_x = 2x$ and $f_y = 6y$, $\nabla f = 2x\mathbf{i} + 6y\mathbf{j}$. In particular

$$\nabla f(2, 1) = 4\mathbf{i} + 6\mathbf{j}.$$

Hence the tangent line at $(2, 1)$ has an equation

$$4(x - 2) + 6(y - 1) = 0 \qquad \text{or} \qquad 4x + 6y = 14.$$

The level curve, normal vector, and tangent line are all shown in Figure 16.5.3. ◇



Normals to a Surface

We can construct a vector perpendicular to a surface $f(x, y, z) = k$ at a given point $P = (a, b, c)$ as easily as we constructed a vector perpendicular to a planar curve. It turns out that the gradient vector ∇f, evaluated at (a, b, c), is perpendicular to the surface $f(x, y, z) = k$. The proof of this result is similar to the proof for normal vectors to a level curve, given earlier in this section.

Before going on, we must state what is meant by a "vector being perpendicular to a surface."

> **DEFINITION** *(Normal vector to a surface)* A vector is perpendicular to a surface at the point (a, b, c) on this surface if the vector is perpendicular to each curve on the surface through the point (a, b, c). Such a vector is called a **normal vector**.

A vector is perpendicular to a curve at a point (a, b, c) on the curve if the vector is perpendicular to a tangent vector to the curve at (a, b, c).

Theorem. *Normal vectors to a level surface The gradient ∇f at (a, b, c) is a normal to the level surface of f passing through (a, b, c).*

Finding a Normal to the surface $f(x, y, z) = k$

SHERMAN: You have notation about using $a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ or $< a, b, c >$. I think we should use both, but I don't have a strong opinion about this.

Proof

Let $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$ be the parameterizations of a curve in the level surface of f that passes through the point (a, b, c). Assume $\mathbf{G}(t_0) = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$. Then $\mathbf{G}'(t_0)$ is the tangent vector to the curve at the point (a, b, c) and the gradient at (a, b, c) is

$$\nabla f = \frac{\partial f}{\partial x}(a, b, c)\mathbf{i} + \frac{\partial f}{\partial y}(a, b, c)\mathbf{j} + \frac{\partial f}{\partial z}(a, b, c)\mathbf{k}.$$

We wish to show that

$$\nabla f \cdot \mathbf{G}'(t_0) = 0;$$

that is

$$\frac{\partial f}{\partial x}(a, b, c)x'(t_0) + \frac{\partial f}{\partial y}(a, b, c)y'(t_0) + \frac{\partial f}{\partial z}(a, b, c)z'(t_0) = 0. \qquad (16.5.2)$$

(See Figure 16.5.4.) Introduce the function $u(t)$ defined by

$$u(t) = f(x(t), y(t), z(t)).$$

By the chain rule,

Once again the chain rule comes to our aid.

$$\left.\frac{du}{dt}\right|_{t=t_0} = \frac{\partial f}{\partial x}(a, b, c)x'(t_0) + \frac{\partial f}{\partial y}(a, b, c)y'(t_0) + \frac{\partial f}{\partial z}(a, b, c)z'(t_0) = 0 \quad (16.5.3)$$

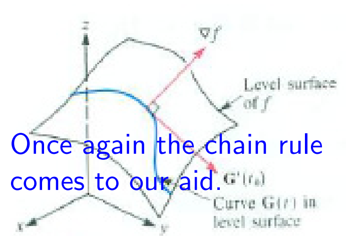


Figure 16.5.4:

However, *since the curve $\mathbf{G}(t)$ lies on a level surface of f, $u(t)$ is constant.* [In fact, $u(t) = f(a, b, c)$.] Thus $du/dt = 0$, and the right side of (16.5.3) is 0, as required. •

A simple check of this result is to see whether it is correct when the level surfaces are just planes. Consider $f(x,y,z) = Ax + By + Cz + D$. The plane $Ax + By + Cz + D = 0$ is the level surface $f(x,y,z) = 0$. According to the theorem, ∇f is perpendicular to this surface. Now, $f_x = A$, $f_y = B$, and $f_z = C$. Hence,

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

This agrees with the fact that $A\mathbf{i} + B\mathbf{j} + C\mathbf{k} = 0$, as we saw in Section 14.4.

EXAMPLE 3 Find a normal vector to the ellipsoid $x^2 + y^2/4 + z^2/9 = 3$ at the point $(1, 2, 3)$.

SOLUTION The ellipsoid is a level surface of the function

$$f(x,y,z) = x^2 + \frac{y^2}{4} + \frac{z^2}{9}.$$

The gradient of f at the point (x,y,z) is

$$\nabla f = 2x\mathbf{i} + \frac{y}{2}\mathbf{j} + \frac{2z}{9}\mathbf{k}.$$

At $(1, 2, 3)$

$$\nabla f = 2\mathbf{i} + \mathbf{j} + 2/3\mathbf{k}.$$

This vector is normal to the ellipsoid at $(1, 2, 3)$. ◇

Tangent Planes to a Surface

Now that we can find a normal to a surface we can define a tangent plane at a point on the surface.

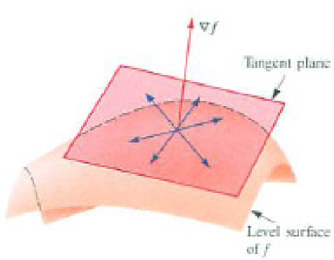


Figure 16.5.5:

DEFINITION *(Tangent plane to a surface)* Consider a surface that is a level surface of a function $u = f(x,y,z)$. Let (a,b,c) be a point on this surface where ∇f is not 0. The tangent plane to the surface at the point (a,b,c) is that plane through (a,b,c) that is perpendicular to the vector ∇f evaluated at (a,b,c).

The tangent plane at (a,b,c) is the plane that best approximates the surface near (a,b,c). It consists of all the tangent lines at (a,b,c) to curves in the surface that pass through the point (a,b,c). See Figure 16.5.5.

Note that an equation of the tangent plane to the surface $f(x,y,z) = k$ at (a,b,c) is

$$\frac{\partial f}{\partial x}(a,b,c)(x-a) + \frac{\partial f}{\partial y}(a,b,c)(y-b) + \frac{\partial f}{\partial z}(a,b,c)(z-c) = 0.$$

EXAMPLE 4 Find an equation of the tangent plane to the ellipsoid $x^2 + y^2/4 + z^2/9 = 3$ at the point $(1, 2, 3)$.

SOLUTION By Example 3, the vector $2\mathbf{i} + \mathbf{j} + 2/3\mathbf{k}$ is normal to the surface at the point $(1, 2, 3)$. The tangent plane consequently has an equation

$$2(x - 1) + 1(u - 2) + 2/3(z - 3) = 0$$

◇

Normals and Tangent Planes to $z = f(x, y)$

A surface may be described explicitly in the form $z = f(x, y)$ rather than implicitly in the form $f(x, y, z) = k$. The techniques already developed enable us to find the normal and tangent plane in the case $z = f(x, y)$ as well.

Finding a normal to the surface $z = f(x, y)$

We need only rewrite the equation $z = f(x, y)$ in the form $z - f(x, y) = 0$. Then define $g(x, y, z)$ to be $z - f(x, y)$. The surface $z - f(x, y)$ is simply the particular level surface of g given by $g(x, y, z) = 0$. There is no need to memorize an extra formula for a vector normal to the surface $z = f(x, y)$. The next example illustrates this advice.

DOUG: I graphed $z = xy$, not $z = x^2 = y^2$. What to do? SHERMAN: I do not see how this graph is incorrect.

EXAMPLE 5 Find a vector perpendicular to the saddle $z = y^2 - x^2$ at the point $(1, 2, 3)$.

SOLUTION In this case, rewrite $z = y^2 - x^2$ as $z + x^2 - y^2 = 0$. The surface in question is a level surface of $g(x, y, z) = z + x^2 - y^2$. Hence $\nabla g = 2x\mathbf{i} - 2y\mathbf{j} + \mathbf{k}$ is perpendicular to the surface at the point $(1, 2, 3)$.

This surface looks like a saddle near the origin. The surface and the normal vector $2\mathbf{i} - 4\mathbf{j} + \mathbf{k}$ are shown in Figure 16.5.6. ◇

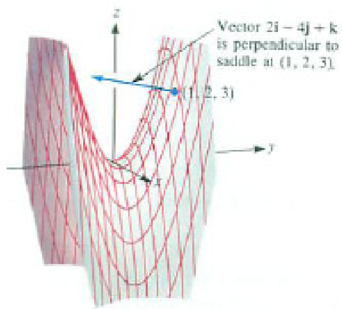


Figure 16.5.6:

Estimates and the Tangent Planes

In the case of a function of one variable, $y = f(x)$, the tangent line at $(a, f(a))$ closely approximates the graph of $y = f(x)$. The equation of the tangent line $y = f(a) + f'(a)(x - a)$ gives us a linear approximation of $f(x)$. (See Section 5.3.)

We can use the tangent plane to the surface $z = f(x, y)$ similarly. To find the equation of the plane tangent at $(a, b, f(a, b))$, we first rewrite the equation of the surface as

$$g(x, y, z) = f(x, y) - z = 0.$$

Then ∇g is a normal to the surface at $(a, b, f(a, b))$. Now,

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

where the partial derivatives are evaluated at (a, b).

The equation of the tangent plane at $(a, b, f(a, b))$ is therefore

$$\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) - (z - f(a,b)) = 0.$$

We can rewrite this equation as

$$z = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b). \qquad (16.5.4)$$

Letting $\Delta x = x - a$ and $\Delta y = y - b$, (16.5.4) becomes

$$z = \frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y.$$

This tells us that the change of the z coordinate on the tangent plane, as the x coordinate changes from a to $a + \Delta x$ and the y coordinate changes from b to $b + \Delta y$ is exactly

$$\frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y.$$

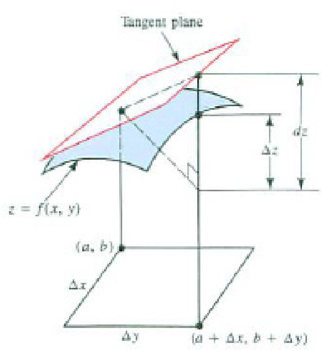


Figure 16.5.7:

By (16.3.1) in Section 16.3, this is an estimate of the change Δf in the function f as its argument changes from (a, b) to $(a + \Delta x, b + \Delta y)$. This is another way of saying that "the tangent plane to the surface $z = f(x, y)$ at $(a, b, f(a, b))$ looks a lot like that surface near that point." See Figure 16.5.7.

The expression $f_x(a, b)\ dx + f_y(a, b)\ dy$ is called the differential of f at (a, b). For small values of dx and dy it is a good estimate of $\Delta f = f(a + dx, b + dy) - f(a, b)$.

EXAMPLE 6 Let $z = f(x, y) = x^2y$. Let $\Delta z = f(1.01, 2.02) - f(1, 2)$ and let

$$dz = \frac{\partial f}{\partial x}(1,2)\cdot 0.0 + \frac{\partial f}{\partial y}(1,2)\cdot 0.02.$$

Compute Δz and dz.

SOLUTION

$$\Delta z = (1.01)^2(2.02) - 1^2 2 = 2.060602 - 2 = 0.060602$$

Since $f_x = 2xy$ and $f_y = x^2$, we have $f_x = 4$ and $f_y = 1$ at $(1, 2)$. Hence,

$$dz = (4)(0.01) + (1)(0.02) = 0.06.$$

Note that dz is a good approximation of Δz. ◇

Function	Level Curve/Surface	Normal	Tangent
$f(x,y)$	level curve: $f(x,y)=k$	$\nabla f = f_x\mathbf{i} + f_y\mathbf{j}$	Tangent line at (a,b) is $f_x(a,b)(x-a)+f_y(a,b)(y-b)=f(a,b)$
$f(x,y,z)$	level surface: $f(x,y,z)=k$	$\nabla f = f_x\mathbf{i} + f_y\mathbf{j} + f_z\mathbf{k}$	Tangent plane at (a,b,c) is $f_x(a,b,c)(x-a)+f_y(a,b,c)(y-b)+f_z(a,b,c)(z-c)=f(a,b)$

Table 16.5.1: t15-5-1

Summary

This table summarizes mot of what we did concerning normal vectors.

To find a normal and tangent plane to a surface given in the form $z = f(x,y)$, treat the surface as a level surface of the function $z - f(x,y)$, normally $z - f(x,y) = 0$.

We also showed that the differential approximation of Δf in Section 16.3 is simply the change along the tangent plane.

DOUG: Must get implicit diff in partials somewhere?? SHERMAN: Exercises?? Maybe back in the Chain Rule section, with a few more exercises in this section. Or, in §16.8.

EXERCISES for Section 16.5 *Key:* R–routine, M–moderate, C–challenging

1.[R] In estimating the value of a right circular cylindrical tree trunk, a lumber jack may make a 5 percent error in estimating the diameter and a 3 percent error in measuring the height. How large an error may he make in estimating the volume?

2.[R] Let T denote the time it takes for a pendulum to complete a back-and-forth swing. If the length of the pendulum is L and g the acceleration due to gravity, then

$$T = 2\pi\sqrt{\frac{L}{g}}.$$

A 3 percent error may be made in measuring L and a 2 percent error in measuring g. How large an error may we make in estimating T?

3.[R] Let $A(x, y) = xy$ be the area of a rectangle of sides x and y. Compute ΔA and dA and show them in Figure 16.5.8

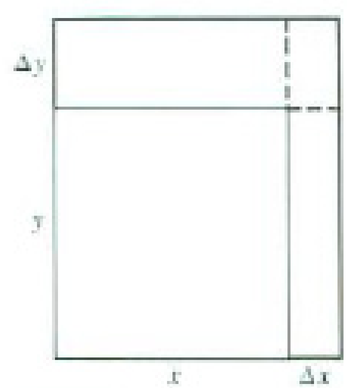


Figure 16.5.8:

The differential of a function $u = f(x, y, z)$ is defined to be $f_x\Delta x + f_y\Delta y + f_z\Delta z$, in analogy with the differential of a function of two variables.

4.[R] Let $V(x, y, z) = xyz$ be the volume of a box of sides x, y, and z. Compute ΔV and dV and show them in Figure 16.5.9.

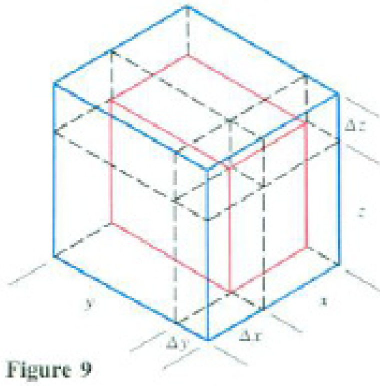


Figure 9

Figure 16.5.9:

5.[R] Let $u = f(x, y, z)$ and $\mathbf{r} = \mathbf{G}(t)$. Then u is a composite function of t. Show that

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

where ∇f is evaluated at $\mathbf{G}(t)$. For instance, let $y = f(x, y, z)$ and let $\mathbf{G}$ describe the journey of a bug. Then the rate of change in the temperature as observed by the but is the dot product of the temperature gradient ∇f and the velocity vector $\mathbf{v} = \mathbf{G}'$.

6.[R] We have found a way to find a normal and a tangent plane to a surface. How would you find a *tangent line* to a surface? Illustrate your method by finding a line that is tangent to the surface $z = xy$ at $(2, 3, 6)$.

7.[R] Suppose you are at the point (a, b, c) on the level surface $f(x, y, z) = k$. At that point $\nabla F = 2\mathbf{i} + 3\mathbf{j} - 4\mathbf{k}$.

(a) If $\mathbf{u}$ is tangent to the surface at (a, b, c), what would $D_{\mathbf{u}}f$ equal?

(b) If $\mathbf{u}$ is normal to the level surface at (a, b, c), what would $D_{\mathbf{u}}$ equal? (There are two such normals.)

8.[R]

(a) Draw three level curves of the function f defined by $f(x, y) = xy$. Include the curve through $(1, 1)$ as one of them.

(b) Draw three level curves of the function g defined by $g(x, y) = x^2 - y^2$. Include the curve through $(1, 1)$ as one of them.

(c) Draw three level curves of the function g defined by $g(x, y) = x^2 - y^2$. Include the curve through $(1, 1)$ as one of them.

(d) Prove that each level curve of f intersects each level curve of g at a right angle.

(e) If we think of f as air pressure, how may we interpret the level curves of g?

9.[R]

(a) Draw a level curve for the function $2x^2 + y^2$.

(b) Draw a level curve for the function y^2/x.

(c) Prove that any level curve of $2x^2+y^2$ crosses any level curve of y^2/x at a right angle.

10.[R] The surfaces $x^2yz = 1$ and $xy + yz + zx = 3$ both pass through the point $(1, 1, 1)$. The tangent planes to these surfaces meet in a line. Find parametric equations for this line.

11.[R] Let $T(x, y, z)$ be the temperature at the point (x, y, z), where ∇T is not $\mathbf{0}$. A level surface $T(x, y, z) - k$ is called an *isotherm*. Show that if you are at the point (a, b, c) and wish to move in the direction in which the temperature changes most rapidly, you would move in a direction perpendicular to the isotherm that passes through (a, b, c).

12.[R] Two surfaces $f(x, y, z) = 0$ and $g(x, y, z) = 0$ both pass through the point (a, b, c). Their intersection is a curve. How would you find a tangent vector to that curve at (a, b, c)?

13.[R] Write a short essay on the wonders of the chain rule. Include a description of how it was used to show that $D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u}$ and in showing that ∇f is a normal to the level surface of f at the point where it is evaluated.

The angle between two surfaces that pass through (a, b, c) is defined as the angle between the two lines through (a, b, c) that are perpendicular to the two surfaces at the point (a, b, c). This angle is taken to be acute. Use this definition in Exercises 14 to 16.

14.[R]

(a) Show that the point $(1, 1, 2)$ lies on the surfaces $xyz = 2$ and $x^3yz^2 = 4$.

(b) Find the angle between the surfaces in (a) at the point $(1, 1, 2)$.

15.[R]

(a) Show that the point $(1, 2, 3)$ lies on the plane

$$2x + 3y - z = 5$$

and the sphere

$$x^2 + y^2 + z^2 = 14.$$

(b) Find the angle between them at the point $(1, 2, 3)$.

16.[R]

(a) Show that the surfaces $z = x^2y^3$ and $z = 2xy$ pass through the point $(2, 1, 4)$.

(b) At what angle do they cross at that point?

17.[R] Let $z = f(x, y)$ describe a surface. Assume that at $(3, 5)$, $z = 7$, $\partial z/\partial x = 2$, and $\partial z/\partial y = 3$.

(a) Find two vectors that are tangent to the surface at $(3, 5, 7)$.

(b) Find a normal to the surface at $(3, 5, 7)$.

(c) Estimate $f(3.02, 4.99)$.

18.[R] This map shows the pressure $p(x, y)$ in terms of level curves called isobars. Where is the gradient of p, ∇p the longest? In what direction does it point? In which direction (approximately) would the wind vector point?

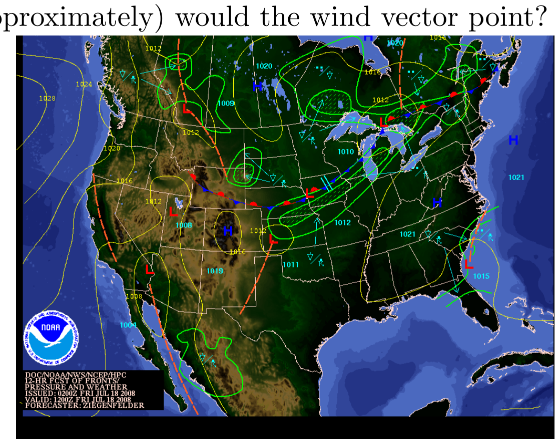


Figure 16.5.10: Source: `http://www.walltechnet.com/b_f/Weather/USAIsobarMap.htm` (18 July 2008)

19.[M] How far is the point $(2,1,3)$ from the tangent plane to $z = xy$ at $(3,4,12)$?

20.[C] The surface $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ is called an ellipsoid. If $a^2 = b^2 = c^2$ it is a sphere. Show that if a^2, b^2, and c^2 are distinct, then there are exactly six normals on the ellipsis that pass through the origin.

21.[C] Let S be a surface with the property that its target planes are always perpendicular to $\mathbf{r}$. Must S be a sphere?

16.6 Critical Points and Extrema

In the case of a function of one variable, $y = f(x)$, the first and second derivatives were of use in searching for relative extrema. First, we looked for critical numbers, that is, solutions of the equation $f'(x) = 0$. Then we checked the value of $f''(x)$ at each such point. If $f''(x)$ were positive, the critical number gave a relative minimum. If $f''(x)$ were negative, the critical number gave a relative maximum. If $f''(x)$ were 0, then anything might happen: a relative minimum or maximum or neither. (For instance, at 0 the functions x^4, $-x^4$, and x^3 have both first and second derivatives equal to 0, but the first function has a relative minimum there, the second has a relative maximum, and the third has neither.) In such a case, we have to resort to other tests.

Recall: $f''(x)$ positive means the graph of f is concave up; $f''(x)$ negative means the graph of f is concave down.

This section extends the idea of a critical point to functions $f(x, y)$ of two variables and shows how to use the second-order partial derivatives f_{xx}, f_{yy}, and f_{xy} to see whether the critical point provides a relative maximum, relative minimum, or neither.

Remember that $\frac{\partial f}{\partial x} = f_x$. The subscript notation is used in text to save space.

Extrema of $f(x, y)$

The number M is called the **maximum** (or **global maximum**) of f over a set R in the plane if it is the largest value of $f(x, y)$ for (x, y) in R. A **relative maximum** (or **local maximum**) of f occurs at a point (a, b) in R if there is a disk around (a, b) such that $f(a, b) \geq f(x, y)$ for all points (x, y) in the disk. **Minimum** and **relative** (or **local) minimum** are defined similarly.

Let us look closely at the surface above a point (a, b) where a relative maximum of f occurs. Assume that f is defined for all points within some circle around (a, b) and possesses partial derivatives at (a, b). Let L_1 be the line $y = b$ in the xy plane; let L_2 be the line $x = a$ in the xy plane. (See Figure 16.6.1. Assume, for convenience, that the values of f are positive.)

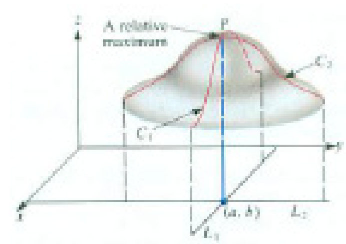


Figure 16.6.1:

Let C_1 be the curve in the surface directly above the line L_1. Let C_2 be the curve in the surface directly above the line L_2. Let P be the point on the surface directly above (a, b).

Since f has a relative maximum at (a, b), no point on the surface near P is higher than P. Thus P is a highest point on the curve C_1 and on the curve C_2 (for points near P). The study of functions of one variable showed that both these curves have horizontal tangents at P. In other words, at (a, b) both partial derivatives of f must be 0:

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

This conclusion is summarized in the following theorem.

Theorem. *Relative Extremum of $f(x, y)$ Let f be defined on a domain that includes the point (a, b) and all points within some circle whose center is (a, b). If f has a relative maximum (or relative minimum) at (a, b) and f_x and f_y exist at (a, b), then both these partial derivatives are 0 at (a, b); that is,*

$$\frac{\partial f}{\partial x}(a, b) = 0 = \frac{\partial f}{\partial y}(a, b),$$

In short, the gradient of f, ∇f is $\mathbf{0}$ at a relative extremum.

A point (a, b) where both partial derivatives f_x and f_y are 0 is clearly of importance. The following definition is analogous to that of a critical point of a function of one variable.

> **DEFINITION** *(Critical point)* If $f_x(a, b) = 0$ and $f_y(a, b) = 0$, the point (a, b) is a **critical point** of the function $f(x, y)$.

You might expect that if (a, b) is a critical point of f and the two second partial derivatives f_{xx} and f_{yy} are both positive at (a, b), then necessarily has a relative minimum at (a, b). The next example shows that *the situation is not that simple.*

EXAMPLE 1 Find the critical points of $f(x, y) = x^2 + 3xy + y^2$ and determine whether there is an extremum there.

SOLUTION First, find any critical points by setting both f_x and f_y equal to 0. This gives the simultaneous equations

$$2x + 3y = 0 \qquad \text{and} \qquad 3x + 2y = 0.$$

Since the only solution of these equations is $(x, y) = (0, 0)$, the function has one critical point, namely $(0, 0)$.

Now look at the graph of f for (x, y) near $(0, 0)$.

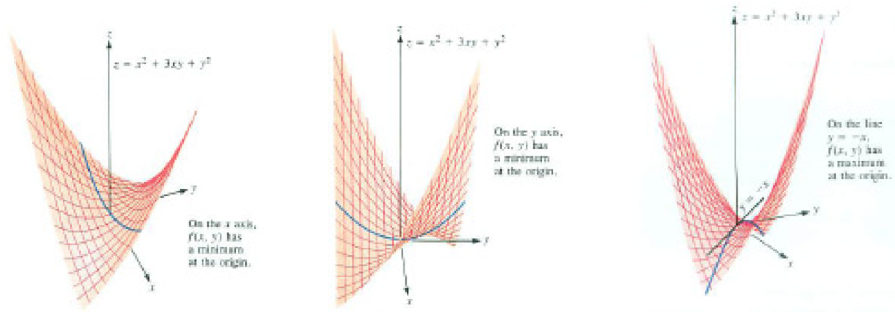
First, consider how f behaves for points on the x axis. We have $f(x, 0) = x^2 + 3 \cdot x \cdot 0 + 0^2 = x^2$. Therefore, considered *only as a function of x*, the function has a minimum at the origin. (See Figure 16.6.2(a).)

On the y-axis, the function reduces to $f(0, y) = y^2$, whose graph is another parabola with a minimum at the origin. (See Figure 16.6.2(b).) Note also that $f_{xx} = 2$ and $f_{yy} = 2$, so both are positive at $(0, 0)$.

So far, the evidence suggests that f has a relative minimum at $(0, 0)$. However, consider its behavior on the line $y = -x$. For points (x, y) on this line

$$f(x, y) = f(x, -x) = x^2 + 2x(-x) + (-x)^2 = -x^2.$$

On this line the function assumes negative values, and its graph is a parabola opening downward, as shown in Figure 16.6.2(c).



(a) (b) (c)

Figure 16.6.2:

Thus $f(x, y)$ has neither a relative maximum nor minimum at the origin. Its graph resembles a saddle. ◇

Example 1 shows that to determine whether a critical point of $f(x, y)$ provides an extremum, it is not enough to look at f_{xx} and f_{yy} The criteria are more complicated and involve the mixed partial derivative f_{xy} as well. Exercise 58 outlines a proof of the following theorem. At the end of this section a proof is presented in the special case when $f(x, y)$ is a polynomial of the form $Ax^2 + Bxy + Cy^2$, where A, B and C are constants.

f_{xx} and f_{yy} describe the behavior of $f(x, y)$ only on lines parallel to the x-axis and y-axis, respectively.

Theorem 16.6.1. *Second-partial-derivative test for $f(x, y)$ Let (a, b) be a critical point of the function $f(x, y)$. Assume that the partial derivatives f_x, f_y, f_{xx}, f_{xy}, and f_{yy} are continuous at and near (a, b). Let*

In subscript notation, $D = f_{xx}f_{yy} - (f_{xy})^2$.

$$D = \frac{\partial^2 f}{\partial x^2}(a, b)\frac{\partial^2 f}{\partial y^2}(a, b) - \left(\frac{\partial^2 f}{\partial x \partial y}(a, b)\right)^2.$$

1. *If $D > 0$ and $f_{xx}(a, b) > 0$, then f has a relative minimum at (a, b).*

2. *If $D > 0$ and $f_{xx}(a, b) < 0$, then f has a relative maximum at (a, b).*

3. *If $D < 0$, then f has neither a relative minimum nor a relative maximum at (a, b). (There is a **saddle point** at (a, b).)*

If $D = 0$, then anything can happen; there may be a relative minimum, a relative maximum, or a saddle. These possibilities are illustrated in Exercise 43.

To see what the theorem says, consider case 1, the test for a relative minimum. It says that $f_{xx}(a, b) > 0$) (which is to be expected) and that

$$\frac{\partial^2 f}{\partial x^2}(a, b)\frac{\partial^2 f}{\partial y^2}(a, b) - \left(\frac{\partial^2 f}{\partial x \partial y}(a, b)\right)^2 > 0,$$

Or equivalently,

$$\left(\frac{\partial^2 f}{\partial x \partial y}(a,b)\right)^2 < \frac{\partial^2 f}{\partial x^2}(a,b)\frac{\partial^2 f}{\partial y^2}(a,b). \tag{16.6.1}$$

Memory aid regarding size of f_{xy}

Since the square of a real number is never negative, and $f_{xx}(a,b)$ is positive, it follows that $f_{yy}(a,b) > 0$, which was to be expected. But inequality (16.6.1) says more. It says that the *mixed partial* $f_{xy}(a,b)$ *must not be too large.* For a relative maximum or minimum, inequality (16.6.1) must hold. This may be easier to remember than "$D > 0$."

EXAMPLE 2 Examine each of these functions for relative extrema:

1. $f(x,y) = x^2 + 3xy + y^2$,
2. $g(x,y) = x^2 + 2xy + y^2$,
3. $h(x,y) = x^2 + xy + y^2$.

SOLUTION

1. The case $f(x,y) = x^2 + 3xy + y^2$ is Example 1. The origin is the only critical point, and it provides neither a relative maximum nor a relative minimum. We can check this by the use of the discriminant. We have

$$\frac{\partial^2 f}{\partial x^2}(0,0) = 2, \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = 3, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = 2.$$

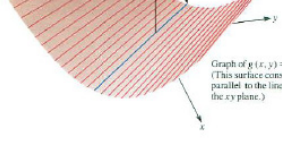
   Hence $D = 2 \cdot 2 - 3^2 = -5$ is negative. By the second-partial-derivative test, there is neither a relative maximum nor a relative minimum at the origin. Instead, there is a saddle there.

2. It is straightforward matter to show that all the points on the line $x+y = 0$ are critical points of $g(x,y) = x^2 + 2xy + y^2$. Moreover,

$$\frac{\partial^2 g}{\partial x^2}(x,y) = 2, \qquad \frac{\partial^2 g}{\partial x \partial y}(x,y) = 2, \qquad \text{and} \qquad \frac{\partial^2 g}{\partial y^2}(x,y) = 2.$$

   Thus the discriminant $D = 2 \cdot 2 - 2^2 = 0$. Since $D = 0$, the discriminant gives no information.

   Note, however, that $x^2 + 2xy + y^2 = (x+y)^2$ and so, being the square of a real number, is always greater than or equal to 0. Hence the origin provides a relative minimum of $x^2 + 2xy + y^2$. (In fact, any point on the line $x + y = 0$ provides a relative minimum. Since $g(x,y) = (x+y)^2$, the function is constant on each line $x + y = c$, for any choice of the constant c. See Figure 16.6.3.)



Graph of $g(x,y) = x^2 + 2xy + y^2$ (This surface consists of lines parallel to the line $x + y = 0$ in the xy plane.)

3. For $h(x,y) = x^2 + xy + y^2$, again the origin is the only critical point and we have

$$\frac{\partial^2 h}{\partial x^2}(0,0) = 2, \qquad \frac{\partial^2 h}{\partial x \partial y}(0,0) = 1, \qquad \text{and} \quad \frac{\partial^2 h}{\partial y^2}(0,0) = 2.$$

In this case, $D = 2 \cdot 2 - 1^2 = 3$ is positive and $h_{xx}(0,0) > 0$. Hence $x^2 + xy + y^2$ has a relative minimum at the origin.

The graph of h is shown in Figure 16.6.4

◇

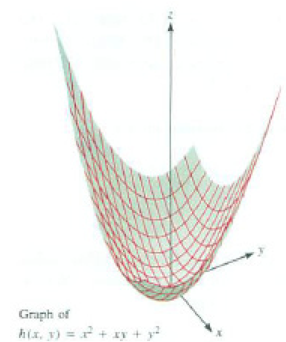


Figure 16.6.4: Function has no global extrema.

EXAMPLE 3 Examine $f(x,y) = x + y + 1/(xy)$ for global and relative extrema.

SOLUTION When x and y are both large positive numbers or small positive numbers, then $F(x,y)$ may be arbitrarily large. There is therefore no global maximum. By allowing x and y to be negative numbers of large absolute values, we see that there is no global minimum.

Any local extrema will occur at a critical point. We have

$$\frac{\partial f}{\partial x} = 1 - \frac{1}{x^2 y} \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 1 - \frac{1}{xy^2}.$$

Setting these derivatives equal to 0 gives

$$\frac{1}{x^2 y} = 1 \qquad \text{and} \qquad \frac{1}{xy^2} = 1 \tag{16.6.2}$$

Hence $x^2 y = xy^2$. Since the function f is not defined when x or y is 0, we may assume $xy \neq 0$. Dividing both sides of $x^2 y = xy^2$ by xy gives $x = y$. By (16.6.2) (either equation), $1/x^3 = 1$; hence $x = 1$. Thus there is only one critical point, namely, $(1,1)$.

To find whether it is a relative extremum, use Theorem 16.6.1. We have

$$\frac{\partial^2 f}{\partial x^2} = \frac{2}{x^3 y}, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{1}{x^2 y^2}, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2} = \frac{2}{xy^3}.$$

Thus at $(1,1)$,

$$\frac{\partial^2 f}{\partial x^2} = 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = 1, \qquad \text{and} \qquad \frac{\partial^2 f}{\partial y^2} = 2.$$

Therefore,

$$D = 2 \cdot 2 - 1^2 = 3 > 0.$$

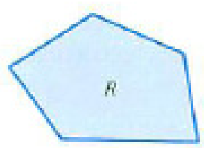
Since $D > 0$ and $f_{xx})(1,1) > 0$, the point $(1,1)$ provides a relative minimum. ◇

Extrema on a Bounded Region

In Section 4.3, we saw how to find a maximum of a differentiable function, $y = f(x)$, on an interval $[a, b]$. The procedure is as follows:

1. First find any numbers x in $[a, b]$ (other than a or b) where $f'(x) = 0$. Such a number is called a critical number. If there are no critical numbers, the maximum occurs at a or b.

2. If there are critical numbers, evaluate f at them. Also find the values of $f(a)$ and $f(b)$. The maximum of f in $[a, b]$ is the largest of the numbers: $f(a)$, $f(b)$, and the values of f at critical numbers.



A continuous function on R (which includes the border) has a maximum value at some point in R.

Figure 16.6.5:

We can similarly find the maximum of $F(x, y)$ in a region R in the plane bounded by some polygon or curve. (See Figure 16.5.7.) It is assumed that R includes its border and is a finite region in the sense that it lies within some disk. (In advanced calculus, it is proved that a continuous function defined on such a domain has a maximum – and a minimum – value.) If f has continuous partial derivatives, the procedure for finding a maximum is similar to that for maximizing a function on a closed interval.

1. First find any points that are in R but not on the boundary of R where both f_x and f_y are 0. These are called **critical points**. (if there are no critical points, the maximum occurs on the boundary.)

2. If there are critical points, evaluate f at them. Also find the maximum of f on the boundary. The maximum of f on R is the largest value of f on the boundary and at critical points.

A similar procedure finds the minimum value on a bounded region.

EXAMPLE 4 Maximize the function $f(x, y) = xy(108-2x-2y) = 108xy - 2x^2y - 2xy^2$ on the triangle R bounded by the x-axis, the y-axis, and the line $x + y = 54$. (See Figure 16.6.6.)

Find all critical points.

SOLUTION First find any critical points. We have

$$\frac{\partial f}{\partial x}(x, y) = 108y - 4xy - 2y^2 = 0 \qquad (16.6.3)$$

$$\frac{\partial f}{\partial y}(x, y) = 108x - 2x^2 - 4xy = 0 \qquad (16.6.4)$$

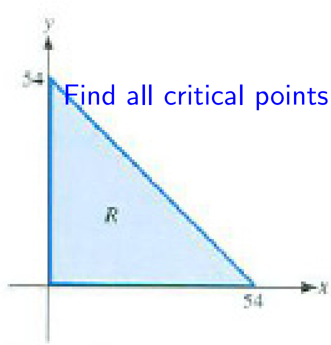


Figure 16.6.6:

which give the simultaneous equations

$$2y(54 - 2x - y) = 0, \qquad (16.6.5)$$

$$2x(54 - x - 2y) = 0. \qquad (16.6.6)$$

By the first equation, $y = 54-2x$. Substitution of this into the second equation gives: $54 - x - 2(54 - 2x) = 0$, or $-54 + 3x = 0$. Hence $x = 18$ and therefore $y = 54 - 2 \cdot 18 = 18$.

The point $(18, 18)$ lies in the interior of R, since it lies above the x-axis, to the right of the y-axis, and below the line $x + y = 54$. Furthermore, $f(18, 18) = 18 \cdot 18(108 - 2 \cdot 18 - 2 \cdot 18) = 11,664$.

Evaluate f at critical points.

Next we examine the function $f(x, y) = xy(108-2x-2y)$ on the boundary of the triangle R. On the base of R, $y = 0$, so $f(x, y) = 0$. On the left edge of R, $x = 0$, so again $f(x, y) = 0$. On the slanted edge, which lies on the same line $x + y = 54$, we have $108 - 2x - 2y = 0$, so $f(x, y) = 0$ on this edge also. Thus $f(x, y) = 0$ on the entire boundary.

Evaluate f on boundary.

Therefore, the local maximum occurs at the critical point $(18, 18)$ and has the value $11, 664$. ⋄

EXAMPLE 5 The combined length and girth (distance around) of a package sent through the mail cannot exceed 108 inches. If the package is a rectangle box, how large can its volume be?

SOLUTION Introduce letters to name the quantities of interest. We label its length (a longest side) z and the other sides x and y, as in Figure 16.6.7. The volume $V = xyz$ is to be maximized, subject to girth plus length at most 108, that is,

$$2x + 2y + z \le 108.$$

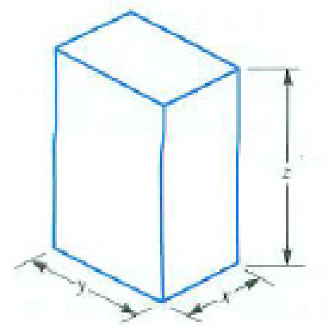


Figure 16.6.7:

Since we want the largest box, we might as well restrict our attention to boxes for which

$$2x + 2y + z = 108. \tag{16.6.7}$$

By (16.6.7), $z = 108 - 2x - 2y$. Thus $V = xyz$ can be expressed as a function of two variables:

$$V = f(x, y) = xy(108 - 2x - 2y).$$

This function is to be maximized on the triangle described by $x \ge 0$, $y \ge 0$, $2x + 2y \le 108$, that is, $x + y \le 54$.

Why is $2x + 2y \le 108$?

These are the same function and region as in the previous example. Hence, the largest box has $x = y = 18$ and $z = 108-2x-2y = 108-2\cdot18-2\cdot18 = 36$; its dimensions are 18 inches by 18 inches by 36 inches and its volume is $11, 664$ cubic inches. ⋄

Remark: In Example 5 we let z be the length of a longest side, an assumption that was never used. So if the Postal Service regulations read "The length of one edge plus the girth around the other edges shall not exceed 108 inches," the effect would be the same. You would not be able to send a larger box by, say, measuring the girth around the base formed by its largest edges.

EXAMPLE 6 Find the maximum and minimum values of $f(x,y) = x^2 + y^2 - 2x - 4y$ on the disk R of radius 3 and center $(0,0)$.

SOLUTION First, find any critical points. We have

$$\frac{\partial f}{\partial x} = 2x - 2 \qquad \text{and} \qquad \frac{\partial f}{\partial y} = 2y = 4.$$

The equations

$$\begin{aligned} 2x - 2 &= 0 \\ 2y - 4 &= 0 \end{aligned}$$

have the solutions $x = 1$ and $y - 2$. This point lies in R (since its distance from the origin is $\sqrt{1^2 + 2^2} = \sqrt{5}$, which is less than 3). At the critical point $(1,2)$, the value of the function is $1^2 + 2^2 - 2(1) - 4(2) = 5 - 2 - 8 = -5$.

Second, find the behavior of f on the boundary, which is a circle of radius 3. We parameterize this circle:

$$\begin{aligned} x &= 3\cos(\theta) \\ y &= 3\sin(\theta). \end{aligned}$$

On this circle,

$$\begin{aligned} f(x,y) &= x^2 + y^2 - 2x - 4y \\ &= (3\cos(\theta))^2 + (3\sin(\theta))^2 - 2(3\cos(\theta)) - 4(3\sin(\theta)) \\ &= 9\cos^2(\theta) + 9\sin^2(\theta) - 6\cos(\theta) - 12\sin(\theta) \\ &= 9 - 6\cos(\theta) - 12\sin(\theta). \end{aligned}$$

We now must find the maximum and minimum of the single-variable function $g(\theta) = 9 - 6\cos(\theta) - 12\sin(\theta)$ for θ in $[0, 2\pi]$.

To do this, find $g'(\theta)$:

$$g'(\theta) = 6\sin\theta - 12\cos\theta.$$

Setting $g'(\theta) = 0$ gives

$$0 = 6\sin(\theta) - 12\cos(\theta)$$

or

$$\sin(\theta) = 2\cos(\theta). \tag{16.6.8}$$

Why is $\cos(\theta)$ not 0?

To solve (16.6.8), divide by $\cos(\theta)$ (which will not be 0), getting

$$\frac{\sin(\theta)}{\cos(\theta)} = 2$$

or

$$\tan(\theta) = 2.$$

There are two angles θ in $[0, 2\pi]$ such that $\tan(\theta) = 2$. One is in the first quadrant, $\theta = \arctan(2)$, and the other is in the third quadrant, $\pi + \arctan(2)$. To evaluate $g(\theta) = 9 - 6\cos(\theta) - 12\sin(\theta)$ at these angles, we must compute their cosine and sine. The right triangle in Figure 16.6.8 helps us do this.

Inspection of Figure 16.6.8 shows that for $\theta = \arctan(2)$,

$$\cos(\theta) = \frac{1}{\sqrt{5}} \qquad \text{and} \qquad \sin(\theta) = \frac{2}{\sqrt{5}}.$$

For this angle

$$g(\arctan(2)) \;=\; 9 - 6\left(\frac{1}{\sqrt{5}}\right) - 12\left(\frac{2}{\sqrt{5}}\right) = 9 - \frac{30}{\sqrt{5}} \approx -4.41641.$$

When $\theta = \pi + \arctan(2)$,

$$\cos(\theta) = \frac{-1}{\sqrt{5}} \qquad \text{and} \qquad \sin(\theta) = \frac{-2}{\sqrt{5}}.$$

So

$$g(\pi + \arctan(2)) \;=\; 9 - 6\left(\frac{-1}{\sqrt{5}}\right) - 12\left(\frac{-2}{\sqrt{5}}\right)$$
$$= 9 + \frac{30}{\sqrt{5}} \approx 22.41641.$$

Figure 16.6.8:

Since $g(2\pi) = g(0) = 9 - 6(1) - 12(0) = 3$, the maximum of f on the border of R is about 22.41641 and the minimum is about -4.41641. (Recall that at the critical point the value of f is -5.)

We conclude that the maximum value of f on R is about 22.41641 and the minimum value is -5 (and it occurs at the point $(1, 2)$, which is not on the boundar)]. See Figure 16.6.9. ◇

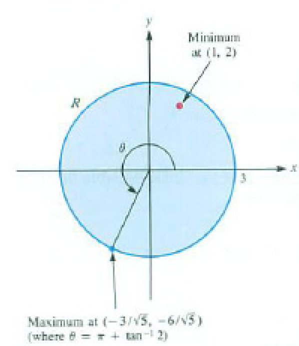


Figure 16.6.9:

Proof of Theorem 16.6.1 in a Special Case

We will prove Theorem 16.6.1 in case $f(x, y)$ is a second-degree polynomial of the form

$$f(x, y) = Ax^2 + Bxy + Cy^2.$$

Theorem 16.6.2. *Let $f(x, y) = Ax^2 + Bxy + Cy^2$, where A, B, and C are constants. Then $(0, 0)$ is a critical point. Let*

$$D = \frac{\partial^2 f}{\partial x^2}(0,0)\frac{\partial^2 f}{\partial y^2}(0,0) - \left(\frac{\partial^2 f}{\partial x \partial y}(0,0)\right)^2.$$

1. *If $D > 0$ and $f_{xx}(0,0) > 0$, then f has a relative minimum at $(0,0)$.*

2. *If $D > 0$ and $f_{xx}(0,0) < 0$, then f has a relative maximum at $(0,0)$.*

3. *If $D < 0$, then f has neither a relative minimum nor a relative maximum at $(0,0)$.*

Proof

We prove Case 1, leaving Cases 2 and 3 for Exercises 60 and 61.

First, compute the first- and second-order partial derivatives of f:

$$\frac{\partial f}{\partial x} = 2Ax + By, \qquad \frac{\partial f}{\partial y} = Bx + 2Cy,$$

$$\frac{\partial^2 f}{\partial x^2} = 2A, \quad \frac{\partial^2 f}{\partial x \partial y} = B, \quad \frac{\partial^2 f}{\partial y^2} = 2C.$$

Note that both f_x and f_y are 0 at $(0,0)$. Hence $(0,0)$ is a critical point and $f(0,0) = 0$. We must show that $f(x,y) \geq 0$ for (x,y) near $(0,0)$. [In fact we will show that $f(x,y) \geq 0$ for all (x,y).]

Next, expressing Case 1 in terms of A, B, and C, we have

$$D = f_{xx}(0,0)f_{yy}(0,0) - f_{xy}^2(0,0) = (2A)(2C) - B^2 = 4AC - B^2 > 0.$$

and $f_{xx}(0,0) = 2A > 0$. In short, we are assuming that $4AC - B^2 > 0$ and $A > 0$, and want to deduce that $f(x,y) = Ax^2 + Bxy + Cy^2 \geq 0$, for (x,y) near $(0,0)$.

Since A is positive, this amounts to showing that

$$A(Ax^2 + Bxy + Cy^2) \geq 0. \tag{16.6.9}$$

We multiply by A to simplify completing the square on the next step.

Now we complete the square,

$$\begin{aligned} A(Ax^2 + Bxy + Cy^2) &= A^2x^2 + ABxy + ACy^2 \\ &= A^2x^2 + ABxy + \frac{B^2}{4}y^2 - \frac{B^2}{4}y^2 + ACy^2 \\ &= (Ax + \frac{B}{2}y)^2 + (AC - \frac{B^2}{4})y^2 \\ &= (Ax + \frac{B}{2}y)^2 + (\frac{4AC - B^2}{4})y^2. \end{aligned}$$

Now, $\left(Ax + \frac{B}{2}y\right)^2 \geq 0$ and $y^2 \geq 0$ since they are squares of real numbers. But by our assumption on D, $4AC - B^2$ is positive. Thus (16.6.9) holds for all (x,y), not just for (x,y) near $(0,0)$ varies Case 1 of the theorem is proved. •

Summary

We defined a critical point of $f(x, y)$ as a point where both partial derivatives f_x and f_y are 0. Even if f_{xx} and f_{yy} are negative there, such a point need not provide a relative maximum. We must also know that f_{xy} is not too large in absolute value.

- If $f_{xx} < 0$ and $f_{xy}^2 < f_{xx}f_{yy}$, then there is indeed a relative maximum at the critical point. (Note that the two inequalities imply $f_{yy} < 0$.)
- Similar criteria hold for a relative maximum: if $f_{xx} > 0$ and $f_{xy}^2 < f_{xx}F_{yy}$, then this critical point is a relative minimum.
- The critical point is a saddle point when $f_{xy} > f_{xx}f_{yy}$.
- When $f_{xy}^2 = f_{xx}f_{yy}$, the critical point may be a relative maximum, relative minimum, or neither.

We also saw how to find extrema of a function defined on a bounded region.

SHERMAN: Have you more varied realistic exercises?

DI: Are there some more complaints about these problems in your earlier books? We can look, but I don't believe it's too critical to be creative here.

EXERCISES for Section 16.6 *Key:* R–routine, M–moderate, C–challenging

Use Theorems 16.6 and 16.6.1 to determine any relative maxima or minima of the functions in Exercises 1 to 10.

1.[R] $x^2 + 3xy + y^2$

2.[R] $f(x,y) = x^2 - y^2$

3.[R] $f(x,y) = x^2 - 2xy + 2y^2 + 4x$

4.[R] $f(x,y) = x^4 + 8x^2 + y^2 - 4y$

5.[R] $f(x,y) = x^2 - xy + y^2$

6.[R] $f(x,y) = x^2 + 2xy + 2y^2 + 4x$

7.[R] $f(x,y) = 2x^2 + 2xy + 5y^2 + 4x$

8.[R] $f(x,y) = -4x^2 - xy - 3y^2$

9.[R] $f(x,y) = 4/x + 2/y + xy$

10.[R] $f(x,y) = x^3 - y^3 + 3xy$

Let f by a function of x and y such that at (a,b) both f_x and f_y equal 0. In each of Exercises 11 to 16. values are specified for f_{xx}, f_{xy}, and f_{yy} at (a,b). Assume that all these partial derivatives are continuous. On the basis of the given information decides whether

1. f has a relative maximum at (a,b),
2. f has a relative minimum at (a,b),
3. f has a saddle point at (a,b),
4. there is inadequate information.

11.[R] $f_{xy} = 4$, $f_{xx} = 2$, $f_{yy} = 8$

12.[R] $f_{xy} = -3$, $f_{xx} = 2$, $f_{yy} = 4$

13.[R] $f_{xy} = 3$, $f_{xx} = 2$, $f_{yy} = 4$

14.[R] $f_{xy} = 2$, $f_{xx} = 3$, $f_{yy} = 4$

15.[R] $f_{xy} = -2$, $f_{xx} = -3$, $f_{yy} = -4$

16.[R] $f_{xy} = -2$, $f_{xx} = 3$, $f_{yy} = -4$

In Exercises 17 to 24 find the critical points and the relative extrema of the given functions.

17.[R] $x + y - \frac{1}{xy}$

18.[R] $3xy - x^3 - y^3$

19.[R] $12xy - x^3 - y^3$

20.[R] $6xy - x^2y - xy^2$

21.[R] $\exp(x^3 + y^3)$

22.[R] 2^{xy}

23.[R] $3x + xy + x^2y - 2y$

24.[R] $x + y + \frac{8}{xy}$

25.[R] Find the dimensions of the open rectangular box of volume 1 of smallest surface area. Use Theorem 16.6.1 as a check that the critical point provides a minimum.

26.[R] The material for the top and bottom of a rectangular box costs 3 cents per square foot, and that for the sides 2 cents per square foot. What is the least expensive box that has a volume of 1 cubic foot? Use Theorem 16.6.1 as a check that the critical point provides a minimum.

27.[R] UPS ships packages whose combined length and girth is at most 165 inches (and weighs at most 150 pounds).

(a) What are the dimensions of the package with the largest volume that it ships?

(b) What are the dimensions of the package with maximum surface are that UPS will ship?

28.[R] Let $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, $P_3 = (x_3, y_3)$, and $P_4 = (x_4, y_4)$. Find the coordinates of the point P that minimizes the sum of the squares of the distances from P to the four points.

29.[R] Find the dimensions of the rectangle box of largest volume if its total surface area is to be 12 square meters.

30.[R] Three nonnegative numbers x, y, and z have the sum 1.

(a) How small can $x^2 + y^2 + z^2$ be?

(b) How large can it be?

31.[R] Each year a firm can produce r radios and t television sets at a cost of $2r^2 + rt + 2t^2$ dollars. It sells a radio for \$600 and a television set for \$900.

(a) What is the profit from the sale of r radios and t television sets? NOTE: Profit is revenue less the cost.

(b) Find the combination of r and t that maximizes profit. Use the discriminant as a check.

32.[R] Find the dimensions of the rectangular box of largest volume that can be inscribed in a sphere of radius 1.

33.[R] For which values of the constant k does $x^2 + kxy + 3y^2$ have a relative minimum at $(0, 0)$?

34.[R] For which values of the constant k does the function $kx^2 + 5xy + 4y^2$ have a relative minimum at $(0, 0)$?

35.[R] Let $f(x, y) = (2x^2 + y^2)e^{-x^2-y^2}$.

(a) Find all critical points of f.

(b) Examine the behavior of f when $x^2 + y^2$ is large.

(c) What is the minimum value of f?

(d) What is the maximum value of f?

36.[R] Find the maximum and minimum values of the function in Exercise 35 on the circle

(a) $x^2 + y^2 = 1$,

(b) $x^2 + y^2 = 4$.

HINT: Express the function in terms of θ.

37.[R] Find the maximum value of $f(x, y) = 3x^2 - 4y^2 + 2xy$ for points (x, y) in the square region whose vertices are $(0, 0)$, $(0, 1)$, $(1, 0)$, and $(1, 1)$.

38.[R] Find the maximum value of $f(x, y) = xy$ for points (x, y) in the triangular region whose vertices are $(0, 0)$, $(1, 0)$, and $(0, 1)$.

39.[R] Maximize the function $-x + 3y + 6$ on the quadrilateral whose vertices are $(1, 1)$, $(4, 2)$, $(0, 3)$, and $(5, 6)$.

40.[M]

(a) Show that $z = x^2 - y^2 + 2xy$ has no maximum and no minimum.

(b) Find the minimum and maximum of z if we consider only (x, y) on the circle of radius 1 and center $(0, 0)$, that is all (x, y) such that $x^2 + y^2 = 1$.

(c) Find the minimum and maximum of z if we consider all (x, y) in the disk of radius 1 and center $(0, 0)$, that is, all (x, y) such that $x^2 + y^2 \leq 1$.

41.[M] Suppose z is a function of x and y with continuous second partial derivatives. If, at the point (x_0, y_0), $z_x = 0 = z_y$, $z_{xx} = 3$, and $z_{yy} = 12$, for what values of z_{xy} is it certain that z has a relative minimum at (x_0, y_0)?

42.[M] Let $U(x, y, z) = x^{1/2}y^{1/3}z^{1/6}$ be the "utility" or "desirability" to a given consumer of the amounts x, y, and z of three different commodities. Their prices are, respectively, 2 dollars, 1 dollar, and 5 dollars, and the consumer has 60 dollars to spend. How much of each product should he buy to maximize the utility?

43.[M] This exercise shows that if the discriminant D is 0, then any of the three outcomes mentioned in Theorem 16.6.1 are possible.

(a) Let $f(x, y) = x^2 + 2xy + y^2$. Show that at $(0, 0)$ both f_x and f_y are 0, f_{xx} and f_{yy} are positive, $D = 0$, and f has a relative minimum.

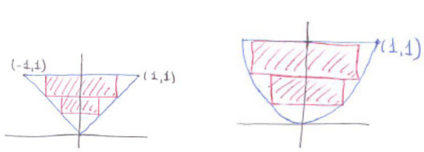
(b) Let $f(x, y) = x^2 + 2xy + y^2 - x^4$ Show that at $(0, 0)$ both f_x and f_y are 0, f_{xx} and f_{yy} are positive, $D = 0$, and f has neither a relative maximum nor a relative minimum at $(0, 0)$.

(c) Give an example of a function $f(x, y)$ for which $(0, 0$ is a critical point and $D = 0$ there, but f has a relative maximum at $(0, 0)$.

44.[M] Let $f(x, y) = ax + by + c$, for constants a, b, and c. Let R be a polygon in the xy plane. Show that the maximum and minimum values of $f(x, y)$ on R are assumed only at vertices of the polygon.

45.[M] Two rectangles are placed in the triangle whose vertices are $(0, 0)$, $(1, 1)$, and $(-1, 1)$ as shown in Figure 16.6.10(a).

SHERMAN:I modified your picture in (b), some. OK? I didn't think it was so bad; see my answer.



(a) (b)

Figure 16.6.10:

Show that they can fill as much as 2/3 of the area of the triangle.

46.[M] Two rectangles are placed in the parabola $y = x^2$ as shown in Figure 16.6.10(b). How large can their total area be?

47.[M] Let $P_0 = (a, b, c)$ be a point not on the surface $f(x, y, z) = 0$. Let P be the point on the surface nearest P_0. Show that $\overrightarrow{PP_0}$ is perpendicular to the surface at P. HINT: Show it is perpendicular to each curve on the surface that passes through P.

48.[C] Let (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) be n points in the plane. Statisticians define the **line of regression** as the line that minimizes the sum of the squares of the differences between y_i and the ordinates of the line at x_i. (See Figure 16.6.11.) Let the typical line in the plane have the equation $y = mx + b$.

(a) Show that the line of regression minimizes the sum $\sum_{i=1}^{n} (y_i - (mx_i + b))^2$ considered as a function of m and b.

(b) Let $f(m, b) = \sum_{i=1}^{n} (y_i - (mx_i + b))^2$. Compute f_m and f_b.

(c) Show that when $f_m = 0 = f_b b$, we have

$$m \sum_{i=1}^{n} x_i^2 + b \sum_{i=1}^{n} x_1 = \sum_{i=1}^{n} x_i y_i$$

and

$$m \sum_{i=1}^{n} x_i + nb = \sum_{i=1}^{n} y_i.$$

(d) When do the simultaneous equations in (c) have a unique solution for m and b?

(e) Find the regression line for the points $(1, 1)$, $(2, 3)$, and $(3, 5)$.

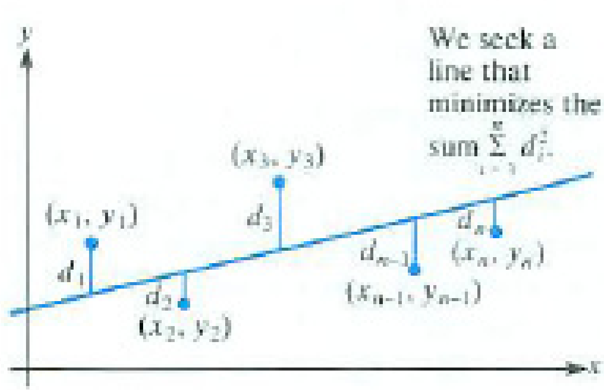


Figure 16.6.11:

49.[C] If your calculator is programmed to compute lines of regression, find and draw the line of regression for the points $(1, 1)$, $(2, 1.5)$, $(3, 3)$, $(4, 2)$ and $(5, 3.5)$.

50.[C] Let $f(x, y) = (y - x^2)(y - 2x^2)$.

(a) Show that f has neither a local minimum nor a local maximum at $(0, 0)$

(b) Show that f has a local minimum at $(0, 0)$ when considered only on any fixed line through $(0, 0)$.

Suggestion for (b): Graph $y = x^2$ and $y = 2x^2$ and show where $f(x, y)$ is positive and where it is negative.

51.[C] Find (a) the minimum value of xyz, and (b) the maximum value of xyz, for all triplets of nonnegative real numbers x, y, z such that $x + y + z = 1$.

52.[C]

(a) Deduce from Exercise 51 that for any three nonegative numbers a, b, and c, $\sqrt[3]{abc} \leq (a+b+c)/3$. NOTE: This exercise asserts that the "geometric mean" of three numbers is not larger than their 'arithmetic mean".

(b) Obtain a corresponding result for four numbers.

53.[C] Prove case 2 of Theorem 16.6.2.

54.[C] Prove case 3 of Theorem 16.6.2.

55.[C] The three dimensions of a box are x, y, and z. The girth plus length are at most 165 inches. If you are free to choose which dimension is the length, which would you choose if you wanted to maximize the volume of the box? Assume $x < y < z$.

56.[C] A surface is called **closed** when it is the boundary of a region R, as a balloon surrounds the air within it. A surface is called **smooth** when it has a continuous outward unit normal vector at each point of the surface. Let S be a smooth closed surface. Show that for any point P_0 in R, there are at least two points on S such that $\overrightarrow{P_0P}$ is normal to S. NOTE: It is conjectured that if P_0 is the centroid of R, then there are at least four points on S such that P_0P is normal to S.

57.[C] Find the point P on the plane $Ax + By + Cz + D = 0$ nearest the point $P_0 = (x_0, y_0, z_0)$, which is not on that plane.

(a) Find P by calculus.

(b) Find P by using the algebra of vectors. (Why is $\overrightarrow{P_0P}$ perpendicular to the plane?)

58.[C] This exercise outlines the proof of Theorem 16.6.2 in the case $f_{xx}(a,b) > 0$ and $f_{xx}(a,b)f_{yy}(a,b) - fxy^2(a,b) > 0$. Assuming that f_{xx}, f_{yy}, and f_{xy} are continuous, we know by the permanence principle that f_{xx} and $f_{xx}f_{yy} - f_{x^2y}$ remain positive throughout some disk R whose center is (a,b). The following steps show that f has a minimum (a,b) on each line L through (a,b). Let $u = \cos(\theta) + \sin(\theta)$ be a unit vector. Show that $D_u(D_uf)$ is positive throughout the part of L that lies in the dark.

(a) Show that $D_uf(a,b) = 0$.

(b) Show that $D_u(D_uf) = f_{xx}\cos^2(\theta) + 2f_{xy}\sin(\theta)\cos(\theta) + f_{yy}\sin^2(\theta)$.

(c) Show that $f_{xx}D_u(D_uf) = (f_{xx}cos(\theta) + fxy\sin(\theta))^2 + (f_{xx}f_{yy} - f_{xy}^2)\sin^2(\theta)$.

(d) Deduce from (b) that f is concave up as the part of each line through (a,b) inside the disk R.

(e) Deduce that f has a relative minimum at (a,b).

59.[C] Let $f(x)$ have period 2π and let

$$S(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k\cos(kx) + \sum_{k=1}^{\infty} b_k\sin(kx)$$

be the series that minimizes the integral

$$\int_{-\pi}^{\pi} (f(x) - S(x))^2 \, dx. \tag{16.6.10}$$

Show that $S(x)$ is the Fourier series associated with $f(x)$. Note: You may assume that in this case you may "differentiate past the integral sign," that is

$$\frac{\partial}{\partial y}\int_a^b g(x,y)\,dx = \int_a^b \frac{\partial g}{\partial y}\,dx.$$

The quantity in (16.6.10) measures the total squared error between $S(x)$ and $f(x)$ over the interval $[-\pi,\pi]$.

60.[C] Prove Case 2 of Theorem 16.6.2.

61.[C] Prove Case 3 of Theorem 16.6.2.

16.7 Lagrange Multipliers

See http://en.wikipedia.org/wiki/Joseph_Louis_Lagrange.

Another method of finding maxima or minima of a function is due to Joseph Louis LaGrange (1736–1813). It makes use of the fact that a gradient of a function is perpendicular to the level curves (or level surfaces) of that function.

The Essence of the Method

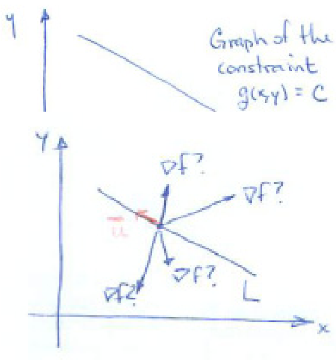


Figure 16.7.2:

We introduce the technique by considering the simplest case. Imagine that you want to find a maxima or a minima of $f(x, y)$ for points (x, y) on the line L that has the equation $g(x, y) = C$. See Figure 16.7.1.

Imagine that $f(x, y)$, for points on L has a maximum or minimum at the point (a, b). Let ∇f be the gradient of f evaluated at (a, b). What can we say about the direction of ∇f? (See Figure 16.7.2

Assume that ∇f is not perpendicular to L. Let $\mathbf{u}$ be a unit vector parallel to L. Then $D_{\mathbf{u}}f = (\nabla f)\cdot\mathbf{u}$ is not 0. If $D_{\mathbf{u}}f$ is positive then $f(x, y)$ is *increasing* in the direction $\mathbf{u}$, which is along L. In the direction $-\mathbf{u}$, $f(x, y)$ is *decreasing*. Therefore the point (a, b) could not provide either a maximum or a minimum of $f(x, y)$ for point (x, y) on L. That means ∇f must be perpendicular to L. But ∇g is perpendicular to L, since $g(x, y) = C$ is a level curve of g. Since ∇f and ∇g are parallel there must be a scalar λ such that

$$\nabla f = \lambda \nabla g \tag{16.7.1}$$

λ, *lambda*, Greek letter L.

The scalar λ is called a **Lagrange multiplier**.

EXAMPLE 1 Find the minimum of x^2y^2 on the line $x + y = 2$.

SOLUTION Since $x^2 + 2y^2$ increases without bound in both directions along the line it must have a minimum somewhere.

Here $f(x, y) = x^2 + 2y^2$ and $g(x, y) = x + y$ so

$$\nabla f = 2x\mathbf{i} + 4y\mathbf{j} \qquad \text{and} \qquad \nabla g = \mathbf{i} + \mathbf{j}$$

At the minimum, the gradients of f and g must be parallel. That is, there is a scalar λ such that

$$\nabla f = \lambda \nabla g,$$

This means

$$2x\mathbf{i} + 4y\mathbf{j} = \lambda(\mathbf{i} + \mathbf{j}). \tag{16.7.2}$$

This single vector equation leads to the 2 equations

$$\begin{cases} 2x = \lambda & \text{equating } \mathbf{i} \text{ components} \\ 4y = \lambda & \text{equating } \mathbf{j} \text{ components} \end{cases} \tag{16.7.3}$$

But we also have the constraint,

$$x + y = 2 \tag{16.7.4}$$

From (16.7.3), $2x = 4y$ or $x = 2y$. Substituting this into (16.7.4) gives $2y+y = 2$ or $y = 2/3$, hence $x = 2y = 4/3$. The minimum is $f\left(\frac{4}{3}, \frac{2}{3}\right) = \left(\frac{4}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \frac{20}{9}$. There is no need to find λ its there just to help us compute. Its task, done, it gracefully departs. ◇

The General Method

Let us see why Lagrange's method works when the constraint not a line, but a curve. Consider this problem:

Maximize or minimize $u = f(x, y)$, given the constraint $g(x, y) = k$.

The graph of $g(x, y) = k$ is in general a curve C, as shown in Figure 16.7.3. Assume that f, *considered only on points of* C, takes a maximum (or minimum) value at the point P_0. Let C be parameterized by the vector function $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$. Let $\mathbf{G}(t_0) = \overrightarrow{OP_0}$. Then u is a function of t:

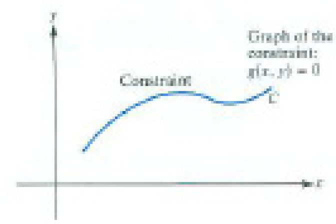


Figure 16.7.3:

$$u = f(x(t), y(t)),$$

and, as shown in the proof of Theorem 16.5 of Section 16.5,

$$\frac{du}{dt} = \nabla f \cdot \mathbf{G}'(t_0). \tag{16.7.5}$$

Since f, considered only on C, has a maximum at $\mathbf{G}(t_0)$,

$$\frac{du}{dt} = 0 \qquad \text{at } t = 0.$$

Thus, by (16.7.5),

$$\nabla f \cdot \mathbf{G}'(t_0) = 0.$$

This means that ∇f is perpendicular to $\mathbf{G}'(t_0)$ at P_0. But ∇g, evaluated at P_0, is also perpendicular to $\mathbf{G}'(t_0)$, since the gradient ∇g is perpendicular to the level curve $g(x, y) = 0$. (We assume that ∇g is not $\mathbf{0}$.) (See Figure 16.7.4.) Thus

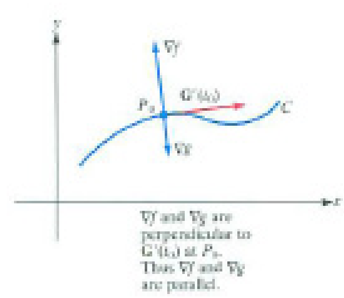


Figure 16.7.4:

∇f is parallel to ∇g.

In other words, there is a scalar λ such that $\nabla f = \lambda \nabla g$.

EXAMPLE 2 Maximize the function x^2y for points (x, y) on the unit circle $x^2 + y^2 = 1$.

SOLUTION We wish to maximize $f(x, y) = x^2y$ for points on the circle $g(x, y) = x^2 + y^2 = 1$. Then

$$\nabla f = \nabla(x^2y) = 2xy\mathbf{i} + x^2\mathbf{j}$$

and

$$\nabla g = \nabla(x^2 + y^2) = 2x\mathbf{i} + 2y\mathbf{j}$$

At an extreme point of f, $\nabla f = \lambda \nabla g$ for some scalar λ. This gives us two scalar equations:

$$2xy = \lambda(2x) \qquad \mathbf{i}\text{ component} \qquad (16.7.6)$$

$$x^2 = \lambda(2y) \qquad \mathbf{j}\text{ component} \qquad (16.7.7)$$

The third equation is the constraint,

$$x^2 + y^2 = 1. \qquad (16.7.8)$$

Since the maximum does not occur when $x = 0$, we may assume x is not 0. Dividing both sides of (16.7.6) by x, we get $2y = 2\lambda$ or $y = \lambda$. Thus (16.7.7) becomes

$$x^2 = 2y^2. \qquad (16.7.9)$$

Combining this with (16.7.8), we have

$$2y^2 + y^2 = 1$$

or

$$y^2 = \frac{1}{3}.$$

Thus

$$y = \frac{\sqrt{3}}{3} \qquad \text{or} \qquad y = -\frac{\sqrt{3}}{3}.$$

By (16.7.9),

$$x = \sqrt{2}y \qquad \text{or} \qquad x = -\sqrt{2}y.$$

There are only four points to be considered on the circle:

$$\left(\frac{\sqrt{6}}{3}, \frac{\sqrt{3}}{3}\right), \left(\frac{-\sqrt{6}}{3}, \frac{\sqrt{3}}{3}\right), \left(\frac{-\sqrt{6}}{3}, \frac{-\sqrt{3}}{3}\right), \left(\frac{\sqrt{6}}{3}, \frac{-\sqrt{3}}{3}\right).$$

At the first and second points x^2y is positive, while at the third and fourth x^2y is negative. The first two points provide the maximum value of x^2y on the circle $x^2 + y^2 = 1$, namely

$$\left(\frac{\sqrt{6}}{3}\right)^2 \frac{\sqrt{3}}{3} = \frac{2\sqrt{3}}{9}.$$

The third and fourth points provide the minimum value of x^2y namely,

$$\frac{-2\sqrt{3}}{9}.$$

◇

More Variables

In the preceding examples we examined the maximum and minimum of $f(x, y)$ on a curve $g(x, y) = k$. But the same method works for dealing with extreme values of $f(x, y, z)$ on a surface $g(x, y, z) = k$. If f has, say, a minimum at (a, b, c), then it does on any level curve on the surface $g(x, y, z) = k$. Thus ∇f is perpendicular to any curve on the surface through P. But so is ∇g. Thus ∇f and ∇g are parallel, and there is a scalar λ such that the $\nabla f = \lambda \nabla g$. So we will have four scalar equations: three from the vector equation $\nabla f = \lambda \nabla g$ and one from the constraint $g(x, y, z) = k$. That gives four equations in four unknowns, x, y, z and λ, but it is not necessary to find λ though it may be useful to determine it. Solving these four simultaneous equations may not be feasible. However, the exercises in this section lead to fairly simple equations that are relatively easy to solve.

EXAMPLE 3 Find the rectangle box with the largest volume, given that its surface area is 96 square feet.

SOLUTION Let the three dimensions be x, y and z and the volume be V, which equals xyz. The surface area is $2xy + 2xz + 2yz$. See Figure 16.7.5.

We wish to maximize $V(x, y, z) = xyz$ subject to the constraint

$$g(x, y, z) = 2xy + 2xz + 2yz = 96. \qquad (16.7.10)$$

Now

$$\nabla V = yz\mathbf{i} + xz\mathbf{j} + xy\mathbf{k}$$

and

$$\nabla g = (2y + 2z)\mathbf{i} + (2x + 2z)\mathbf{j} + (2x + 2y)\mathbf{k}.$$

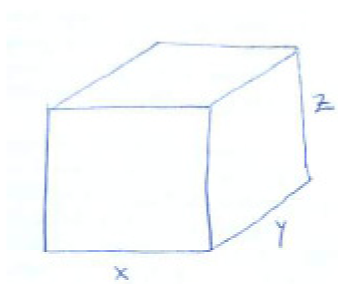


Figure 16.7.5:

The vector equation $\nabla V = \lambda \nabla g$ provides three scalar equations

$$\begin{aligned} yz &= \lambda(2y+2z) \\ xz &= \lambda(2x+2z) \\ xy &= \lambda(2x+2y) \end{aligned}$$

The fourth equation is the constraint,

$$2xy + 2xz + 2yz = 96.$$

Solving for λ in (16.7.11) and in (16.7.11), and equating the results gives

$$\frac{yz}{2y+2z} = \frac{xz}{2x+2z}.$$

Why not?

Since z will not be 0, we have

$$\frac{y}{2y+2z} = \frac{x}{2x+2z}.$$

Clearing denominators gives

$$\begin{aligned} 2xy + 2yz &= 2xy + 2xz \\ 2yz &= 2xz. \end{aligned}$$

Since $z \neq 0$, we reach the conclusion that

$$x = y.$$

Since x, y and z play the same roles in both the volume xyz and in the surface area, $2(xy + xz + yz)$, we conclude also that

$$x = z.$$

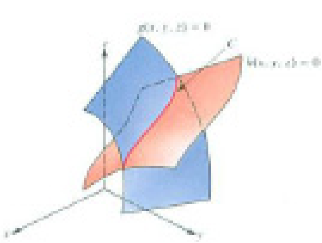
Then $x = y = z$. The box of maximum volume is a cube.

To find its dimensions we return to the constraint, which tells us that $6x^2 = 96$ or $x = 4$. Hence y and z are 4 also. ◇

More Constraints

Lagrange multipliers can also be used to maximize $f(x, y, z)$ subject to more than one constraint; for instance, the constraints may be

$$g(x,y,z) = k_1 \qquad \text{and} \qquad h(x,y,z) = k_2. \tag{16.7.11}$$



The two surfaces (16.7.11) in general meet in a curve C, as shown in Figure 16.7.6. Assume that C is parameterized by the function $\mathbf{G}$. Then at a maximum (or minimum) of f at a point $P_0(x_0, y_0, z_0)$ on C,

$$\nabla f \cdot \mathbf{G}'(t_0) = 0.$$

Thus ∇f, evaluated at P_0, is perpendicular to $\mathbf{G}'(t_0)$. But ∇g and ∇h, being normal vectors at P_0 to the level surfaces $g(x, y, z) = K_1$ and $h(x, y, z) = K_2$, respectively, are both perpendicular to $\mathbf{G}'(t_0)$. Thus

∇f, ∇g, and ∇h are all perpendicular to $\mathbf{G}'(t_0)$ at (x_0, y_0, z_0).

(See Figure 16.7.7.) Consequently, ∇f lies in the plane determined by the vectors ∇g and ∇h (which we assume are not parallel). Hence there are scalars λ and μ such that

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

mu, *mew*, is Greek for the letter M.

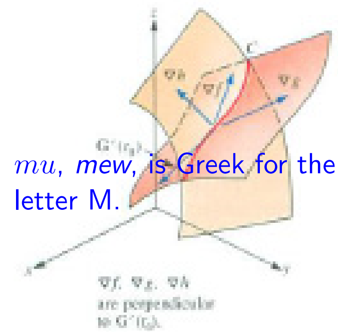


Figure 16.7.7:

This vector equation provides three scalar equations in λ, μ, x, y, z. The two constraints give two more equations. All told: five equations in five unknowns. (Of course we find λ and μ only if they assist the algebra.)

A rigorous development of the material in this section belongs in an advanced calculus course. If a maximum occurs at an endpoint of the curves in question or if the two surfaces do not meet in a curve or if the ∇g and ∇h are parallel, this method does not apply. We will content ourselves by illustrating the method with an example in which there are two constraints.

EXAMPLE 4 Minimize the quantity $x^2 + y^2 + z^2$ subject to the constraints $x + 2y + 3z = 6$ and $x + 3y + 9z = 9$.

SOLUTION There are three variables and two constraints. Each of the two constraints mentioned describes a plane. Thus the two constraints together describe a *line*. The function $x^2 + y^2 + z^2$ is the square of the distance from (x, y, z) to the origin. So the problem can be rephrased as "How far is the origin from a certain line?" (It could be solved by vector algebra. See Exercises 19 and 20.) When viewed this way, the problem certainly has a solution; that is, there is clearly a minimum.

In this case

$$f(x, y, z) = x^2 + y^2 + z^2 \quad (16.7.12)$$
$$g(x, y, z) = x + 2y + 3z \quad (16.7.13)$$
$$h(x, y, z) = x + 3y + 9z. \quad (16.7.14)$$

Thus

$$\nabla f = 2x\mathbf{i} + 2y\mathbf{j} + 2z\mathbf{k} \quad (16.7.15)$$
$$\nabla g = \mathbf{i} + 2\mathbf{j} + 3\mathbf{k} \quad (16.7.16)$$
$$\nabla h = \mathbf{i} + 3\mathbf{j} + 9\mathbf{k}. \quad (16.7.17)$$

There are constants λ and μ so

$$\nabla f = \lambda \nabla g + \mu \nabla h.$$

Therefore, the five equations for x, y, z, λ, and μ are

$$2x = \lambda + \mu \qquad (16.7.18)$$
$$2y = 2\lambda + 3\mu \qquad (16.7.19)$$
$$2z = 3\lambda + 9\mu \qquad (16.7.20)$$
$$x + 2y + 3z = 6 \qquad (16.7.21)$$
$$x + 3y + 9z = 9 \qquad (16.7.22)$$

One way is to use software programs that solve simultaneous linear equations.

There are several ways to solve these equations.

One way is to use the first three of the five equations: to express x, y, and z in terms of λ and μ. Then substitute these values in the last two equations, getting an old friend from high school "two simultaneous equations in two unknowns"

By (16.7.18), (16.7.19), and (16.7.20),

$$x = \frac{\lambda + \mu}{2}, \qquad y = \frac{2\lambda + 3\mu}{2}, \qquad z = \frac{3\lambda + 9\mu}{2}.$$

Equations (16.7.21) and (16.7.22) then become

$$\frac{\lambda + \mu}{2} + \frac{2(2\lambda + 3\mu)}{2} + \frac{3(3\lambda + 9\mu)}{2} = 6$$

and

$$\frac{\lambda + \mu}{2} + \frac{3(2\lambda + 3\mu)}{2} + \frac{9(3\lambda + 9\mu)}{2} = 9,$$

which simplify to

$$14\lambda + 34\mu = 12 \qquad (16.7.23)$$

and

$$34\lambda + 91\mu = 18. \qquad (16.7.24)$$

Solving (16.7.23) and (16.7.24) gives

$$\lambda = \frac{240}{59} \qquad \mu = -\frac{78}{59}.$$

Thus

$$x = \frac{\lambda + \mu}{2} = \frac{81}{59} \approx 1.37288,$$
$$y = \frac{2\lambda + 3\mu}{2} = \frac{123}{59} \approx 2.08475,$$
$$z = \frac{3\lambda + 9\mu}{2} = \frac{9}{59} \approx 0.15254.$$

The minimum of $x^2 + y^2 + z^2$ is this

Since there is no maximum, this must be a minimum. Why?

$$\left(\frac{81}{59}\right)^2 + \left(\frac{123}{59}\right)^2 + \left(\frac{9}{59}\right)^2 = \frac{21,771}{3,481} = \frac{369}{59} \approx 6.24542.$$

◇

In Example 4 there were three variables, x, y, and z, and two constraints. There may, in some cases, be many variables, x_1, x_2, $\ldots x_n$, and many constraints. If there are m constraints, g_1, $g_2 \ldots g_m$ introduce Lagrange multipliers λ_1, λ_2, $\ldots \lambda_m$, one for each constraints. So there would be $m+n$ equations, n from the equation

$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2 + \cdots + \lambda_m \nabla g_m$$

and m more equations from the m constraints. There would be $m+n$ unknowns, λ_1, λ_2, $\ldots$, λ_m, x_1, x_2, $\ldots$, x_n.

Summary

The basic idea of Lagrange multipliers is that if $f(x,y,z)$ (or $f(x,y)$) has an extreme value on a curve that lies on the surface $g(x,y,z) = C$ (or the curve $g(x,y) = k$), then ∇f and ∇g are both perpendicular to the curve at the point where the extreme value occurs. If there is only one constraint, then ∇f and ∇g are parallel. If there are two constraints $g(x,y,z) = k_1$ and $h(x,y,z) = k_2$, then ∇f lies on the plane of ∇g and ∇h. In the first case there is a scalar λ such that $\nabla f = \lambda \nabla g$. In the second case, there are scalars λ and μ such that $\nabla f = \lambda \nabla g + \mu \nabla h$. These vector equations, together with the constraints, provide simultaneous scalars equations, which must then be solved.

EXERCISES for Section 16.7 *Key:* R–routine, M–moderate, C–challenging

In the exercises use Lagrange multipliers unless otherwise suggested.

1.[R] Maximize xy for points on the circle $x^2 + y^2 = 4$.

2.[R] Minimize $x^2 + y^2$ for points on the line $2x + 3y = 6$.

3.[R] Minimize $2x+3y$ on the portion of the hyperbola $xy = 1$ in the first quadrant.

4.[R] Maximize $x + 2y$ on the ellipse $x^2 + y^2 = 8$.

5.[R] Find the largest area of all rectangles whose perimeters are 12 centimeters.

6.[R] A rectangular box is to have a volume of 1 cubic meter. Find its dimensions if its surface area is minimal.

7.[R] Find the point on the plane $x + 2y + 3z = 6$ that is closest to the origin. HINT: Minimize the square of the distance in order to avoid square roots.

8.[R] Maximize $x + y + 2z$ on the sphere $x^2 + y^2z^2 = 9$.

9.[R] Minimize the distance from (x, y, z) to $(1, 3, 2)$ for points on the plane $2x + y + z = 5$.

10.[R] Find the dimensions of the box of largest volume whose surface area is to be 6 square inches.

11.[R] Maximize $x^2y^2z^2$ subject to the constraint $x^2 + y^2 + z^2 = 1$.

12.[R] Find the points on the surface $xyz = 1$ closest to the origin.

13.[R] Minimize $x^2+y^2+z^2$ on the line common to the two planes $x+2y+3z = 0$ and $2x + 3y + z = 4$.

14.[R] The plane $2y + 4z - 5 = 0$ meets the cone $z^2 = 4(x^2 + y^2)$ in a curve. Find the point on this curve nearest the origin.

In Exercises 15 to 18 solve the given exercise in Section 16.5 by Lagrange multipliers.

15.[R] Exercise 25

16.[R] Exercise 26

17.[R] Exercise 29

18.[R] Exercise 30

19.[R] Solve Example 4 by vector algebra.

20.[R] Solve Exercise 13 by vector algebra.

21.[R]

(a) Sketch the elliptical paraboloid $z = x^2 + 2y^2$.

(b) Sketch the plane $x + y + z = 1$.

(c) Sketch the intersection of the surfaces in (a) and (b).

(d) Find the highest point on the intersection in (c).

22.[R]

(a) Sketch the ellipsoid $x^2 + y^2/4 + z^2/9 = 1$ and the point $P(2, 1, 3)$.

(b) Find the point Q on the ellipsoid that is nearest P.

(c) What is the angle between PQ and the tangent plane at Q?

23.[R]

(a) Sketch the hyperboloid $x^2 - y^2/4 - z^4/9 = 1$. (How many sheets does it have?)

(b) Sketch the point $(1, 1, 1)$. (Is it "inside" or "outside" the hyperboloid?)

(c) Find the point on the hyperboloid nearest P.

24.[R] Maximize $x^3 + y^3 + 2z^3$ on the intersection of the surfaces $x^2 + y^2 + z^2 = 4$ and $(x - 3)^2 + y^2 + z^2 = 4$.

25.[R] Show that a triangle in which the product of the sines of the three angles is maximized is equilateral. HINT: Use Lagrange multipliers.

26.[R] Solve Exercise 25 by labeling the angles x,y, and $\pi - x - y$ and minimizing a function of x and y by the method of Section 16.6.

27.[R] Maximize $x + 2y + 3z$ subject to the constraints $x^2 + y^2 + z^2 = 2$ and $x + y + z = 0$.

28.[C]

(a) Maximize $x_1x_2, \cdots x_n$ subject to the constraints that $\sum_{i=1}^{n} x_i = 1$ and all $x_i \geq 0$.

(b) Deduce that for nonnegative numbers a_1, a_2, ..., a_n, $\sqrt[n]{a_1a_2\cdots a_n} \leq (a_1+a_2+\cdots+a_n)/n$. (The **geometric mean** is less than or equal to the **arithmetic mean**.)

29.[C]

(a) Maximize $\sum_{i=1}^{n} x_iy_i$ subject to the constraints $\sum_{i=1}^{n} x_i^2 = 1$ and $\sum_{i=1}^{n} y_i^2 = 1$.

(b) Deduce that for any numbers a_1, a_2, ..., a_n and $b_1, b_2, \dots b_n$, $\sum_{i=1}^{n} a_ib_i \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2}\left(\sum_{i=1}^{n} b_i^2\right)^{1/2}$, which is called the Schwarz inequality. HINT: Let $x_i = \frac{a_i}{(\sum_{i=1}^{n} a_i^2)^{1/2}}$ and $y_1 = \frac{b_i}{(\sum_{i=1}^{n} b_i^2)^{1/2}}$.

(c) How would you justify the inequality in (b), for $n = 3$, by vectors?

30.[C] Let $a_1, a_2 \dots a_n$ be fixed nonzero numbers. Maximize $\sum_{i=1}^{n} a_ix_i$ subject to $\sum_{i=1}^{n} x_i^2 = 1$.

31.[C] Let p and q be positive numbers that satisfy the equation $1/p + 1/q = 1$. Obtain Holder's inequality for nonnegative numbers a_i and b_i,

$$\sum_{i=1}^{n} a_ib_i \leq \left(\sum_{i=1}^{n} a_i^p\right)^{1/p}\left(\sum_{i=1}^{n} b_i^q\right)^{1/q},$$

as follows.

(a) Maximize $\sum_{i=1}^{n} x_iy_i$ subject to $\sum_{i=1}^{n} x_i^p = 1$ and $\sum_{i=1}^{n} y_i^q = 1$.

(b) By letting $x_i = \frac{a_i}{(\sum_{i=1}^{n} a_i^p)^{1/p}}$ and $y_i = \frac{b_i}{(\sum_{i=1}^{n} b_i^q)^{1/q}}$, obtain Holder's inequality.

Note that Holder's inequality, with $p = 2$ and $q = 2$, reduces to the Schwarz inequality in Exercise 29.

32.[C] A consumer has a budget of B dollars and my purchase n different items. The price of the ith item is p_1 dollars. When the consumer buys x_1 units of the ith item, the total cost is $\sum_{i=1}^{n} p_ix_i$. Assume that $\sum_{i=1}^{n} p_ix_i = B$ and that the consumer wishes to maximize her utility $u(x_1, x_2 \dots x_n)$.

(a) Show that when $x_1, \ldots, x_n$, are chosen to maximize utility, then

$$\frac{\partial u/\partial x_i}{p_i} = \frac{\partial u/\partial x_j}{p_j}.$$

(b) Explain the result in (a) using just economic intuition. HINT: Consider a slight change in x_i and x_j, with the other x_k's held fixed.

33.[C] The following is quoted from Colin W. Clark in *Mathematical Bioeconomics*, Wiley, New York, 1976:

> [S]uppose there are N fishing grounds. Let $H^i = H^i(R^i, E^i)$ denotes the production function for the total harvest H^i on the ith ground as a function of the recruited stock level R^i and effort E^i on the ith ground. The problem is to determine the least total cost $\sum_{i=1}^{N} c_i E^i$ at which a given total harvest $H = \sum_{i=1}^{n} H^i$ can be achieved. This problem can be easily solved by Lagrange multipliers. The result is simply
>
> $$\frac{1}{c_i}\frac{\partial H^i}{\partial E^i} = \text{constant}$$
>
> [independent of i].

Verify his assertion. The c_i's are constants. The superscripts name the functions; they are not exponents.

34.[C] (*Computer science*) This exercise is based on J. D. Ullman, *Principles of Database Systems*, pp. 82–83, Computer Science Press, Potomac, Md., 1980. It arises in the design of efficient "bucket" sorts. (A *bucket sort* is a particular way of rearranging information in a database.) Let p_1, p_2, …, p_k and B be positive constants. Let b_1, b_2, …, b_k be k nonnegative variables satisfying $\sum_{j=1}^{k} b_j = B$. The quantity $\sum_{j=1}^{k} p_j \cdot 2^{B-b_j}$ represents the expected search time. What values of $b_1, b_2, \ldots, b_k$ does the method of Lagrange multipliers suggest provide the minimum expected search time?

35.[C] Assume that $f(x, y, z)$ has an extreme value at P_0 on the level surface $g(x, y, z) = k$

(a) Why is ∇g evaluated at P_0 perpendicular to the surface at P_0?

(b) Why is ∇f evaluated at P_0 perpendicular to the surface at P_0?

36.[C] Solve Example 35 by vector algebra (or just algebra).

16.8 What Everyone Who Will Study Thermodynamics Needs to Know

Review the Chain Rule, if necessary.

The basic equations of thermodynamics follow from the Chain Rule and the equality of the mixed partial derivatives. We will describe the mathematics within the thermodynamics context.

Implications of The Chain Rule

We start with a function of three variables, $f(x, y, z)$, which we assume has first partial derivatives

$$\left.\frac{\partial f}{\partial x}\right|_{y,z} \qquad \left.\frac{\partial f}{\partial y}\right|_{x,z} \qquad \left.\frac{\partial f}{\partial z}\right|_{x,y}.$$

This notation is standard practice in thermodynamics, though it offends some mathematicians.

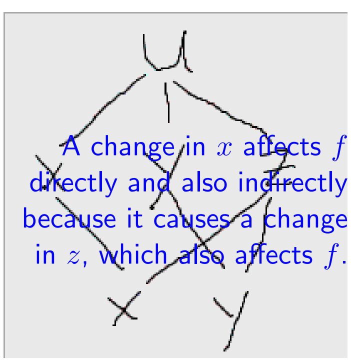
The subscripts denote the variables held fixed.

Without this explicit reminder it is necessary to remember the other variables. At this point this is not difficult. But, when additional information is included, it can become more difficult to keep track of all of the variables in the problem.

Now assume that z is a function of x and y, $z = g(x, y)$. Then $f(x, y, z) = f(x, y, g(x, y))$ is a function of only two variables. This new function we name $h(x, y)$: $h(x, y) = f(x, y, g(x, y))$. There are only two first partial derivatives of h:

$$\left.\frac{\partial h}{\partial x}\right|_{y} \quad \text{and} \quad \left.\frac{\partial h}{\partial y}\right|_{x}.$$

Let the value of $f(x, y, z)$ be called u, $u = f(x, y, z)$. But x, y, and z are functions of x and y: $x = x$, $y = y$, and $z = g(x, y)$.



A change in x affects f directly and also indirectly because it causes a change in z, which also affects f.

Figure 16.8.1:

Figure 16.8.1 provides a pictorial view of the relationship between the different variables. Both x and y appear as middle and independent variables. We have $u = f(x, y, z)$ and also $u = h(x, y)$. By the Chain Rule Then

$$\left.\frac{\partial h}{\partial x}\right|_{y} = \left.\frac{\partial f}{\partial x}\right|_{y,z}\left.\frac{\partial x}{\partial x}\right|_{y} + \left.\frac{\partial f}{\partial y}\right|_{x,z}\left.\frac{\partial y}{\partial x}\right|_{y} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_{y}.$$

Since x and y are independent variables, $\partial x/\partial x = 1$ and $\partial y/\partial x = 0$ and we have

$$\left.\frac{\partial h}{\partial x}\right|_{y} = \left.\frac{\partial f}{\partial x}\right|_{y,z} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_{y}, \tag{16.8.1}$$

or simply

$$\frac{\partial h}{\partial x} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial x}. \tag{16.8.2}$$

When the subscripts are omitted we have to look back at the definitions of f, g, and h to see which variables are held fixed.

EXAMPLE 1 Let's check (16.8.2) when

$$f(x,y,z) = x^2y^3z^5 \quad \text{and} \quad g(x,y) = 2x+3y.$$

SOLUTION We have $h(x,y) = f(x,y,g(x,y)) = x^2y^3(2x+3y)^5$. Then $\frac{\partial f}{\partial x} = 2xy^3z^5$ and $\frac{\partial f}{\partial z} = 5x^2y^3z^4$. Also $\frac{\partial g}{\partial x} = 2$.

Computing $\partial h/\partial x$ directly gives

$$\begin{aligned}
\frac{\partial h}{\partial x} &= \frac{\partial}{\partial x}\left(x^2y^3(2x+3y)^5\right) \\
&= y^3\frac{\partial}{\partial x}\left(x^2(2x+3y)^5\right) \\
&= y^3\left(2x(2x+3y)^5 + x^2\left(5(2x+3y)^4(2)\right)\right) \\
&= 2xy^3(2x+3y)^5 + 10x^2y^3(2x+3y)^4. \qquad (16.8.3)
\end{aligned}$$

On the other hand, by (16.8.2), we have

$$\begin{aligned}
\frac{\partial h}{\partial x} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial z}\frac{\partial g}{\partial x} \\
&= 2xy^3z^5 + (5x^2y^3z^4)(2) \\
&= 2xy^3(2x+3y)^5 + 10x^2y^3(2x+3y)^4,
\end{aligned}$$

which agrees with (16.8.3). ◇

What If $z = g(x,y)$ Makes $f(x,y,z)$ Constant?

Next, assume that when z is replaced by $g(x,y)$, the function $h(x,y) = f(x,y,g(x,y))$ is constant: $h(x,y) = f(x,y,g(x,y)) = C$. This happens when we use the equation $f(x,y,z) = C$ to determine z implicitly as a function of x and y.

Then

$$\left.\frac{\partial h}{\partial x}\right|_y = 0 \quad \text{and} \quad \left.\frac{\partial h}{\partial y}\right|_x = 0.$$

In this case, which occurs frequently in thermodynamics, (16.8.1) becomes

(16.8.4) will be the foundation for deriving (16.8.9) and (16.8.10), key mathematical relationships used in thermodynamics.

$$0 = \left.\frac{\partial f}{\partial x}\right|_{y,z} + \left.\frac{\partial f}{\partial z}\right|_{x,y}\left.\frac{\partial g}{\partial x}\right|_y. \qquad (16.8.4)$$

Solving (16.8.4) for $\left.\frac{\partial g}{\partial x}\right|_y$ we obtain

$$\left.\frac{\partial g}{\partial x}\right|_y = \frac{-\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}}. \tag{16.8.5}$$

Equation (16.8.5) expresses the partial derivative of $g(x,y)$ with respect to x in terms of the partial derivatives of the original function $f(x,y,z)$.

EXAMPLE 2 Let $f(x,y,z) = x^3y^5z^7$. Define $g(x,y)$ implicitly by $x^3y^5(g(x,y))^7 = 1$. That is, $g(x,y) = x^{-3/7}y^{-5/7}$. Verify (16.8.5).

SOLUTION First of all, $\left.\frac{\partial g}{\partial x}\right|_y = \frac{-3}{7}x^{-10/7}y^{-5/7}$. Then

$$\left.\frac{\partial f}{\partial x}\right|_{y,z} = 3x^2y^5z^7 \quad \text{and} \quad \left.\frac{\partial f}{\partial z}\right|_{x,x} = 7x^3y^5z^6.$$

Substituting in (16.8.5), we have

$$\begin{aligned}
\frac{-\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}} &= \frac{-\left(3x^2y^5z^7\right)}{7x^3y^5z^6} \\
&= -\frac{3}{7}x^{-1}z \\
&= -\frac{3}{7}x^{-1}x^{-3/7}y^{-5/7} && \text{because } x^3y^5z^7 = 1 \\
&= -\frac{3}{7}x^{-10/7}y^{-5/7} \\
&= \left.\frac{\partial g}{\partial x}\right|_y && \text{so (16.8.5) is satisfied.}
\end{aligned}$$

◇

The Reciprocity Relations

In a thermodynamics text you will see equations of the form

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{1}{\left.\frac{\partial z}{\partial x}\right|_y}. \tag{16.8.6}$$

We will explain where this equation comes from, presenting the mathematical details often glossed over in the applied setting. There is a function $f(x,y,z)$ with constant value C, $f(x,y,z) = C$. It is assumed that this equation determines z as a function of x and y, or, similarly, determines x as a function of y and z, or y as a function of x and z. There are six first partial derivatives:

See Exercise 5.

$$\left.\frac{\partial z}{\partial x}\right|_y, \quad \left.\frac{\partial z}{\partial y}\right|_x, \quad \left.\frac{\partial x}{\partial y}\right|_z, \quad \left.\frac{\partial x}{\partial z}\right|_y, \quad \left.\frac{\partial y}{\partial x}\right|_z, \quad \left.\frac{\partial y}{\partial z}\right|_x. \tag{16.8.7}$$

An equation analogous to (16.8.5) holds for each of them. For instance,

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}}. \tag{16.8.8}$$

Combining (16.8.5) and (16.8.8) verifies that

This is to be expected, for $\frac{\Delta z}{\Delta x}$ is the reciprocal of $\frac{\Delta x}{\Delta z}$.

$$\left.\frac{\partial x}{\partial z}\right|_y = \frac{1}{\left.\frac{\partial z}{\partial x}\right|_y}. \tag{16.8.9}$$

Equation (16.8.9) is an example of a **reciprocity relation**: The partial derivative of one variable with respect to a second variable is the reciprocal of the partial derivative of the second variable with respect to the first variable.

EXAMPLE 3 Let $f(x,y,z) = 2x + 3y + 5z = 12$. Verify that $\partial z/\partial x$ is the reciprocal of $\partial x/\partial z$.

SOLUTION Since $2x + 3y + 5z = 12$, $z = (12 - 2x - 3y)/5$. Then $\partial z/\partial x = -2/5$.

Also, $x = (12 - 3y - 5z)/2$, so $\partial x/\partial z = -5/2$, which is, as predicted, the reciprocal of $\partial z/\partial x$. ◇

The Cyclic Relations

With the aid of equations like (16.8.8) it is easy to establish the surprising relation

The **Cyclic Relation**, also known as the Triple Product Rule, the Cyclic Chain Rule, or Euler's Chain Rule. See `http://en.wikipedia.org/wiki/Triple_`

$$\left.\frac{\partial x}{\partial y}\right|_z \left.\frac{\partial y}{\partial z}\right|_x \left.\frac{\partial z}{\partial x}\right|_y = -1. \qquad (16.8.10)$$

Equation (16.8.10) results from the use of three versions of (16.8.8). The left-hand side of (16.8.10) can be expressed as

$$\left(\frac{-\left.\frac{\partial f}{\partial y}\right|_{x,z}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}} \right) \left(\frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial y}\right|_{x,z}} \right) \left(\frac{-\left.\frac{\partial f}{\partial z}\right|_{x,y}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}} \right) \qquad (16.8.11)$$

Cancellation reduces (16.8.11) to -1.

EXAMPLE 4 Let $f(x, y, z) = 2x+3y+5z = 12$. This equation determines implicitly each of the variables in terms of the two others. Verify (16.8.10) in this case.

SOLUTION By the equation $2x + 3y + 5z = 12$,

$$x = \frac{12 - 3y - 5z}{2} \qquad y = \frac{12 - 2x - 5z}{3} \qquad z = \frac{12 - 2x - 3y}{5}$$

Then $\partial x/\partial y = -3/2$, $\partial y/\partial z = -5/3$, and $\partial z/\partial x = -2/5$, and we have

$$\frac{\partial x}{\partial y}\frac{\partial y}{\partial z}\frac{\partial z}{\partial x} = \left(\frac{-3}{2}\right)\left(\frac{-5}{3}\right)\left(\frac{-2}{5}\right) = -1$$

◇

If two of the three partial derivatives in (16.8.10) are easy to calculate, then we can use (16.8.10) to find the third, which may otherwise be hard to calculate. We illustrate this use of the cyclic relationship with an example from thermodynamics. In this context T denotes temperature, p, pressure, and v the mass per unit volume.

v is the reciprocal of density

Equations (16.8.4, (16.8.9), and (16.8.10) are the 15sential mathematical relationships used in thermodynamics. We now show their use in a few typical thermodynamics problems.

EXAMPLE 5 In van der Waal's equation p, T, and v are all related by the relation

van der Waal's equation is only one example of an **equation of state**. See also Exercises 11 and 12.

$$p = \frac{RT}{v - b} - \frac{a}{v^2}; \qquad (16.8.12)$$

R, a and b are constants. Use a cyclic relation to find $(\partial v/\partial T)_p$.

SOLUTION We use the cyclic relation

Exercises 13 and 14 describe other ways to solve Example 5.

$$\left.\frac{\partial v}{\partial T}\right|_p \left.\frac{\partial T}{\partial p}\right|_v \left.\frac{\partial p}{\partial v}\right|_T = -1. \tag{16.8.13}$$

Looking at (16.8.12), we see that $(\partial p/\partial T)_v$ is easier to calculate than $(\partial T/\partial p)_v$. So (16.8.13) becomes

$$\frac{\left.\frac{\partial v}{\partial T}\right|_p \left.\frac{\partial p}{\partial v}\right|_T}{\left.\frac{\partial p}{\partial T}\right|_v} = -1$$

and therefore

$$\left.\frac{\partial v}{\partial T}\right|_p = -\frac{\left.\frac{\partial p}{\partial T}\right|_v}{\left.\frac{\partial p}{\partial v}\right|_T}. \tag{16.8.14}$$

Since p is given as a function of v and T, it is easy to calculate the numerator and denominator in (16.8.14):

$$\left(\frac{\partial p}{\partial T}\right)_v = \frac{R}{v-b} \quad \text{and} \quad \left(\frac{\partial p}{\partial v}\right)_T = \frac{-RT}{(v-b)^2} + \frac{2a}{v^3}.$$

Thus, by (16.8.14),

$$\left(\frac{\partial v}{\partial T}\right)_p = \frac{-R/(v-b)}{-RT/(v-b)^2 + 2a/v^3}.$$

◇

Using the Equality of the Mixed Partial Derivatives

Having shown how the Chain Rule provides some of the basic equations in thermodynamics, let us show how the equality of the mixed partials leads to other basic equations.

We resume our consideration of a thermodynamic process in which the pressure is denoted by P, the temperature by T, and the volume per unit mass by v. Other common variables are

u	thermal energy per unit mass
s	entropy per unit mass
a	Helmholtz free energy per unit mass
g	Gibbs free energy per unit mass
h	enthalpy per unit mass

That is a total of 8 variables of interest. If they were independent, the possible states would be part of an eight-dimensional space. However, they are very *inter*dependent. In fact *any two* determine all the others.

For instance, u may be viewed as a function of s and v, and we have $\left.\frac{\partial u}{\partial s}\right|_v$, which is the *definition* of temperature, T. Thermodynamic texts either state or derive the "Gibbs relation"

When you look at your thermometer, remember that you are gazing at the value of a partial derivative.

$$du = T\ ds - P\ dv.$$

This equation involving differentials tells us that u is viewed as a function of s and v, and that

$$\left.\frac{\partial u}{\partial s}\right|_v = T \quad \text{and} \quad \left.\frac{\partial u}{\partial v}\right|_s = -P.$$

Equating the mixed second partial derivatives then gives us

$$\begin{aligned}
\frac{\partial^2 u}{\partial v \partial s} &= \frac{\partial^2 u}{\partial s \partial v} && \text{equality of mixed partials of } u(s,v) \\
\frac{\partial}{\partial v}\left(\tfrac{\partial u}{\partial s}\right) &= \frac{\partial}{\partial s}\left(\tfrac{\partial u}{\partial v}\right) && \\
\left.\frac{\partial T}{\partial v}\right|_s &= \left.\frac{\partial(-P)}{\partial s}\right|_v && \text{because } \left.\frac{\partial u}{\partial s}\right|_v = T \text{ and } \left.\frac{\partial u}{\partial v}\right|_s = -P \\
\left.\frac{\partial T}{\partial v}\right|_s &= -\left.\frac{\partial P}{\partial s}\right|_v.
\end{aligned}$$

Several thermodynamic statements that equate two partial derivatives are obtained this way. The starting point is an equation of the form

In other contexts we will say that $dz = M dx + N dy$ is an **exact differential**.

$$dz = M\ dx + N\ dy$$

where M is $\left.\frac{\partial z}{\partial x}\right|_y$ and N is $\left.\frac{\partial z}{\partial y}\right|_x$. Then, because

$$\frac{\partial z}{\partial x \partial y} = \frac{\partial z}{\partial y \partial x},$$

it is found that

$$\left.\frac{\partial M}{\partial y}\right|_x = \left.\frac{\partial N}{\partial x}\right|_y.$$

Summary

We showed how the Chain Rule in the special case where an intermediate variable is also a final variable justifies certain identities, namely, the *reciprocal* and *cyclic relations* used in thermodynamics. Then we showed how the equality of the mixed partial derivatives is used to derive other equations linking various partial derivatives.

EXERCISES for Section 16.8 *Key:* R–routine, M–moderate, C–challenging

1.[R] Let $u = x^2 + y^2 + z^2$ and let $z = x + y$.

(a) The symbol $\frac{\partial u}{\partial x}$ has two interpretations. What are they?

(b) Evaluate $\frac{\partial u}{\partial x}$ in both cases identified in (a).

2.[R] Let $z = rst$ and let $r = st$.

(a) The symbol $\frac{\partial z}{\partial t}$ has two interpretations. What are they?

(b) Evaluate $\frac{\partial z}{\partial t}$ in both cases identified in (a).

3.[R] Let $u = f(x, y, z)$ and $z = g(x, y)$. Then u is indirectly a function of x and of y. Express $\left.\frac{\partial u}{\partial x}\right|_y$ in terms of partial derivatives of f. (Supply all the steps.)

4.[R] Assume that the equation $f(x, y, z) = C$, a constant, determines x as a function of y and z: $x = h(y, z)$. Express $\left.\frac{\partial x}{\partial y}\right|_z$ in terms of partial derivatives of f. (Supply all the steps.)

5.[R] What is the product of the six partial derivatives in (16.8.7)?

6.[R] Using the function f from Example 2, verify the analog of (16.8.8) for $\left.\frac{\partial z}{\partial y}\right|_x$.

7.[R] Let $f(x, y, z) = 2x + 4y + 3z$. The equation $f(x, y, z) = 7$ determines any variable as a funciton of the other two. Verify (16.8.8), where z is viewed as a function of x and y.

8.[R] Obtain the cyclic relation

$$\left.\frac{\partial x}{\partial z}\right|_y \left.\frac{\partial z}{\partial y}\right|_x \left.\frac{\partial y}{\partial x}\right|_z = -1.$$

HINT: Duplicate the steps leading to (16.8.10).

9.[R] Verify (16.8.10) in the case $f(x, y, z) = x^3y^5z^7 = 1$.

10.[R] Verify (16.8.10) in the case $f(x, y, z) = 2x + 4y + 3z = 7$.

11.[R] The equation of state for an ideal gas is $pv = RT$. Find $(\partial v/\partial T)_p$.

12.[R] The Redlich-Kwang equation

$$p = \frac{RT}{v - b} - \frac{a}{v(v + b)T^{1/2}}.$$

is an improvement upon the van der Waal's equation of state (16.8.12) for gases and liquids. Find $(\partial v/\partial T)_p$. NOTE: Do a Google search for "Redlich Kwang equation", or visit `http://en.wikipedia.org/wiki/Equation_of_state`.

13.[R] Find $(\partial v/\partial T)_p$ in Example 5 by differentiating both sides of (16.8.12) with respect to T, holding p constant.

14.[R] One might try to find $(\partial v/\partial T)_p$ in Example 5 by first finding an equation that expresses v in terms of T and p. What unpleasantness happens when you try this approach?

15.[R] In Example 5, find $(\partial v/\partial p)_T$, $(\partial T/\partial v)_p$, and $(\partial T/\partial p)_v$.

16.[M] In thermodynamics there is the Gibbs relation

$$dh = T\ ds + v\ dP.$$

It is understood that $\left.\frac{\partial h}{\partial s}\right|_p = T$ and $\left.\frac{\partial h}{\partial p}\right|_s = v$. Deduce that $\left.\frac{\partial T}{\partial P}\right|_s = \left.\frac{\partial v}{\partial s}\right|_P$.

17.[R] Consider the thermodynamic equation

$$\left.\frac{\partial E}{\partial T}\right|_v = \left.\frac{\partial E}{\partial T}\right|_P + \left.\frac{\partial E}{\partial P}\right|_T \left.\frac{\partial P}{\partial T}\right|_v. \tag{16.8.15}$$

(a) What is the dependent variable?

(b) What are the independent variables?

(c) What are the intermediate variables?

(d) Draw a diagram showing all the paths from the dependent variables to the independent variables.

(e) Use the Chain Rule to complete the derivation of (16.8.15).

18.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v = \dfrac{-\left.\frac{\partial v}{\partial T}\right|_P}{\left.\frac{\partial v}{\partial P}\right|_T}$.

19.[M] Show that

(a) $\left.\frac{\partial E}{\partial v}\right|_P = \left.\frac{\partial E}{\partial T}\right|_P \left.\frac{\partial T}{\partial v}\right|_P$

(b) $\left.\frac{\partial E}{\partial P}\right|_v = \left.\frac{\partial E}{\partial T}\right|_P \left.\frac{\partial T}{\partial P}\right|_v + \left.\frac{\partial E}{\partial P}\right|_T$.

20.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v \left.\frac{\partial T}{\partial P}\right|_v = 1$. HINT: Express each of the partial derivaties as a quotient of partial derivatives, as in Exercise 18.

21.[M] Show that $\left.\frac{\partial P}{\partial T}\right|_v \left.\frac{\partial T}{\partial v}\right|_P \left.\frac{\partial v}{\partial P}\right|_T = -1$.

22.[M] Let $u = F(x, y, z)$ and $z = f(x, y)$. Thus u is a (composite) function of x and y: $u = G(x, y) = F(x, y, f(x, y))$. Assume that $G(x, y) = x^2y$. Obtain a formula for $\frac{\partial f}{\partial x}$ in terms of $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, and $\frac{\partial F}{\partial z}$. (All three need not appear in your answer.)

23.[M] Let $u = F(x, y, z)$ and $x = f(y, z)$. Thus u is a (composite) function of y and z: $u = G(y, z) = F(f(y, z), y, z)$. Assume that $G(y, z) = 2y + z^2$. Obtain a formula for $\frac{\partial f}{\partial z}$ in terms of $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, and $\frac{\partial F}{\partial z}$. (All three need not appear in your answer.)

24.[C] Two functions u and v of the variables x and y are defined implicitly by the two simultaneous equations

$$F(u, v, x, y) = 0 \quad \text{and} \quad G(u, v, x, y) = 0.$$

Assuming all necessary differentiability, find a formula for $\frac{\partial u}{\partial x}$ in terms of the partial derivatives of F and of G.

16.S Chapter Summary

This chapter extends to functions of two or more variables the notions of rate of change and derivative originally in Chapter 3. For a function of several variables a "partial derivative" is simply the derivative with respect to one of the variables, when all the other variables are held constant.

The precise definition rests on a limit. For instance, the partial derivative with respect to x of $f(x, y)$ at (a, b) is

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

Just as there are higher-order derivatives, there are higher-order partial derivatives, for instance:

$$\frac{\partial^2 f}{\partial x \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \text{ and } \frac{\partial^2 f}{\partial y \partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

For functions usually encountered in applications, the two "mixed partialas," $\partial^2 f/\partial x \partial y$ and $\partial^2 f/\partial y \partial x$, are equal; we can therefore not worry about the order of the differentiation.

Also, for common functions "differentiation under the integral sign" is legal:

$$\text{if } g(y) = \int_a^b f(x,y)\, dx, \text{ then } \frac{dg}{dy} = \int_a^b \frac{\partial f}{\partial y}(x,y)\, dx.$$

For a function of one variable, $f(x)$, with a continuous derivative,

$$\Delta f = f(a+\Delta x) - f(a) = f'(c)\Delta x = (f'(a)+\epsilon)\Delta x = f'(a)\Delta x + \epsilon \Delta x. \quad (16.S.1)$$

Here c is in $[a, a + \Delta x]$ and $\epsilon \to 0$ as $\Delta x \to 0$. The analog of (16.S.1) for a function of two or more variables is the basis for the chain rule for functions of several variables:

$$\Delta f = f(a+\Delta x, b+\Delta y) - f(a,b) = (f(a+\Delta x, b+\Delta y) - f(a, b+\Delta y)) + (f(a, b+\Delta y) \quad (16.S.2)$$

See the CIE section on Maxwell's equations at the end of Chapter 18.

where ϵ_1 and $\epsilon_2 \to 0$ as $Deltax$ and $\Delta y \to 0$.

The chain rule showed then, if $g(u)$ and $h(u)$ are differentiable functions, then $y = g(x + kt) + h(x - kt)$, k constant, satisfies the partial differential equation (PDE) $\partial^2 y/\partial t^2 = k^2 \partial^2 y/\partial x^2$. This PDE was the key to Maxwell's conjecture that light is an electro-magnetic phenomenon.

The gradient, a vector function, was defined in terms of partial derivatives: $\nabla f = \langle f_x, f_y \rangle$ or, for a function of three variables: $\nabla f = \langle f_x, f_y, f_z \rangle$. The gradient points in the direction a function increases most rapidly. The rate at which $f(x,y)$ changes in the direction of a unit vector $\mathbf{u}$ is $\nabla f \cdot \mathbf{u}$. The gradient is perpendicular to the level curve (or level surface) passing through a given point. At a critical point the gradient vanishes.

For a function of one variable the sign of the second derivative helps tell whether a critical point is a maximum or a minimum. For a function of two variables, the test also involves all three second derivatives. In particular, the signs of f_{xx} and $f_{xx}f_{yy} - (f_{xy})^2$ are important.

The number λ is called a Lagrange multiplier.

Maximizing a function f subject to a constraint g depends on the observation that at an extremum ∇f is parallel to ∇g. Hence there is a number λ such that $\nabla f = \lambda \nabla g$.

The final section showed that the chain rule is the bases of two facts in thermodynamics. It also shows how to apply the chain rule when a middle variable is also a final variable.

EXERCISES for 16.S *Key:* R–routine, M–moderate, C–challenging

1.[R] Let $f(x,y) = x^2 - y^2$ and $g(x,y) = 2xy$. Show that

(a) $\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y}$

(b) $\frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}$

(c) $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$

(d) $\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 0$

2.[R] Repeat Exercise 1 for $f(x,y) = \ln\left(\sqrt{x^2+y^2}\right)$ and $g(x,y) = \arctan(y/x)$.

3.[M] Let f and g be functions of x and y that have continuous second derivatives. Assume the first partial derivatives of f and g satisfy:

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial y} \quad \text{and} \quad \frac{\partial f}{\partial y} = -\frac{\partial g}{\partial x}. \tag{16.S.3}$$

Show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0 \quad \text{and} \quad \frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} = 0. \tag{16.S.4}$$

NOTE: The two equations in (16.S.3) are known as the Cauchy–Riemann equations. A pair of functions that satisfy (16.S.4) are called a **conformal pair** of functions.

In Exercises 4 to 12 assume the functions have continuous partial derivatives throughout the xy plane.

4.[R] If $f_x(x,y) = 0$ for all points (x,y) in the plane, must f be constant? If not, describe f.

5.[R] If $f_x(x,y) = 0$ and $f_y(x,y) = 0$ for all points (x,y) in the plane, must f be constant? If not, describe f.

6.[R] The function $3x+g(y)$, for any differentiable function $g(y)$ satisfies the partial differential equation $\partial f/\partial x = 3$. Are there any other solutions to that equation? Explain your answer.

7.[R] Find all functions f such that $\partial f/\partial x = 3$ and also $\partial y/\partial x = 3$ are satisfied.

8.[R] Show that there is no function f such that $\partial f/\partial x = 3y$ and $\partial f/\partial y = 4x$.

9.[R] Find all functions such that $f_{xx}(x,y) = 0$.

10.[R] Find all functions such that $f_{xx}(x,y) = 0$ and $f_{yy}(x,y) = 0$.

11.[R] Find all functions such that $f_{xy}(x,y) = 0$.

12.[R] Find all functions such that $f_{xy}(x,y) = 1$.

13.[M] A hiker is at the origin on a hill whose equation is $z = x$. If he walks south, along the positive x-axis the slope of his path would be steep, 1, with angle $\pi/4$. If he walked along the y-axis, the slope would be 0.

(a) If he walked NE what would the slope of his path be?

(b) In what direction should he walk in order that his path would have a slope of 0.2?

14.[C] This exercise outlines a proof that the two mixed partials of $f(x,y)$ are generally equal. It suffices to show that $f_{xy}(0,0) = f_{yx}(0,0)$. We assume that all the first and second partial derivatives are continuous in some disk with center $(0,0)$.

(a) Why is $f_{xy}(0,0)$ equal to

$$\lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k}? \tag{16.S.5}$$

(b) Why is (16.S.5) equal to

$$\lim_{k\to 0}\left(\lim_{h\to 0} \frac{(f(h,k) - f(0,k)) - (f(h,0) - f(0,0))}{hk}\right)? \tag{16.S.6}$$

(c) Let $u(y) = f(h, y) - f(0, y)$. Show that the fraction in (16.S.6) equals

$$\frac{u(k) - u(0)}{hk},$$

and this fraction equals $u'(k)/h$ for some k between 0 and k.

(d) Why is $u'(k) = f_y(h, k) - f_y(0, k)$?

(e) Why is $u'(k)/h$ equal to $(f_y)_x(H, K)$ for some H between 0 and h?

(f) Deduce that $f_{xy}(0, 0) = f_{yx}(0, 0)$.

(g) Did this derivation use the continuity of f_{yx}? If so, how?

(h) Did this derivation use the continuity of f_{xy}? If so, how?

(i) Did we need to assume f_{xy} exists? If so, where was this assumption used?

(j) Did we need to assume f_{yx} exists? If so, where was this assumption used?

15.[C] The assertion that it is safe to "differentiate across the integral sign," amounts to the statement that two definite integrals are equal. To illustrate this, translate the assertion into the language of limits:

$$\frac{d}{dt}\int_a^b f(x, t)\ dx = \int_a^b \frac{\partial}{\partial t} f(x, t)\ dx. \tag{16.S.7}$$

(a) Why is the derivative on the left an ordinary derivative, $d()/dt$, but the derivative on the right is a partial derivative?

(b) Using the definitions of ordinary derivatives and partial derivatives as limits, show what (16.S.7) says about limits.

(c) Verify (16.S.7) for $f(x, t) = x^7 t^4$.

(d) Verfiy (16.S.7) for $f(x, t) = \cos(xt)$.

Exercise 16 provides another motivation for the definition of the Fourier series of a function f defined on the interval $[0, 2\pi]$.

16.[C] For a particular integer n consider all functions $S(x)$ of the form

$$S(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos(kx) + b_k \sin(kx)\right),$$

Let $f(x)$ be a continuous function defined on $[0, 2\pi]$. The definite integral

$$\int_0^{2\pi} (f(x) - S(x))^2 \ dx$$

is a measure of how close $S(x)$ is to $f(x)$ on the interval $[0, 2\pi]$. The integral can never be negative. (Why?) The smaller the integral, the better S approximates f on $[0, 2\pi]$. Show that the $S(x)$ that minimizes the integral is precisely a front-end of the Fourier series associated with $f(x)$.

Calculus is Everywhere # 21

The Wave in a Rope

We will develop what may be the most famous partial differential equation. In the CIE of the next chapter we will solve that equation and, then, use it in the final chapter to show how it helped Maxwell discover that light is an electrical-magnetic phenomenon.

As Morris Kline writes in *Mathematical Thought from Ancient to Modern Times*, "The first real success with partial differential equations came in renewed attacks on the vibrating string problem, typified by the violin string. The approximation that the vibrations are small was imposed by d'Alembert (1717-1783) in his papers of 1746."

Imagine shaking the end of a rope up and down gently, as in Figure C.21.1.

Figure C.21.1:

That motion starts a wave moving along the rope. The individual molecules in the rope move up and down, while the wave travels to the right. In the case of a sound wave, the wave travels at 700 miles per hour, but the air just vibrates back and forth. (When someone says "good morning" to us, we are not struck with a hurricane blast of wind.)

To develop the mathematics of the wave in a weightless rope, we begin with some simplifying assumptions. First, each molecule moves only up and down. Second, the distance each one moves is very small and the slope of the curve assumed by the rope remains close to zero. (Think of a violin string.)

At time t the vertical position of the molecule whose x-coordinate is x is $y = y(x,t)$, for it depends on both x and t. Consider a very short section of the rope at time t, shown as PQ in Figure C.21.2.

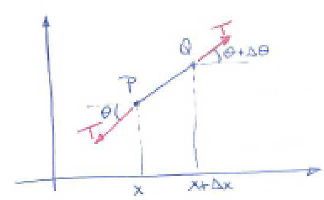


Figure C.21.2:

We assume that the tension T is the same throughout the rope. Apply Newton's Second Law, "force equals mass times acceleration," to the mass in PQ.

If the linear density of the rope is λ, the mass of the segment is λ times the length of the segment. Because we are assuming small displacements, we will approximate that length by Δx. The upward force exerted by the rope on the segment is $T\sin(\theta+\Delta\theta)$ and the downward force is $T\sin(\theta)$. The net vertical force is $T\sin(\theta+\Delta\theta) - T\sin(\theta)$. Thus

$$\underbrace{T sin(\theta+\Delta\theta) - T\sin(\theta)}_{\text{net vertical force}} = \underbrace{\lambda\Delta x}_{\text{mass}} \quad \underbrace{\frac{\partial^2 y}{\partial t^2}}_{\text{acceleration}} \quad . \tag{C.21.1}$$

(Because y is a function of x and t, we have a partial derivative, not an ordinary derivative.)

Next we express $\sin(\theta)$ and $\sin(\theta + \Delta\theta)$ in terms of the partial derivative $\partial y/\partial x$.

First of all, because θ is near 0, $\cos(\theta)$ is near 1. Thus $\sin(\theta)$ is approximately $\sin(\theta)/\cos(\theta) = \tan(\theta)$, the slope of the rope at time t above (or below) x, which is $\partial y/\partial x$ at x and t. Similarly, $\sin(\theta + \Delta\theta)$ is approximately $\partial y/\partial x$ at $x + \Delta x$ and t. So (C.21.1) is approximated by

$$T\frac{\partial y}{\partial x}(x + \Delta x, t) - T\frac{\partial y}{\partial x}(x, t) = \lambda \Delta x \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.2}$$

Dividing both sides of (C.21.2) by Δx gives

$$\frac{T\left(\frac{\partial y}{\partial x}(x + \Delta x, t) - \frac{\partial y}{\partial x}(x, t)\right)}{\Delta x} = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.3}$$

Letting Δx in (C.21.3) approach 0, we obtain

$$T\frac{\partial^2 y}{\partial x^2}(x, t) = \lambda \frac{\partial^2 y}{\partial t^2}(x, t). \tag{C.21.4}$$

Since both T and λ are positive, we can write (C.21.4) in the form

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \tag{C.21.5}$$

This is the famous **wave equation**. It relates the acceleration of the molecule to the geometry of the curve; the latter is expressed by $\partial^2 y/\partial x^2$. Since we are assuming that the slope of the rope remains near 0, $\frac{\partial^2 y}{\partial x^2}$ is approximately

$$\frac{\frac{\partial^2 y}{\partial x^2}}{\left(\sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2}\right)^3}$$

which is the curvature at a given location and time. At the curvier part of the rope, the acceleration is greater.

As the CIE in the next chapter shows, the constant c turns out to be the velocity of the wave.

EXERCISES

1.[M] Figure C.21.3 shows a vibrating string whose ends are fixed at A and B. Assume that each part of the string moves parallel to the y-axis (a reasonable approximation of the vibrations are small.) Let $y = f(x, t)$ be the height of the string at the point with abscissa x at time t, as shown in the figure. In this case, the partial derivatives are denoted $\partial y/\partial x$ and $\partial y/\partial t$.

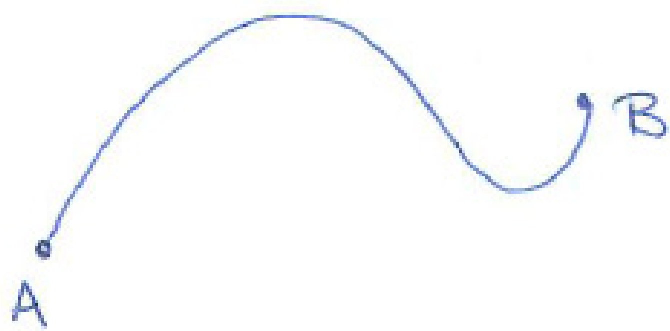


Figure C.21.3:

(a) What is the meaning of y_x?

(b) What is the meaning of y_t?

Chapter 17

Plane and Solid Integrals

In Chapter 2 we introduced the derivative, one of the two main concepts in calculus. Then in Chapter 15 we extended the idea to higher dimensions. In the present chapter, we generalize the concept of the definite integral, introduced in Chapter 6, to higher dimensions.

Take a moment to review the definite integral. Instead of using the notation of Chapter 6, we will restate the definition in a notation that easily generalizes to higher dimension.

We started with an interval $[a, b]$, which we will call I, and a continuous function f defined at each point P of I. Then we cut I into n short intervals $I_1, I_2, \ldots, I_u$, chose a point P_1 in I_1, P_2 in I_2, $\ldots$, P_n in I_n. See Figure 17.0.1. Denoting the length of I_i by L_i, we formed the sum

$$\sum_{i=1}^{n} f(P_i)L_i.$$

Figure 17.0.1:

The limit of these sums as all the subintervals are chosen shorter and shorter is the definite integral of f over interval I. We denoted it $\int_b^a f(x)\,dx$. We now denote it $\int_I f(P)dL$. This notation tells us that we are integrating a function, f, over an interval I. The dL reminds us that the integral is the limit of approximations formed as the sum of products of the function value and the length of an interval.

We will define integrals of functions over plane regions, such as square and disks, over solid regions, such as tubes and balls, and over surfaces such as the surface of a ball, in the same way. You can probably conjecture already what the definition will be. These integrals are needed to compute total mass if we know the density at each point, or total gravitational attraction, or center of gravity, and so on.

It is one thing to define these higher-dimensional integrals. It is another to calculate them. Most of our attention will be devoted to seeing how to compute them with the aid of so-called "iterated integrals," which involve integrals over intervals, the type defined in Chapter 6.

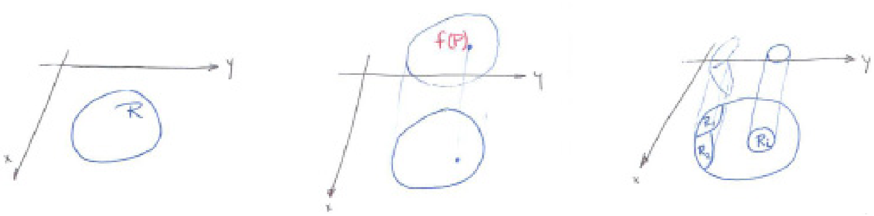
17.1 The Double Integral: Integrals Over Plane Areas

We suggest you re-read the introduction to this chapter and the definition of the definite integral $\int_a^b f(x)\,dx$ before going on.

The goal of this section is to define the integral of a function defined in a region of a plane. With only a slight tweaking of this definition, we will define later in the chapter integrals over surfaces and solids.

Volume Approximated by Sums

Let R be a region in the xy plane, bounded by curves. For convenience, assume R is convex (no dents), for example, an ellipse, a disk, a parallelogram, a rectangle, or a square. We draw R in perspective in Figure 17.1.1(a). Imagine



(a) (b) (c)

Figure 17.1.1:

that there is a surface above R (perhaps an umbrella). The height of the surface above point P on R is $f(P)$, as shown in Figure 17.1.1(b)

If you know $f(P)$ for every point P how would you estimate the volume, V, of the solid under the surface and above R?

Just as we used rectangles to estimate the area of regions back in Section 6.1, we will use cylinders to estimate the volume of a solid. Recall, from Section 7.4, that the volume of a cylinder is the product of its height and the area of its base.

Inspired by the approach in Section 6.1, we cut R into n small regions R_1, R_2, …, R_n. Each R_i has area A_i. Choose points P_1 in R_1, P_2 in R_2, …, P_n in R_n. Then we build a cylinder over each little region R_i. Its height will be $f(P_i)$. There will then be n cylinders. The total volume of these cylinders is

$$\sum_{i=1}^{n} f(P_i)A_i. \tag{17.1.1}$$

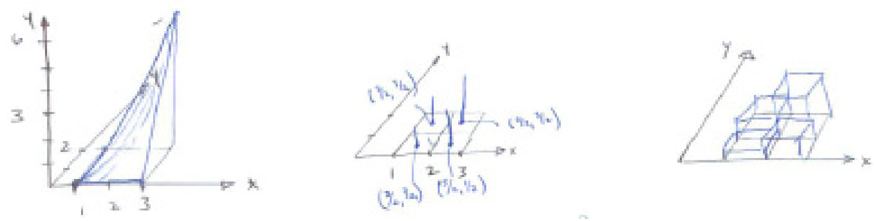
As we choose the regions R_1, R_2, …, R_n, smaller and smaller, the sum (17.1.1) approaches the volume V, if f is a continuous function.

EXAMPLE 1 Estimate the volume of the solid under the saddle $z = xy$

SHERMAN: Changed left edge from 0 to 1 so that base and height are not the same.

and above the rectangle R whose vertices are $(1,0)$, $(2,0)$, $(2,3)$, and $(1,3)$.

SOLUTION Figure 17.1.2(a) shows the solid region in question.



(a) (b) (c)

Figure 17.1.2:

The highest point is above $(2,3)$, where $z = 6$. So the solid fits in a box whose height is 6 and whose base has area 4. So we know the volume is at most $4 \cdot 6 = 24$.

To estimate the volume we cut the rectangular box into four 1 by 1 squares and evaluate $z = xy$ at, say, the center of the squares, as shown in Figure 17.1.2(b).

Then we form a cylinder for each square. The base is the square and the height is determined by the value of xy at the center of the square. These are shown in Figure 17.1.2(c). (The cylinder over rectangle boxes.)

Then the total volume is

$$\underbrace{\frac{3}{4}}_{\text{height}} \cdot \underbrace{1}_{\text{area of base}} + \underbrace{\frac{5}{4}}_{\text{height}} \cdot \underbrace{1}_{\text{area of base}} + \underbrace{\frac{9}{4}}_{\text{height}} \cdot \underbrace{1}_{\text{area of base}} + \underbrace{\frac{15}{4}}_{\text{height}} \cdot \underbrace{1}_{\text{area of base}} = 8 \tag{17.1.2}$$

This estimate is then 8 cubic units. We know this is an overestimate (Why?) By cutting the base into smaller pieces and using more cylinders we could make a more accurate estimate of the volume of the solid. ◇

Density

Before we consider a "total mass" problem we must define the concept of "density." Consider a piece of sheet metal, which we view as part of a plane. It is homogeneous, "the same everywhere." Let R be any region in it, of area A and mass m. The quotient m/A is the same for all regions R, and is called the "density."

It may happen that the material, unlike sheet metal is not uniform. For instance, a towel that was just used to dry dishes. As R varies, the quotient

m/A, or "average density in R," also varies. Physicists define the **density at a point** as follows.

They consider a small disk R of radius r and center at P, as in Figure 17.1.4. Let $m(r)$ be the mass in that disk and $A(r)$ be the area of the disk (πr^2). The

$$\text{"Density at } P\text{"} = \lim_{r\to 0} \frac{m(r)}{A(r)}.$$

Thus density is denoted $\sigma(P)$, "sigma of P,"

With the physicists, we will assume the density $\sigma(P)$ exists at each point and that it is a continuous function. In addition, we will assume that if R is a very small region of area A and P is a point in that region then the product $\sigma(P)A$ is an approximation of the mass in R.

σ is Greek for our letter "s", the initial letter of "surface." $\sigma(P)$ denotes the density of a surface or "lamina" at P.

Figure 17.1.3:

Total Mass Approximated by Sums

Assume that a flat region R is occupied by a material of varying density. The density at point P in R is $\sigma(P)$. Estimate M, the total mass in R.

As expected, we cut R into n small regions R_1, R_2, …, R_i has area A_i. We next choose points P_1 in R_1, P_2 in $R_2, \dots, P_u$ in R_n. Then we estimate the mass in each little region R_i, as shown in Figure 17.1.4. The mass in R_i is approximately

$$\underbrace{\sigma(P_i)}_{\text{density}} \cdot \underbrace{A_i}_{\text{area}}$$

Thus

$$\sum_{i=1}^{n} \sigma(P_1)A_i \qquad (17.1.3)$$

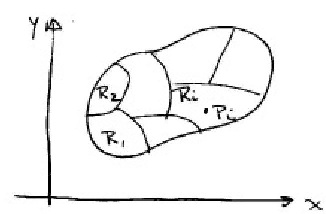


Figure 17.1.4: This example has $i = 7$ subregions.

is the total estimate. As we divide R into smaller and smaller regions, , the sums (17.1.2) approaches the total mass M, if σ is a continuous function.

EXAMPLE 2 A rectangular **lamina**, of varying density occupies the rectangle with corners at $(0,0)$, $(2,0)$, $(2,3)$, and $(0,3)$ in the xy plane. Its density at (x,y) is xy grams per square cm. Estimate its mass by cutting it into six 1 by 1 squares and evaluating the density at the center of each square.

SOLUTION One such square is shown in Figure 17.1.5. The density at its center is $\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$. Since its area is $1\times 1 = 1$, an estimate of σ, its mass, is

$$\underbrace{\frac{1}{4}}_{\text{density}} \cdot \underbrace{1}_{\text{area}} = \frac{1}{4} \text{ grams.}$$

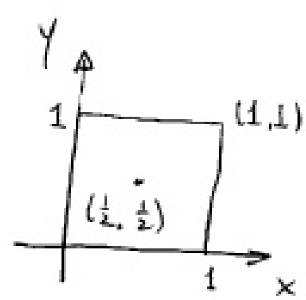


Figure 17.1.5:

Similar estimates for the remaining six small squares gives a total estimate of

$$\frac{1}{4} \cdot 1 + \frac{3}{4} \cdot 1 + \frac{3}{4} \cdot 1 + \frac{9}{4} \cdot 1 + \frac{5}{4} \cdot 1 + \frac{15}{4} \cdot 1 = 9 \text{ grams}$$

Thus sum is identical to the sum (17.1.2), which estimates a volume. $\diamond$

The arithmetic in Examples 1 and 2 show that totally unrelated problems, one in volume, the other in mass, lead to the same estimates. Moreover, as the rectangle is cut into smaller pieces, the estimate would become closer and closer to the volume or the mass. These estimates, similar to the estimates $\sum_{i=1}^{n}(f(c_i)\Delta x_i$ that appears in the definition of the definite integral $\int_a^b f(x)\,dx$, brings us to the definition of "double integral". It is called the double integral because the domain of the function is in the two-dimensional plane.

———————-

The Double Integral

The definition of the double integral is almost the same as that of $\int_a^b f(x)\,dx$, the integral over an interval. The only differences are:

1. instead of dividing an interval into smaller intervals, we divide a planar region into smaller planar regions,

2. instead of a function defined on an interval, we have a function defined on a planar region, and

3. we need a quantitative way to say that a "little" region is "small."

To meet the need described in (3) we define the "diameter" of a planar region. The **diameter** of a region bounded by a curve is the maximum distance between two points in the region. For instance, the diameter of a square of side s is $s\sqrt{2}$ and the diameter of a disk is the same as its traditional diameter that we know from geometry.

With that aside taken care of, we are ready to define a double integral.

DEFINITION *(Double Integral)* Let R be a region in a plane bounded by curves and f a continuous numerical function defined at least on R. Partition R into smaller regions R_1, R_2, …, R_n of respective areas A_1, A_2, …, A_n. Choose a point P_1 in R_1, P_2 in R_2, …, P_n in R_n and form the approximating (Riemann) sum

$$\sum_{i=1}^{n} f(P_i)A_i. \tag{17.1.4}$$

Form a sequence of such partitions such that as one goes out in the sequence of partitions, the sequence of diameters of the largest region in each partition approaches 0. Then the sums (17.1.4) approach a limit, which is called "the integral of f over R" or the "double integral" of f over R. It is denoted

$$\int_R f(P)\, dA.$$

Before looking at some examples, we make four brief remarks:

1. It is called a double integral because R lies in a plane, which has dimension 2.

2. We use the notion of a diameter of a region only to be able to define the double integral.

3. It is proved in advanced calculus that the sums do indeed approach a limit.

4. Other notations for a double integral are discussed near the end of this section.

Our discussion of integrals over a plane region started with two important illustrations . The rest of this section is devoted to these applications in the context of double integrals.

Volume Expressed as a Double Integral

Consider a solid S and its projections ("shadows") R on a plane, as in Figure 17.1.6. Assume that for each point P in R the line through P perpendicular to R intersects S in a line segment of length $C(P)$. Then

"The double integral of cross-section is the volume."

$$\text{Volume of } S = \int_R C(P)\, dA.$$

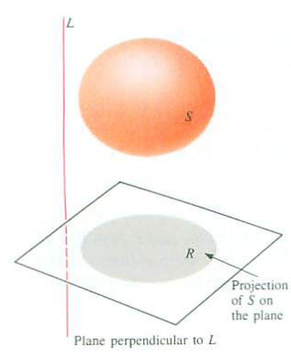


Figure 17.1.6: ARTIST: Delete the line L, and the current caption. Add a point P in R and draw the vertical line through P, highlighting the part that is

Mass Expressed as a Double Integral

Consider a plane distribution of mass through a region R, as shown in Figure 17.1.7. The density may vary throughout the region. Denote the density at P by $\sigma(P)$ (in grams per square centimeters). Then

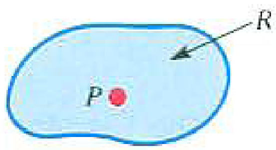


Figure 17.1.7:

"The double integral of density is the total mass."

$$\text{Mass in } R = \int_R \sigma(P)\ dA$$

Average Value as a Double Integral

The average value of $f(x)$ for x is the interval $[a, b]$ was defined in Section 6.3 as

$$\frac{\int_a^b f(x)\ dx}{\text{length of interval.}}$$

We make a similar definition for a function defined on a two-dimensional region.

DEFINITION *(Average value)* The **average value** of f over the region R is

$$\frac{\int_R f(P)\ dA}{\text{Area of } R}.$$

If $f(P)$ is positive for all P in R, there is a simple geometric interpretation of the average of f over R. Let S be the solid situated below the graph of f (a surface) and above the region R. The average value of f over R is the height of the cylinder whose base is R and whose volume is the same as the volume of S. (See Figure 17.1.8. The integral $\int_R f(P)\ dA$ is called "an integral over a plane region" to distinguish it from $\int_a^b f(x)\ dx$, which, for contrast, is called, "an integral over an interval."

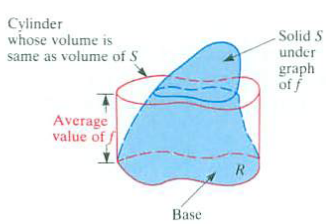


Figure 17.1.8:

/mnoteSHERMAN: Duplicitous? Or needed? Shorten to margin note? Recall that $\int_R f(P)\ dA$ is often denoted $\int\int_R f(P)\ dA$, with the two integral signs emphasizing that the integral is over a plane set. However, the symbol dA, which calls to mind areas, is an adequate reminder.

The integral of the function $f(P) = 1$ over a region is of special interest. The typical approximating sum $\sum_{i=1}^n f(P_i)A_i$ then equals $\sum_{i=1}^n 1 \cdot A_i = A_1 + A_2 + \cdots + A_n$, which is the area of the region R that is being partitioned. Since *every* approximating sum has this same value, it follows that

$$\lim_{n\to\infty} \sum_{i=1}^n f(P_i)A_i = \text{Area of } R.$$

Integral	Interpretation
$\int_R 1\ dA$	Area of R
$\int_R \sigma(P)\ dA$, $\sigma(P) = \text{density}$	Mass of R
$\int_R c(P)\ dA$, $c(P) = \text{length of cross section of solid}$	Volume of R

Table 17.1.1:

Consequently

The integral of a constant function, 1, gives area.

$$\int\limits_R 1\ dA = \text{Area of } R.$$

This formula will come in handy on several occasions. The 1 is often omitted, in which case we write $\int_R\ dA = \text{Area of } R$. This table summarizes some of the main applications of the double integral $\int_R\ dA$:

Properties of Double Integrals

Integrals over plane regions have properties similar to those of integrals over intervals:

1. $\int_R cf(P)\ dA = c\int_R f(P)\ dA$ for any constant c.
2. $\int_R [f(P) + g(P)]\ dA = \int_R f(P)\ dA + \int_R g(P)\ dA$.
3. If $f(P) \leq g(P)$ for all points P in R, then $\int_R f(P)\ dA \leq \int_R g(P)\ dA$.
4. If R is broken into two regions, R_1 and R_2, overlapping at most on their boundaries, then
$$\int\limits_R f(P)\ dA = \int\limits_{R_1} f(P)\ dA + \int\limits_{R_2} f(P)\ dA.$$

For instance, consider **3** when $f(P)$ and $g(P)$ are both positive. Then $\int_R f(P)\ dA$ is the volume under the surface $z = f(P)$ and above R in the xy plane. Similarly $\int_R g(P)\ dA$ is the volume under $z = f(P)$ and above R. Then **3** asserts that the volume of a solid is not larger than the volume of a solid that contains it. (See Figure 17.1.9.)

SHERMAN: This summary needs to be written.

Summary

A Word about 4-Dimensional Space

We can think of 2-dimensional space as the set of ordered pairs (x, y) of real numbers. The set of ordered triplets of real numbers (x, y, z) represents 3-dimensional space. The set of ordered quadruplets of real numbers (x, y, z, t) represents 4-dimensional space.

It is easy to show that 4-dimensional space is a very strange place.

In 2-dimensional space the set of points of the form $(x, 0)$, the x-axis, meets the set of points of the form $(0, y)$, the y-axis, in a point, namely the origin $(0, 0)$. Now watch what can happen in 4-space. The set of points of the form $(x, y, 0, 0)$ forms a plane congruent to our familiar xy-plane. The set of points of the form $(0, 0, z, t)$ forms another such plane. So far, no surprise. But notice what the intersection of those two planes is. Their intersection is just the point (0,0,0,0). Can you picture two endless planes meeting in a single point? If so, tell us how.

EXERCISES for Section 17.1 *Key:* R–routine, M–moderate, C–challenging

1.[R] In the estimates for the volume in Example 1, the centers of the squares were used as the P_i's. Make an estimate for the volume in Example 1 by using the same partition but taking as P_i

(a) the lower left corner of each R_i,

(b) the upper right corder of each R_i.

(c) What do (a) and (b) tell about the volume of the solid?

2.[R] Estimate the mass in Example 2 using the partition of R into six squares and taking as the P_i's

(a) upper left corners,

(b) lower right corners.

3.[R] Let R be a set in the plane whose area is A. Let f be the function such that $f(P) = 5$ for every point P in R.

(a) What can be said about any approximating sum $\sum_{i=1}^{n} f(P_i)A_i$ formed for this R and this f?

(b) What is the value of $\int_R f(P)\ dA$?

4.[R] Let R be the square with vertices $(1,1)$, $(5,1)$, $(5,5)$, and $(1,5)$. Let $f(P)$ be the distance from P to the y-axis.

(a) Estimate $\int_R f(P)\ dA$ by partitioning R into four squares and using midpoints as sampling points.

(b) Show that $16 \leq \int_R f(P)\ dA \leq 80$.

5.[R] Let f and R be as in Example 1. Use the estimate of $\int_R f(P)\ dA$ obtained in the text to estimate the average of f over R.

6.[R] Assume that for all P in R, $m \leq f(P) \leq M$, where m and M are constants.

Let A be the area of R. By examining approximating sums, show that

$$mA \leq \int_R f(P)\, dA \leq MA.$$

7.[R]

(a) Let R be the rectangle with vertices $(0,0)$, $(2,0)$, $(2,3)$, and $(0,3)$. Let $f(x,y) = \sqrt{x+y}$. Estimate $\int_R \sqrt{x+y}\, dA$ by participating R into six squares and choosing the sampling points to be their centers.

(b) Use (a) to estimate the average value of f over R.

8.[R]

(a) Let R be the square with vertices $(0,0)$, $(0.8,0)$, $(0.8,0.8)$, and $(0,0.8)$. Let $f(P) = f(x,y) = e^{xy}$. Estimate $\int_R e^{xy}\, dA$ by partitioning R into 16 squares and choosing the sampling points to be their centers.

(b) Use (a) to estimate the average value of $f(P)$ over R.

(c) Show that $0.64 \leq \int_R f(P)\, dA \leq 0.64e^{0.64}$.

9.[R]

(a) Let R be the triangle with vertices $(0,0)$, $(4,0)$, and $(0,4)$ shown in Figure 17.1.10. Let $f(x,y) = x^2y$. Use the partition into four triangles and sampling points shown in the diagram to estimate $\int_R f(P)\, dA$.

(b) What is the maximum value of $f(x,y)$ in R?

(c) From (b) obtain an upper bound on $\int_R f(P)\, dA$.

Figure 17.1.10:

10.[R]

(a) Sketch the surface $z = \sqrt{x^2 + y^2}$.

(b) Let $\mathcal{V}$ be the region in space below the surface in (a) and above the square R with vertices $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$. Let V be the volume of $\mathcal{V}$. Show that $V \leq \sqrt{2}$.

(c) Using a partition of R with 16 squares, find an estimate for V that is too large.

(d) Using the partition in (c), find an estimate for V that is too small.

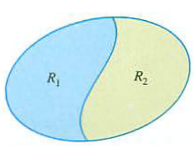
11.[R] The amount of rain that falls at point P during one year is $f(P)$ inches. Let R be some geographic region, and assume areas are measured in square inches.

(a) What is the meaning of $\int_R f(P)\, dA$?

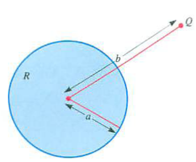
(b) What is the meaning of

$$\frac{\int_R f(P)\, dA}{\text{Area of } R}?$$

12.[M] A region R in the plane is divided into two regions R_1 and R_2. The function $f(P)$ is defined throughout R. Assume that you know the areas of R_1 and R_2 (they are A_1 and A_2) and the average of f over R_1 and the average of f over R_2 (they are f_1 and f_2). Find the average of f over R. (See Figure 17.1.11(a).)



(a)



(b)

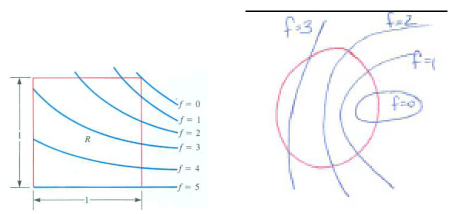
Figure 17.1.11:

13.[M] A point Q on the xy plane is at a distance b from the center of a disk R of radius $a(a < b)$ in the xy plane. For P in R let $f(P) = 1/\overrightarrow{PQ}$. Find positive numbers c and d such that:

$$c < \int_R f(P)\, dA < d.$$

(The numbers c and d depend on a and b.) See Figure 17.1.11(b).

14.[M] Figure 17.1.12(a) shows the parts of some level curves of a function $z = f(x, y)$ and a square R. Estimate $\int_R f(P)\ dA$, and describe your reasoning.



(a) (b)

Figure 17.1.12:

15.[M] Figure 17.1.12(b) shows the parts of some level curves of a function $z = f(x, y)$ and a unit circle R. Estimate $\int_R f(P)\ dA$, and describe your reasoning.

16.[C]

(a) Let R be a disk of radius 1. Let $f(P)$, for P in R, be the distance from P to the center of the disk. By cutting R into narrow circular rings with center at the center of the disk, evaluate $\int_R f(P)\ dA$.

(b) Find the average of $f(P)$ over R.

Exercises 17 and 18 introduce an idea known as **Monte Carlo methods** for estimating a double integral using randomly chosen points. These methods tend to be rather inefficient because the error decreases on the order of $1/\sqrt{n}$, where n is the number of random points. That is a slow rate. These methods are used only when it's possible to choose n very large.

17.[C] This exercise involves estimating an integral by choosing points randomly. A computing machine can be used to generate random numbers and thus random points in the plane which can be used to estimate definite integrals, as we now show. Say that a complicated region R lies in the square whose vertices are $(0, 0)$, $(2, 0)$, $(2, 2)$, and $(0, 2)$, and a complicated function f is defined in R. The machine generated 100 random points (x, y) in the square. Of these, 73 lie in R. The average value of f for these 73 points is 2.31.

(a) What is a reasonable estimate of the area of R?

(b) What is a reasonable estimate of $\int_R f(P)\ dA$?

18.[C] Let R be the disk bounded by the unit circle $x^2 + y^2 = 1$ in the xy plane. Let $f(x, y) = e^{x^2 y}$ be the temperature at (x, y).

(a) Estimate the average value of f over R by evaluating $f(x, y)$ at twenty random points in R. (Adjust your program to select each of x and y randomly in the interval $[-1, 1]$. In this way you construct a random point (x, y) in the square whose vertices are $(1, 1)$, $(-1, 1)$, $(-1, -1)$, $(1, -1)$. Consider only those points that lie in R.)

(b) Use (a) to estimate $\int_R f(P)\, dA$.

(c) Show why $\pi/e \leq f_R f(P)\, dA \leq \pi e$.

19.[C] Sam is heckling again. "As usual, the authors made this harder than necessary. They didn't need to introduce "diameters." Instead they could have used good old area. They could have taken the limit as all the areas of the little pieces approached 0. I'll send them a note."
Is Sam right?

In making finer and finer partitions as $n \to \infty$ we saw that each R_i is small in the sense it fits in a disk of radius r_n, where $r_n \to 0$ as $n \to \infty$. The Exercises 20 to 23 in this section explore another way to control the size of a region.
20.[C] Consider a region R in the plane. The diameter, d of R, is defined as the greater distance between two points in R. Find the diameter of

(a) a disk of radius r,

(b) and equilateral triangle of side length s,

(c) a square whose sides have length s.

21.[C]

(a) Show that a region of diameter d can always fit into a disk of diameter $2d$.

(b) Can it alway fit into a disk of diameter d?

22.[C] If a region has diameter d,

(a) how small can its area be?

SHERMAN: Is this in polar coordinate area? If so, move to Section 17.3 or Chapter Summary.

(b) show that area is less than or equal to $\pi d^2/2$.

23.[C] The unit square can be partitioned with nine congruent squares.

(a) What is the diameter of each of these small squares?

(b) It is possible to partition that square into nine regions whose largest diameter is 5/11. Show that 5/11 is smaller than the diameter in (a).

24.[R] Some practice differentiates.

25.[R] Some practice integrals, e.g. $\int \frac{x^2+1}{x^3}\,dx$, etc.

17.2 Computing $\int_R f(P)\ dA$ Using Rectangular Coordinates

In this section, we will show how to use rectangular coordinates to evaluate the integral of a function f over a plane region R, $\int_R f(P)\ dA$. This method requires that both R and f be described in rectangular coordinates. We first show how to describe plane regions R in rectangular coordinates.

Describing R in Rectangular Coordinates

Some examples illustrate how to describe planar regions by their cross sections in terms of rectangular coordinates.

EXAMPLE 1 Describe a disk R of radius a in a rectangular coordinates.

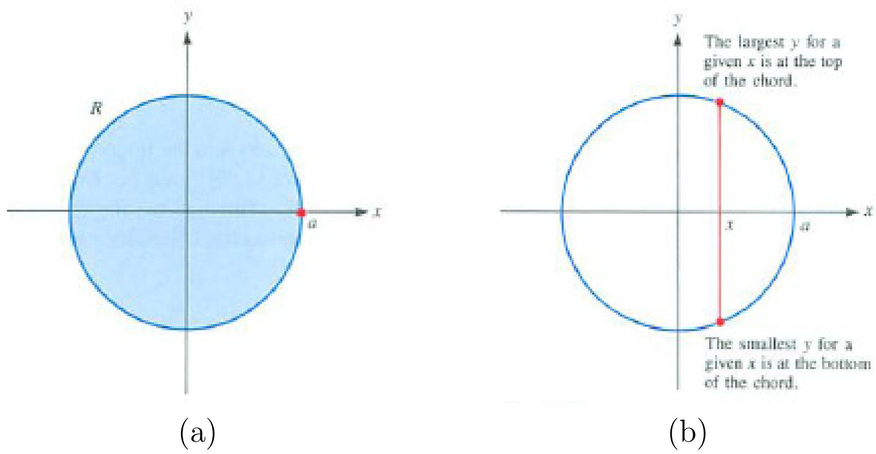


Figure 17.2.1:

SOLUTION Introduce an xy coordinate system with its origin at the center of the disk, as in Figure 17.2.1(a). A glance at the figure shows that x ranges from $-a$ to a. All that remains is to tell how y varies for each x in $[-a, a]$.

Figure 17.2.1(b) shows a typical x in $[-a, a]$ and corresponding cross section. The circle has the equation $x^2+y^2 = a^2$. The top half has the description $y = \sqrt{a^2 - x^2}$ and the bottom half, $y = -\sqrt{z^2 - y^2}$. So, for each x in $[-a, a]$, y varies from $-\sqrt{a^2 - x^2}$ to $\sqrt{a^2 - x^2}$. (As a check, test $x = 0$. Does y vary from $-\sqrt{a^2 - 0^2} = -a$ to $\sqrt{a^2 - 0^2} = a$? It does, as an inspection of Figure 17.2.1(b) shows.)

All told, this is the description of R by vertical cross sections:

$$-a \leq x \leq a, \qquad -\sqrt{a^2 - x^2} \leq y \leq \sqrt{a^2 - x^2}.$$

◇

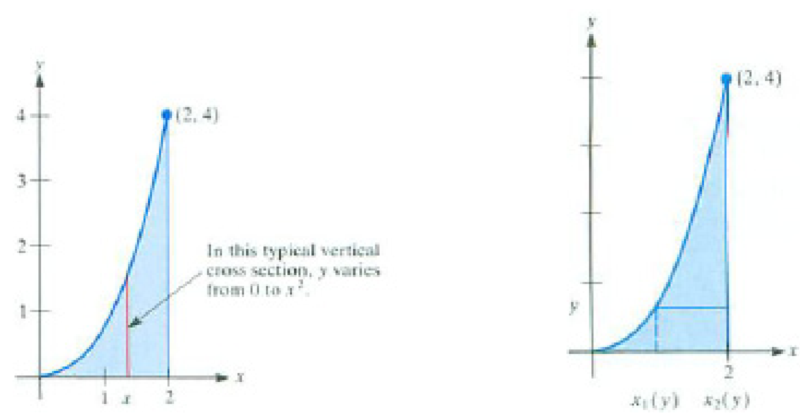
EXAMPLE 2 Let R be the region bounded by $y = x^2$, the x-axis, and the line $x = 2$. Describe R in terms of cross sections parallel to the y-axis.

SOLUTION A glance at R in Figure 17.2.2(a) shows that for points (x, y) in R, x ranges from 0 to 2. To describe R completely, we shall describe the behavior of y for any x in the interval $[0, 2]$.

Hold x fixed and consider only the cross section above the point $(x, 0)$. It extends from the x-axis to the curve $y = x^2$; for any x, the y coordinate varies from 0 to x^2. The compact description of R by vertical cross sections is:

$$0 \le x \le 2, \qquad 0 \le y \le x^2.$$

◇



(a) (b)

Figure 17.2.2:

EXAMPLE 3 Describe the region R of Example 2 by cross sections parallel to the x-axis, that is, horizontal cross sections.

SOLUTION A glance at R in Figure 17.2.2(b) shows that y varies from 0 to 4. For any y in the interval $[0, 4]$, x varies from a smallest value $x_1(y)$ to a largest value $x_2(y)$. Note that $x_2(y) = 2$ for each value of y in $[0, 4]$. To find $x_1(y)$, utilize the fact that the point $(x_1(y), y)$ is on the curve $y = x^2$, that is,

$$x_1(y) = \sqrt{y}.$$

The compact description of R in terms of horizontal cross sections is

$$0 \le y \le 4, \qquad \sqrt{y} \le x \le 2.$$

$$0 \le x \le 4, \quad 0 \le y \le 2$$

and

$$4 \le x \le 6, \quad 0 \le y \le 6 - x.$$

◇

EXAMPLE 4 Describe the region R whose vertices are $(0,0)$, $(0,6)$, $(4,2)$, and $(0,2)$ by vertical cross sections and then by horizontal cross sections. (See Figure 17.2.3.)

SOLUTION Clearly, x varies between 0 and 6. For any x in the interval $[0,4]$, y ranges from 0 to 2 (independently of x). For x in $[4,6]$, y ranges from 0 to the value of y on the line through $(4,2)$ and $(6,0)$. This line has the equation $y = 6 - x$. The description of R by vertical cross sections therefore requires two separate statements:

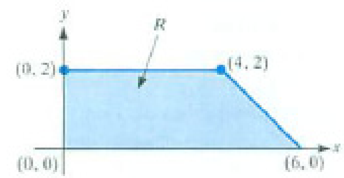


Figure 17.2.3:

Use of horizontal cross sections provides a simpler description. First, y goes from 0 to 2. For each y in $[0,2]$, x goes from 0 to the value of x on the line $y = 6 - x$. Solving this equation for x yields $x = 6 - y$.

The compact description in terms of horizontal cross-sections is much shorter:

$$0 \le y \le 2, \qquad 0 \le x \le 6 - y.$$

◇

These examples are typical. First, determine the range of one coordinate, and then see how the other coordinate varies for any fixed value of the first coordinate.

Evaluating $\int_R f(P)\ dA$ by Iterated Integrals

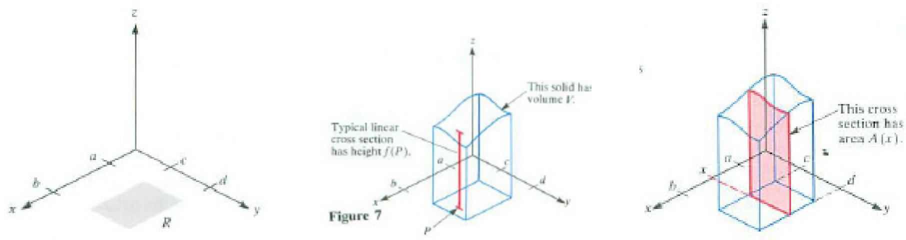
We will offer an intuitive development of a formula for computing double integrals over plane regions.

We first develop a way for computing a double integral over a rectangle. After applying this formula in Example 5, we make the slight modification needed to evaluate double integrals over more general regions.

Consider a rectangular region R whose description by cross sections is

$$a \le x \le b, \qquad c \le y \le d,$$

as shown in Figure 17.2.4(a). If $f(P) \le 0$ for all P in R, then $\int_R f(P)\ dA$ is the volume V of the solid whose base is R and which has, above P, height $f(P)$. (See Figure 17.2.4(b).) Let $A(x)$ be the area of the cross section made by a



(a) (b) (c)

Figure 17.2.4:

plane perpendicular to the x-axis and having abscissa x, as in Figure 17.2.4(c). As was shown in Section 5.1,

$$V = \int_b^a A(x)\ dx.$$

But the area $A(x)$ is itself expressible as a definite integral:

$$A(x) = \int_c^d f(x,y)\ dy.$$

Note that x is held fixed throughout the integration to find $A(x)$. This reasoning provides an iterated integral whose value is $V = \int_R f(P)\ dA$, namely,

$$\int_R f(P)\ dA = V = \int_a^b A(x)\ dx = \int_a^b \left(\int_c^d f(x,y)\ dy \right)\ dx.$$

In short

$$\int_R f(P)\ dA = \int_a^b \left(\int_c^d f(x,y)\ dy \right)\ dx.$$

An integral over a rectangle expressed an iterated integral

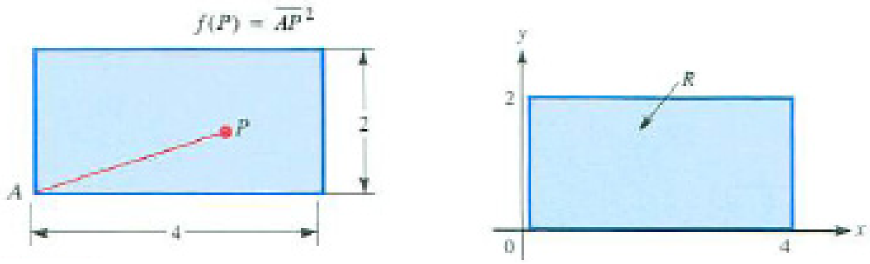
Of course, cross sections by planes perpendicular to the y-axis could be used. Then similar reasoning shows that

$$\int_R f(P)\ dA = \int_c^d \left(\int_a^b f(x,y)\ dx \right)\ dy.$$

The quantities $\int_a^b \left(\int_c^d f(x,y)\ dy\right)\ dx$ and $\int_c^d \left(\int_a^b f(x,y)\ dx\right)\ dy$ are called **iterated integrals**. Usually the brackets are omitted and are written $\int_a^b \int_c^d f(x,y)\, dy\, dx$ and $\int_c^d \int_a^b f(x,y)\ dx\ dy$.

The order of dx and dy matters; the differential that is on the left tells which integration is performed first.

EXAMPLE 5 Compute the double integral $\int_R f(P)\ dA$, where R is the rectangle shown in Figure 17.2.5(a) and the function f is defined by $f(P) = \overline{AP}^2$.



(a) (b)

Figure 17.2.5:

SOLUTION Introduce xy coordinates in the convenient manner depicted in Figure 17.2.5(b). Then f has this description in rectangular coordinates:

$$f(x,y) = \overline{AP}^2 = x^2 + y^2.$$

To describe R, observe that x takes all values from 0 to 4 and that for each x the number y takes all values between 0 and 2. Thus

$$\int_R f(P)\ dA = \int_0^4 \left(\int_0^2 (x^2+y^2)\ dy\right)\ dx.$$

We must first compute the inner integral

The cross-sectional area $A(x)$.

$$\int_0^2 (x^2+y^2)\ dy, \qquad \text{where } x \text{ is fixed in } [0,4].$$

To apply the Fundamental Theorem of Calculus, first find a function $F(x,y)$ such that

$$\frac{\partial F}{\partial y} = x^2 + y^2.$$

Keep in mind that x is constant during this first integration.

$$F(x,y) = x^2y + \frac{y^3}{3}$$

is such a function. The appearance of x in this formula should not disturb us, since x is fixed for the time being. By the Fundamental Theorem of Calculus,

$$\int_0^2 (x^2+y^2)\ dy \ = \ \left(x^2y+\frac{y^3}{3}\right)\Bigg|_{y=0}^{y=2} = \left(x^2\cdot 2+\frac{2^3}{3}\right) - \left(x^2\cdot 0+\frac{0^3}{3}\right) = 2x^2+\frac{8}{3}.$$

The notation $|_{y=0}^{y=2}$ reminds us that y is replaced by 0 and 2.

The formula $2x^2 + \frac{8}{3}$ is the area $A(x)$ discussed earlier in this section. Now compute

$$\int_0^4 A(x)\ dx = \int_0^4 (2x^2+\frac{8}{3})\ dx.$$

By the Fundamental Theorem of Calculus,

$$\int_0^4 \left(2x^2+\frac{8}{3}\right)\ dx = \left(\frac{2x^3}{3}+\frac{8x}{3}\right)\Bigg|_0^4 = \frac{160}{3}.$$

How do these compare with the estimates in Section 17.1?

Hence the two-dimensional definite integral has the value $\frac{160}{3}$. The volume of the region in Problem 1 of Sec. 16.1 is $\frac{160}{3}$ cubic inches. The mass in Problem 2 is $\frac{160}{3}$ grams. ◇

If R is not a rectangle, the repeated integral that equals $\int_R f(P)\ dA$ differs from that for the case where R is a rectangle only in the intervals of integration. If R has the description

$$a \leq x \leq b \qquad y_1(x) \leq y \leq y_2(x),$$

by cross sections parallel to the y-axis, then

$$\int_R f(P)\ dA = \int_a^b \left[\int_{y_1(x)}^{y_2(x)} f(x,y)\ dy\right]\ dx.$$

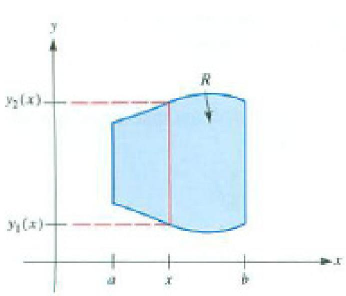


Figure 17.2.6:

Similarly, if R has the description

$$c \leq y \leq d \qquad x_1(y_{\leq} x \leq x_2(y),$$

by cross sections parallel to the x-axis, then

$$\int_R f(P)\, dA = \int_c^d \left(\int_{x_1(y)}^{x_2(y)} f(x,y)\, dx \right) dy.$$

The intervals of integration are determined by R; the function f influences only the integrand. (See Figure 17.2.7.)

In the next example R is the region bounded by $y = x^2$, $x = 2$, and $y = 0$; the function is $f(x,y) = 3xy$. The integral $\int_R 3xy\, dA$ has at least three interpretations:

1. If at each point $P = (x,y)$ in R we erect a line segment above P of length $3xy$, then the integral is the volume of the resulting solid. (See Figure 17.2.8.)
2. If the density of matter at (x,y) in R is $3xy$, then $\int_R 3xy\, dA$ is the total mass in R.
3. If the temperature at (x,y) in R is $3xy$ then $\int_R 3xy\, dA$ divided by the area of R is the average temperature in R.

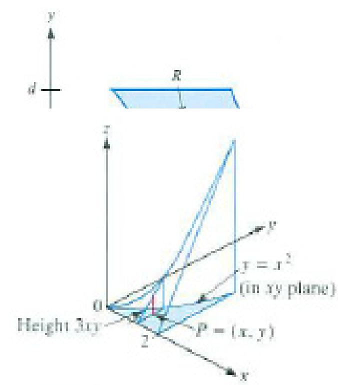


Figure 17.2.8:

EXAMPLE 6 Evaluate $\int_R 3xy\, dA$ over the region R shown in Figure 17.2.9.

This is the same R as in Examples 2 and 3.

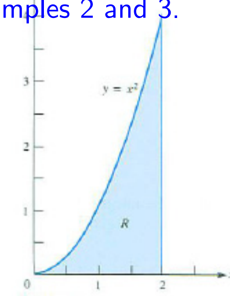


Figure 17.2.9:

SOLUTION If cross sections parallel to the y-axis are used, then R is described by

$$0 \le x \le 2 \qquad 0 \le y \le x^2.$$

Thus

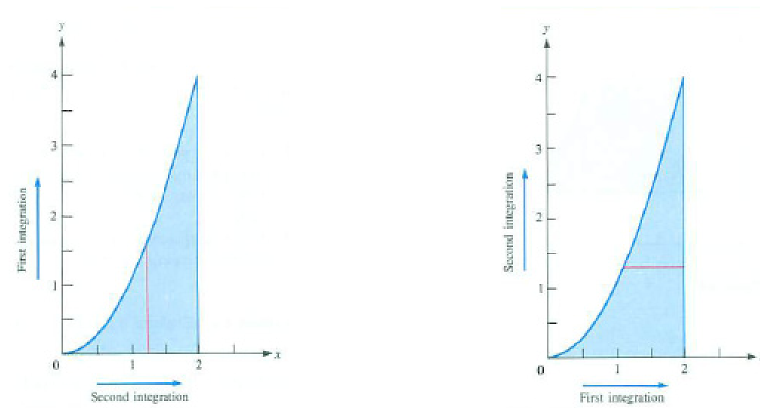
$$\int_R 3xy\, dA = \int_0^2 \left(\int_0^{x^2} 3xy\, dy \right) dx,$$

which is easy to compute. First, with x fixed,

$$\int_0^{x^2} 3xy\, dy = \left(3x\frac{y^2}{2} \right)\Big|_{y=0}^{y=x^2} = 3x\frac{(x^2)^2}{2} - 3x\frac{0^2}{2} = \frac{3x^5}{2}.$$

Then,

$$\int_0^2 \frac{3x^5}{2}\, dx = \frac{3x^6}{12}\Big|_0^2 = 16.$$



(a) (b)

Figure 17.2.10:

Figure 17.2.10(a) shows which integration is performed first.

The region R can also be described in terms of cross sections parallel to the x-axis:

$$0 \leq y \leq 4 \qquad \sqrt{y} \leq x \leq 2.$$

In this case, the double integral is evaluated as:

$$\int_R 3xy\ dA = \int_0^4 \left(\int_{\sqrt{y}}^2 3xy\ dx \right)\ dy,$$

which, as the reader may verify, equals 16. See Figure 17.2.10(b). ◇

In Example 6 we could evaluate $\int_R f(P)\ dA$ by cross sections in either direction. In the next example we don't have that choice.

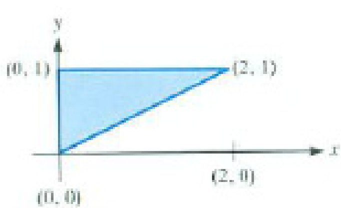


Figure 17.2.11:

EXAMPLE 7 A triangular lamina is located as in Figure 17.2.11. Its density at (x, y) is e^{y^2}. Find its mass, that is $\int_R f(P)\ dA$, where $f(x, y) = e^{y^2}$.

SOLUTION The description of R by vertical cross sections is

$$0 \leq x \leq 2, \qquad \frac{x}{2} \leq y \leq 1.$$

Hence

$$\int_R f(P)\ dA = \int_0^2 \left(\int_{x/2}^1 e^{y^2}\ dy \right)\ dx.$$

Since e^{y^2} does not have an elementary antiderivative, the Fundamental Theorem of Calculus is useless in computing

$$\int_{x/2}^{1} e^{y^2}\ dy.$$

So we try horizontal cross sections instead. The description of R is now

$$0 \le y \le 1, \qquad 0 \le x \le 2y.$$

This leads to a different iterated integral, namely:

$$\int_R f(P)\ dA = \int_0^1 \left(\int_0^{2y} e^{y^2}\ dx \right)\ dy.$$

The first integration, $\int_0^2 e^{y^2}\ dx$, is easy, since y is fixed; the integrand is constant. Thus

Note that the integrand does not depend on x.

$$\int_0^{2y} e^{y^2}\ dx \quad = \quad e^{y^2} \int_0^{2y} 1\ dx = e^{y^2} x \Big|_{x=0}^{x=2y} = e^{y^2} 2y.$$

The second definite integral in the repeated integral is thus $\int_0^1 e^{y^2} 2y\ dy$, which can be evaluated by the Fundamental Theorem of Calculus, since $d(e^{y^2})/dy = e^{y^2} 2y$:

$$\int_0^1 e^{y^2} 2y\ dy = e^{y^2}\Big|_0^1 = e^{1^2} - e^{0^2} = e - 1.$$

The total mass is $e - 1$. ◇

Notice that computing a definite integral over a plane region R involves, first, a wise choice of an xy-coordinate system; second, a description of R and f relative to this coordinate system; and finally, the computation of two successive definite integrals over intervals. The order of these integrations should be considered carefully since computation may be much simpler in one than the other. This order is determined by the description of R by cross sections.

Summary

We showed that the integral of $f(P)$ over a plane region R can be evaluated by an iterated integral, where the limits of integration are determined by R

(not by f). If each line parallel to the y-axis meets R in at most two points then R has the description

$$a \leq x \leq b, \qquad y_1(x) \leq y \leq y_2(x)$$

and

$$\int_R f(P)\ dA = \int_a^b \left(\int_{y_1(x)}^{y_2(x)} f(x,y)\ dy \right)\ dx.$$

If each line parallel to the x-axis meets R in at most two points, then, similarly, R can be described in the form

$$c \leq y \leq d \qquad x_1(y) \leq x \leq x_2(y)$$

and

$$\int_R f(P)\ dA = \int_c^d \left(\int_{x_1(y)}^{x_2(y)} f(x,y)\ dx \right)\ dy.$$

A Few Words on Notation

We use the notation $\int f(P)\ dA$ or $\int_R f(P)\ dA$ for a (double) integral over a plane region, $\int f(P)\ dS$ or $\int_{\mathcal{S}} f(P)\ dS$ for an integral over a surface, and $\int f(P)\ dV$ or $\int_R f(P)\ dV$ for a (triple) integral over a region in space. The symbols dA, dS, and dV indicate the type of set over which the integral is defined.

Many people traditionally use repeated integral signs to distinguish dimensions. For instance they would write $\int f(P)\ dA$ as $\iint f(P) dA$ or $\iint f(x,y)\ dxdy$. Similarly, they denote a triple integral by $\iiint f(P)\ dxdydz$. We use the single-integral-sign notation for all integrals for three reasons:

1. it is free of any coordinate system
2. it is compact (uses the fewest symbols): $\int$ for "integral", $f(P)$ or f for the integrand, and dA, dS, or dV for the set
3. it allows the symbols $\iint$ and $\iiint$ to be reserved for use exclusively for iterated integrals.

Iterated integrals are a completely different mathematical object. Each integral in an iterated integral is an integral over an interval. Note that this means we we write dx (or dy or dz) only when we are talking about an integral over an interval.

EXERCISES for Section 17.2 *Key:* R–routine, M–moderate, C–challenging

Exercises 1 to 12 provide practice in describing plane regions by cross sections in recangular coordinates. In each exercise, describe the region by (a) vertical cross sections and (b) horizontal cross sections.

1.[R] The triangle whose vertices are $(0,0)$, $(2,1)$, $(0,1)$.

2.[R] The triangle whose vertices are $(0,0)$, $(2,0)$, $(1,1)$.

3.[R] The parallelogram with vertices $(0,0)$, $(1,0)$, $(2,1)$, $(1,1)$.

4.[R] The parallelogram with vertices $(2,1)$, $(5,1)$, $(3,2)$, $(6,2)$.

5.[R] The disk of radius 5 and center $(0,0)$.

6.[R] The trapezoid with vertices $(1,0)$, $(3,2)$, $(3,3)$,$(1,6)$.

7.[R] The triangle bounded by the lines $y = x$, $x + y = 2$, and $x + 3y = 8$.

8.[R] The region bounded by the ellipse $4x^2 + y^2 = 4$.

9.[R] The triangle bounded by the lines $x = 0$, $y = 0$, and $2x + 3y = 6$.

10.[R] The region bounded by the curves $y = e^x$, $y = 1 - x$, and $x = 1$.

11.[R] The quadrilateral bounded by the lines $y = 1$, $y = 2$, $y = x$, $y = x/3$.

12.[R] The quadrilateral bounded by the lines $x = 1$, $x = 2$, $y = x$, $y = 5 - x$.

In Exercises 13 to 16 draw the regions and describe them by horizontal cross sections.

13.[R] $0 \le x \le 2$, $2x \le y \le 3x$

14.[R] $1 \le x \le 2$, $x^3 \le y \le 2x^2$

15.[R] $0 \le x \le \pi/4$, $0 \le y \le \sin x$ and $\pi/4 \le x \le \pi/2$, $0 \le y \le \cos x$

16.[R] $1 \le x \le e$, $(x - 1)/(e - 1? \le y \le \ln x$

In Exercises 17 to 22 evaluate the iterated integrals.

17.[R] $\int_0^1 \left(\int_0^x (x + 2y)\ dy\right)\ dx$

18.[R] $\int_1^2 \left(\int_x^{2x}\ dy\right)\ dx$

19.[R] $\int_0^2 \left(\int_0^{x^2} xy^2\ dy\right)\ dx$

20.[R] $\int_1^2 \left(\int_0^y e^{x+y}\ dx\right)\ dy$

21.[R] $\int_1^2 \left(\int_0^{\sqrt{y}} yx^2\ dx\right)\ dy$

22.[R] $\int_0^1 \left(\int_0^x y\sin(\pi x)\ dy\right)\ dx$

23.[R] Complete the calculation of the second iterated integral in Example 6.

24.[R]

(a) Sketch the solid region S below the plane $z = 1 + x + y$ and above the triangle R in the place with vertices $(0,0)$, $(1,0)$, $(0,2)$.

(b) Describe R in terms of coordinates.

(c) Set up an iterated integral for the volume of S.

(d) Evaluate the expression in (c), and show in the manner of Figure 17.2.10(a) and 17.2.10(b) which integration you performed first.

(e) Carry out (c) and (d) in the other order of integration.

25.[R] Let S be the solid region below the paraboloid $z = x^2 + 2y^2$ and above the rectangle in the xy plane with vertices $(0,0)$, $(1,0)$, $(1,2)$, $(0,2)$. Carry out the steps of Exercise 24 in this case.

26.[R] Let S be the solid region below the saddle $z = xy$ and above the triangle in the xy plane with vertices $(1,1)$, $(3,1)$, and $(1,4)$. Carry out the steps of Exercise 24 in this case.

27.[R] Let S be the solid region below the saddle $z = xy$ and above the region n the first quadrant of the xy plane bounded by the parabolas $y = x^2$ and $y = 2x^2$ and the line $y = 2$. Carry out the steps of Exercise 24 in this case.

28.[R] Find the mass of a thin lamina occupying the finite region bounded by $y = 2x^2$ and $y = 5x - 3$ and whose density at (x,y) is xy.

29.[R] Find the mass of a thin lamina occupying the triangle whose vertices are $(0,0)$, $(1,0)$, $(1,1)$ and whose density at (x,y) is $1/(1+x^2)$.

30.[R] The temperature at (x,y) is $T(x,y) = \cos(x+2y)$. Find the average temperature in the triangle with vertices $(0,0)$, $(1,0)$, $(0,2)$.

31.[R] The temperature at (x,y) is $T(x,y) = e^{x-y}$. Find the average temperature in the region in the first quadrant bounded by the triangle with vertices $(0,0)$, $(1,1)$, and $(3,1)$.

In each of Exercises 32 to 35 replace the given iterated integral by an equivalent one with the order of integration reversed. First sketch the region R of integration.

32.[R] $\int_0^2 \left(\int_0^{x^2} x^3 y \, dy \right) \, dx$

33.[R] $\int_0^{\pi/2} \left(\int_0^{\cos x} x^2 \, dy \right) \, dx$

34.[R] $\int_0^1 \left(\int_{x/2}^{x} xy \, dy \right) \, dx + \int_1^2 \left(\int_{x/2}^{1} xy \, dy \right) \, dx$

35.[R] $\int_{-1/\sqrt{2}}^{0} \left(\int_{-x}^{\sqrt{1-x^2}} x^3 y \, dy \right) \, dx + \int_0^1 \left(\int_0^{\sqrt{1-x^2}} x^3 y \, dy \right) \, dx$

In Exercises 36 to 39 evaluate the iterated integrals. First sketch the region of integration.

36.[R] $\int_0^1 \left(\int_x^1 \sin(y^2)\ dy \right)\ dx$

37.[R] $\int_0^1 \left(\int_{\sqrt{x}}^1 \frac{dy}{\sqrt{1+y^3}} \right)\ dx$

38.[R] $\int_0^1 \left(\int_{\sqrt[3]{y}}^1 \sqrt{1+x^4}/dx \right)\ dy$

39.[R] $\int_1^2 \left(\int_1^y \frac{\ln x}{x}\ dx \right)\ dy + \int_2^4 \left(\int_{y/2}^2 \frac{\ln x}{x}\ dx \right)\ dy$

40.[R] Let $f(x,y) = y^2 e^{y^2}$ and let R be the triangle bounded by $y = a$, $y = x/2$, and $y = x$. Assume that a is positive.

(a) Set up two repeated integrals for $\int_R f(P)\ dA$.

(b) Evaluate the easier one.

41.[R] Let R be the finite region bounded by the curve $y = \sqrt{x}$ and the line $y = x$. Let $f(x,y) = (\sin(y))/y$ if $y \neq 0$ and $f(x,0) = 1$. Compute $\int_R f(P)\ dA$.

17.3 Computing $\int_R f(P)\, dA$ Using Polar Coordinates

This section shows how to evaluate $\int_R f(P)\, dA$ by using polar coordinates. This method is especially appropriate when the region R has a simple description in polar coordinates, for instance, if it is a disk or cardioid. As in Section 17.2, we first examine how to describe cross sections in polar coordinates. Then we describe the iterated integral in polar coordinates that equals $\int_R f(P)\, dA$.

Describing R in Polar Coordinates

In describing a region R in polar coordinates, we first determine the range of θ and then see how r varies for any fixed value of θ. (The reverse order is seldom useful.) Some examples show how to find how r varies for each θ.

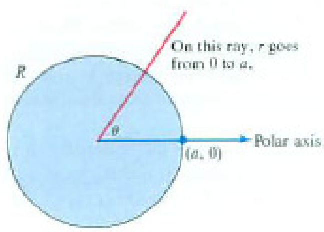


Figure 17.3.1:

EXAMPLE 1 Let R be the disk of radius a and center at the pole of a polar coordinate system. (See Figure 17.3.1.) Describe R in terms of cross sections by rays emanating from the pole.
SOLUTION To sweep out R, θ goes from 0 to 2π. Hold θ fixed and consider the behavior of r on the ray of angle θ. Clearly, r goes from 0 to a, independently of θ. (See Figure 17.3.1.) The complete description is

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a.$$

◇

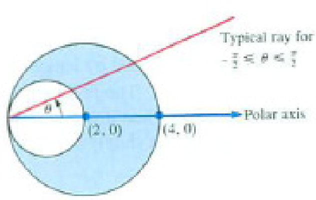


Figure 17.3.2:

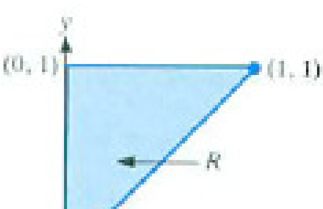
EXAMPLE 2 R Let R be the region between the circles $r = 2\cos\theta$ and $r = 4\cos\theta$. Describe R in terms of cross sections by rays from the pole. (See Figure 17.3.2.)
SOLUTION To sweep out this region, use the rays from $\theta = -\pi/2$ to $\theta = \pi/2$. for each such θ, r varies from $2\cos\theta$ to $4\cos\theta$. The complete description is

$$-\frac{\pi}{2} \le \theta \le \frac{\pi}{2}, \qquad 2\cos\theta \le r \le 4\cos\theta.$$

◇

As Examples 1 and 2 suggest, polar coordinates provide simple descriptions for regions bounded by circles. The next example shows that polar coordinates may also provide simple descriptions of regions bounded by straight lines, especially if some of the lines pass through the origin.

EXAMPLE 3 Let R be the triangular region whose vertices, in rectangular coordinates, are $(0,0)$, $(1,1)$, and $(0,1)$. Describe R in polar coordinates.



SOLUTION Inspection of R in Figure 17.3.3 shows that θ varies from $\pi/4$ to $\pi/2$. For each θ, r goes from 0 until the point (r, θ) is on the line $y = 1$, that is, on the line $r \sin(\theta) = 1$. Thus the upper limit of r for each θ is $1/\sin(\theta)$. The description of R is

$$\frac{\pi}{4} \le \theta \le \frac{\pi}{2}, \qquad 0 \le r \le \frac{1}{\sin(\theta)}.$$

◇ In general, cross sections by rays lead to descriptions of plane regions of the form:

$$\alpha \le \theta \le \beta, \qquad r_1(\theta) \le r \le r_2(\theta).$$

A Basic Difference Between Rectangular and Polar Coordinates

Before we can set up an iterated integral in polar coordinates for $\int_R f(P)\, dA$ we must contrast certain properties of rectangular and polar coordinates.

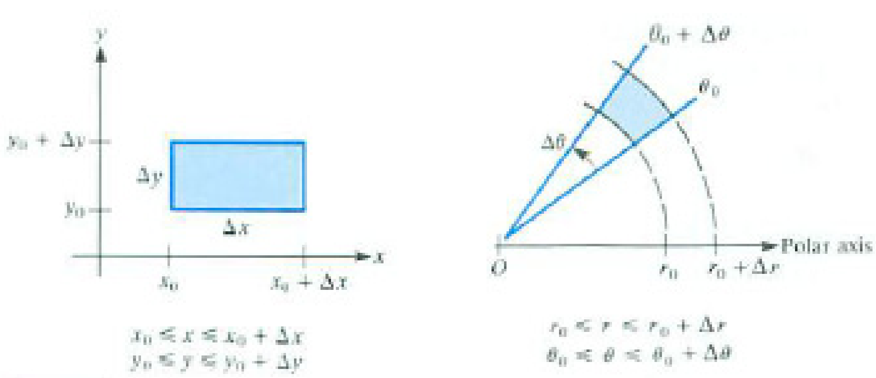
Consider all points (x, y) in the plane that satisfy the inequalities

$$x_0 \le x \le x_0 + \Delta x \qquad \text{and} \qquad y_0 \le y \le y_0 + \Delta y,$$

where x_0, Δx, y_0 and Δy are fixed numbers with Δx and Δy positive. The set is a rectangle of sides Δx and Δy shown in Figure 17.3.4(a). The area of this rectangle is simply the product of Δx and Δy; that is,

$$\text{Area} = \Delta x \Delta y. \tag{17.3.1}$$

This will be contrasted with the case of polar coordinates.



(a) (b)

Figure 17.3.4:

Consider the set in the plane consisting of the points (r, θ) such that

$$r_0 \le r \le r_0 + \Delta r \qquad \text{and} \qquad \theta_0 \le \theta \le \theta_0 + \Delta\theta,$$

where r_0, Δr, θ_0 and $\Delta\theta$ are fixed numbers, with r_0, Δr, θ_0 and $\Delta\theta$ all positive, as shown in Figure 17.3.4(b).

The exact area is found in Exercise 32.

When Δr and $\Delta\theta$ are small, the set is approximately a rectangle, one side of which has length Δr and the other, $r_0\Delta\theta$. So its area is approximately $r_0\Delta r\Delta\theta$. In this case,

$$\text{Area} \approx r_0 \Delta r \Delta\theta. \tag{17.3.2}$$

The area is *not* the product of Δr and $\Delta\theta$. (It couldn't be since $\Delta\theta$ is in radians, a dimensionless quantity – "arc length subtended on a circle divided by length of radius" – so $\Delta r\Delta\theta$ has the dimension of length, not of area.) The presence of this extra factor r_0 will be reflected in the integrand we use when integrating in polar coordinates.

It is necessary to replace dA by $r\ dr\ d\theta$, not simply by $dr\ d\theta$.

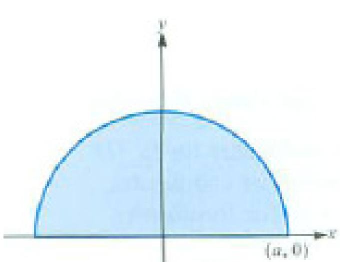
How to Evaluate $\int_R f(P)\ dA$ by an Iterated Integral in Polar Coordinates

The method for computing $\int_R f(P)\,dA$ with polar coordinates involves an iterated integral where the dA is replaced by $r\ dr\ d\theta$. A more detailed explanation of why the r must be added is given at the end of this section.

Notice the factor r in the integrand.

Evaluating $\int_R f(P)\ dA$ in Polar Coordinates

1. Express $f(P)$ in terms of r and θ: $f(r, \theta)$.
2. Describe the region R in polar coordinates:
$$\alpha \le \theta \le \beta, \qquad r_1(\theta) \le r \le r_2(\theta).$$
3. Evaluate the iterated integral:
$$\int_\alpha^\beta \int_{r_1(\theta)}^{r_2(\theta)} f(r, \theta) r\ dr\ d\theta.$$



EXAMPLE 4 Let R be the semicircle of radius a shown in Figure 17.3.5. Let $f(P)$ be the distance from a point P to the x-axis. Evaluate $\int_R f(P)\ dA$ by an iterated integral in polar coordinates.
SOLUTION In polar coordinates, R has the description

$$0 \leq \theta \leq \pi, \qquad 0 \leq r \leq a.$$

The distance from P to the x-axis is, in rectangular coordinates, y. Since $y = r\sin(\theta)$, $f(P) = r\sin(\theta)$. Thus,

Notice the extra r in the integrand.

$$\int_R f(P)\ dA = \int_0^\pi \left(\int_0^a (r\sin(\theta))r\ dr \right)\ d\theta.$$

The calculation of the iterated integral is like that for an iterated integral in rectangular coordinates. First, evaluate the inside integral:

From here on the calculation are like those in the preceding section.

$$\int_0^a r^2 \sin(\theta)\ dr = \sin(\theta) \int_0^a r^2\ dr = \sin(\theta) \left(\frac{r^3}{3} \right) \Big|_0^a = \frac{a^3 \sin(\theta)}{3}.$$

The outer integral is therefore

$$\begin{aligned} \int_0^\pi \frac{a^3 \sin\theta}{3}\ d\theta &= \frac{a^3}{3} \int_0^\pi \sin\theta\ d\theta = \frac{a^3}{3}(-\sin\theta)\Big|_0^\pi \\ &= \frac{a^3}{3}[(-\cos\pi) - (-\cos 0)] = \frac{a^3}{3}(1+1) = \frac{2a^3}{3}. \end{aligned}$$

Thus

$$\int_R y\ dA = \frac{2a^3}{3}.$$

◇

Example 5 refers to a ball of radius a. Generally, we will distinguish between a **ball**, which is a solid region, and a **sphere**, which is only the surface of a ball.

EXAMPLE 5 A ball of radius a has its center at the pole of a polar coordinate system. Find the volume of the part of the ball that lies above the plane region R bounded by the curve $r = a\cos(\theta)$. (See Figure 17.3.6.)
SOLUTION It is necessary to describe R and f in polar coordinates, where $f(P)$ is the length of a cross section of the solid made by a vertical line through P. R is described as follows: r goes 0 to $a\cos(\theta)$ for each θ in $[-\pi/2, \pi/2]$, that is,

$$-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}, \qquad 0 \leq r \leq a\cos\theta.$$

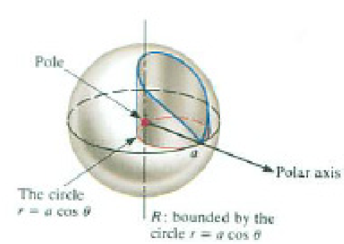


Figure 17.3.6:

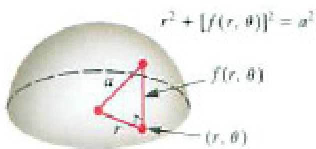


Figure 17.3.7:

To express $f(P)$ in polar coordinates, consider Figure 17.3.7, which shows the top half of a ball of radius a. By the Pythagorean Theorem,

$$r^2 + (f(r, \theta))^2 = a^2.$$

Thus

$$f(r, \theta) = \sqrt{a^2 - r^2}.$$

Consequently,

$$\text{Volume} = \int_R f(P)\, dA = \int_{-\pi/2}^{\pi/2} \left(\int_0^{a\cos(\theta)} \sqrt{z^2 - r^2} r\, dr \right) d\theta.$$

Remember to double.

Exploiting symmetry, compute half the volume, keeping θ in $[0, \pi/2]$, and then double the result:

$$\begin{aligned}
\int_0^{a\cos(\theta)} \sqrt{a^2 - r^2} r\, dr &= \left. \frac{-(a^2 - r^2)^{3/2}}{3} \right|_0^{a\cos(\theta)} = -\left(\frac{(a^2 - a^2\cos^2(\theta))^{3/2}}{3} - \frac{(a^2)^{3/2}}{3} \right) \\
&= \frac{a^3}{3} - \frac{(a^2 - a^2\cos^2(\theta))^{3/2}}{3} = \frac{a^3}{3} - \frac{a^3(1 - \cos^2(\theta))^{3/2}}{3} \\
&= \frac{a^3}{3}(1 - \sin^3(\theta)).
\end{aligned}$$

(The trigonometric formula used above, $\sin(\theta) = \sqrt{1 - \cos^2(\theta)}$, is true when $0 \le \theta \le \pi/2$ but not when $-\pi/2 \le \theta \le 0$.)

Then comes the second integration:

$$\begin{aligned}
\int_0^{\pi/2} \frac{a^3}{3}(1 - \sin^3(\theta))\, d\theta &= \frac{a^3}{3} \int_0^{\pi/2} (1 - (1 - \cos^2(\theta))\sin(\theta))\, d\theta \\
&= \frac{a^3}{3} \int_0^{\pi/2} 1 - \sin(\theta) - \cos^2(\theta)\sin(theta)\, d\theta \\
&= \frac{a^3}{3} \left. \left(\theta + \cos(\theta) - \frac{\cos^3(\theta)}{3} \right) \right|_0^{\pi/2} \\
&= \frac{a^3}{3} \left[\frac{\pi}{2} - \left(1 - \frac{1}{3} \right) \right] = a^3 \left(\frac{3\pi - 4}{18} \right).
\end{aligned}$$

We remembered.

The total volume is twice is large:

$$a^3 \left(\frac{3\pi - 4}{9} \right).$$

◇

EXAMPLE 6 A circular disk of radius a is formed of a material which had a density at each point equal to the distance from the point to the center.

(a) Set up an iterated integral in rectangular coordinates for the total mass of the disk.

(b) Set up an iterated integral in polar coordinates for the total mass of the disk.

(c) Compute the easier one.

SOLUTION The disk is shown in Figure 17.3.8.

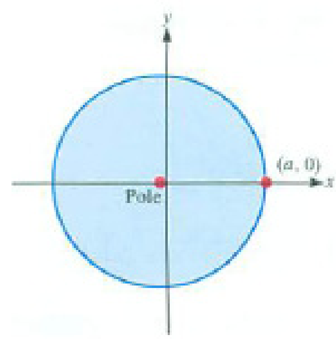


Figure 17.3.8:

(a) (Rectangular coordinates) The density $\sigma(P)$ at the point $(P) = (x, y)$ is $\sqrt{x^2 + y^2}$. The disk has the description

$$-a \le x \le a, \qquad -\sqrt{a^2 - x^2} \le y \le \sqrt{a^2 - x^2}.$$

Thus

$$\text{Mass} = \int_R \sigma(P)\,dA = \int_{-a}^{a} \left(\int_{-\sqrt{a^2-x^2}}^{\sqrt{a^2-x^2}} \sqrt{x^2 + y^2}\,dy \right) dx.$$

(b) (Polar coordinates) The density $\sigma(P)$ at $P = (r, \theta)$ is r. The disk has the description

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a.$$

Thus

$$\text{Mass} = \int_R \sigma(P)\,dA = \int_0^{2\pi} \left(\int_0^a r \cdot r\,dr \right) d\theta = \int_0^{2\pi} \left(\int_0^a r^2\,dr \right) d\theta.$$

(c) Even the first integration in the iterated integral in (a) would be tedious. However, the iterated integral in (b) is a delight: The first integration gives

$$\int_0^a r^2\,dr = \left.\frac{r^3}{3}\right|_0^a = \frac{a^3}{3}.$$

The second integration gives

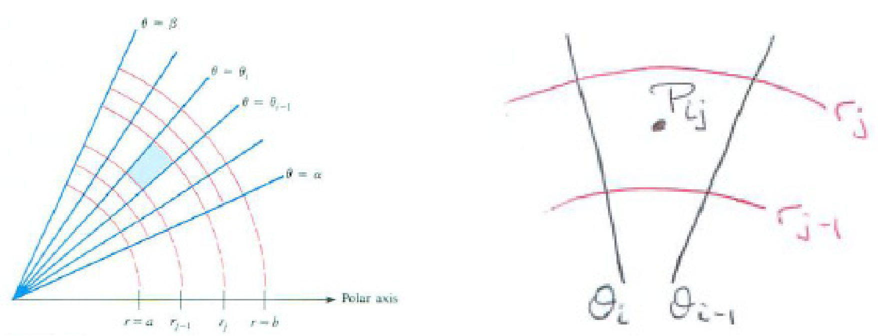
$$\int_0^{2\pi} \frac{a^3}{3}\, d\theta = \left.\frac{a^3\theta}{3}\right|_0^{2\pi} = \frac{2\pi a^3}{3}.$$

The total mass is $2\pi a^3/3$.

◇

A Fuller Explanation of the Extra r in the Integrand

Consider $\int_R f(P)\, dA$ as the region in the plane bound by the circle $r = a$ and $r = b$ and the range $\theta = \alpha$ and $\theta = \beta$. Break it into n^2 little pieces with the aid of the partitions $r_0 = a, r_1, r_i, r_n = b$ and $\theta_0 = \alpha, \theta_1, \theta_j, \theta_n = \beta$. For convenience, assume that all $r_i - r_{i-1}$ are equal to Δr and all $\theta_j - \theta_{j-1}$ are equal to $\Delta\theta$. (See Figure 17.3.9(a).) The typical patch, shown in Figure 17.3.9(b),



(a) (b)

Figure 17.3.9: (b) P_{ij} is $\left(\frac{r_j+r_{j+1}}{2}, \frac{\theta_j+\theta_{i-1}}{2}\right)$

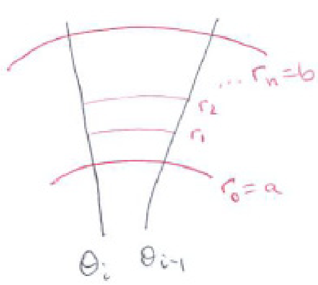
has area, exactly

$$A_ij = \frac{(r_j + r_{j-1})}{2}(r_j - r_{j-1})(\theta_i - \theta_{i-1}),$$

as shown in Exercise 6.

Then the sum of the n^2 terms of the form $f(P_{ij})A_{ij}$ is an estimate of $\int_R f(P)\, dA$.

Let us look closely at the summand for the n patches between the rays $\theta = \theta_{i-1}$ and $\theta = \theta_i$, as shown in Figure 17.3.10.



The sum is

$$\sum_{j=1}^{n} f\left(\frac{r_j + r_{j-1}}{2}, \frac{\theta_i + \theta_{i-1}}{2}\right) \frac{r_j + r_{j+1}}{2} \Delta r \Delta\theta. \qquad (17.3.3)$$

In (17.3.3), θ_i, θ_{i-1}, and $\Delta\theta$ are constants. If we define $g(r, \theta)$ to be $f(r, \theta)r$, then the sum is

$$\left(\sum_{i-1}^{n} g\left(\frac{r_j + r_{j+1}}{2}, \frac{\theta_i + \theta_{i-1}}{2}\right) \Delta r\right)\ \Delta\theta. \qquad (17.3.4)$$

The sum in brackets in (17.3.4) is an estimate of

$$\int_a^b g\left(r, \frac{\theta_j + \theta_{j-1}}{2}\right)\ dr.$$

Thus the sum, corresponding to the region between the rays $\theta = \theta_i$ and $\theta = \theta_{i-1}$, is

$$\sum_{i=1}^{n} \int_a^b g\left(r, \frac{\theta_i + \theta_{i-1}}{2}\right)\ dr\ \Delta\theta. \qquad (17.3.5)$$

Now let $h(\theta) = \int_a^b g(r, \theta)\ dr$. Then (17.3.5) equals

$$\sum_{i=1}^{n} h\left(\frac{\theta_i + \theta_{i-1}}{2}\right)\ \Delta\theta.$$

This is an estimate of $\int_a^b f(\theta)\ d\theta$. Hence the sum of all n^2 little terms of the form $f(P_{ij})A_{ij}$ is an approximation of

$$\int_\alpha^\beta h(\theta)\ d\theta \ = \ \int_\alpha^\beta \left(\int_a^b g(r, \theta)\ dr\right)\ d\theta = \int_\alpha^\beta \left(\int_a^b f(r, \theta) r\ dr\right)\ d\theta.$$

The extra factor r appears as we obtained the first integral, $\int_a^b f(r, \theta)r\ dr$. The sum of the n^2 terms A_{ij}, which we knew approximated the double integral $\int_R f(P)\ dA$, we now see approximate also the iterated integral (17.3.6). Taking limits as $n \to \infty$ show that the iterated integral equals the double integral.

Summary

We saw how to calculate an integral $\int_R f(P)\ dA$ by introducing polar coordinates. In this case, the plane region R can be described, in polar coordinates, as

$$\alpha \le \theta \le \beta, \qquad r_1(\theta) \le r \le r_2(\theta)$$

then

$$\int_R f(P)\ dA = \int_\alpha^\beta \int_{r_1(\theta)}^{r_2(\theta)} f(r,\theta) r\ dr\ d\theta.$$

The extra r in the integrand is due to the fact that a small region corresponding to changes dr and $d\theta$ has area area approximately $r\ dr\ d\theta$ (not $dr\ d\theta$). Polar coordinates are convenient when either the function f or the region R has a simple description in terms of r and θ.

EXERCISES for Section 17.3 *Key:* R–routine, M–moderate, C–challenging

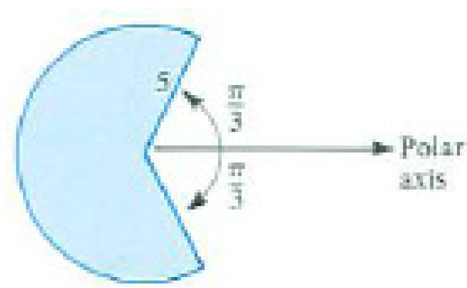
In Exercises 1 to 6 draw and describe the given regions in the form $\alpha \le \theta \le \beta$, $r_1(\theta) \le r \le r_2(\theta)$.

1.[R] The region inside the curve $r = 3 + \cos(\theta)$.

2.[R] The region between the curve $r = 3 + \cos(\theta)$ and the curve $r = 1 + \sin(\theta)$.

3.[R] The triangle whose vertices have the rectangular coordinates $(0, 0)$,$(1, 1)$, and $(1, \sqrt{3})$.

4.[R] The circle bounded by the curve $r = 3\sin(\theta)$.



5.[R] The region shown in Figure 17.3.11.

Figure 17.3.11:

6.[R] The region in the loop of the three-leaved rose, $r = \sin(3\theta)$, that lies in the first quadrant.

7.[R]

(a) Draw the region R bounded by the lines $y = 1$, $y = 2$, $y = x$, $y = x/\sqrt{3}$.

(b) Describe R in terms of horizontal cross sections,

(c) Describe R in terms of vertical cross sections,

(d) Describe R in terms of cross sections by polar rays.

8.[R]

(a) Draw the region R whose description is given by

$$-2 \le y \le 2, \qquad -\sqrt{4 - y^2} \le x \le \sqrt{4 - y^2}.$$

(b) Describe R by vertical cross sections.

(c) Describe R by cross sections obtained using polar rays.

9.[R] Describe in polar coordinates the square whose vertices have rectangular coordinates $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$.

10.[R] Describe the trapezoid whose vertices have rectangular coordinates $(0,1)$, $(1,1)$, $(2,2)$, $(0,2)$.

(a) in polar coordinates,

(b) by horizontal cross sections,

(c) by vertical cross sections.

In Exercises 5 to 14 draw the regions and evaluate $\int_R r^2\, dA$ for the given regions R.

11.[R] $-\pi/2 \le \theta \le \pi/2$, $0 \le r \le \cos(\theta)$

12.[R] $0 \le \theta \le \pi/2$, $0 \le r \le \sin^2(\theta)$

13.[R] $0 \le \theta \le 2\pi$, $0 \le r \le 1 + \cos(\theta)$

14.[R] $0 \le \theta \le 0.3$, $0 \le r \le \sin 2(\theta)$

In Exercises 15 to 18 draw R and evaluate $\int_R y^2\, dA$ for the given regions R.

15.[R] The circle of radius a, center at the pole.

16.[R] The circle of radius a with center at $(a,0)$ in polar coordinates.

17.[R] The region within the cardioid $r = 1 + \sin\theta$.

18.[R] The region within one leaf of the four-leaved rose $r = \sin 2\theta$.

In Exercises 19 and 20, use iterated integrals in polar coordinates to find the given point.

19.[R] The center of mass of the region within the cardioid $r = 1 + \cos(\theta)$.

20.[R] The center of mass of the region within the leaf $r = \cos 3(\theta)$ that lies along the polar axis.

The average of a function $f(P)$ over a region R in the plane is defined as $\int_R f(P)\, dA$ divided by the area of R. In each of Exercises 21 to 24, find the average of the given function over the given region.

21.[R] $f(P)$ is the distance from P to the pole; R is one leaf of the three-leaved rose, $r = \sin(3\theta)$.

22.[R] $f(P)$ is the distance from P to the x-axis; R is the region between the rays $\theta = \pi/6$, $\theta = \pi/4$, and the circles $r = 2$, $r = 3$.

23.[R] $f(P)$ is the distance from P to a fixed point on the border of a disk R of radius a. (HINT: Choose the pole wisely.)

24.[R] $f(P)$ is the distance from P to the x-axis; R is the region within the cardioid $r = 1 + \cos(\theta)$.

In Exercises 25 to 28 evaluate the given iterated integrals using polar coordinates. Pay attention to the elements of each exercise that makes it appropriate for evaluation in polar coordinates.

25.[R] $\int_0^1 \left(\int_0^x \sqrt{x^2+y^2}\ dy\right)\ dx$

26.[R] $\int_0^1 \left(\int_0^{\sqrt{1-x^2}} x^3\ dy\right)\ dx$

27.[R] $\int_0^1 \left(\int_x^{\sqrt{1-x^2}} xy\ dy\right)\ dx$

28.[R] $\int_1^2 \left(\int_{x/\sqrt{3}}^{\sqrt{3}x} (x^2+y^2)^{3/2}\ dy\right)\ dx$

29.[R] Evaluate the integrals over the given regions.

(a) $\int_R \cos(x^2+y^2)\ dA$; R is the portion in the first quadrant of the disk of radius a centered at the origin.

(b) $\int_R \sqrt{x^2+y^2}\ dA$; R is the triangle bounded by the line $y = x$, the line $x = 2$, and the x-axis.

30.[R] Find the volume of the region above the paraboloid $z = x^2 + y^2$ and below the plane $z = x + y$.

31.[R] The area of a region R is equal to $\int_R 1\ dA$. Use this to find the area of a disk of radius a. (Use an iterated integral in polar coordinates.)

32.[R] Find the area of the shaded region in Figure 17.3.4(b) as follows:

(a) Find the area of the ring between two circles, one of radius r_0, the other of radius $r_0 + \Delta r$.

(b) What fraction of the area in (a) is included between two rays whose angles differ by $\Delta\theta$?

(c) Show that the area of the shaded region in Figure 17.3.4(b) is precisely

$$\left(r_0 + \frac{\Delta r}{2}\right)\Delta r \Delta\theta.$$

33.[R] Evaluate the repeated integral

$$\int_{-\pi/2}^{\pi/2} \left(\int_0^{a\cos(\theta)} \sqrt{a^2 - r^2}\; r\, dr \right) d\theta$$

directly. The result should still be $a^3(3\pi - 4)/9$. (In Example 5 we computed half the volume and doubled the result.)

Caution: Use trigonometric formulas with care.

Prior to beginning Exercise 34, consider the following two quotes:

> Once when lecturing to a class he [the physicist Lord Kelvin] used the word "mathematician" and then interrupting himself asked the class: "Do you know what a mathematician is?" Stepping to his blackboard he wrote upon it: $\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}$. Then putting his finger on what he had written, he turned to his class and said, "A mathematician is one to whom this is as obvious as that twice two makes four is to you."

S. P. Thompson, in *Life of Lord Kelvin* (Macmillan, London, 1910).

> Many things are not accessible to intuition at all, the value of $\int_0^{\infty} e^{-x^2}\, dx$ for instance.

J. E. Littlewood, "Newton and the Attraction of the Sphere", *Mathematical Gazette*, vol. 63, 1948.

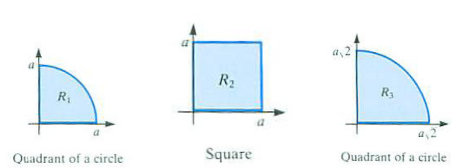
34.[M] This exercise shows that $\int_0^{\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}$. Let R_1, R_2, and R_3 be the three regions indicated in Figure 17.3.12, and $f(P) = e^{-r^2}$ where r is the distance from P to the origin. Hence, $f(r,\theta) = e^{-r^2}$ in polar coordinates and in rectangular coordinates $f(x,y) = e^{-x^2-y^2}$. NOTE: Observe that R_1 is inside R_2 and R_2 is inside R_3.

(a) Show that $\int_{R_1} f(P)\, dA = \frac{\pi}{4}\left(1 - e^{-a^2}\right)$ and that $\int_{R_3} f(P)\, dA = \frac{\pi}{4}\left(1 - e^{-2a^2}\right)$.

(b) By considering $\int_{R_2} f(P)\, dA$ and the results in (a), show that

$$\frac{\pi}{4}\left(1 - e^{-a^2}\right) < \left(\int_0^{\infty} e^{-x^2}\, dx \right)^2 < \frac{\pi}{4}\left(1 - e^{-2a^2}\right).$$

(c) Show that $\int_0^{\infty} e^{-x^2}\, dx = \frac{\sqrt{\pi}}{2}$.



(a) (b) (c)

Figure 17.3.12:

35.[R] Figure 17.3.13 shows the "bell curve" or "normal curve" often used to assign grades in large classes. Using the fact established in Exercise 34 that $\int_0^\infty e^{-x^2}\, dx = \sqrt{\pi/2}$, show that the area under the curve in Figure 17.3.13 is 1.

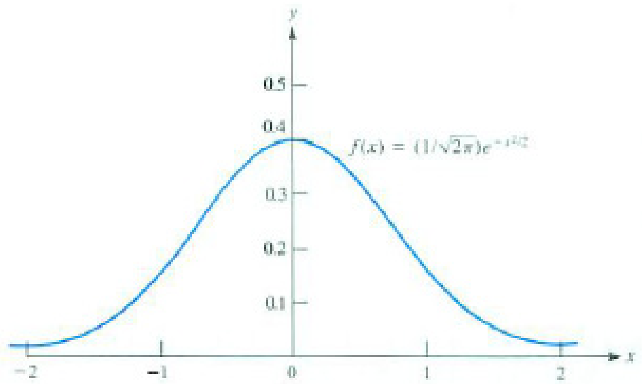


Figure 17.3.13:

36.[R] (The spread of epidemics.) In the theory of a spreading epidemic it is assumed that the probability that a contagious individual infects an individual D miles away depends only on D. Consider a population that is uniformly distributed in a circular city whose radius is 1 mile. Assume that the probability we mentioned is proportional to $2 - D$. For a fixed point Q let $f(P) = 2 - \overline{PQ}$. Let R be the region occupied by the city.

(a) Why is the exposure of a person residing at Q proportional to $\int_R f(P)\, dA$, assuming that contagious people are uniformly distributed throughout the city?

(b) Compute this definite integral when Q is the center of town and when Q is on the edge of town.

(c) In view of (b), which is the safer place?

Transportation problems lead to integrals over plane sets, as Exercises 37 to 40 illustrate.

37.[R] Show that the average travel distance from the center of a disk of area A to points in the disk is precisely $2\sqrt{A}/(3\sqrt{})\pi \approx 0.376\sqrt{A}$.

38.[R] Show that the average travel distance from the center of a regular hexagon of area A to points in the hexagon is

$$\frac{\sqrt{2A}}{3^{3/4}}\left(\frac{1}{3} + \frac{\ln 3}{4}\right) \approx 0.377\sqrt{A}.$$

39.[R] Show that the average travel distance from the center of a square of area A to points in the square is $(\sqrt{2} + \ln(\tan(3\pi/8)))\sqrt{A/6} \approx 0.383\sqrt{A}$.

40.[R] Show that the average travel distance from the centroid of an equilateral triangle of area A to points in the triangle is

$$\frac{\sqrt{A}}{3^{9/4}}\left(2\sqrt{3} + \ln(\tan(\frac{5\pi}{12}))\right) \approx 0.404\sqrt{A}$$

NOTE: The centroid of a triangle is its center of mass.

In Exercises 37 to 40 the distance is the ordinary straight-line distance. In cities the usual street pattern suggests that the "metropolitan" distance between the points (x_1, y_1) and (x_2, y_2) should be measured by $|x_1 - x_2| + |y_1 - y_2|$.

41.[M] Show that if in Exercise 37 metropolitan distance is used, then the average is $8\sqrt{A}/(3\pi^{3/2}) \approx 0.470\sqrt{A}$.

42.[M] Show that if in Exercise 40 metropolitan distance is used, then the average is $\sqrt{A/2}$. In most cities the metropolitan average tends to be about 25 percent larger than the direct-distance average.

43.[C]

Sam: The formula in this section for integrating in polar coordinates is wrong. I'll get the right formula. We don't need the factor r.

Jane: But the book's formula gives the correct answers.

Sam: I don't care. Let $f(r, \theta)$ be positive and I'll show how to integrate over the set R bounded by $r = b$ and $r = a$, $b > a$, and $\theta = \beta$ and $\theta = \alpha$. We have $\int_R f(P)\ dA$ is the volume under the graph of f and above R. Right?

Jane: Right.

Sam: The area of the cross-section corresponds to a fixed angle θ is $\int_a^b f(r, \theta)\ dr$. Right?

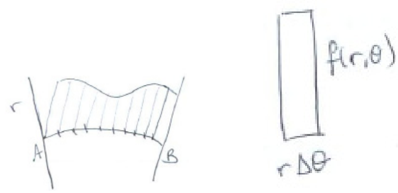
Jane: Right.

Sam: So I, just integrate cross-sectional areas as θ goes from α to β, and the volume is therefore $\int_\alpha^\beta (\int_a^b f(r, \theta)\ dr)\ d\theta$. Perfectly straightforward. I hate to overthrow a formula that's been around for three centuries.

What does Jane say next?

44.[C]

Jane: I won't use a partition. Instead, look at the area under the graph of f and above the circle of radius r. I'll draw this fence for you (see Figure 17.3.14(a).



(a) (b)

Figure 17.3.14:

To estimate its area I'll cut the arc AB into n sections of equal length by angle $\theta_0 = a \ldots$.

Then break AB into n short area, each of length $r\Delta\theta$. (Remember, Sam, how radians are defined.) The typical small approach to the shaded area looks like Figure 17.3.14(b). That's just an estimate of $\int_\alpha^\beta f(r,\theta) r\, d\theta$. Here r is fixed. Then I integrate the cross-sectional area as r goes from a to b. The total volume is then $\int_a^b \int_\alpha^\beta f(r,\theta) r\, d\theta\, dr$. But $\int_R f(r,\theta)\, dA$ is the volume.

Sam: All right.

Jane: At least it gives the r factor.

Sam: But you had to assume f is positive.

Jane: Well, if it isn't just add a big positive number k to f, then $g = f + k$ is positive. From then on its easy. If it's so far g it's so far f.

Check that Jane is right about g and f.

17.4 The Triple Integral: Integrals Over Solid Regions

In this section we define integrals over solid regions in space and show how to compute them by iterated integrals using rectangular coordinates. Throughout we assume the regions are bounded by smooth surfaces and the functions are continuous.

The Triple Integral

Let R be a region in space bounded by some surface. For instance, R could be a ball, a cube, or a tetrahedron. Let f be a function refined at least on R.

For each positive integer n break R into n small region $R_1, R_2, \ldots R_n$. Choose a point P, in R_1, P_2 in R_2, $\ldots$, P_n in R_n. Let the volume of R_i be V_i. Then

$$\lim_{n\to\infty} \sum_{i=1}^{n} f(P_i)V_i$$

exists. It is denoted

$$\int\limits_R f(P)\ dV \tag{17.4.1}$$

and is called the integral of f over R or the **triple integral** of f over R.

NOTE:

1. As in the preceding section, we define small. For each n let r_n be the smallest number such that each R_i in the partition fits inside a ball of radius r_n. We assume that $r_n \to 0$ as $n \to \infty$.

2. The notation $\int\int\int_R f(P)\ dV$ is commonly used, but, we stick to using one integral sign, $\int_R f(P)\ dV$ to emphasize that the triple integral is not a repeated integral.

3. The notation $\int\int\int f(x, y, z)\ dV$ is also used, but, again, we prefer not to refer to a particular coordinate system.

EXAMPLE 1 If $f(P) = 1$ for each point P in a solid region R, compute $\int_R f(P)\ dV$.

SOLUTION Each approximating sum $\sum_{i=1}^{n} f(P_i)V_i$ has the value

$$\sum_{i=1}^{n} 1 \cdot V_i = V_1 + V_2 + \cdots + V_n = \text{Volume of } R.$$

Hence

$$\int_R f(P)\ dV = \text{Volume of } R,$$

a fact that will be useful for computing volumes. ◇

Average of a function

The **average value** of a function f defined on a region R in space is defined as

$$\frac{\int_R 1\ dV}{\text{Volume of } R}.$$

This is the analog of the definition of the average of a function over an interval (Section 6.3) or the average of a function over a plane region (Section 17.1). If f describes the density of matter in R, then the average value of f is the density of a *homogeneous* solid occupying R and having the same total mass as the given solid.

Think about it. If the number

$$\frac{\int_R f(P)\ dV}{\text{Volume of } R}.$$

is multiplied by the volume of R, the result is

$$\int_R f(P)\ dV,$$

which is the total mass.

SHERMAN: I have a feeling I've read this before, but didn't find it in a quick search. Is this a repeat? If, should one be removed?

"Density" at a point is defined for lamina; with balls replacing disks. For a positive number r, let $m(r)$ be the mass in a ball with center P and radius r. Let $V(r)$ be the volume of the ball of radius r. Then the density at P is defined as

$$\lim_{r\to 0} \frac{m(r)}{V(r)}.$$

An Interpretation of $\int_R f(P)\ dV$.

Triple integrals appear in the study of gravitation, rotating bodies, centers of gravity, and electro-magnetic theory. The simplest way to think of them is to interpret $f(P)$ as the density at P of some disturbance of matter and, then, $\int_R f(P)\ dV$ is the total mass in a region R.

We can't picture $\int_R f(P)\ dV$ as measuring the volume of something. We could do this for $\int_R f(P)\ dA$, because we could use two dimensions for describing the region of integration and then the third dimension for the values of the function, obtaining a surface in three-dimensional space. However, with

$\int_R f(P)\, dV$, we use up three dimensions just describing the region of integration. We need four-dimensional space to show the values of the function. But it's hard to visualize such a space, no matter how hard we squint.

A Word about Four-Dimensional Space

We can think of 2-dimensional space as the set of ordered pairs (x, y) of real numbers. The set of ordered triplets of real numbers (x, y, z) represents 3-dimensional space. The set of ordered quadruplets of real numbers (x, y, z, t) represents 4-dimensional space.

It is easy to show 4-D space is a very strange place.

In 2-dimensional space the set of points of the form $(x, 0)$, the y-axis, meets the set of points of the form $(0, y)$, the y-axis, in a point, namely the origin $(0, 0)$. Now watch what can happen in 4-space. The set of points of the form $(x, y, 0, 0)$ forms a plane congruent to our familiar xy-plane. The set of points of the form $(0, 0, z, t)$ forms another such plane. So far, no surprise. But notice what the intersection of those two planes is. Their intersection is just the point $(0, 0, 0, 0)$. Can you picture two endless planes meeting in a single point? If so, please tell us how.

Describing a Solid Region

In order to evaluate triple integrals, it is necessary to describe solid regions in terms of coordinates.

A description of a typical solid region in rectangular coordinates has the form

This is the order x, y, then z. There are six possible orders, as you may check.

$$a \leq x \leq b, \qquad y_1(x) \leq y \leq y_2(x), \qquad z_1(x, y) \leq z \leq z_2(x, y).$$

The inequalities on x and y describe the "shadow" or projection of the region on the xy plane. The inequalities for z then tell how z varies on a line parallel to the z-axis and passing through the point (x, y) in the projection. (See Figure 17.4.1.)

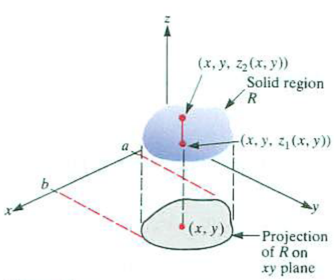
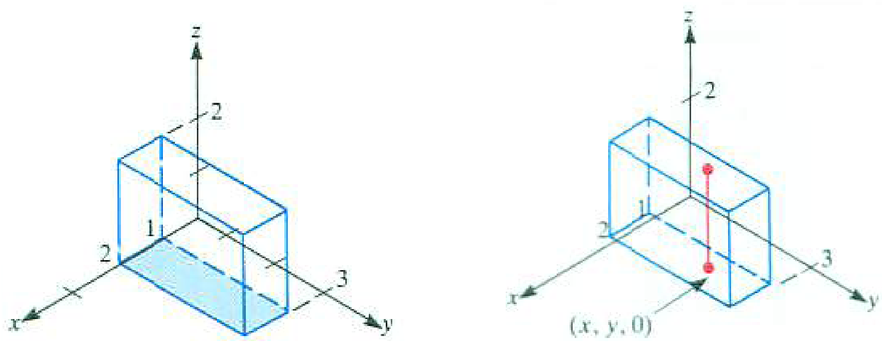


Figure 17.4.1:

EXAMPLE 2 Describe in terms of x, y, and z the rectangular box shown in Figure 17.4.2(a).

SOLUTION The shadow of the box on the xy plane has a description $1 \leq x \leq 2$, $0 \leq y \leq 3$. For each point in this shadow, z varies from 0 to 2, as shown in Figure 17.4.2(b). So the description of the box is

$$1 \leq x \leq 2, \qquad 0 \leq y \leq 3, \qquad 0 \leq z \leq 2,$$



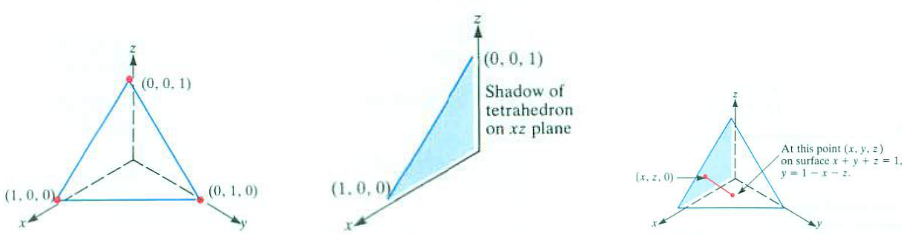
(a) (b)

Figure 17.4.2:

which is read from left to right as "x goes from 1 to 2; for each such x, the variable y goes from 0 to 3; for each such x and y, the variable z goes from 0 to 2."

Of course, we could have changed the order of x and y in the description of the shadow or projected the box on one of the other two coordinate planes. (All told, there are six possible descriptions.) ◇

EXAMPLE 3 Describe by cross sections the tetrahedron bounded by the planes $x = 0$, $y = 0$, $z = 0$, and $x + y + z = 1$, as shown in Figure 17.4.3(a).



(a) (b) (c)

Figure 17.4.3:

SOLUTION For the sake of variety, project the tetrahedron onto the xz plane. The shadow is shown in Figure 17.4.3(b). A description of the shadow is

$$0 \le x \le 1, \qquad 0 \le z \le 1 - x,$$

since the slanted edge has the equation $x + z = 1$. For each point (x, z) in this shadow, y ranges from 0 up to the value of y that satisfies the equation

$x+y+z=1$, that is, up to $y=1-x-z$. (See Figure 17.4.3(c).) A description of the tetrahedron is

$$0 \le x \le 1, \qquad 0 \le z \le 1-x, \qquad 0 \le y \le 1-x-z.$$

That is, x goes from 0 to 1; for each x, z goes from 0 to $1-x$; for each x and z, y goes from 0 to $1-x-z$. ◇

EXAMPLE 4 Describe in rectangular coordinates the ball of radius 4 whose center is at the origin.

SOLUTION The shadow of the ball on the xy plane is the disk of radius 4 and center $(0,0)$. Its description is

$$-4 \le x \le 4, \qquad -\sqrt{16-x^2} \le y \le \sqrt{16-x^2}.$$

Hold (x,y) fixed in the xy plane and consider the way z varies on the line parallel to the z-axis that passes through the point $(x,y,0)$. Since the sphere that bounds the ball has the equation

$$x^2+y^2+z^2=16,$$

for each appropriate (x,y), z varies from

$$-\sqrt{16-x^2-y^2} \qquad \text{to} \qquad \sqrt{16-x^2-y^2}.$$

This describes the line segment shown in Figure 17.4.4.

The ball, therefore, has a description

$$-4 \le x \le 4, \qquad -\sqrt{16-x^2} \le y \le \sqrt{16-x^2}, \qquad \sqrt{16-x^2-y^2} \le z \le \sqrt{16-x^2-$$

◇

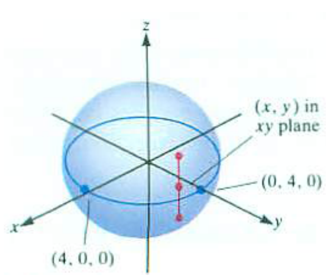


Figure 17.4.4:

Iterated Integrals for $\int_R f(P)\ dV$

The iterated integral in rectangular coordinates for $\int_R f(P)\ dV$ is similar to that for evaluating integrals over plane sets. It involves three integrations instead of two. The limits of integration are determined by the description of R in rectangular coordinates. If R has the description

$$a \le x \le b, \qquad y_1(x) \le y \le y_2(x), \qquad z_1(x,y) \le z \le z_2(x,y),$$

then

$$\int_R f(P)\ dV = \int_a^b \int_{y_1(x)}^{y_2(x)} \left(\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\ dz \right) dy\ dx.$$

An example illustrates how this formula is applied. In Exercise 31 an argument for its plausibility is presented.

EXAMPLE 5 Compute $\int_R z\, dV$, where R is the tetrahedron in Example 3.

SOLUTION A description of the tetrahedron is

$$0 \leq y \leq 1, \qquad 0 \leq x \leq 1-y, \qquad 0 \leq z \leq 1-x-y.$$

Hence

$$\int_R z\ dV = \int_0^1 \left(\int_0^{1-y} \left(\int_0^{1-x-y} z\ dz \right) dx \right) dy.$$

Compute the inner integral first, treating x and y as constants. By the Fundamental Theorem,

$$\int_0^{1-x-y} z\ dz = \left.\frac{z^2}{2}\right|_{z=0}^{z=1-x-y} = \frac{(1-x-y)^2}{2}.$$

The next integration, where y is fixed, is

$$\int_0^{1-y} \frac{(1-x-y)^2}{2}\ dx = -\left.\frac{(1-x-y)^3}{6}\right|_{x=0}^{x=1-y} = -\frac{0^3}{6} + \frac{(1-y)^3}{6} = \frac{(1-y)^3}{6}.$$

The third integration is

$$\int_0^1 \frac{(1-y)^3}{6}\ dy = -\left.\frac{(1-y)^4}{24}\right|_0^1 = -\frac{0^4}{24} + \frac{1^4}{24} = \frac{1}{24}.$$

This completes the calculation that

$$\int_R z\ dV = \frac{1}{24}.$$

◇

Summary

We defined $\int_R f(P)\ dV$, where R is a region in space. The volume of a solid region R is $\int_R\ dV$ and, if $f(P)$ is the density of matter near P, then $\int_R f(P)\ dV$ is the total mass. We also showed how to evaluate these integrals by introducing rectangular coordinates.

There are six possible orders.

The general approach is to, first, describe R, for instance, as

$$a \leq x \leq b, \qquad y_1(x) \leq y \leq y_2(x), \qquad z_1(x,y) \leq z \leq z_2(x,y).$$

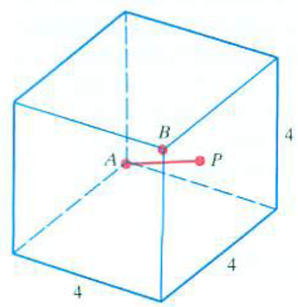
Then

$$\int_R f(P)\ dV = \int_a^b \left(\int_{y_1(x)}^{y_2(x)} \left(\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\ dx \right)\ dy \right)\ dx.$$

EXERCISES for Section 17.4 *Key:* R–routine, M–moderate, C–challenging

Exercises 1 to 4 concern the definition of $\int_R f(P)\, dV$.

1.[R] A cube of side 4 centimeters is made of a material of varying density. Near one corner A it is very light; at the opposite corner it is very dense. In fact, the density $f(P)$ (in grams per cubic centimeter) at any point P in the cube is the square of the distance from A to P (in centimeters). See Figure 17.4.5.



The density at P is the square of the distance $\overline{AP}$. P is a typical point in the cube.

Figure 17.4.5:

(a) Find upper and lower estimates for the mass of the cube by partitioning it into eight cubes.

(b) Using the same partition as in (a), estimate the mass of the cube, but select as the P_i's the centers of the four rectangular boxes.

(c) Estimate the mass of the cube described in the opening problem by cutting it into eight congruent cubes and using their centers as the P_i's.

(d) What does (c) say about the average density in the cube?

2.[R] How would you define the average distance from points of a certain set in space to a fixed point P_0?

3.[R] If R is a ball of radius r and $f(P) = 5$ for each point in R, compute $\int_R f(P)\, dV$ by examining approximating sums. Recall that the ball has volume $4/3\pi r^3$.

4.[R] If R is a three-dimensional set and $f(P)$ is never more than 8 for all P in R.

(a) what can we say about the maximum possible value of $\int_R f(P)\, dV$?

(b) what can we say about the average of f over R?

In Exercises 5 to 10 draw the solids described.

5.[R] $1 \le x \le 3$, $0 \le y \le 2$, $0 \le z \le x$

6.[R] $0 \le x \le 1$, $0 \le y \le 1$, $1 \le z \le 1 + x + y$

7.[R] $0 \le y \le 1$, $0 \le x \le y^2$, $y \le z \le 2y$

8.[R] $0 \le y \le 1$, $y^2 \le x \le y$, $0 \le z \le x + y$

9.[R] $-1 \le z \le 1$, $-\sqrt{1-z^2} \le x \le \sqrt{1-z^2}$, $-\frac{1}{2} \le y \le \sqrt{1-x^2-z^2}$

10.[R] $0 \le z \le 3$, $0 \le y \le \sqrt{9-z^2}$, $0 \le x \le \sqrt{9-y^2-z^2}$

In Exercises 11 to 14 evaluate the iterated integrals.

11.[R] $\int_0^1 \left(\int_0^2 \left(\int_0^x z\, dz \right)\, dy \right)\, dx.$

12.[R] $\int_0^1 \left(\int_{x^3}^{x^2} \left(\int_0^{x+y} z\, dz \right)\, dy \right)\, dx.$

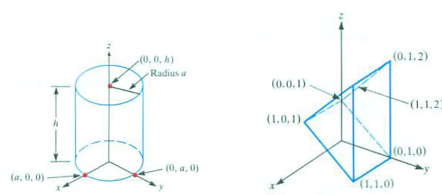
13.[R] $\int_2^3 \left(\int_x^{2x} \left(\int_0^1 (x+z)\, dz \right)\, dy \right)\, dx.$

14.[R] $\int_0^1 \left(\int_0^x \left(\int_0^3 (x^2+y^2)\, dz \right)\, dy \right)\, dx.$

15.[R] Describe the solid cylinder of radius a and height h shown in Figure 17.4.6(a) in rectangular coordinates

(a) in the order first x, then y, then z,

(b) in the order first x, then z, then y.



(a) (b)

Figure 17.4.6:

16.[R] Describe the prism shown in Figure 17.4.6(b) in rectangular coordinates, in two ways:

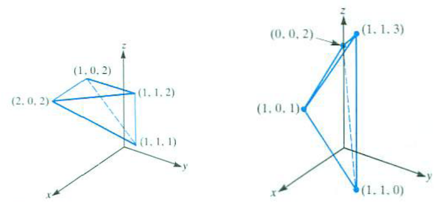
(a) First project it onto the xy plane.

(b) First project it onto the xz plane.

17.[R] Describe the tetrahedron shown in Figure 17.4.7(a) in rectangular coordinates in two ways:

(a) First project it onto the xy plane.

(b) First project it onto the xz plane.



(a) (b)

Figure 17.4.7:

18.[R] Describe the tetrahedron whose vertices are given in Figure 17.4.7(b) in rectangular coordinates as follows:

(a) Draw its shadow on the xy plane.

(b) Obtain equations of its top and bottom planes.

(c) Give a parametric description of the tetrahedron.

19.[R] Let R be the tetrahedron whose vertices are $(0,0,0)$, $(a,0,0)$, $(0,b,0)$, and $(0,0,c)$, where a, b, and c are positive.

(a) Sketch the tetrahedron.

(b) Find the equation of its top surface.

(c) Compute $\int_R z\, dV$.

20.[R] Compute $\int_R z\, dV$, where R is the region above the rectangle whose vertices are $(0,0,0)$, $(2,0,0)$, $(2,3,0)$, and $(0,3,0)$ and below the plane $z = x + 2y$.

21.[R] Find the mass of the cube in Exercise 1. (See Figure 17.4.1.)

22.[R] Find the average value of the square of the distance from a corner of a cube of side a to points in the cube.

23.[R] Find the average of the square of the distance from a point P in a cube of side a to the center of the cube.

24.[R] A solid consists of all points below the surface $z = xy$ that are above the triangle whose vertices are $(0, 0, 0)$, $(1, 0, 0)$, and $(0, 2, 0)$. If the density at (x, y, z) is $x + y$, find the total mass.

25.[R] Compute $\int_R xy\ dV$ for the tetrahedron of Example 3.

26.[R]

(a) Describe in rectangular coordinates the right circular cone of radius r and height h if its axis is on the positive z-axis and its vertex is at the origin. Draw the cross sections for fixed x and fixed x and y.

(b) Find the z coordinate of its centroid.

27.[R] The temperature at the point (x, y, z) is e^{-x-y-z}. Find the average temperature in the tetrahedron whose vertices are $(0, 0, 0)$, $(1, 1, 0)$, $(0, 0, 2)$, and $(1, 0, 0)$.

28.[R] The temperature at the point (x, y, z), $y > 0$, is $e^{-x}/\sqrt{y}$. Find the average temperature in the region bounded by the cylinder $y = x^2$, the plane $y = 1$, and the plane $z = 2y$.

29.[R] Without using a repeated integral, evaluate $\int_R x\ dV$, where R is a spherical ball whose center is $(0, 0, 0)$ and whose radius is a.

30.[R] The work done in lifting a weight of w pounds a vertical distance of x feet is wx foot-pounds. Imagine that through geological activity a mountain is formed consisting of material originally at sea level. Let the density of the material near point P in the mountain be $g(P)$ pounds per cubic foot and the height of P be $h(P)$ feet. What definite integral represents the total work expended in forming the mountain? This type of problem is important in the geological theory of mountain formation.

31.[R] In Section 17.2 an intuitive argument was presented for the equality

$$\int_R f(P)\ dA = \int_a^b \left(\int_{y_1(x)}^{y_2(x)} f(x, y)\ dy \right) dx.$$

Here is an intuitive argument for the equality

$$\int_R f(P)\ dV = \int_{x_1}^{x_2} \left(\int_{y_1(x)}^{y_2(x)} \left(\int_{z_1(x,y)}^{z_2(x,y)} f(x, y, z)\ dz \right) dy \right) dx.$$

To start, interpret $f(P)$ as "density."

(a) Let $R(x)$ be the plane cross section consisting of all points in R with abscissa x. Show that the average density in $R(x)$ is

$$\frac{\int^{y_2(x)}{}_{y_1(x)}[\left(\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\ dz\right)\ dy}{\text{Area of } R(x)}$$

(b) Show that the mass of R between the plane sections $R(x)$ and $R(x+\Delta x)$ is approximately

$$\int_{y_1(x)}^{y_2(x)} \left(\int_{z_1(x,y)}^{z_2(x,y)} f(x,y,z)\ dz \right)\ dy\ \Delta x.$$

(c) From (b) obtain a repeated integral in rectangular coordinates for $\int_R f(P)\ dV$.

17.5 Cylindrical and Spherical Coordinates

Rectangular coordinates provide convenient descriptions of solids bounded by planes. In this section we describe two other coordinate systems, cylindrical – ideal for describing circular cylinders — and spherical — ideal for describing spheres, balls, and cones. Both will be used in the next section to evaluate multiple integrals by iterated integrals.

CYLINDRICAL COORDINATES

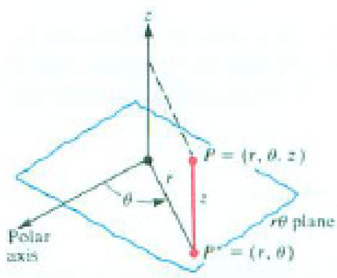
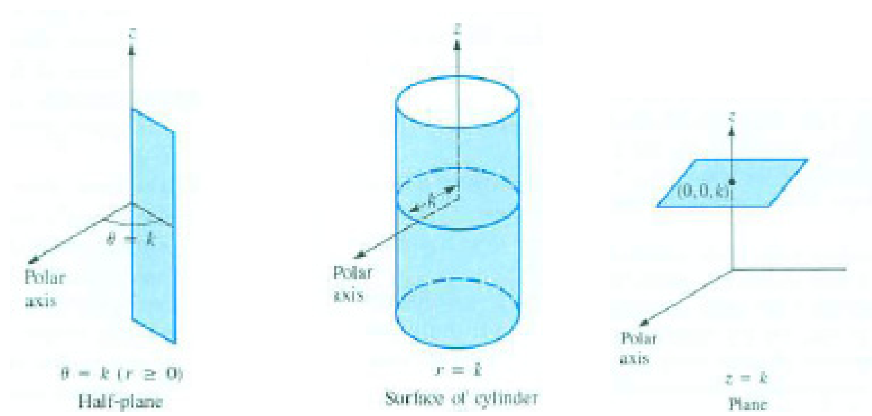


Figure 17.5.1:

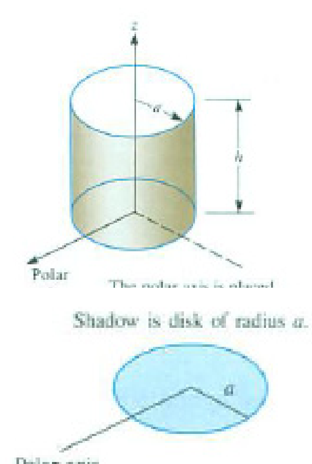
Cylindrical coordinates combine polar coordinates in the plane with the z of rectangular coordinates in space. Each point P in space receives the name (r, θ, z) as in Figure 17.5.1. We are free to choose the direction of the polar axis; usually it will coincide with the x-axis of an (x, y, z) system. Note that (r, θ, z) is directly above (or below) $P^* = (r, \theta)$ in the $r\theta$ plane. Since the set of all points $P = (r, \theta, z)$ for which r is some constant is a circular cylinder, this coordinate system is convenient for describing such cylinders. Just as with polar coordinates, cylindrical coordinates of a point are not unique.



(a) (b) (c)

Figure 17.5.2:

Figure 17.5.2 shows the surfaces $\theta = k$, $r = k$, and $z = k$, where k is a positive number.



EXAMPLE 1 Describe a solid cylinder of radius a and height h in cylindrical coordinates. Assume that the axis of the cylinder is on the positive z-axis and the lower base has its center at the pole, as in Figure 17.5.3.

SOLUTION The shadow of the cylinder on the $r\theta$ plane is the disk of radius

a with center at the pole shown in Figure 17.5.4. Its description is

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a.$$

For each point (r, θ) in the shadow, the line through the point parallel to the z-axis intersects the cylinder in a line segment. On this segment z varies from 0 to h for every (r, θ). (See Figure 17.5.5.) Thus a description of the cylinder is

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a, \qquad 0 \le z \le h.$$

◇

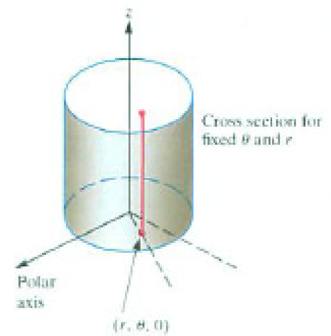


Figure 17.5.5:

EXAMPLE 2 Describe in cylindrical coordinates the region in space formed by the intersection of a solid cylinder of radius 3 with a ball of radius 5 whose center in on the axis of the cylinder. Place the cylindrical coordinate system as shown in Figure 17.5.6.

SOLUTION Note that the point $P = (r, \theta, z)$ is a distance $\sqrt{r^2 + z^2}$ from the origin O, for, by the pythagorean theorem, $r^2 + z^2 = \bar{OP}^2$. (See Figure 17.5.7.) We will use this fact in a moment.

Now consider the description of the solid. First of all, θ varies from 0 to 2π and r from 0 to 3, bounds determined by the cylinder. For fixed θ and r, the cross section of the solid is a line segment determined by the sphere that bounds the ball, as shown in Figure 17.5.7(b). Now, since the sphere has radius 5, for any point (r, θ, z) on it,

$$r^2 + z^2 = 25 \qquad \text{or} \qquad z \pm \sqrt{25 - r^2}.$$

Thus, on the line segment determined by fixed r and θ, z varies from $-\sqrt{25 - r^2}$ to $\sqrt{25 - r^2}$.

The solid has this description:

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le 3, \qquad -\sqrt{25 - r^2} \le z \le \sqrt{25 - r^2}.$$

◇

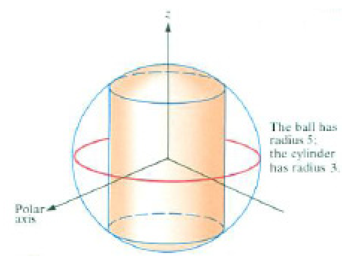


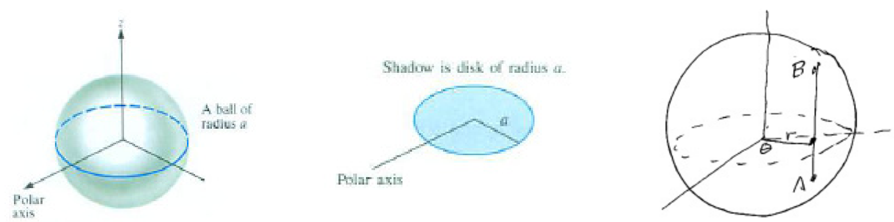
Figure 17.5.6:

EXAMPLE 3 Describe a ball of radius a in cylindrical coordinates.

SOLUTION Place the origin at the center of the ball, as in Figure 17.5.7(a). The shadow of the ball on the (r, θ) plane is a disk of radius a, shown in Figure 17.5.7(b) in perspective. This shadow is described by the equations

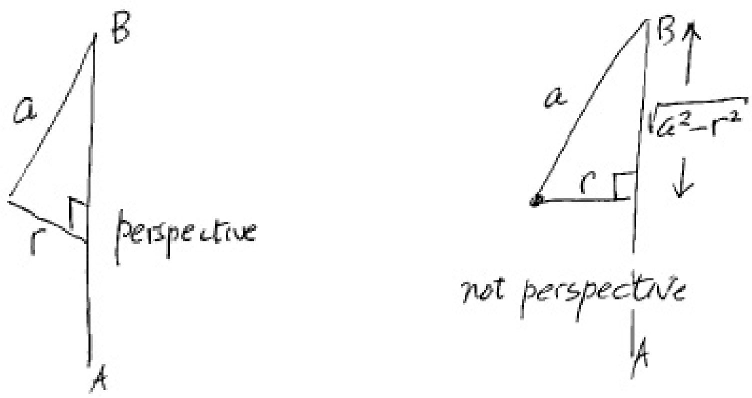
$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a.$$

All that is left is to see how z varies for a given r and θ. In other words, how does z vary on the line AB in Figure 17.5.7(c)?



(a) (b) (c)

Figure 17.5.7:



(a) (b)

Figure 17.5.8:

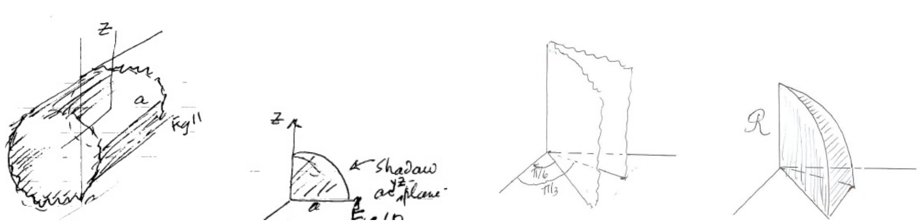
If r is a, then z "varies" from 0 to 0, as Figure 17.5.7(c) shows. If r is 0, then z varies from $-a$ to a. The bigger r is, the shorter AB is. Figure 17.5.8 presents the necessary geometry, first in perspective. With the aid of Figure 17.5.8, we see that z varies from $-\sqrt{a^2-r^2}$ to $\sqrt{a^2-r^2}$. You can check this by testing the easy cases, $r=0$ and $r=a$. All told,

$$\underbrace{0 \le \theta \le 2\pi, 0 \le r \le a}_{\text{The shadow}}, \underbrace{-\sqrt{a^2-r^2} \le z \le \sqrt{a^2-r^2}}_{\text{Range of } z \text{ for each } \theta \text{ and } r}$$

◇

EXAMPLE 4 Draw the region R bounded by the surfaces $r^2+z^2=a^2$, $\theta=\pi/6$, and $\theta=\pi/3$, situated in the first octant.

SOLUTION In the rz-plane, $r^2+z^2=a^2$ describes a circle of radius a, center at the origin. There is no restriction on θ. Thus it is a circular cylinder with its axis along the polar axis, as shown in Figure 17.5.9(a) in perspective. The shadow of R, which lies in the first octant, on the rz-plane is a quarter circle, shown in Figure 17.5.9(b).



(a) (b) (c) (d)

Figure 17.5.9:

Next we draw the half planes $\theta=\pi/6$ and $\theta=\pi/3$, as in Figure 17.5.9(c) showing at least the part in the first octant.

Finally we put Figure 17.5.9(a) and (c) together in (d), to show R.

R has three planar surfaces and one curved surface. The two curved edges are parts of ellipses, not parts of circles.

The description of R is

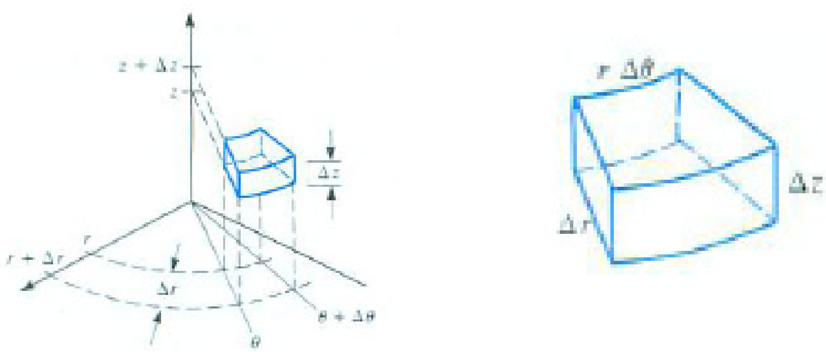
$$0 \le r \le a, \qquad 0 \le z \le \sqrt{a^2-r^2}, \qquad \pi/6 \le \theta \le \pi/3.$$

◇

Note that the shading and dashed hidden line help make the diagram clearer.

THE VOLUME SWEPT OUT BY Δr, Δ, and $\Delta\theta$

To use polar coordinates to evaluate an integral over a plane set we needed to know that the area of the little region corresponding to small changes Δr and $\Delta\theta$ is roughly $r\Delta r\Delta\theta$. In order to evaluate integrals over solids using an iterated integral in cylindrical coordinates, we will need to estimate the volume of the small region correspond to small changes Δr, $\Delta\theta$, Δz in the three coordinates.



(a) (b)

Figure 17.5.10:

The set of all points (r, θ, z) whose r coordinates are between r and $r + \Delta r$, whose θ coordinates are between θ and $\theta + \Delta\theta$, and whose z coordinates are between z and $z + \Delta z$ is shown in Figure 17.5.10(a). It is a solid with four flat surfaces and two curved surfaces.

When Δr is small, the area of the flat base of the solid is approximately $r\Delta r\Delta\theta$, as shown in Section 9.2 and as we saw when working with polar coordinates in the plane. Thus, when Δr, $\Delta\theta$, and Δz are small, the volume ΔV of the solid in Figure 17.5.10(b) is

$$\Delta V = (\text{Area of base})(\text{height}) \approx r\Delta r\Delta\theta\Delta z.$$

That is,

$$\Delta V \approx r\Delta r\Delta\theta\Delta z.$$

Just as the factor r appears in iterated integrals in polar coordinates, the same factor appears in iterated integrals in cylindrical coordinates.

SPHERICAL COORDINATES

The third standard coordinate system in space is **spherical coordinates**, which combines the θ of cylindrical coordinates with two other coordinates.

In spherical coordinates a point P is described by three numbers:

ρ is pronounced "row" or "roe"; it is the Greek letter for r. The letter ϕ is pronounced "fee" or "fie."

> ρ the distance from P to the origin O, θ the same angle as in cylindrical coordinates, ϕ the angle between the positive z-axis and the ray from O to P.

In physics and engineering r is used instead of ρ.

The point P is denoted $P = (\rho, \theta, \phi)$. Note the order: first ρ, then θ, then ϕ. See Figure 17.5.11. Note that ϕ is the same as the direction angle of $\overline{OP}$ with k, $0 \leq \phi \leq \pi$. The surfaces $\rho = k$ (a sphere), $\phi = k$ (a cone), and $\theta = k$ (a half plane) are shown in Figure 17.5.12.

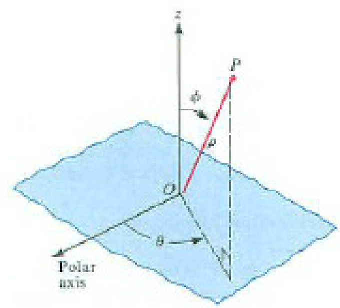
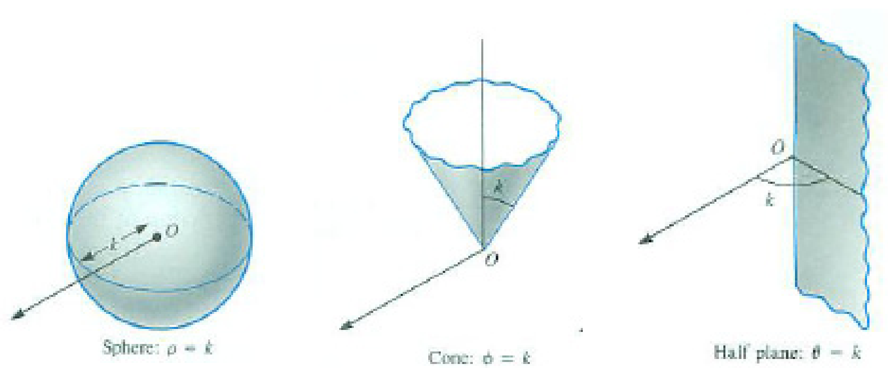


Figure 17.5.11:



(a) (b) (c)

Figure 17.5.12: (a) θ and ϕ vary, (b) ρ and θ vary, (c) ρ and ϕ vary.

When ϕ and θ are fixed and ρ varies, we describe a ray, as shown in Figure 17.5.13.

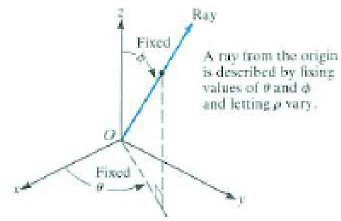


Figure 17.5.13:

RELATION TO RECTANGULAR COORDINATES

Figure 17.5.14 displays the relation between spherical and rectangular coordinates of a point $P = (\rho, \theta, \phi) = (x, y, z)$.

Note, in particular, right triangle OSP has hypotenuse OP and a right angle at S, and right triangle OQR has a right angle at Q.

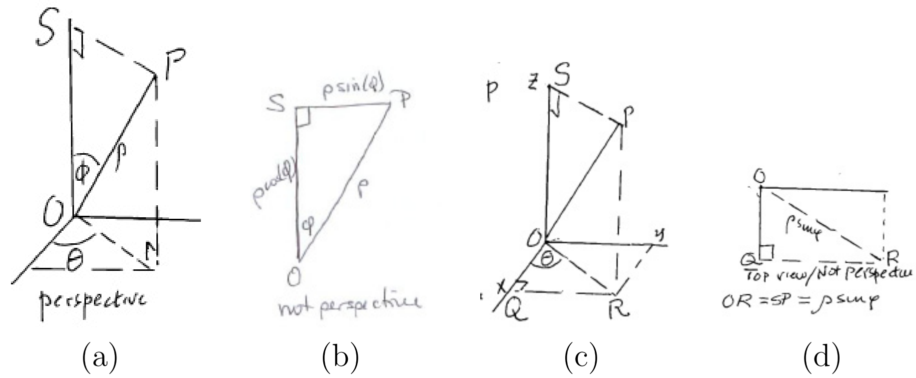


Figure 17.5.14:

First of all, $z = \rho\cos(\phi)$. Then $OR = \rho\sin(\phi)$. Finally $x = OR\cos(\theta) = \rho\sin(\phi)\cos(\theta)$ And $y = OR\sin(\theta) = \rho\sin(\phi)\sin(\theta)$.

EXAMPLE 5 Figure 17.5.15 shows a point given in spherical coordinates. Find its rectangular coordinates.

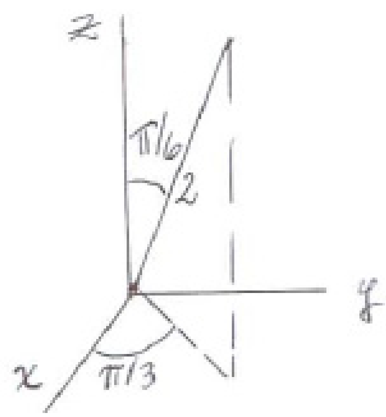


Figure 17.5.15:

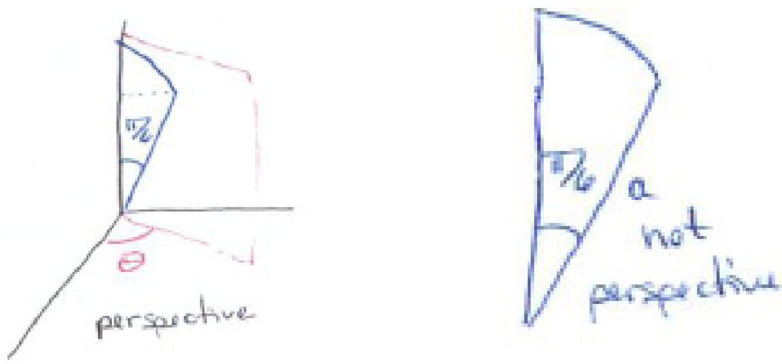
SOLUTION In this case, $\rho = 2$, $\theta = \pi/3$, $\phi = \pi/6$. Thus

$$\begin{aligned} x &= 2\sin(\pi/6)\cos\frac{\pi}{3} = 2\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{2} \\ y &= 2\sin(\frac{\pi}{6})\sin(\frac{\pi}{3}) = 2\cdot\frac{1}{2}\cdot\frac{\sqrt{3}2}{=}\frac{\sqrt{3}}{2} \\ z &= 2\cos(\frac{\pi}{6} = 2\frac{\sqrt{3}}{2} = \sqrt{3}. \end{aligned}$$

As a check, $x^2 + y^2 + z^2$ should equal ρ^2, and it does, for $(1/2)^2 + (\frac{\sqrt{3}}{2}) + (\sqrt{3})^2 = \frac{1}{4} + \frac{3}{4} + 3 = 4 = 2^2$. ◇

The next example exploits spherical coordinates to describe a cone and a ball.

EXAMPLE 6 The region R consists of the portion of a ball of radius a that lies within a cone of half angle $\pi/6$. The vertex of the cone is at the center of the ball.



(a) (b)

Figure 17.5.16:

SOLUTION R is shown in Figure 17.5.17. It resembles an ice cream cone, the dry cone topped with spherical ice cream.

Because R is a solid of revolution (around the z-axis), $0 \leq \theta \leq 2\pi$. The section of R corresponding to a fixed angle θ is the intersection of R with a half plane, shown in Figure 17.5.16.

In this sector of a disk, ϕ goes from 0 to $\pi/6$, independent of θ. Finally, a fixed θ and ϕ determine a ray on which ρ goes from 0 to a, as in Figure 17.5.18. ◇

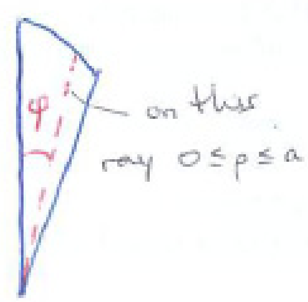


Figure 17.5.18:

The next example describes a ball in rectangular and spherical coordinates.

EXAMPLE 7 Describe a ball of radius a in rectangular and spherical coordinates.

SOLUTION In each case we put the origin of the coordinate system at the center of the ball.

Rectangular coordinates: The shadow of the ball on the xy-plane is a disk of radius a, described by

$$-a \leq x \leq a, \qquad -\sqrt{a^2 - x^2} \leq y \leq \sqrt{a^2 - x^2}.$$

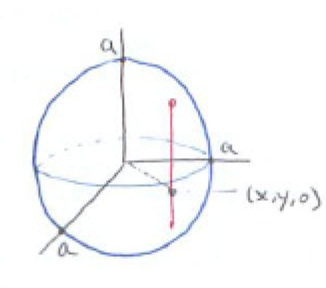


Figure 17.5.19:

For each point (x, y) in that projection, z varies along the line AB in Figure 17.5.19.

Since the equation of the sphere is $x^2+y^2+z^2 = a^2$ at A, z is $-\sqrt{z^2 - x^2 - y^2}$, and at B is $\sqrt{a^2 - x^2 - y^2}$. The entire description is

$$-a \leq x \leq a, \qquad -\sqrt{a^2 - x^2} \leq y \leq \sqrt{a^2 - x^2}, \qquad -\sqrt{a^2 - x^2 - y^2} \leq z \leq \sqrt{z^2 - x^2$$

Spherical coordinates: This time the shadow on the xy-plane plays no role. Instead, we begin with

$$0 \leq \theta \leq 2\pi, \qquad 0 \leq \phi \leq \pi,$$

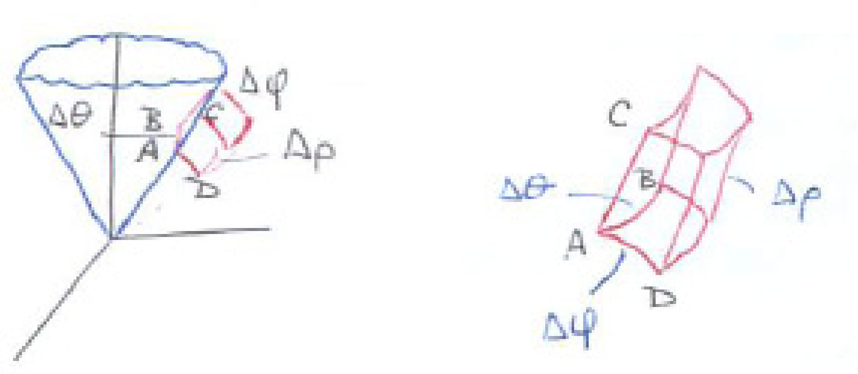
which sweeps out all the rays from the origin. On each such ray ρ goes from 0 to a. The complete description involves only constants as bounds:

$$0 \leq \theta \leq 2\pi, \qquad 0 \leq \phi \leq \pi, \qquad 0 \leq \rho \leq a.$$

Since the range of each variable is not influenced by other variables, the three restraints can be given in any order. ◇

THE VOLUME SWEPT OUT BY Δrho, Δphi, and $\Delta\theta$

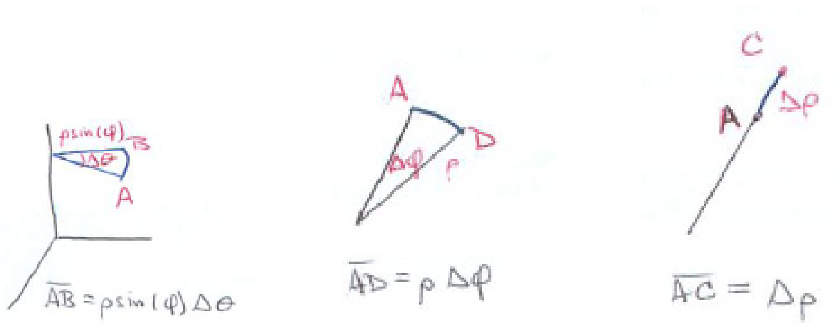
In the next section we will need an estimate of the volume of the little curvy "box-like" region bounded by spheres with radii ρ and $\rho + \Delta\rho$, the half-planes with angles θ and $\theta + \Delta\theta$, and the cones with half-angles ϕ and $\phi + \Delta\phi$. This region is shown in Figure 17.5.20. Two of its surfaces are flat, two are spherical, and two are patches on cones.



(a) (b)

Figure 17.5.20:

AB and AD are arcs of circles, while AC is straight



(a) (b) (c)

Figure 17.5.21:

The product of the length of AB, AC and AD is an estimate of the volume of the little box. Figure 17.5.21 shows how to find these lengths.

Therefore the volume of the small box is approximately $(\rho\sin(\phi)\Delta\theta)(\rho\Delta\phi)(\Delta\rho)$:

$$\Delta V \approx \rho^2 \sin(\phi)\Delta\rho\Delta\phi\Delta\theta$$

Just as we added an r to an integrand in polar coordinates, we must, in the next section, and the factor $\rho^2\sin(\phi)$ to an integrand when using an iterated integral in spherical coordinates.

Summary

This section described cylindrical and spherical coordinates. The volume of the small box corresponding to small changes in the three cylindrical coordinates is approximately $r\Delta r\Delta\theta\Delta z$. Because of the presence of the factor r, we must adjoin an r to the integrand when using an iterated integral in cylindrical coordinates.

Similarly, $\rho^2\sin(\phi)$ must be added to an integrand when using an iterated integral in spherical coordinates.

The next section illustrates the computations using these coordinates.

EXERCISES for Section 17.5 *Key:* R–routine, M–moderate, C–challenging

DOUG: Perhaps there should be examples and exercises with the bounds involving variables more?? (See Stewart)

1.[R] On the region in Example 2 draw the set of points described by (a) $z = 2$, (b) $z = 3$, (c) $z = 4.5$.

2.[R] For the cylinder in Example 1 draw the set of points described by (a) $r = a/2$, (b) $\theta = \pi/4$, (c) $z = h/3$.

3.[R]

(a) In the formula $\Delta V \approx r\Delta r\Delta\theta\Delta z$, which factors have the dimension of length?

(b) Why would you expect three such factors?

4.[R]

(a) In the formula $\Delta V \approx \rho^2\Delta\rho\Delta\theta\Delta\phi$, which factors have the dimension of length?

(b) Why would you expect three such factors?

5.[R] Drawing one clear, large diagram, show how to express rectangular coordinates in terms of cylindrical coordinates.

6.[R] Drawing one clear, large diagram, show how to express rectangular coordinates in terms of spherical coordinates.

7.[R] Find the cylindrical coordinates of $(x, y, z) = (3, 3, 1)$, including a clear diagram.

8.[R] Find the spherical coordinates of $(x, y, z) = (3, 3, 1)$,including a clear diagram.

In Exercises 9 to 11 (a) draw the set of points described, and (b) describe that set in words.

9.[R] ρ and ϕ fixed, θ varies.

10.[R] ρ and θ fixed, ϕ varies.

11.[R] θ and ϕ fixed, ρ varies.

12.[R] What is the equation of a sphere of radius a centered at the origin in

(a) spherical,

(b) cylindrical ,

(c) rectangular coordinates?

13.[R] Explain why if $P = (x, yz) = (\rho, \theta, \phi)$, in spherical coordinates, that $x^2 + y^2 + z^2 = \rho^2$. HINT: Draw a box.

14.[R] Describe the region in Example 6 in cylindrical coordinates in the order $\alpha \leq \theta \leq \beta$, $r_1(\theta) \leq r \leq r_2(\theta)$, $z_1(r, \theta) \leq z \leq z_2(r, \theta)$.

15.[R] Like Exercise 14, but in the order $a \leq z \leq b$, $\theta_1(z) \leq \theta \leq \theta_2(z)$, $r_1(\theta, z) \leq r \leq r_2(\theta, z)$.

16.[R] Sketch the region in the first octant bounded by the planes $\theta = \frac{\pi}{6}$ and $\theta = \frac{\pi}{3}$ and the sphere $\rho = a$.

17.[R] Estimate the area of the bottom face of the curvy box shown in Figure 17.5.20. It lies on the sphere of radius ρ.

18.[A] cone of half-angle $\pi/6$ is cut by a plane perpendicular to its axis at a distance 4 from its vertex.

(a) Place it conveniently on a cylindrical coordinate system.

(b) Describe it in cylindrical coordinates.

19.[R] Like the preceding exercise, but use spherical coordinates.

20.[R] A cone has its vertex at the origin and its axis along the positive z-axis. It is made by revolving a line through the origin that has an angle A with the z-axis, about the z-axis. Describe it in

(a) spherical coordinates,

(b) cylindrical coordinates, and

(c) rectangular coordinates.

21.[R] Use spherical coordinates to describe the surface in Figure 17.5.22. It is part of a cone of half vertex angle B with the z-axis as its axis, situated within a sphere of radius a centered at the origin.

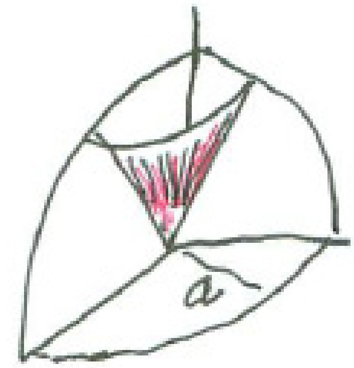


Figure 17.5.22:

22.[R] A triangle ABC is inscribed in a circle, with AB a diameter of the circle.

(a) Using elementary geometry, show that angle ACB is a right angle.

(b) Instead, using the equation of a circle in rectangular coordinates, show that AC and BC are perpendicular.

(c) Use (a) or (b) to show that the graph in the plane of $r = b\cos(\theta)$ is a circle of diameter b.

(d) In view of the preceding exercise, show that the equation of the circle in Figure 17.5.22 is $r = 2a\cos(\theta)$.

23.[R] (See Exercise 22.) A ball of radius a has a diameter coinciding with the interval $[0, 2a]$ on the x-axis. Describe the ball in spherical coordinates.

24.[R] The ray described in spherical coordinates by $\theta = \frac{\pi}{6}$ and $\phi = \frac{\pi}{4}$ makes an angle A with the x-axis.

(a) Draw a picture that shows the three angles.

(b) Find $\cos(A)$.

25.[R]

(a) If you describe the region in Example 2 in the order $0 \le \theta \le 2\pi$, $z_1(\theta) \le z \le z_2(\theta)$, $r_1(\theta, z) \le r \le r_2(\theta, z)$, what complication arises?

(b) Describe the region using the order given in (a).

By differentiating, verify the equations in Exercises 26 to 27.

26.[R] $\int |fracdx(x^3\sqrt{z^2+x^2} = -\frac{\sqrt{a^2+x^2}}{2a^2+x^2} + \frac{1}{2a^3}\ln|\frac{a+\sqrt{a^2+x^2}}{x}|$.

27.[R] $\int \frac{x^2]dx}{a^4-x^4} = \frac{1}{4a}\ln|\frac{a+x}{a-x}| - \frac{1}{2a}\arctan\frac{x}{a}$.

28.[R] What is the distance between $P_1 = (\rho_1, \theta_1, \phi_1)$ and $P_2 = (\rho_2, \theta_2, \phi_2)$?

29.[R] The points $P_1 = (\rho_1, \theta_1, \phi_1)$ and $P_2 = (\rho_1, \theta_2, \phi_2)$ both lie on a sphere of radius ρ_1. Assuming that both are in the first octant, find the great circle distance between them. NOTE: If the sphere is the earth's surface, ρ is approximately 3960 miles, ϕ is the complement of the latitude, and θ is related to longitude.

30.[R] At time t a particle moving along a curve is at the point $(\rho(t), \theta(t), \phi(t))$ What is its speed?

31.[R] How far apart are the points (r_1, θ_1, z_1) and (r_2, θ_2, z_2) in the first octant?

(a) Draw a large clear diagram.

(b) Find the distance.

32.[R] A bug is wandering on the surface of a cylinder whose description is $0 \le \theta \le 2\pi, 0 \le r \le 3, 0 \le z \le 2$. It is at the point $(3, 0, 2)$ and wants the shortest route on the surface to $(3, \pi, 0)$. The bug plans to go straight down, keeping $\theta = 0$, and then taking a straight path on the base along a diameter. Is that the shortest path? If not, what is?

17.6 Iterated integrals for $\int_R f(P)\,dV$ in Cylindrical or Spherical Coordinates

In Section 17.2 we evaluated an integral of the form $\int_R f(P)\,dA$ by an iterated integral in polar coordinates. In this method it is necessary to multiply the integrand by an "r." This is necessary because the small patch determined by small increments in r and θ is not $\Delta r\Delta\theta$ but $r\Delta r\Delta\theta$. Similarly, when developing iterated integrals using cylindrical coordinates, an extra r must be adjoined to the integrand. In the case of spherical coordinates one must adjoin $\rho^2\rho\sin(\phi)$. These adjustments are based on the estimates of the volumes of the small curvy boxes made in the previous section.

A few examples will illustrate the method, which is: Describe the solid R and the integrand in the most convenient coordinate system. Then use that description to set up an iterated integral, being sure to include the appropriate extra factor in the integrand.

ITERATED INTEGRALS IN CYLINDRICAL COORDINATES

To evaluate $\int_R f(P)\,dV$ in cylindrical coordinates we express the integrand in cylindrical coordinates and describe the region R in cylindrical coordinates. It must be kept in mind that dV is replaced by $r\ dz\ dr\ d\theta$. There are six possible orders of integration, but the most common one is: z varies first, then r, finally θ:

$$\int_R f(P)\,dV = \int_\alpha^\beta \int_{r_1(\theta)}^{r_2(\theta)} \left(\int_{z_1(r,\theta)}^{z_2(r,\theta)} f(r,\theta,z) r\ dz\right)\ dr\ d\theta.$$

EXAMPLE 1 Find the volume of a ball R of radius a using cylindrical coordinates.

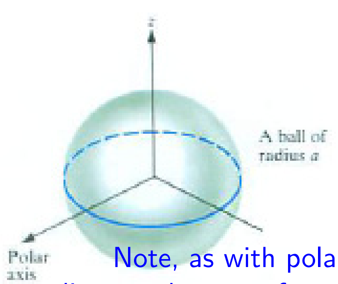


Figure 17.6.1: Note, as with polar coordinates, the extra factor r.

SOLUTION Place the origin of a cylindrical coordinate system at the center of the ball, as in Figure 17.6.1.

The volume of the ball is $\int_R 1\ dV$. The description of R in cylindrical coordinates is

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le a, \qquad -\sqrt{a^2-r^2} \le z \le \sqrt{a^2-r^2}.$$

The iterated integral for the volume is thus

$$\int_R 1\ dV = \int_0^{2\pi} \left(\int_0^a \left(\int_{-\sqrt{a^2-r^s}}^{\sqrt{a^2-r^2}} 1 \cdot r\ dz \right) dr \right) d\theta.$$

Evaluation of the first integral, where r and θ are fixed, yields

$$\int_{-\sqrt{a^2-r^2}}^{\sqrt{a^2-r^2}} r\ dz \quad = \quad rz|_{z=-\sqrt{z^2-r^2}}^{z=\sqrt{a^2-r^2}} = 2r\sqrt{a^2-r^2}.$$

Note that the order of integration is determined by the order of the variables in describing R.

Evaluation of the second integral, where θ is fixed, yields

$$\int_0^a 2r\sqrt{a^2-r^2}\ dr = \left. \frac{-2(a^2-r^2)^{3/2}}{3} \right|_{r=0}^{r=a} = \frac{2a^3}{3}.$$

Finally, evaluation of the third integral gives

$$\int_0^{2\pi} \frac{2a^3}{3}\ d\theta \quad = \quad \frac{2a^3}{3} \int_0^{2\pi} d\theta = \frac{2a^3}{3}\ cdot 2\pi = \frac{4}{3}\pi a^3.$$

◇

EXAMPLE 2 Find the volume of the region R inside the cylinder $x^2+y^2 = a$, above the xy-plane, and below the plane $z = x + 2y + 9$. Use cylindrical coordinates.

SOLUTION We wish to evaluate $\int_R 1\ dV$ over the region R described in cylindrical coordinates R by

$$0 \le \theta \le 2\pi, \qquad 0 \le r \le 3, \qquad 0 \le z \le r\cos(\theta) + 2r\sin(\theta) + 9.$$

(Here we replace the equation $z = x + 2y + 9$ by $z = r\cos(\theta) + 2r\sin(\theta) + 9$.)

Note, as with polar coordinates, the extra factor r.

The iterated integral takes the form

$$\int_0^{2\pi} \left(\int_0^3 \left(\int_0^{r\cos(\theta)+2r\sin(\theta)+9} 1 \cdot r\ dz \right) dr \right) d\theta.$$

Integration with respect to z gives

r and θ are constant

$$\int_0^{r\cos(\theta)+2r\sin(\theta)+9} r\ dz \quad = \quad r \int_0^{r\cos(\theta)+2r\sin(\theta)+9} dz = r^2\ cos(\theta) + 2r^2\sin(\theta) + 9r.$$

Then comes integration with respect to r, with θ constant:

$$\int_0^3 \left(r^2\cos(\theta)+2r^2\sin(\theta)+9r\right)\,dr = \frac{r^3}{3}\cos(\theta)+\frac{2r^3}{3}\sin(\theta)+\frac{9r^2}{2}\Big|_0^3 = 9\cos(\theta)+18\sin$$

Finally, integration with respect to θ gives

$$\int_0^{2\pi}\left(9\cos(\theta)+18\sin(\theta)+\frac{81}{2}\right)\,d\theta. \qquad (17.6.1)$$

Because $\int_0^{2\pi}\cos(\theta)\,d\theta = 0 = \int_0^{2\pi}\sin(\theta)\,d\theta$, (17.6.1) reduces to $\int_0^{2\pi}\frac{81}{2}\,d\theta = 81\pi$.
The volume is 81π. ◇

Computing $\int_R f(P)\,dV$ in Spherical Coordinates

To evaluate a triple integral $\int_R f(P)\,dV$ in spherical coordinates, first describe the region R in spherical coordinates. Usually this will be in the order:

$$\alpha\le\theta\le\beta,\qquad \phi_1(\theta)\le\phi\le\phi_2(\theta),\qquad \rho_1(\theta,\phi)\le\rho\le\rho_2(\rho,\theta).$$

Sometimes the order of ρ and ϕ is switched:

$$\alpha\le\theta\le\beta\qquad \rho_1(\theta)\le\rho\le\rho_2(\theta)\qquad \phi_1(\rho,\theta)\le\phi\le\phi_2(\rho,\theta).$$

Then set up an iterated integral, being sure to express dV as $\rho^2\sin(\phi)\,d\rho\,d\phi\,d\theta$ (or $\rho^2\sin(\phi)\,d\phi\,d\rho\,d\theta$).

EXAMPLE 3 Find the volume of a ball of radius a, using spherical coordinates.

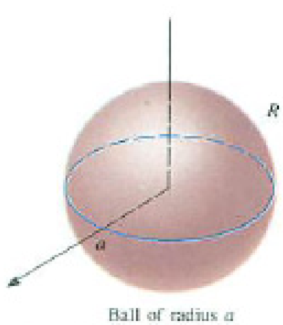


Figure 17.6.2:

SOLUTION Place the origin of spherical coordinates at the center of the ball, as in Figure 17.6.2. The ball is described by

$$0\le\theta\le 2\pi,\qquad 0\le\phi\le\pi,\qquad 0\le\rho\le a.$$

Hence

$$\text{Volume of ball} = \int_R 1\,dV = \int_0^{2\pi}\int_0^{\pi}\int_0^{a}\rho^2\sin(\phi)\,d\rho\,d\phi\,d\theta.$$

The inner integral is

$$\int_0^a \rho^2\sin(\phi)\,d\rho = \sin\phi\int_0^a\rho^2\,d\rho = \frac{a^3\sin(\phi)}{3}.$$

The next integral is

$$\int_0^\pi \frac{a^3\sin(\phi)}{3}\,d\phi \;=\; \left.\frac{-a^3\sin(\phi)}{3}\right|_0^\pi = \frac{-a^3(-1)}{3} - \frac{-a^3(1)}{3} = \frac{2a^3}{3}.$$

The final integral is

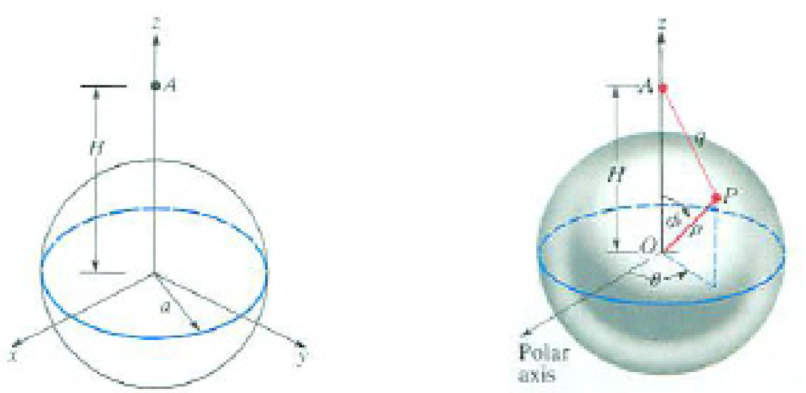
$$\int_0^{2\pi} \frac{2a^3}{3}\,d\theta \;=\; \frac{2a^3}{3}\int_0^{2\pi} d\theta = \frac{2a^3}{3}2\pi = \frac{4\pi a^3}{3}.$$

◇

An Integral in Gravity

The next example is of importance in the theory of gravitational attraction. It implies that a homogeneous ball attracts a particle (or satellite) as if all the mass of the ball were at its center.

EXAMPLE 4 Let A be a point at a distance H from the center of the ball, $H > a$. Compute $\int_R (\delta/q)\ dV$, where δ is density and q is the distance from a point P in R to A. (See Figure 17.6.3.)



(a) (b)

Figure 17.6.3:

SOLUTION First, express q in terms of spherical coordinates. To do so, choose a spherical coordinate system whose origin is at the center of the sphere and such that the ϕ coordinate of A is 0. (See Figure 17.6.3(b).)

Let $P = (\rho, \theta, \phi)$ be a typical point in the ball. Applying the law of cosines to triangle AOP, we find that

$$a^2 = H^2 + \rho^2 - 2\rho H \cos(\phi).$$

Hence

$$q = \sqrt{H^2 + \rho^2 - 2\rho H \cos(\phi)}.$$

Since the ball is homogeneous,

$$\delta = \frac{M}{\frac{4}{3}\pi a^3} = \frac{3M}{4\pi a^3}.$$

Hence

$$\int_R \frac{\delta}{q}\, dV = \int_R \frac{3M}{4\pi a^3 q}\, dV = \frac{3M}{4\pi a^3} \int_R \frac{1}{q}\, dV. \tag{17.6.2}$$

Now evaluate

$$\int_R \frac{1}{q}\, dV$$

A case where integration with respect to ρ is not first

by an iterated integral in spherical coordinates:

$$\int_R \frac{1}{q}\, dV = \int_0^{2\pi} \left(\int_0^a \left(\int_0^\pi \frac{\rho^2 \sin(\phi)}{\sqrt{H^2 + \rho^2 - 2\rho H \cos(\phi)}}\, d\phi \right) d\rho \right) d\theta.$$

We integrate with respect to ϕ first, rather than ρ, because it is easier in this case.

Evaluation of the first integral, where ρ and θ are constants, is accomplished with the aid of the fundamental theorem:

$$\begin{aligned} \int_0^\pi \frac{\rho^2 \sin\phi}{\sqrt{H^2 + \rho^2 - 2\rho H \cos(\phi)}}\, d\phi &= \left. \frac{\rho\sqrt{H^2 + \rho^2 - 2\rho H \cos(\phi)}}{H} \right|_{\phi=0}^{\phi=\pi} \\ &= \frac{\rho}{H}(\sqrt{H^2 + \rho^2 + 2\rho H} - \sqrt{H^2 + \rho^2 - 2\rho H}. \end{aligned}$$

Now, $\sqrt{H^2 + \rho^2 + 2\rho H} = H + \rho$. Since $\rho \leq a < H$, $H - \rho$ is positive and $\sqrt{H^2 + \rho^2 - 2\rho H} = H - \rho$.

Thus the first integral equals

$$\frac{\rho}{H}[H + \rho) - (H - \rho)] = \frac{2\rho^2}{H}.$$

Evaluation of the second integral yields

$$\int_0^a \frac{2\rho^2}{H}\, d\rho = \frac{2a^3}{3H}.$$

Evaluation of the third integral yields

$$\int_0^{2\pi} \frac{2a^3}{eH}\, d\theta = \frac{4\pi a^3}{3H}.$$

Hence

$$\int_R \frac{1}{q}\, dV = \frac{4\pi a^3}{3H}.$$

By (17.6.2)

$$\int_R \frac{\delta}{q}\, dV = \frac{3M}{4\pi a^3}\frac{4\pi a^3}{3H} = \frac{M}{H}.$$

This result, M/H, is exactly what we would get if all the mass were located at the center of the ball. ◇

Newton obtained this remarkable result in 1687.

SHERMAN: Compare this with your version. What is your M? I thought it was the object, i.e., the region together with its density, but you are using it also as the mass of the object. If you want the latter, you can't talk about "a mass M occupies a region R." I prefer to say we have an object that occupies a region R, with density δ, and mass M where $M = \int_R \delta(P)\, dV$. See what I have written. This may need to be changed elsewhere. That should not be difficult to do, but I want to see your comments first.

THE MOMENT OF INERTIA ABOUT A LINE

In the study of rotation of a object about an axis, one encounters the "moment of inertia", I of the object. It is defined as follows. The object occupies a region R. The density of the object at a typical point P is $\delta(P)$, so the mass of the object is $M = \int_R \delta(P)\, dV$. Usually the density is constant, in which case it is M divided by the volume of R (or M divided by the area of R if R is planar). Let $r(P)$ be the distance from P to a fixed line L. Then, by definition,

An object with constant density is called homogeneous.

$$I = \text{Moment of Inertia} = \int_R (r(P))^2 \delta(P)\, dV.$$

A similar definition holds for objects distributed on a planar region. The only difference is that dV is replaced by dA.

Exercise 31 shows that I plays the same role in a rotating body (such as a spinning skater) as mass does in an object moving

EXAMPLE 5 Compute the moment of inertia of a uniform mass M in the form of a ball of radius a around a diameter L.

SOLUTION The density $\delta(P)$, being constant, is $M/(\frac{4}{3}\pi a^3)$. We place the diameter L along the z-axis, as in Figure 17.6.4

As a check on our answer

Because the distance $r(P)$ is just r in cylindrical coordinates, we will first work in those coordinates. Then we will calculate the moment of inertia in spherical coordinates.

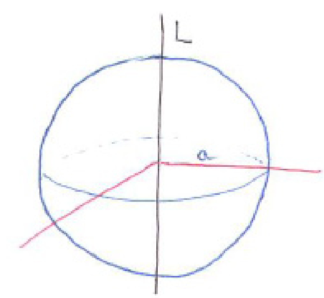


Figure 17.6.4:

One description of the ball is

$$0 \le \theta \le 2\pi, \qquad -a \le z \le a, \qquad 0 \le r \le \sqrt{a^2 - z^2}.$$

Then

$$\begin{aligned} I &= \int_R \frac{M}{\frac{4}{3}\pi a^3} r^2 \, dV = \frac{3M}{4\pi a^3} \int_R r^2 \, dV \qquad \text{Note the introduction of the extra } r \\ &= \frac{3M}{4\pi a^3} \int_0^{2\pi} \int_{-a}^{a} \int_0^{\sqrt{a^2-z^2}} r^3 \, dr \, dz \, d\theta \end{aligned}$$

The first integration is

$$\int_0^{\sqrt{a^2-z^2}} r^3 \, dr = \left.\frac{r^4}{4}\right|_0^{\sqrt{a^2-z^2}} = \frac{(a^2-z^2)^2}{4}.$$

The second is

$$\begin{aligned} \int_{-a}^{a} \frac{(a^2-z^2)^2}{4} \, dz &= \int_{-a}^{a} \frac{a^4 - 2a^2z^2 + z^4}{4} \, dz = \frac{1}{4}\left(a^4 z - \frac{2a^2z^3}{3} + \frac{z^5}{5}\right) |_{-a}^{a} \\ &= \frac{1}{4}\left(a^5 - \frac{2a^5}{3} + \frac{a^5}{5}\right) - \frac{1}{4}\left(-a^5 + \frac{2a^5}{3} - \frac{a^5}{5}\right) = \frac{4}{15}a^5. \end{aligned}$$

The third is

$$\int_0^{2\pi} \frac{4}{15} a^5 \, d\theta = \frac{8\pi}{15} a^5.$$

Then remembering to include the factor $3M/4\pi a^3$, we have

$$I = \frac{3M}{4\pi a^3} \cdot \frac{8\pi}{15} a^5 = \frac{2}{5} M a^2.$$

Because spherical coordinates provide a simple description of the ball, we will also use them to find I to see if the computations are easier. Now the distance $r(P)$ has a more complicated form, $\delta(P) = \delta(\rho, \theta, \phi) = \rho sin(\phi)$. The integral for the moment of inertia is

$$I = \frac{3M}{4\pi a^3} \int_R (\rho \sin(\phi)^2 \, dV.$$

The iterated integral for this multiple integral is

$$\int_0^{2]pi} \left(\int_0^{\pi} \left(\int_0^{a} (\rho \sin(\phi))^2 \rho^2 \sin(\phi) \, d\rho \right) \, d\phi \right) \, d\theta.$$

The first integration is

$$\int_0^{a} \rho^4 \sin^3(\phi) \, d\rho = \left.\frac{\rho^5}{5} \sin^3(\phi)\right|_{\rho=0}^{\rho=a} = \frac{a^5}{5\sin^3(\phi)}.$$

The second is

$$\int_0^\pi \frac{a^5}{5}\sin^3(\phi)\ d\phi = \frac{a^5}{5}\int_0^\pi \sin^3(\phi)\ d\phi.$$

Since the exponent, 3, is odd, we write $sin^3(\phi)$ as $(1-\cos^2(\phi))\sin(\phi)$ and have

$$\begin{aligned}\int_0^\pi \sin^3(\phi)\ d\phi &= \int_0^\pi (\sin(\phi)-\cos^2(\phi)\sin(\phi)) = (-\cos(\phi)+\frac{\cos^3(\phi)}{3}\Big|_0^\pi \\ &= (-(-1)+\frac{(-1)^3}{3})-(-1+\frac{1}{3}) = \frac{4}{3}.\end{aligned}$$

The final integration is just

$$\int_0^{2\pi} \frac{a^5}{5}\cdot\frac{4}{3}\ d\theta = \frac{8}{15}\pi.$$

And, as expected, gives, again

$$I = (2/5)Ma^2.$$

Note that this is 2/5 of our upper estimate, hence is plausible.

◇

Summary

A multiple integral $\int_R f(P)\ dV$ may be evaluated by an iterated integral in cylindrical or spherical coordinates. In cylindrical coordinates the iterated integral takes the form

$$\int_{\theta_1}^{\theta_2}\left(\int_{r_1(\theta)}^{r_2(\theta)}\left(\int_{z_1(r,\theta)}^{z_2(r,\theta)} rf(r,\theta,z)\ dz\right)\ dr\right)\ d\theta.$$

The description of the region determines the range of integration on each of the three integrals over intervals. (Changing the order of the description of R changes the order of the integrations.) The factor r must be inserted into the integrand.

In spherical coordinates the iterated integral usually takes the form

$$\int_{\theta_1}^{\theta_2}\left(\int_{\phi_1(\theta)}^{\phi_2(\theta)}\left(\int_{\rho_1(\theta,\phi)}^{\rho_2(\theta,\phi)} f(r,\theta,\phi)\rho^2\sin(\phi)\ d\phi\right)\ d\phi\right)\ d\theta.$$

In this form, integration with respect to ρ is first, but as Example 4 illustrates, it may be convenient to integrate first with respect to ϕ. The factor $\rho^2\sin(\phi)$ must be inserted in the integrand.

EXERCISES for Section 17.6 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4: (a) draw the region, (b) set up an iterated integral in cylindrical coordinates for the given multiple integrals, and (c) evaluate the iterated integral.

1.[R] $\int_R r^2\ dV$, R is bounded by the cylinder $r = 3$ and the planes $z = 2x$ and $z = 3x$.

2.[R] $\int_R z\ dV$, R is bounded by the sphere $z^2 + r^2 = 25$, the $r\theta$ coordinate plane, and the plane $z = 2$.

3.[R] $\int_R rz\ dV$, R is the part of the ball bounded by $r^2+z^2 = 16$ in the first octant.

4.[R] $\int_R \cos\theta / dV$, R is bounded by the cylinder $r = 2\cos(\theta)$ and the paraboloid $z = r^2$.

5.[R] Compute the volume of a right circular cone of height h and radius r using (a) spherical coordinates, (b) cylindrical coordinates, and (c) using rectangular coordinates.

6.[R] Find the volume of the region above the xy plane and below the paraboloid $z = 9 - r^2$ using cylindrical coordinates.

7.[R] A right circular cone of radius a and height h has a density at point P equal to the distance from P to the base of the cone. Find its mass, using spherical coordinates.

In Exercises 8 to 9 draw the region R and give a formula for the integrand $f(P)$ such that $\int_R\ dV$ is described by the given iterated integrals.

8.[R] $\int_0^{\pi/2}[\int_0^{\pi/4}(\int_0^{\cos\phi} \rho^3 \sin^2(\theta)\sin(\phi)\ d\rho)\ d\phi]\ d\theta.$

9.[R] $\int_0^{\pi/4}[\int_{\pi/6}^{\pi/2}(\int_0^{\sec\theta} \rho^3 \sin(\theta)\cos(\phi)\ d\rho)\ d\phi]\ d\theta.$

10.[R] Let R be the solid region inside both the sphere $x^2 + y^2 + z^2 = 1$ and the cone $z = \sqrt{x^2 + y^2}$. Let the density at (x, y, z) be $f(x, y, z) = z$. Set up iterated integrals for the mass in R using (a) rectangular coordinates, (b) cylindrical coordinates, (c) spherical coordinates. (d) Evaluate the iterated integral in (c).

11.[R] Find the average temperature in a ball of radius a if the temperature is the square of the distance from a fixed equatorial plane.

In each of Exercises 12 to 13 evaluate the iterated integral.

12.[R] $\int_0^{2\pi}\left(\int_0^1\left(\int_r^1 zr^3\cos^2\theta\ dz\right)\ dr\right)\ d\theta$

13.[R] $\int_0^{2\pi}\left(\int_0^1\left(\int_{-\sqrt{a^2-r^2}}^{\sqrt{a^2-r^2}} z^2 r\ dz\right)\ dr\right)\ d\theta$

14.[R] Let R be the solid region inside both the sphere $x^2+y^2+z^2=1$ and the cone $z=\sqrt{x^2+y^2}$. Let the density at (x,y,z) be $f(x,y,z)=z$. Using cylindrical coordinates, find the mass of R.

15.[R] Using cylindrical coordinates, find the volume of the region below the plane $z=y+1$ and above the circle in the xy plane whose center is $(0,1,0)$ and whose radius is 1. (Include a drawing of the region.) HINT: What is the equation of the circle in polar coordinates when the polar axis is along the positive x-axis?

16.[R] Find the average distance from the center of a ball of radius a to other points of the ball by setting up appropriate iterated integrals in the three types of coordinate systems and evaluating the easiest.

17.[R] A solid consists of that part of a ball of radius a that lies within a cone of half-vertex angle $\phi=\pi/6$, the vertex being at the center of the ball. Set up iterated integrals for $\int_R z\ dV$ in all three coordinate systems and evaluate the simplest.

In Exercises 18 to 23 evaluate the multiple integrals over a ball of radius a with center at the origin, without using an iterated integral (ϕ,θ, and z are cylindrical or spherical coordinates).

18.[R] $\int_R \cos(\theta)\ dV$

19.[R] $\int_R \cos^2\theta\ dV$

20.[R] $\int_R z\ dV$

21.[R] $\int_R (3+2\sin(\theta)\ dV$

22.[R] $\int_R \sin^2(\phi)\ dV$

23.[R] $\int_R \sin(\phi)\ dV$

24.[R] In polar, cylindrical, and spherical coordinates one must introduce an extra factor in the integrand when using an iterated integral. Why is that not necessary when using rectangular coordinates?

25.[R] Is $\sqrt{a^2}$ always equal to a?

26.[R] Using the method of Example 4 find the average value of q for all points P in the ball. Note that it is *not* the same as if the entire ball were placed at its center.

27.[C] Show that the result of Example 4 holds if the density $\delta(P)$ depends only on ρ, the distance to the center. (This is approximately the case with the planet Earth, which is not homogeneous.) Let $g(\rho)$ denote $\delta(\rho,\theta,\phi)$.

In Exercises 28 to 29check the equations by differentiation.

28.[R] $\tan\left(\frac{x}{2}\right) = \int \frac{dx}{1+\cos(x)}$

29.[R] $x\tan\left(\frac{x}{2}\right) + 2\ln\left|\cos\left(\frac{x}{2}\right)\right| = \int \frac{x\ dx}{1+\cos(x)}$

30.[R]

(a) Find the exact volume of the little curvy box corresponding to the changes $\Delta\rho$, $\Delta\theta$, $\Delta\phi$.

(b) One hopes that the ratio between that exact volume and our estimate, $\rho^2\sin(\phi)\Delta\rho\Delta\theta\Delta$ approaches 1 as $\Delta\rho$, $\Delta\theta$, $\Delta\phi$ approach 0. Show that it does. HINT: Recall the definition of a derivative.

(c) Show that the exact volume in (a) can be written in the form $(\rho^* t)^2\sin(\phi^*)\Delta\rho\Delta\phi\Delta\theta$, where ρ^* is between ρ and $\rho + \Delta\rho$ and ϕ^* is between ϕ and $\phi + \Delta\phi$.

31.[R] The kinetic energy of an object with mass m moving at the velocity v is $mv^2/2$. An object moving in a circle of radius r at the angular speed ω radians per unit time has velocity $r\omega$. (Why?) Thus its kinetic energy is $(mr^2/2)\omega^2$. Now consider a mass M that occupies the region R in space. Its density is $\Delta(P)$, which may vary from point to point. (If it is constant, it equals $M/(\text{Volume of R})$.) Let $f(P)$ be the distance from P to a fixed line L. If the mass is spinning around the axis L at the angular rate ω, show that its total kinetic energy is

$$\int_R \frac{1}{2}(f(P))^2\delta(P)w^2\ dv.$$

This can be written as

$$\text{Kinetic Energy} = (\frac{1}{2}I\omega^2.$$

Thus I plays the same role in rotational motion that mass m plays in linear motion in the formula $(\frac{1}{2}mv^2$.

Every spinning ice skater knows this. When spinning with her arms extended she has a certain amount of kinetic energy. If she suddenly puts her arms to her sides she decreases her moment of inertia but has not destroyed her kinetic energy. That forces her angular speed to increase. The larger the mass m is, the harder it is to start it moving and to stop it when it is moving. Similarly, the larger I is, the harder it is to stop the mass from spinning and to stop it when it is spinning.

In Exercises 32 to 36 the objects have a homogeneous (constant density) mass M. Find I.

32.[R] A rectangular box of dimensions, a, b, c around a line through its center and perpendicular to the face of dimensions a and b.

33.[R] A solid cylinder of radius a and height h around its axis.

34.[R] A solid cylinder of radius a and height h around a line on its surface.

35.[R] A hollow cylinder of height h, inner radius a, and outer radius b, about its axis.

36.[R] A solid cylinder of radius a and height h around a diameter in its base.

37.[R] In Example 2 what unpleasantness occurs when you describe the region in the order of the form $a \le \theta \le b$, $z_1(\theta) \le z \le z_2(\theta)$, $r_1(\theta, z) \le r \le r_2(\theta, z)$?

38.[R] Solve Example 2 using rectangular coordinates.

39.[R] Evaluate the moment of inertia in Example 5 using the description $0 \le \theta \le 2\pi$, $0 \le r \le a$, $-\sqrt{a^2 - r^2} \le z \le \sqrt{a^2 - r^2}$.

40.[R] Let R be a solid ball of radius a with center at the origin of the coordinate system

(a) Explain why $\int_R x^2\ dV = \frac{1}{3} \int_R (x^2 + y^2 + z^2)\ dV$.

(b) Evaluate the second integral by spherical coordinates.

(c) Use (b) to find $\int_R x^2\ dV$.

41.[M] Show that $\int_R (x^3 + y^3 + z^3)\ dV = 0$, where R is a ball whose center is the origin of a rectangular coordinate system. NOTE: Do not use an iterated integral. HINT: Use symmetry.

42.[R] A homogeneous object with mass M occupies the region R between concentric spheres of radii a and b, $a < b$. Let A be a point at a distance H from their center, $H < a$. Evaluate $\int_R \frac{\delta}{q}\ dV$, where δ is the density and $q = q(P)$ is the distance from H to any point P in R. (That the value of the integral does not involve H has an important consequence: A uniform hollow sphere exerts no gravitational force on objects in its interior.)

43.[R] In Example 4, H is greater than a. Solve the same problem for H less than a. NOTE: For some ρ, $\sqrt{H^2 + \rho^2 A - 2\rho H}$ equals $H - \rho$ and for some it equals $\rho - H$.

44.[C] (See Example 43.) Let A be a point in the plane of a disk but outside the disk. Is the average of the reciprocal of the distance from A to points in the disk equal to the reciprocal of the distance to the center of the disk?

45.[C] A certain ball of radius a is *not* homogeneous. However, its density at P depends only on the distance from P to the center of the ball. That is, there is a function $f(\rho)$ such that the density at $P = (\rho, \theta, \phi)$ is $f(\rho)$. Using an iterated integral, show that the mass of the ball is

$$4\pi \int_0^a f(\rho^2_\rho \, d\rho.$$

46.[C] Let R be the part of a ball of radius a removed by a cylindrical drill of diameter a whose edge passes through the center of the sphere.

(a) Sketch R.

(b) Notice that R consists of four congruent pieces. Find the volume of one of these pieces using cylindrical coordinates. Multiply by four to get the volume of R.

47.[C] Let R be the ball of radius a. For any point P in the ball other than the center of the ball, define $f(P)$ to be the reciprocal of the distance from P to the origin. The average value of r over R involves an improper integral, since the function blows up near the origin. Does this improper integral converge or diverge? What is the average value of f over R? *Suggestion:* Examine the integral over the region between concentric spheres of radii a and t, and let $t \to 0^+$.

17.7 Integrals Over Surfaces

In this section we define an integral over a surface and then show how to compute it by an iterated integral.

Definition of a Surface Integral

Consider a surface $\mathcal{S}$ such as the surface of a ball or part of the saddle $z = xy$. If f is a numerical function defined at least on $\mathcal{S}$, we will define the integral $\int_{\mathcal{S}} f(P)\ dS$. The definition is practically identical with the definition of the double integral, which is the special case when the surface is a plane.

We assume that the surfaces we deal with are smooth, or composed of a finite number of smooth pieces, and that the integrals we define exist.

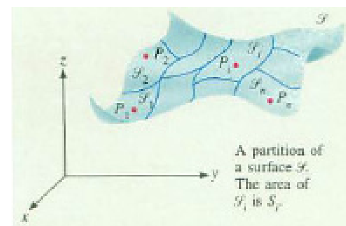


Figure 17.7.1:

DEFINITION *(Definite integral of a function f over a surface $\mathcal{S}$.)* Let f be a function that assigns to each point P in a surface $\mathcal{S}$ a number $f(P)$. Consider the typical sum

$$f(P_1)S_1 + f(P_2)S_2 + \cdots + f(P_n)S_n,$$

formed from a partition of $\mathcal{S}$, where S_i is the area of the ith region in the partition and P_i is a point in the ith region. (See Figure 17.7.1.) If these sums approach a certain number as the S_i are chosen smaller and smaller, the number is called the **integral of f over $\mathcal{S}$** and is written

$$\int_{\mathcal{S}} f(P)\ dS.$$

Surface integrals are also denoted $\iint_{\mathcal{S}} f(P)\ dS$.

If $f(P)$ is 1 for each point P in $\mathcal{S}$ then $\int_{\mathcal{S}} f(P)\ dS$ is the area of $\mathcal{S}$. If $\mathcal{S}$ is occupied by material of density $\sigma(P)$ at P then $\int_{\mathcal{S}} \sigma(P)\ dS$ is the total mass of $\mathcal{S}$.

First we show how to integrate over a sphere.

Integrating over a Sphere

If $\mathcal{S}$ is a sphere or part of a sphere, it is often convenient to evaluate an integral over it with the aid of spherical coordinates.

See Section 17.6 for a similar argument, where ρ was not constant.

If the center of a spherical coordinate system (ρ, θ, ϕ) is at the center of a sphere of radius a, then ρ is constant on the sphere $\rho = a$. As Figure 17.7.2 suggests, the area of the small region on the sphere corresponding to slight changes $d\theta$ and $d\phi$ is approximately

$$(a\ d\phi)\,(a\sin(\phi)\ d\theta) = a^2\sin(\phi)\ d\theta\ d\phi.$$

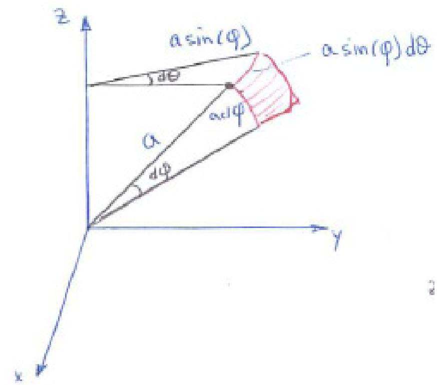


Figure 17.7.2:

Thus we may write

$$dS = a^2 \sin(\phi)\ d\theta\ d\phi$$

and evaluate

$$\int_{\mathcal{S}} f(P)\ dS$$

in terms of a repeated integral in ϕ and θ. Example 1 illustrates this technique.

EXAMPLE 1 Let $\mathcal{S}$ be the top half of the sphere with radius a. Evaluate $\int_{\mathcal{S}} z\ dS$.

SOLUTION Since the sphere has radius a, $\rho = a$. The top half of the sphere is described by $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi/2$. And, in spherical coordinates, $z = \rho\cos(\phi) = a\cos(\phi)$. Thus

$$\int_{\mathcal{S}} z\ dS = \int_{\mathcal{S}} (a\cos(\phi))\ dS = \int_0^{2\pi} \left(\int_0^{\pi/2} (a\cos(\phi)) a^2 \sin(\phi)\ d\phi \right)\ d\theta.$$

Now,

$$\begin{aligned} \int_0^{\pi/2} (a\cos(\phi)) a^2 \sin(\phi)\ d\phi &= a^3 \int_0^{\pi/2} \cos(\phi)\sin(\phi)\ d\phi = a^3 \left. \frac{(-\cos^2(\phi))}{2} \right|_0^{\pi/2} \\ &= \frac{a^3}{2}[-0 - (-1)] = \frac{a^3}{2}. \end{aligned}$$

so that

$$\int\limits_{\mathcal{S}} z\ dS = \int\limits_{0}^{2\pi} \frac{a^3}{2}\ d\theta = \pi a^3.$$

◇

We can interpret the result in Example 1 in terms of average value. The **average value** of $f(P)$ over a surface $\mathcal{S}$ is defined as

$$\frac{\int_{\mathcal{S}} f(P)\ dS}{\text{Area of } \mathcal{S}}.$$

Geometric interpretation

Example 1 shows that the average value of z over the given hemisphere is

$$\frac{\int_{\mathcal{S}} z\ dS}{\text{Area of } \mathcal{S}} = \frac{\pi a^3}{2\pi a^2} = \frac{a}{2}.$$

"The average height above the equator is exactly half the radius."

A General Technique

When we faced an integral over a curve, $\int_C f\ ds$, we evaluated it by replacing it with $\int_a^b f\ \frac{ds}{dt}\ dt$, an integral over an interval $[a, b]$.

We will do something similar for an integral over a surface: We will replace an surface integral by a double integral over a set in a coordinate plane.

The basic idea is to replace a small patch on the surface $\mathcal{S}$ by its projection (shadow) or, say, the xy-coordinate plane. The area of the shadow is not the same as the area of the patch. With the aid of Figure 17.7.3 we will express the area of the shadow in terms of the tilt of the patch.

The unit normal vector to the patch is $\mathbf{n}$. The angle between $\mathbf{n}$ and $\mathbf{k}$ is γ. Call the area of the patch, dS, and the area of its projection, dA. Then

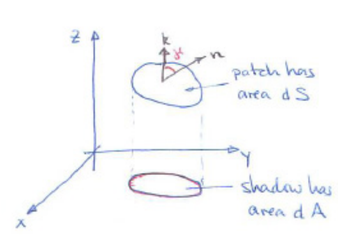


Figure 17.7.3:

$$dA \approx |\cos(\gamma)|\ dS.$$

Recall the discussion of direction angles and direction cosines in Section 14.4.

Notice that the angle γ is one of the direction angles of the unit normal vector, $\mathbf{k}$.

For instance, if $\gamma = 0$, then $dA = dS$. If $\gamma = \pi/2$, then $dA = 0$. We use the absolute value of $\cos(\gamma)$, since γ could be larger than $\pi/2$.

It follows, if $\cos(\gamma)$ is not 0, that

$$dS = \frac{dA}{|\cos(\gamma)|} \tag{17.7.1}$$

With the aid of (17.7.1), we replace an integral over $\mathcal{S}$ with an integral over its shadow in the xy plane.

The replacement is visible in the approximating sums involved in the integral over a surface.

Let $\mathcal{S}$ be a surface that meets each line parallel to the z-axis at most once. Let f be a function whose domain includes $\mathcal{S}$.

Consider an approximating sum for $\int_{\mathcal{S}} f(P)\ dS$, namely $\sum_{i=1}^{n} f(p_i)\Delta S_i$. The partition is shown in Figure 17.7.4.

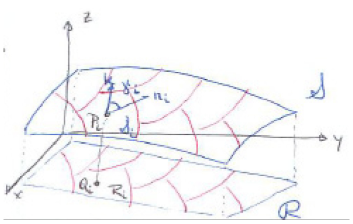


Figure 17.7.4:
Replacing an integral over a surface with an integral over a planar region.

Let R be the projection of $\mathcal{S}$ in the xy plane. The patch $\mathcal{S}_i$ with area S_i, projects down to R_i, of area A_i, and the point P_i on $\mathcal{S}_i$ points down to Q_i in R_i. Let γ_i be the angle between the normal at P_i and $\mathbf{k}$.

Then $f(P)S_i$ is approximately $\frac{f(P_i)}{|\cos(\gamma_i)|}A_i$. Thus an approximation of $\int_{\mathcal{S}} f(P)\, dS$ is

$$\sum_{i=1}^{n} \frac{f(P_i)}{|\cos\gamma_i|} A_i. \tag{17.7.2}$$

Theorem 17.7.1. *Let $\mathcal{S}$ be a surface and let $\mathcal{A}$ be its projection on the xy plane. Assume that for each point Q on $\mathcal{A}$ the line through Q parallel to the z-axis meets $\mathcal{S}$ in exactly one point P. Let f be a function defined on $\mathcal{S}$. Define a function h on $\mathcal{A}$ by*

$$h(Q) = f(P).$$

Then

$$\int_{\mathcal{S}} f(P)\ dS = \int_{\mathcal{A}} \frac{h(Q)}{|\cos(\gamma)|}\ dA.$$

In this equation γ denotes the angle between $\mathbf{k}$ and a vector normal to the surface of $\mathcal{S}$ at P. (See Figure 17.7.5.)

In order to apply this result, we need to be able to compute $\cos(\gamma)$.

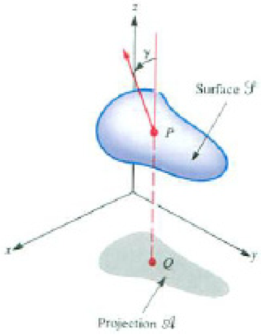


Figure 17.7.5:

Computing $\cos(\gamma)$

We find a vector perpendicular to the surface in order to compute $\cos(\gamma)$. If $\mathcal{S}$ is the level surface of $g(x, y, z)$, that is $g(x, y, z) = c$, for some constant c, then the gradient ∇g is such a vector.

If the surface $\mathcal{S}$ is given in the form $z = f(x,y)$, rewrite it as $z - f(x,y) = 0$. That means that $\mathcal{S}$ is a level surface of $g(x,y,z) = z - f(x,y)$, Theorem 17.7.2 shows what the formulas for $\cos(\gamma)$ look like. However, it is unnecessary, even distracting, to memorize them. Just remember that *a gradient provides a normal* to a level surface.

Theorem 17.7.2. *(a) If the surface $\mathcal{S}$ is part of the level surface $g(x,y,z) = c$, then*

$$|\cos(\gamma)| = \frac{|\frac{\partial g}{\partial z}|}{\sqrt{(\frac{\partial g}{\partial x})^2 + (\frac{\partial g}{\partial y})^2 + (\frac{\partial g}{\partial z})^2}}.$$

(b) If the surface $\mathcal{S}$ is given in the form $z = f(x,y)$, then

$$|\cos(\gamma)| = \frac{1}{\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 + 1}}.$$

Proof

(a) A normal vector to $\mathcal{S}$ at a given point is provided by the gradient

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

The cosine of the angle between $\mathbf{k}$ and ∇g is

$$\frac{\mathbf{k}\cdot\nabla g}{\|\mathbf{k}\|\|\nabla g\|} = \frac{\mathbf{k}\cdot(\frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k})}{(1)\left(\cdot\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial g}{\partial y})^2 + (\frac{\partial g}{\partial z})^2}\right)};$$

hence

$$|\cos(\gamma)| = \frac{|\frac{\partial g}{\partial z}|}{\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial g}{\partial y})^2 + (\frac{\partial g}{\partial z})^2}}.$$

(b) Rewrite $z = f(x,y)$ as $z - f(x,y) = 0$. The surface $z = f(x,y)$ is thus the level surface $g(x,y,z) = 0$ of the function $g(x,y,z) = z - f(x,y)$. Note that

$$\frac{\partial g}{\partial x} = -\frac{\partial f}{\partial x}, \qquad \frac{\partial g}{\partial y} = -\frac{\partial f}{\partial y} \qquad \text{and} \qquad \frac{\partial g}{\partial z} = 1.$$

By the formula in (a),

$$|\cos(\gamma)| = \frac{1}{\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 + 1}}$$

Theorem 17.7.2 is stated for projections on the xy plane. Similar theorems hold for projections on the xz or yz plane. The direction angle γ is then replaced by the corresponding direction angle, β or α, and the normal vector is dotted into $\mathbf{j}$ or $\mathbf{i}$. Just draw a picture in each case; there is no point in trying to memorize formulas for each situation.

EXAMPLE 2 Find the area of the part of the saddle $z = xy$ inside the cylinder $x^2 + y^2 = a^2$.

SOLUTION Let $\mathcal{S}$ be the part of the surface $z = xy$ inside $x^2 + y^2 = a^2$. Then

$$\text{Area of } \mathcal{S} = \int_{\mathcal{S}} 1 \; dS.$$

The projection of $\mathcal{S}$ on the xy plane is a disk of radius a and center $(0,0)$. Call it $\mathcal{A}$, as in Figure 17.7.6. Then

$$\text{Area of } \mathcal{S} = \int_{\mathcal{S}} 1 \; dS = \int_{\mathcal{A}} \frac{1}{|\cos(\gamma)|} \; dA. \qquad (17.7.3)$$

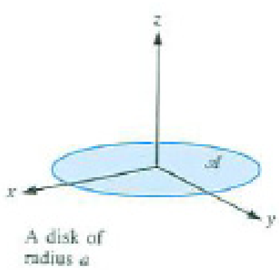


Figure 17.7.6:

To find the normal to $\mathcal{S}$ rewrite $z = xy$ as $z - xy = 0$. Thus $\mathcal{S}$ is a level surface of the function $g(x,y,z) = z - xy$. A normal to $\mathcal{S}$ is therefore

$$\begin{aligned} \nabla g &= \frac{\partial g}{\partial x}\mathbf{i} + \frac{\partial g}{\partial y}\mathbf{j} + \frac{\partial g}{\partial z}\mathbf{k} \\ &= -y\mathbf{i} - x\mathbf{j} + \mathbf{k}. \end{aligned}$$

Then

$$\cos(\gamma) = \frac{\mathbf{k} \cdot \nabla g}{\|\mathbf{k}\| \|\nabla g\|} = \frac{\mathbf{k} \cdot (-y\mathbf{i} - x\mathbf{j} + \mathbf{k})}{\sqrt{y^2 + x^2 + 1}} = \frac{1}{\sqrt{y^2 + x^2 + 1}}.$$

The area of $\mathcal{S}$ is $\int_{\mathcal{A}} \sqrt{(\partial f/\partial x)^2 + (\partial f/\partial y)^2 + 1} \, dA$

By (17.7.3),

$$\text{Area of } \mathcal{S} = \int_{\mathcal{A}} \sqrt{y^2 + x^2 + 1} \; dA. \qquad (17.7.4)$$

Use polar coordinates to evaluate the integral in (17.7.4):

$$\int_{\mathcal{A}} \sqrt{y^2 + x^2 + 1} \; dA = \int_0^{2\pi} \int_0^{a} \sqrt{r^2 + 1} r \; dr \; d\theta.$$

The inner integration gives

$$\int_0^a \sqrt{r^2+1}f\ dr \quad = \quad \left.\frac{(r^2+1)^{3/2}}{3}\right|_0^a = \frac{(1+a^2)^{3/2}-1}{3}.$$

The second integration gives

$$\int_0^{2\pi} \frac{(1+a^2)^{3/2}-1}{3}\ d\theta = \frac{2\pi}{3}\left((1+a^2)^{3/2}-1\right).$$

◇

Summary

After defining $\int_{\mathcal{S}} f(P)\ dS$, an integral over a surface, we showed how to compute it when the surface is part of a sphere.

Replace dS by $a^2\sin(\phi)\ d\phi\ d\theta$, where a is the radius of the sphere.

If each line parallel to the z-axis meets the surface $\mathcal{S}$ in at most one point, an integral over $\mathcal{S}$ can be replaced by an integral over $\mathcal{A}$, the projection of $\mathcal{S}$ on the xy plane:

$$\int_{\mathcal{S}} f(P)\ dS = \int_{\mathcal{A}} \frac{h(Q)}{|\cos(\gamma)|}\ dA.$$

To find $\cos(\gamma)$, use a gradient. If the surface is a level surface of, $g(x,y,z)=c$, use ∇g. If it has the equation $z=f(x,y)$, rewrite the equation as $z-f(x,y)=0$. As a special case, if $\mathcal{S}$ is the graph of $z=f(x,y)$, then the area of $\mathcal{S}$

$$\text{Area of } \mathcal{S} = \int_{S} dS = \int_{\mathcal{A}} \sqrt{(\partial f/\partial x)^2+(\partial f/\partial y)^2+1}\ dA.$$

EXERCISES for Section 17.7 *Key:* R–routine, M–moderate, C–challenging

1.[R] A small patch of a surface makes an angle of $\pi/4$ with the xy plane. Its projection on that plane has area 0.05. Estimate the area of the patch.

2.[R] A small patch of a surface makes an angle of $25°$ with the yz plane. Its projection on that plane has area 0.03. Estimate the area of the patch.

3.[R]

(a) Draw a diagram of the part of the plane $x + 2y + 3z = 12$ that lies inside the cylinder $x^2 + y^2 = 9$.

(b) Find as simply as possible the area of the part of the plane $x + 2y + 3z = 12$ that lies inside the cylinder $x^2 + y^2 = 9$.

4.[R]

(a) Draw a diagram of the part of the plane $z = x+3y$ that lies inside the cylinder $r = 1 + \cos\theta$.

(b) Find as simply as possible the area of the part of the plane $z = x + 3y$ that lies inside the cylinder $r = 1 + \cos\theta$.

5.[R] Let $f(P)$ be the square of the distance from P to a fixed diameter of a sphere of radius a. Find the average value of $f(P)$ for points on the sphere.

6.[R] Find the area of that part of the sphere of radius a that lies within a cone of half-vertex angle $\pi/4$ and vertex at the center of the sphere, as in Figure 17.7.7.

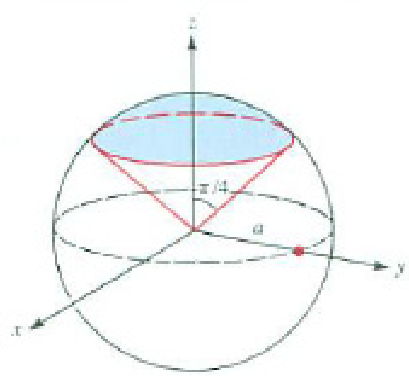


Figure 17.7.7:

In Exercises 7 and 8 evaluate $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,dS$ for the given spheres and vectors fields ($\mathbf{n}$ is the outward unit normal.)

7.[R] The sphere $x^2+y^2+z^2=9$ and $\mathbf{F}=x^2\mathbf{i}+y^2vj+z^2\mathbf{k}$.

8.[R] The sphere $x^2+y^2+z^2=1$ and $\mathbf{F}=x^3\mathbf{i}+y^2\mathbf{j}$.

9.[R] Find the area of the part of the spherical surface $x^2+y^2+z^2=1$ that lies within the vertical cylinder erected on the circle $r=\cos\theta$ and above the xy plane.

10.[R] Find the area of that portion of the parabolic cylinder $z=\frac{1}{2}x^2$ between the three planes $y=0$, $y=x$, and $x=2$.

11.[R] Evaluate $\int_{\mathcal{S}} x^2y\,dS$, where $\mathcal{S}$ is the portion in the first octant of a sphere with radius a and center at the origin, in the following way:

(a) Set up an integral using x and y as parameters.

(b) Set up an integral using ϕ and θ as parameters.

(c) Evaluate the easier of (a) and (b).

12.[R] A triangle in the plane $z=x+y$ is directly above the triangle in the xy plane whose vertices are $(1,2)$, $(3,4)$, and $(2,5)$. Find the area of

(a) the triangle in the xy plane,

(b) the triangle in the plane $z=x+y$.

13.[R] Let $\mathcal{S}$ be the triangle with vertices $(1,1,1)$, $(2,3,4)$, and $(3,4,5)$.

(a) Using vectors, find the area of $\mathcal{S}$.

(b) Using the formula

$$\text{Area of } \mathcal{S} = \int_{\mathcal{S}} 1\,dS,$$

find the area of $\mathcal{S}$.

14.[R] Find the area of the portion of the cone $z^2=x^2+y^2$ that lies above one loop of the curve $r=\sqrt{\cos 2(\theta)}$.

15.[R] Let $\mathcal{S}$ be the triangle whose vertices are $(1,0,0)$, $(0,2,0)$, and $(0,0,3)$. Let

$f(x, y, z) = 3x + 2y + 2z$. Evaluate $\int_{\mathcal{S}} f(P)\ dS$.

In Exercises 16 and 17 let $\mathcal{S}$ be a sphere of radius a with center at the origin of a rectangular coordinate system.

16.[R] Evaluate each of these integrals with a minimum amount of labor.

(a) $\int_{\mathcal{S}} x\ dS$

(b) $\int_{\mathcal{S}} x^3\ dS$

(c) $\int_{\mathcal{S}} \frac{2x+4y^5}{\sqrt{2+x^2+3y^2}}\ dS$

17.[R]

(a) Why is $\int_{\mathcal{S}} x^2\ dS = \int_{\mathcal{S}} y^2\ dS$?

(b) Evaluate $\int_{\mathcal{S}}(x^2 + y^2 + z^2)\ dS$ with a minimum amount of labor.

(c) In view of (a) and (b), evaluate $\int_{\mathcal{S}} x^2\ dS$.

(d) Evaluate $\int_{\mathcal{S}}(2x^2 + 3y^3)\ dS$.

18.[R] An electric field radiates power at the rate of $k(\sin^2(\phi)/\rho^2$ units per square meter to the point $P = (\rho, \theta, \phi)$. Find the total power radiated to the sphere $\rho = a$.

19.[R] A sphere of radius $2a$ has its center at the origin of a rectangular coordinate system. A circular cylinder of radius a has its axis parallel to the z-axis and passes through the z-axis. Find the area of that part of the sphere that lies within the cylinder and is above the xy plane.

Consider a distribution of mass on the surface $\mathcal{S}$. Let its density at P be $\sigma(P)$. The **moment of inertia** of the mass around the z-axis is defined as $\int_{\mathcal{S}}(x^2+y^2)\sigma(P)\ dS$. Exercises 20 and 21 concern this integral.

20.[R] Find the moment of inertia of a homogeneous distribution of mass on the surface of a ball of radius a around a diameter. Let the total mass be M.

21.[R] Find the moment of inertia about the z-axis of a homogeneous distribution of mass on the triangle whose vertices are $(a, 0, 0)$, $(0, b, 0)$, and $(0, 0, c)$. Take a, b, and c to be positive. Let the total mass be M.

22.[R] Let $\mathcal{S}$ be a sphere of radius a. Let A be a point at distance $b > a$ from the center of $\mathcal{S}$. For P in $\mathcal{S}$ let $\delta(P)$ be $1/q$, where q is the distance from P to A. Show that the average of $\delta(P)$ over $\mathcal{S}$ is $1/b$.

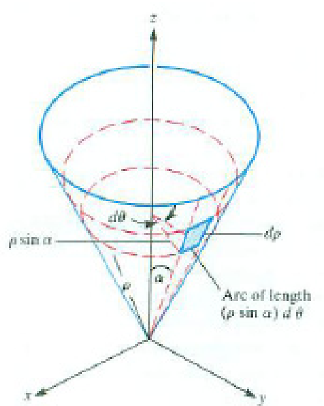
23.[R] The data are the same as in Exercise 22 but $b < a$. Show that in this case the average of $1/q$ is $1/a$. (The average does *not* depend on b in this case.)

Exercises 24 to 26 concern integration over the curved surface of a cone. Spherical coordinates are also useful for integrating over a right circular cone. Place the origin at the vertex of the cone and the "$\phi = 0$" ray along the axis of the cone, as shown in Figure 17.7.8(a). Let α be the half-vertex angle of the cone.
On the surface of the cone ϕ is constant, $\phi = \alpha$, but ρ and θ vary. A small "rectangular" patch on the surface of the cone corresponding to slight changes $d\theta$ and $d\rho$ has area approximately

$$(\rho \sin(\alpha)\ d\theta)\ d\rho = \rho \sin(\alpha)\ d\rho\ d\theta.$$

(See Figure 17.7.8.) So we may write

$$dS = \rho \sin\alpha\ d\rho\ d\theta.$$



ch16/f16-7-9

Figure 17.7.8:

24.[R] Find the average distance from points on the curved surface of a cone of radius a and height h to its axis.

25.[R] Evaluate $\int_{\mathcal{S}} z^2\ dS$, where $\mathcal{S}$ is the *entire* surface of the cone shown in Figure 17.7.8(b), including its base.

26.[R] Evaluate $\int_{\mathcal{S}} x^2\ dS$, where $\mathcal{S}$ is the curved surface of the right circular cone of radius 1 and height 1 with axis along the z-axis.

Integration over the curved surface of a right circular cylinder is easiest in cylindrical coordinates. Consider such a cylinder of radius a and axis on the z-axis. A small patch on the cylinder corresponding to dz and $d\theta$ has area approximately $dS = a\ dz\ d\theta$. (Why?) Exercises 27 and 28 illustrate the use of these coordinates.

27.[R] Let $\mathcal{S}$ be the *entire* surface of a solid cylinder of radius a and height h. For P in $\mathcal{S}$ let $f(P)$ be the square of the distance from P to one base. Find $\int_{\mathcal{S}} f(P)\ dS$. Be sure to include the two bases in the integration.

28.[R] Let $\mathcal{S}$ be the curved part of the cylinder in Exercise 27. Let $f(P)$ be the square of the distance from P to a fixed diameter in a base. Find the average value of $f(P)$ for points in $\mathcal{S}$.

29.[R] The areas of the projections of a small flat surface patch on the three coordinate planes are 0.01, 0.02, and 0.03. Is that enough information to find the area of the patch? If so, find the area. If not, explain why not.

30.[R] Let $\mathbf{F}$ describe the flow of a fluid in space. (See Section 16.3 for fluid flow in a planar region.) $\mathbf{F}(P) = \delta(P)\mathbf{v}(P)$, where $\delta(P)$ is the density of the fluid at P and $\mathbf{v}(P)$ is the velocity of the fluid at P. Making clear, large diagrams, explain why the rate at which the fluid is leaving the solid region enclosed by a surface $\mathcal{S}$ is $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS$, where $\mathbf{n}$ denotes the unit outward normal to $\mathcal{S}$.

31.[R] Let $\mathcal{S}$ be the smooth surface of a convex body. Show that $\int_{\mathcal{S}} z\cos(\gamma)\, dS$ is equal to the volume of the solid bounded by $\mathcal{S}$. HINT: Break $\mathcal{S}$ into two parts. In one part $\cos(\gamma)$ is positive; and the other it negative.

32.[M] Let $R(x, y, z)$ be a scalar function defined over a closed surface $\mathcal{S}$. (See Figure 17.7.9.)

(a) Show that

$$\int_{\mathcal{S}} R(x,y,z)\cos(\gamma)\, dS = \int_{\mathcal{A}} \left(P(x,y,z_2) - P(x,y,z_1)\right)\, dA,$$

where $\mathcal{A}$ is the projection of $\mathcal{S}$ on the xy plane and the line through $(x, y, 0)$ parallel to the z-axis meets $\mathcal{S}$ at (x, y, z_1) and (x, y, z_2), with $z_1 \le z_2$.

(b) Let $\mathcal{S}$ be a surface of the type in (a). Evaluate $\int_{\mathcal{S}} x\cos\gamma\, dS$.

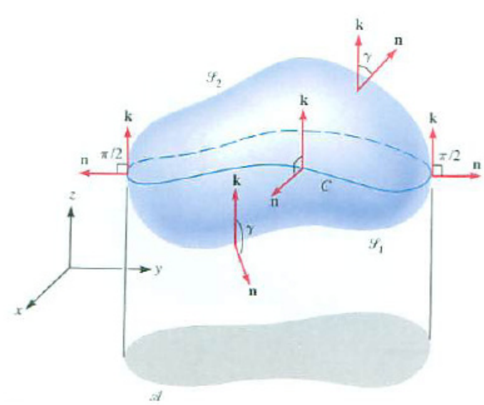


Figure 17.7.9:

33.[C]

(a) Let g be a differentiable function such that $g((x+y)/2) = ((g(x)+g(y))/2$ for all x and y. Show that $g(x) = kx + c$ for some constraints k and c. HINT: Differentiate.

(b) Let f be a differentiable function such that $(x+y)f(x+y)+(x-y)f(x-y) = 2xf(x)$ for all x and y. Deduce that there are constraints k and c such that $f(x) = k + c/x$.

34.[C] (Suggested by Exercises 22 and 23.) The function $f(x) = 1/x$ has the remarkable property that the average value of $f(d(P))$ over a sphere is the same as $f(H)$. Here $d(P)$ is the distance from P to a fixed point at a distance H for the center of a sphere, of radius a, $a < H$. Show that the only functions with this property have the form $k + c/x$ for some constraints k and c. HINT: Use part of the Fundamental Theorem of Calculus to remove integration. Then the Exercise 33 many come in handy.

17.8 Magnification, Jacobian, and Change of Coordinates

We now consider functions whose domain and range are parts of planes, curved surfaces, or spatial regions. Of particular interest is how much they magnify or shrink the areas (or volumes) of small regions. This magnifying factor will be used in Chapter 17 to simplify some definite integrals over two- and three-dimensional sets.

Throughout we assume the functions have continuous derivatives.

Mappings

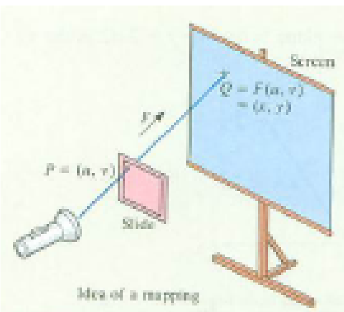


Figure 17.8.1:

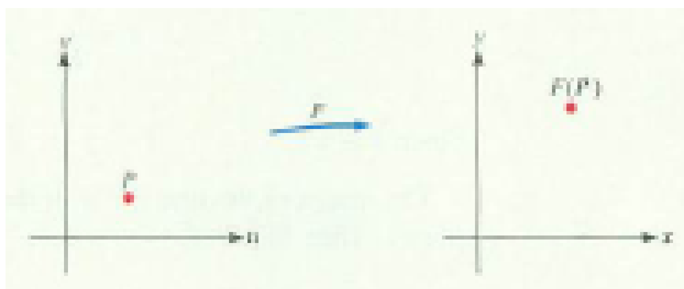


Figure 17.8.2:

EXAMPLE 1 Let F be the mapping that assigns to the point (u, v) the point $(2u, 3v)$.

(a) Describe the mapping geometrically.

(b) Find the image of the line $v = u$.

(c) Find the image of the square in the uv-plane whose vertices are $(0, 0)$, $(1, 0)$, $(1, 1)$, and $(0, 1)$.

SOLUTION

(a) In this case, $x = 2u$ and $y = 3v$. The table below records the effect of the mapping on the points listed in (c):

(u, v)	$(0, 0)$	$(1, 0)$	$(1, 1)$	$(0, 1)$
$(2u, 3v)$	$(0, 0)$	$(2, 0)$	$(2, 3)$	$(0, 3)$

In the notation $F(u, v) = (x, y)$, these data read

$F(0, 0) = (2{\cdot}0, 3{\cdot}0) = (0, 0)$; $F(1, 0) = (2{\cdot}1, 3{\cdot}0) = (2, 0)$; $F(1, 1) = (2{\cdot}1, 3{\cdot}1) = (2$

Note that the first coordinate of $(x, y) = F(u, v)$ is $x = 2u$, twice the first coordinate of (u, v). Thus the mapping magnifies horizontally by a factor of 2. Similarly, it stretches vertically by a factor of 3. This causes a six-fold magnification of areas.

(b) Let $P = (u, v)$ be on the line $v = u$. Then $F(P) = F(u, v) = (x, y)$, with $x = 2u$ and $y = 3v$. Thus

$$u = \frac{x}{2} \qquad \text{and} \qquad v = \frac{y}{3}.$$

Since $v = u$,

$$\frac{y}{3} = \frac{x}{2} \qquad \text{or} \qquad y = \frac{3}{2}x.$$

The image of the line $v = u$ in the uv-plane is the line $y = 3x/2$ in the xy-plane. (See Figure 17.8.3.)

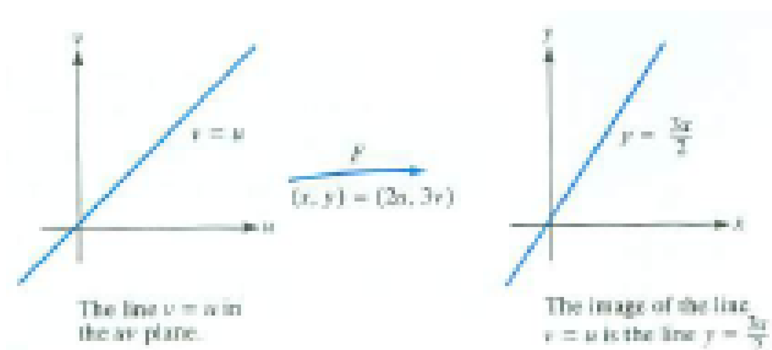


Figure 17.8.3:

A similar argument shows that for this mapping the image of any line $Au + Bv + C = 0$ in the uv-plane is a line in the xy-plane, namely, the line $Ax/2 + By/3 + C = 0$.

(c) If P is a point in the square R whose vertices are

$$(0, 0) \quad (1, 0), \quad (1, 1), \quad (0, 1),$$

then the image of P is a point in the rectangle S whose vertices are

$$(0, 0) \quad (2, 0), \quad (2, 3), \quad (0, 3).$$

(See Figure 17.8.4.)

Think of (u, v) as a point on a slide and $(2u, 3v)$ as its image on the screen. Then the mapping F projects the square R on the slide onto a rectangle S on the screen. (See Figure 17.8.5.)

◇

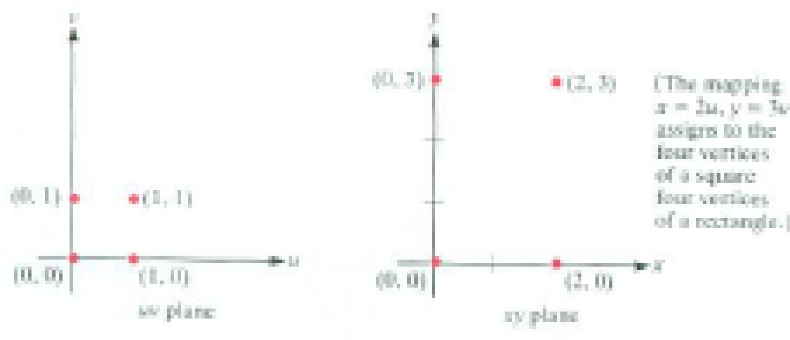


Figure 17.8.4:

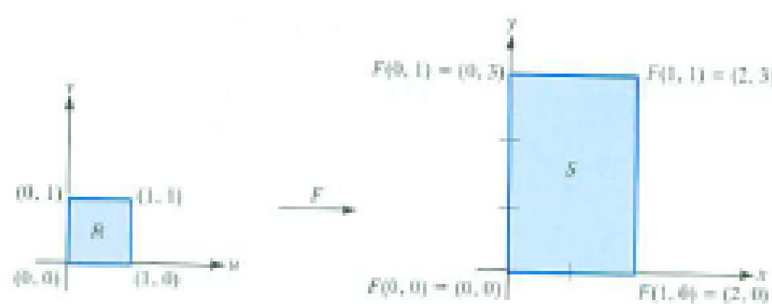


Figure 17.8.5:

Summary

EXERCISES for Section 17.8 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4 compute m_F by the mappings at the given points in the uv-plane.

1.[R] $F(u,v) = (uv, v^2)$, u, $v > 0$, at (a) $(1,2)$ and (b) $(3,1)$.

2.[R] $F(u,v) = (1/u, 1/v)$, u, $v > 0$, at (a) $(2,3)$ and (b) $(\frac{1}{2}, 4)$.

3.[R] $F(u,v) = (e^u \cos v, e^u \sin v)$, $0 \le v\mathcal{L}\,[2]\,\pi$, at (a) $(1, \pi/4)$ and (b) $(2, \pi/6)$.

4.[R] $F(u,v) = (u/(u^2+v^2), v/(u^2+v^2))$, $u^2+v^2 \ne 0$, at (a) $(3,1)$ and (b) $(1,0)$.

5.[R] Let a, b, c, and d be constants such that $ad - bc \ne 0$. Let

$$x = au + bv, \qquad y = cu + dv.$$

Show that the determinant Jacobian of the mapping is $ad - bc$ at all points.

6.[R] The magnification of a mapping is 3 at $(2,4)$. Let R be a small region around $(2,4)$ of area 0.05. Approximately, how large is the image of R under the mapping?

17.9 Moments, Centers of Mass, and Centroids

Now that we can integrate over planar regions, surfaces, and solid regions, we can define and calculate the center of mass of a physical object. The center of mass is important in the eyes of a naval architect, who wants his ships not to tip over easily. A pole vaulter hopes that as she clears the bar her center of mass goes under it. Archimedes, the first person to study the center of mass, was interested in the stability of floating paraboloids.

The Center of Mass

A small boy on one side of a seesaw (which we regard as weightless) can balance a bigger boy on the other side. For example, the two boys in Figure 17.9.1 balance. (According to physical laws, each boy exerts a force on the seesaw, due to gravitational attraction, proportional to his mass.)

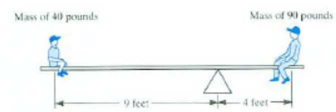


Figure 17.9.1:

The small mass with the long lever arm balances the large mass with the small lever arm. Each boy contributes the same tendency to turn–but in opposite directions.

This tendency is called the **moment**:

$$\text{Moment} = (\text{Mass}) \cdot (\text{Lever arm}),$$

where the lever arm can be positive or negative. To be more precise, introduce on the seesaw an x-axis with its origin 0 at the fulcrum, the point on which the seesaw rests. Define the moment about 0 of a mass m located at the point x on the x-axis to be the product mx. Then the bigger boy has a moment (90)(4), which the smaller boy has a moment (40)(−9). The total moment of the lever-mass system is 0, and the masses balance. (See Figure 17.9.2.)

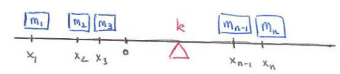


Figure 17.9.2:

If a mass m is located on a line with coordinate x, we define its moment about the point having coordinate k as the product $m(x-k)$.

Now consider several point masses $m_1, m_2, \dots, m_i$. If mass m_i is located at x_i, with $i = 1, 2, \dots, n$, then $\sum_{i=1}^{n} m_i(x_i - k)$ is the total moment of all the masses about the point k. If a fulcrum is placed at k, then the seesaw rotates clockwise if the total moment is greater that 0, rotates counterclockwise if it is less than 0, and is in equilibrium if the total moment is 0. See Figure 17.9.3.

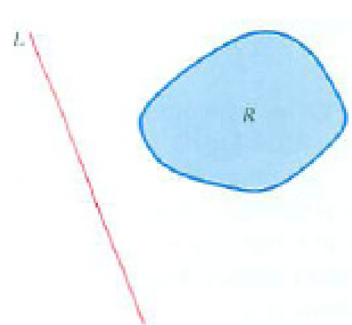


Figure 17.9.3:

To find where to place the fulcrum so that the total tendency to turn is 0, we find k such that

$$\sum_{i=1}^{n} m_i(x_i - k) = 0.$$

Writing this as

$$k \sum_{i=1}^{n} m_i = \sum_{i=1}^{n} m_i - x_i,$$

we see that

$$k = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i}. \tag{17.9.1}$$

The number k given by (17.9.1) is called the **center of mass** or center of gravity of the system of masses. It is the point about which all the masses balance. *The center of mass is found by dividing the total moment about* 0 *by the total mass.* It is usually denoted $\bar{x}$.

(a) (b)

Figure 17.9.4:

Finding the center of mass of a finite number of "point masses" involves only arithmetic, no calculus. For example, suppose three masses are placed on a seesaw as in Figure 17.9.4(a). Introduce an x-axis with origin at mass $m_1 = 20$ pounds. Two additional masses are located at $x_2 = 4$ feet and $x_3 = 14$ feet with masses $m_2 = 10$ pounds and $m_3 = 50$ pounds, respectively. The total moment about $x = k$ is

$$M = 20(0-k) + 10(4-k) + 50(14-k) = 740 - 80k.$$

This moment vanishes when $M = 0$, that is, when $k = 740/80 = 9.25$.

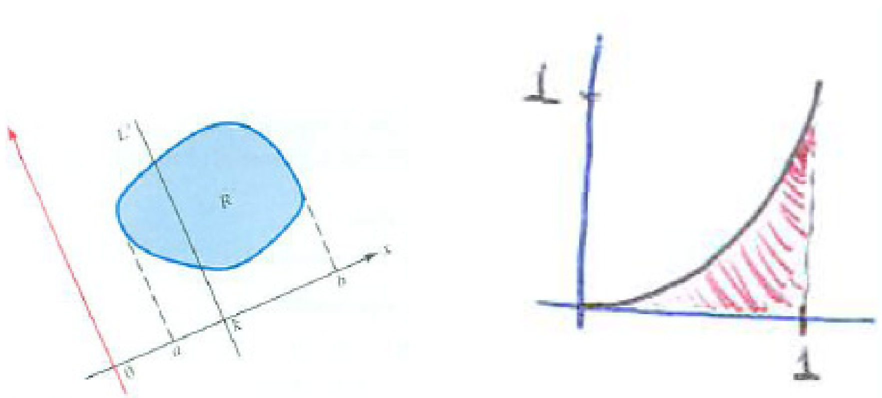
This is consistent with the formula for the center of mass:

$$\bar{x} = \frac{m_1x_1 + m_2x_2 + m_3x_3}{m_1 + m_2 + m_3} = \frac{0 + 40 + 700}{10 + 20 + 50} = \frac{740}{80} = 9.25.$$

The seesaw balances when the fulcrum is placed 9.25 feet from the first mass. (See Figure 17.9.4(b).)

Now let us turn our attention to finding the center of mass of a continuous distribution of matter in a plane region. For this purpose, we consider double integrals.

Let R be a region in the plane occupied by a thin piece of metal whose density, $\sigma(P)$, varies. Let L be a line in the plane, as shown in Figure 17.9.5(a). We will find a formula for the unique line parallel to L, around which the mass in R balances.



(a) (b)

Figure 17.9.5:

To begin, let L' be *any* line parallel to L. We will compute the moment about L' and then see how to choose L' to make that moment equal to 0. To compute the moment of R about L', introduce an x-axis perpendicular to L with its origin at its intersection with L. Assume that L' passes through the x-axis at the point$x = k$, as in Figure 17.9.5(b). In addition, assume that each line parallel to L meets R either in a line segment or at a point on the boundary of R. The lever arm of the mass distributed throughout R varies from point to point.

We partition R into n small regions $R_1, R_2, \dots, R_n$. Call the area of R_i, A_i. In each of these regions the lever arm around L' varies only a little. So, if we pick a point P_1 in R_1, P_2 in R_2, $\dots$, P_n, in R_n, and the x-coordinate of P_i is x_i, then

$$\underbrace{(x_i - k)}_{\text{lever arm}}\underbrace{\sigma(P_i)A_i}_{\text{mass in } R_i}$$

is a local estimate of the turning tendency.

Thus

$$\sum_{i=1}^{n}(x_i - k)\sigma(P_i)A_i \tag{17.9.2}$$

would presumably be a good estimate of the total turning tendency around L'. Taking the limit of (17.9.2) as all R_i are chosen smaller and smaller, we expect

$$\int\limits_R (x - k)\sigma(P)\; dA \tag{17.9.3}$$

to represent the turning tendency of the total mass around L'. The quantity

(17.9.3) is called the **moment of torque** of the mass distribution around L'.

EXAMPLE 1 Let R be the region under $y = x^2$ and above $[0, 1]$ DOUG with the density $\sigma(x, y) = xy$. Find its moment around the line $x = 1/2$.
SOLUTION R is shown in Figure 17.9.6. The moment (17.9.3) equals

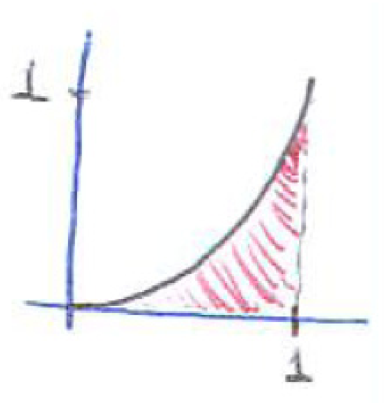


Figure 17.9.6:

$$\int\limits_R \left(x - \frac{1}{2}\right) xy \, dA. \tag{17.9.4}$$

We evaluate this double integral by the iterated integral

$$\int\limits_0^1 \left(\int\limits_0^{x^2} \left(x - \frac{1}{2}\right) xy \, dy \right) \, dx.$$

See Exercise 2.

The first integration gives

$$\int\limits_0^{x^2} \left(x - \frac{1}{2}\right) xy \, dy \quad = \quad (x - 1/2)x \int\limits_0^{x^2} y \, dy = \frac{\left(x - \frac{1}{2}\right) x^5}{2}.$$

The second integration is

$$\int\limits_0^1 \frac{\left(x - \frac{1}{2}\right) x^5}{2} = \int\limits_0^1 \frac{2x^6 - x^5}{4} \, dx = \frac{5}{168}.$$

Since the total moment (17.9.4) is positive, the object would rotate clockwise around the line $x = \frac{1}{2}$. ◇

Now that we have a way to find the moment around any line parallel to the y-axis we can find the line around which the moment is zero, the so-called "balancing line." We just solve for k in the equation

$$\int\limits_R (x-k)\sigma(P)\ dA = 0.$$

Thus

$$\int\limits_R x\sigma(P)\ dA = k\int\limits_R \sigma(P)\ dA,$$

from which we find that

$$k = \frac{\int_R x\sigma(P)\ dA}{\int_R \sigma(P)\ dA}. \tag{17.9.5}$$

The denominator is the total mass. The numerator is the total torque. So we can think of k as "the average lever arm as integrated by the density."

That is therefore a unique balancing line parallel to the y axis. Call its x-coordinate $\overline{x}$ (read: "x bar"). Similarly, there is a unique balancing line parallel to the x axis. Call its y-coordinate $\overline{y}$. The point $(\overline{x}, \overline{y})$ is called the center of mass of the region R. We have:

The center of mass of a region R with density $\sigma(P)$ has coordinates $(\bar{x}, \bar{y})$ where

$$\overline{x} = \frac{\int_R x\sigma(P)\ dA}{\int_R \sigma(P)\ dA} \qquad \text{and} \qquad \overline{y} = \frac{\int_R y\sigma(P)\ dA}{\int_R \sigma(P)\ dA}.$$

The integral $\int_R x\sigma(P)\ dA$ is called the **moment of R around the y-axis**, and is denoted M_y. Similarly, $M_x = \int_R y\sigma(P)\ dA$.

If the density $\sigma(P)$ is constant, say, equal to 1 everywhere in R, then the two equations reduce to

$$\overline{x} = \frac{\int_R x\ dA}{\int_R\ dA} \qquad \text{and} \qquad \overline{y} = \frac{\int_R y\ dA}{\int_R\ dA}.$$

In this case the center of mass R is also called the **centroid** of the region, a purely geometric concept:

The centroid of the plane region R has the coordinates $(\bar{x}, \bar{y})$ where

$$\overline{x} = \frac{\int_R x\ dA}{\text{Area of } R} \qquad \text{and} \qquad \overline{y} = \frac{\int_R y\ dA}{\text{Area of } R)}. \tag{17.9.6}$$

EXAMPLE 2 Find the center of mass of the region in Example 1.

SOLUTION The density at (x, y) in R is given by $\sigma = xy$. We compute three double integrals: the mass $\int_R xy\ dA$ and the two moments $M_y = \int_R x(xy)\ dA$ and $M_x = \int_R y(xy)\ dA$.

We have

$$\int_R x^2y\ dA = \int_0^1 \left(\int_0^{x^2} x^2y\ dy \right)\ dx = \int_0^1 \frac{x^6}{2}\ dx = \frac{1}{14}.$$

Then

$$\int_R xy\ dA = \int_0^1 \left(\int_0^1 xy\ dy \right)\ dx = \int_0^1 \frac{x^5}{2}\ dx = \frac{1}{12}.$$

Finally,

$$\int_R xy^2\ dA = \int_1^0 \left(\int_1^0 xy^2\ dy \right)\ dx = \int_0^1 \frac{x^7}{3}\ dx = \frac{1}{24}.$$

Thus

$$\overline{x} = \frac{\frac{1}{14}}{\frac{1}{12}} = \frac{6}{7} \qquad \text{and} \qquad \overline{y} = \frac{\frac{1}{24}}{\frac{1}{12}} = \frac{1}{2}.$$

It is not surprising that $\overline{x}$ is greater than 1/2, since in Example 17.9.1 we found that the object rotates clockwise around the line $x = 1/2$. ◇

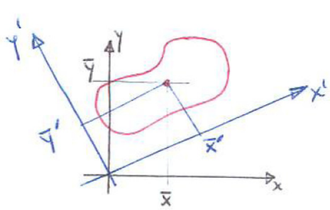


Figure 17.9.7:

An Important Point About an Important Point

We defined the center of mass $(\overline{x}, \overline{y})$ by first choosing an xy coordinate system. What if we choose an $x'y'$ coordinate system at an angle to the xy coordinate system? Would the center of mass computed in this system, $(\overline{x}', \overline{y}')$ be the same point as $(\overline{x}, \overline{y})$? See Figure 17.9.7. Fortunately, it is, as Exercise 59 shows.

Shortcuts for Computing Centroids

Assume that R is the region under $y = f(x)$ for x in $[a, b]$. Then the moment about the x-axis is

$$M_x = \int_R y\ dA.$$

Thus

$$M_y \quad = \int_a^b \left(\int_0^{f(x)} y\ dy \right) dx = \int_a^b \frac{(f(x))^2}{2}\ dx = \frac{1}{2}\int_a^b (f(x))^2\ dx.$$

Thus, by (17.9.6)

$$\overline{y} = \frac{\frac{1}{2}\int_a^b (f(x))^2\ dx}{\text{Area of } R}. \qquad (17.9.7)$$

EXAMPLE 3 Find the centroid of the semicircular region of radius a shown in Figure 17.9.8.

SOLUTION By symmetry, $\overline{x} = 0$.

To find $\overline{y}$, use (17.9.7). The function f in this case is given by the formula $f(x) = \sqrt{a^2 - x^2}$, an even function. The moment of R about the x-axis is

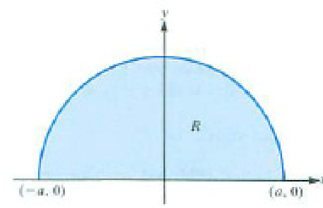


Figure 17.9.8:

$$\begin{aligned} \int_{-a}^{a} \frac{(\sqrt{z^2 - x^2})^2}{2}\ dx &= \int_{-a}^{a} \frac{a^2 - x^2}{2}\ dx = 2\int_0^a \frac{a^2 - x^2}{2}\ dx \\ &= \int_0^a (a^2 - x^2)\ dx = \left(a^2 x - \frac{x^3}{3} \right)\Bigg|_0^a \\ &= (a^3 - \frac{a^3}{3}) - 0 = \frac{2}{3}a^3. \end{aligned}$$

Thus

$$\overline{y} = \frac{\frac{2}{3}a^3}{\text{Area of } R} = \frac{\frac{2}{3}a^3}{\frac{1}{2}\pi a^2} = \frac{4a}{3\pi}.$$

◇

Since $4/(3\pi) \approx 0.42$, the center of gravity of R is at a height of about $0.42a$.

Centers of Other Masses

We developed the ideas of moments and centers of mass for masses situation in a plane. The definition generalizes easily to masses distributed on a curve (such as a wire) or in space (such as a potato).

In the case of a curve, the curve would have a linear density $\lambda(P)$. A short piece around P of length Δs would have mass approximately $\lambda(P)\Delta s$. Thus, the mass and moments of the curve would be

$$M = \int\limits_C \lambda(P)\ ds, \qquad M_y = \int\limits_C x\lambda(P)\ ds, \quad \text{and} \quad M_x = \int\limits_C y\lambda(P)\ ds.$$

We state the definition in the case of a solid object of density $\lambda(P)$ occupying the region R. We assume an xyz-coordinate system. The total mass is

$$M = \int\limits_R \delta(P)\ dV.$$

Now, there are three moments — one around each of the three coordinate planes:

$$M_{yz} = \int\limits_R x\delta(P)\ dV, M_{xz} = \int\limits_R y\delta(P)\ dV, M_{xy} \int\limits_R z\rho\ dV.$$

The center of mass is $(\overline{x}, \overline{y}, \overline{z})$, where

$$overlinex = \frac{\int_R x\delta(P)\ dV}{M}, \qquad \overline{y} = \frac{\int_R y\delta(P)\ dv}{M}, \qquad \overline{z} = \frac{\int_R z\delta(P)\ dV}{M}.$$

If $\delta(P) = 1$ for all P in R, the center of mass is called the centroid. In this case the mass is the same as the volume.

EXAMPLE 4 Find the centroid of a hemisphere of radius a.

SOLUTION We place the origin of an xyz-coordinate system at the center of the hemisphere, as in Figure 17.9.9.

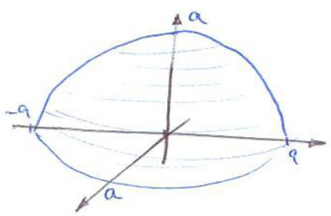


Figure 17.9.9:

First of all, by symmetry, the centroid must be at the z-axis. mnoteIf the centroid were not at the z-axis, you would get two centroids for the same object. (If you spin the hemisphere about the z-axis you get the same hemisphere back, which must have the same centroid.)

So $\overline{x} = \overline{y} = 0$. Calling the hemisphere R, we have

$$\overline{z} = \frac{\int_R z\ dV}{\text{Volume of } R}.$$

The volume of the hemisphere is half that of a ball, $(2/3)\pi a^3$. To evaluate the moment $\int_R z\ dV$, we bring in an iterated integral in spherical coordinates:

$$\int\limits_R z\ dV = \int\limits_0^{2\pi} \int\limits_0^{\pi/2} \int\limits_0^{a} (\rho\cos(\phi))\rho 2\sin(\phi)\ d\rho\ d\phi\ d\theta.$$

See Exercise 3.

Straightforward computations show that

$$\int_R z \, dV = \frac{\pi a^4}{4}.$$

Thus

$$\overline{z} = \frac{\frac{\pi a^4}{4}}{\frac{2}{3}\pi a^3} = \frac{3a}{8}.$$

The centroid is $(0, 0, \frac{3a}{8})$. ◇

EXAMPLE 5 Find the centroid of a homogeneous cone of height h and radius a.

SOLUTION As we just saw for the sphere in Example 4, symmetry tells us the centroid lies on the axis of the cone.

Introduce a spherical coordinate system with the origin at the vertex of the cone and with the axis of the cone lying on the ray $\phi = 0$, as in Figure 17.9.10.

The half-vertex angle is $\arctan(a/h)$. The plane of the base of the cone is $z = h$ (in rectangular coordinates), hence

$$\rho\cos(\phi) = h.$$

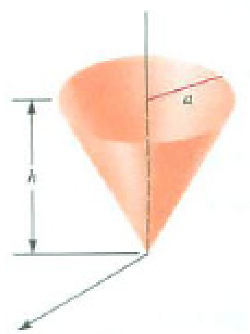


Figure 17.9.10:

In spherical coordinates, the cone's description is

$$0 \le \theta \le 2\pi, \qquad 0 \le \phi \le \arctan(a/h), \qquad 0 \le \rho \le h/\cos(\phi).$$

To find the centroid of the cone we compute $\int_R z \, dV$ and divide the results by the volume of the cone, which is $\frac{1}{3}\pi a^2 h$.

Now

$$\int_R z \, dV = \int_0^{2\pi} \int_0^{\arctan(a/h)} \int_0^{h/\cos(\phi)} \rho\cos(\phi)(\rho^2\sin(\phi)) \, d\rho \, d\phi \, d\theta.$$

See Exercise 4

For the first integration, ϕ and θ are constant; hence

$$\int_0^{h/\cos(\phi)} \rho\cos(\phi)\rho^2\sin(\phi) \, d\rho = \cos(\phi)\sin(\phi) \int_0^{h/\cos(\phi)} \rho^3 \, d\rho = \frac{h^4\sin(\phi)}{4\cos^3(\phi)}.$$

The second integration is

$$\int_0^{\arctan(a/h)} \frac{h^4\sin(\phi)}{4\cos^3(\phi)} \, d\phi = \frac{h^4}{4} \int_0^{\arctan(a/h)} \frac{\sin(\phi)}{\cos^3(\phi)} \, d\phi = \frac{a^2h^2}{8}.$$

	curve (C)	solid (R)
density	$\lambda(P)$	$\delta(P)$
M	$\int_C \lambda(P)\ ds$	$\int_S \delta(P)\ dV$
M_{yz}	$\int_C x\lambda(P)\ ds$	$\int_S x\delta(P)\ dV$
M_{xz}	$\int_C y\lambda(P)\ ds$	$\int_S y\delta(P)\ dV$
M_{xy}	$\int_C z\lambda(P)\ ds$	$\int_S z\delta(P)\ dV$

The final integral is simply:

$$\int_0^{2\pi} \frac{a^2h^2}{8}\ d\theta = \frac{a^2h^2}{8}2\pi = \frac{\pi a^2h^2}{4}.$$

Thus,

$$\overline{z} = \frac{\int_R z\ dV}{\text{Volume of } R} = \frac{\left(\frac{\pi a^2h^2}{4}\right)}{\left(\frac{\pi a^2 h}{3}\right)} = \frac{3h}{4}.$$

The centroid of a cone is three-fourths of the way from the vertex to the base. ◇

Summary

We defined the moment about a line and used this concept to define the center of mass for a plane distribution of mass. The moment of a mass about a line L indicates the tendency of the mass to rotate about the line L. The center of mass for a region R is the point in the region where the region balances.

- The moment about the y-axis, M_y, is $\int_R x\delta(P)\ dA$.
- The moment about the x-axis, M_x, is $\int_R y\delta(P)\ dA$.

Then, the center of mass is $(\bar{x}, \bar{y})$ where

$$\overline{x} = \frac{M_y}{\text{Mass}}, \overline{y} = \frac{M_x}{\text{Mass}}.$$

If the density is constant, we have a purely geometric concept,

$$\overline{x} = \frac{\int_R x\ dA}{\text{Area of } R}, \overline{y} = \frac{\int_R y\ dA}{\text{Area of } R}.$$

These definitions generalize to curves and solids.

EXERCISES for Section 17.9 *Key:* R–routine, M–moderate, C–challenging

1.[R]

(a) How would you define the centroid of a curve? Call its (linear) density $\lambda(P)$.

(b) Find the centroid of a semicircle of radius a.

2.[R] Carryout the integrations in Example 1.

3.[R] Carryout the "straightforward calculations" in Example 4.

4.[R] Provide the details needed to complete the integrals in Example 5.

5.[R] Example 4 showed that the centroid of a hemisphere is less than halfway from the center to its surface. Why is that to be expected?

6.[M] If R is the region below $y = f(x)$ and above $[a, b]$, show that

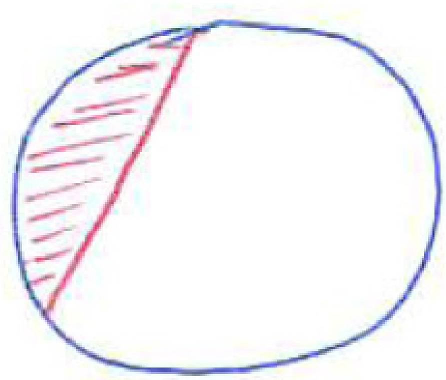
$$\overline{x} = \frac{\int_a^b x f(x)\ dx}{\text{Area of } R}.$$

7.[M] The corners of a triangular piece of metal of constant density 1 are $(0, 0)$, $(1, 0)$, and $(0, 2)$.

(a) Is the line $y = 11x/5$ a balancing line?

(b) If not, if the metal rests on this line which way would it rotate?

DEFINITION *(Section of a region)* Let R be a convex set in the plane. A **section** of R is a part of R that is bounded by a chord and part of the



the boundary, as shown as Figure 17.9.11

Figure 17.9.11:

8.[C] Consider a convex set R in the plane furnished with a density. Show that different sections have different centers of gravity.

9.[C] (See Exercise 8.) Is every point in R that is not on the boundary the center of mass of some section of R?

10.[C] Archimedes (287-212 B.C.) investigated the centroid of a section of a parabola. Consider the parabola $y = x^2$. The typical section is shown in Figure 17.9.12. M is the midpoint of the chord and N is the point on the parabola directly below M.

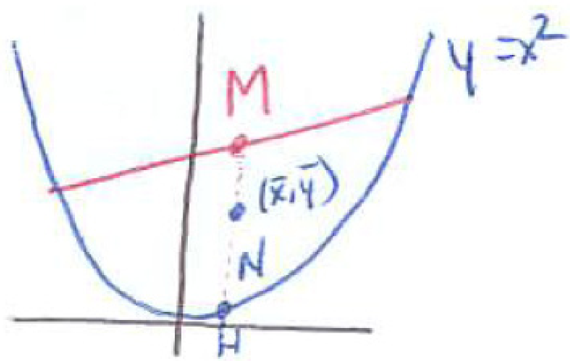


Figure 17.9.12:

He showed, without calculus, that the centroid is on the line MN, three-fifths of the way from N and M. Obtain his result with the aid of calculus.

11.[C] (See Exercise 10.) Is every point in the region bounded by the parabola the centroid of some section?

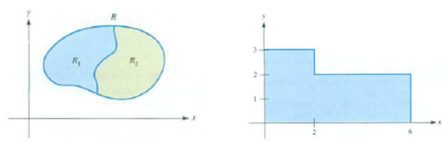
12.[R] Find the centroid of a solid paraboloid of revolution. This is the region above $z = x^2 + y^2$ and below the plane $z = c$. Archimedes solved this problem without calculus and used the result to analyze the equilibrium of a floating paraboloid. (If it is slightly tilted, will it come back to the vertical or topple over?) For details as how he did this 2200 years ago see S. Stein, *Archimedes: What Did He Do Besides Cry Eureka?*, Math. Assoc. America, 1999.

13.[C] (See Exercise 12.) The plane $z = c$ in Exercise 12 is perpendicular to the axis of the paraboloid. Archimedes was also interested in the case when the plane is not perpendicular to the axis. Find the centroid of the region below the tilted plane $z = cy$ and above the paraboloid $z = x^2 + y^2$.

14.[R] Using cylindrical coordinates, find $\bar{z}$ for the region below the paraboloid $z = x^2 + y^2$ and above the disk in the $r\theta$ plane bounded by the circle $r = 2$. (Include a drawing of the region.)

15.[R] Find the z coordinate, $\overline{z}$, of the centroid of the part of the saddle $z = xy$ that lies above the portion of the disk bounded by the circle $x^2 + y^2 = a^2$ in the first quadrant.

16.[M] A plane distribution of matter occupies the region R. It is cut into two pieces, occupying regions R_1 and R_2, as in Figure 17.9.13(a). The part in R_1 has mass M_1 and centroid $(\overline{x}_1, \overline{y}_1)$. The part in R_2 has mass M_2 and centroid $(\overline{x}_2, \overline{y}_2)$. Find the centroid $(\overline{x}, \overline{y})$ of the entire mass, which occupies R. [Express $(\overline{x}, \overline{y})$ in terms of M_1, M_2, $\overline{x}_1$, $\overline{x}_2$, $\overline{y}_1$ and $\overline{y}_2$.]



(a) (b)

Figure 17.9.13:

17.[M] Use the formula in Exercise 16 to find the center of mass of the homogeneous lamina shown in Figure 17.9.13(b).

In Exercises 18 to 25 find the centroid of the given regions R. (Exercises 22 to 25 require integral tables or techniques of Chapter 8.)

18.[R] R is bounded by $y = x^2$ and $y = 4$.

19.[R] R is bounded by $y = x^4$ and $y = 1$.

20.[R] R is bounded by $y = 4x - x^2$ and the x-axis.

21.[R] R is bounded by $y = x$, $x + y = 1$, and the x-axis.

22.[R] The region bounded by $y = e^x$ and the x-axis, between the lines $x = 1$ and $x = 2$.

23.[R] The region bounded by $y = \sin(2x)$ and the x-axis, between the lines $x = 0$ and $x = \pi/2$.

24.[R] The region bounded by $y = \sqrt{1 + x}$ and the x-axis, between the lines $x = 0$ and $x = 3$.

25.[R] The region bounded by $y = \ln(x)$ and the x-axis between the lines $x = 1$ and $x = e$.

Exercises 26 to 28 concern Pappus's Theorem, which relates the volume of a solid of revolution to the centroid of the planar region R that is revolved to form the solid.

Theorem 17.9.1 (Pappus)**.** *Let R be a region in the plane and L a line in the plane that does not cross R (though it can touch R at its border). Then the volume of the solid formed by revolving R about L is equal to the product*

$$(\textit{Distance the centroid of } R \textit{ is rotated}) \cdot (\textit{Area of } R)\,.$$

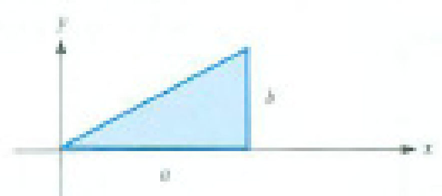
26.[C]

(a) Prove Pappus's Theorem

(b) Use Pappus's Theorem to find the volume of the torus or "doughnut" formed by revolving a circle of radius 3 inches about a line 5 inches from its center.

27.[C] Use Pappus's Theorem to find the centroid of the half disk R of radius a.

28.[C] Use Pappus's Theorem to find the centroid of the right triangle in Fig-



ure 17.9.14.

Figure 17.9.14:

29.[M] Consider a distribution of mass in a plane region R with density $\sigma(P)$ at P. Use the following steps to show that any line in the plane that passes through the center of the mass is a balancing line.

(a) For convenience, place the origin of the xy-coordinate system at the center of mass. That is, assume $(\bar{x}, \bar{y}) = (0, 0)$. Show that $\int_R x\sigma(P)\, dA = 0$ and $\int_R y\sigma(P)\, dA = 0$.

(b) Let L be any line $ax + by = 0$ through the origin. Show that the moment of the mass about L is

$$\int_R \frac{ax + by}{\sqrt{a^2 + b^2}}\sigma(P)\, dA.$$

Hint: What is the distance from a point (x, y) in R to the line $ax + by = 0$?

(c) From (a) and (b) deduce that the moment of the mass about L is 0. Thus all balancing lines for the mass pass through a single point. Any two of them therefore determine that point, which is called the center of mass. It is customary to use the two lines parallel to the x and y axes to determine that point.

30.[M] (See Exercise 29.) Show that the moment of a mass occupying a solid region R about any plane through its center of mass is 0.

31.[C] This exercise concerns hydrostatic pressure. (See Section 7.6.)

(a) Show that the pressure of water against a submerged vertical surface occupying the plane region R equals the pressure at the centroid of R times the area of R.

(b) Is the assertion in (a) correct if R is not vertical?

In each of Exercises 32 to 39 find the center of mass of the lamina occupying the given region and having the given density.

32.[R] The triangle with vertices $(0,0)$, $(1,0)$, $(0,1)$; density at (x,y) is $x+y$.

33.[R] The triangle with vertices $(0,0)$, $(2,0)$, $(1,1)$; density at (x,y) is y.

34.[R] The square with vertices $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$; density at (x,y) equals to $y\arctan(x)$.

35.[R] The finite region bounded by $y=1+x$ and $y=2^x$; density at (x,y) is $x+y$.

36.[R] The triangle with vertices $(0,0)$, $(1,2)$, $(1,3)$; density at (x,y) is xy.

37.[R] The finite region bounded by $y=x^2$, the x-axis, and $x=2$; density at (x,y) is e^x

38.[R] The finite region bounded by $y=x^2$ and $y=x+6$, situated to the right of the y-axis; density at (x,y) is $2x$.

39.[R] The trapezoid with vertices $(0,0)$, $(3,0)$, $(2,1)$, $(0,1)$; density at (x,y) is $\sin(x)$.

40.[C] Let R be a region in a plane and P a point a distance $h>0$ from the plane. P and R determine a cone with base R and vertex P, as shown in Figure 17.9.15. Let the area of R be A. What can be said about the distance of the centroid of the cone from the plane of R?

(a) What is that distance in the case of a right circular cone?

(b) Experiment with another cone with any convenient base of your choice.

(c) Make a conjecture.

(d) Explain why the conjecture is true.

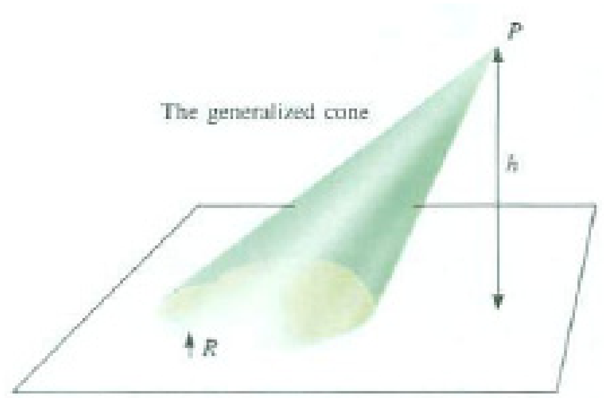


Figure 17.9.15:

In Exercises 41 and 42 find $\overline{z}$ for the given surfaces.

41.[M] The portion of the paraboloid $2z = x^2 + y^2$ below the plane $z = 9$.

42.[M] The portion of the plane $x + 2y + 3z = 6$ above the triangle in the xy plane whose vertices are $(0, 0)$, $(4, 0)$, and $(0, 1)$.

43.[R] In a letter of 1680 Leibniz wrote:

> Huygens, as soon as he had published his book on the pendulum, gave me a copy of it; and at that time I was quite ignorant of Cartesian algebra and also of the method of indivisibles, indeed I did not know the correct definition of the center of gravity. For, when by chance I spoke of it to Huygens, I let him know that I thought that a straight line drawn through the center of gravity always cut a figure into two equal parts; since that clearly happened in the case of a square, or a circle, an ellipse, and other figures that have a center of magnitude. I imagine that it was the same for all other figures. Huygens laughed when he heard this, and told me that nothing was further from the truth.

(Quoted in C.H. Edwards, *The Historical Development of the Calculus*, p. 239, Springer-Verlag, New York, 1979.)
Give an example showing that "nothing is further from the truth."

44.[R] Let a be a constant that is not less than 1. Let R be the region below $y = x^a$, above the x-axis, and between the lines $x = 0$ and $x = 1$.

(a) Sketch R for a large value of a.

(b) Compute the centroid $(\overline{x}, \overline{y})$ of R.

(c) Find $\lim_{a\to\infty} \overline{x}$ and $\lim_{a\to\infty} \overline{y}$.

(d) For large a, does the centroid of R lie in R?

45.[C] (Contributed by Jeff Lichtman) Let f and g be two continuous functions such that $f(x) \geq g(x) \geq 0$ for x in $[0,1]$. Let R be the region under $y = f(x)$ and above $[0,1]$; let R^* be the region under $y = g(x)$ and above $[0,1]$.

(a) Do you think the center of mass of R is at least as high as the center of mass of R^*? (An opinion only.)

(b) Let $g(x) = x$. Define $f(x)$ to be $\frac{1}{3}$ for $0 \leq x \leq \frac{1}{3}$ and $f(x)$ to be x if $\frac{1}{3} \leq x \leq 1$. (Note that f is continuous.) Find $\overline{y}$ for R and also for R^*. (Which is larger?)

(c) Let a be a constant, $0 \leq a \leq 1$. Let $f(x) = a$ for $0 \leq x \leq a$ and let $f(x) = x$ for $a \leq x \leq 1$. Find $\overline{y}$ for R.

(d) Show that the number a for which $\overline{y}$ defined in part (c) is a minimum is a root of the equation $x^3 + 3x - 1 = 0$.

(e) Show that the equation in (d) has only one real root q.

(f) Find q to four decimal places.

46.[M] This exercise shows that the three medians of a triangle meet at the centroid of the triangle. (A **median** of a triangle is a line that passes through a vertex and the midpoint of the opposite edge.)
Let R be a triangle with vertices A, B, and C. It suffices to show that the centroid of R lies on the median through C and the midpoint M of the edge AB. Introduce an xy coordinate system such that the origin is at A, and B lies on the x-axis, as in Figure 17.9.16.

(a) Compute $(\overline{x}, \overline{y})$.

(b) Find the equation of the median through C and M.

(c) Verify that the centroid lies on the median computed in (b).

(d) Why would you expect the centroid to lie on each median? (Just use physical intuition.)

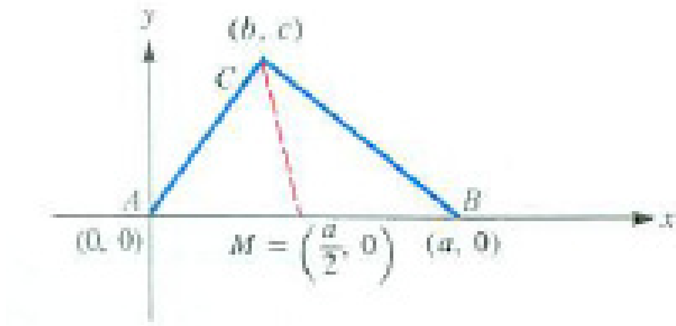


Figure 17.9.16:

47.[R] Cut an irregular shape out of cardboard and find three balancing lines for it experimentally. Are they concurrent; that is, do they pass through a common point?

48.[R] Let f and g be continuous functions such that $f(x) \geq g(x) \geq 0$ for x in $[a, b]$. Let R be the region above $[a, b]$ which is bounded by the curves $y = f(x)$ and $y = g(x)$.

(a) Set up a definite integral (in terms of f and g) for the moment of R about the y-axis.

(b) Set up a definite integral with respect to x (in terms of f and g) for the moment of R about the x-axis.

In Exercises 49 to 52 find (a) the moment of the given region R about the y-axis, (b) the moment of R about the x-axis, (c) the area of R, (d) $\overline{x}$, (e) $\overline{y}$. Assume the density is 1. (See Exercise 48.)

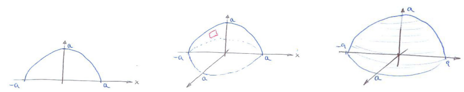
49.[R] R is bounded by the curves $y = x^2$, $y = x^3$.

50.[R] R is bounded by $y = x$, $y = 2x$, $x = 1$, and $x = 2$.

51.[R] R is bounded by the curves $y = 3^x$ and $y = 2^x$ between $x = 1$ and $x = e$.

52.[R] (Use a table of integrals or techniques from Chapter 8.) R is bounded by the curves $y = x - 1$ and $y = \ln(x)$, between $x = 1$ and $x = e$.

53.[M] Which do you think would have the highest centroid? The semicircular wire of radius a, shown in Figure 17.9.17(a); the top half of the surface of a ball of radius a, shown in Figure 17.9.17(b); the top half of a ball of radius a, shown in Figure 17.9.17(c).



(a) (b) (c)

Figure 17.9.17:

54.[C] Consider the parabolic surface $z = x^2 + y^2$ below the plane za^2.

(a) Set up a double integral in the xy-plane for the moment about the xy plane.

(b) Express this integral as an iterated integral in polar coordinates.

(c) Evaluate the integral.

(d) Find the centroid of the (curved) surface.

SHERMAN: This set of exercises was moved from Section 17.4.

Exercises 55 to 58 concern the moment of inertia. Note that if the object is homogeneous, has mass M and volume V, its density $\delta(P)$ is M/V.

55.[R] A homogeneous rectangular solid box has mass M and sides of lengths a, b, and c. Find its moment of inertia about an edge of length a.

56.[R] A rectangular homogeneous box of mass M has dimensions a, b and c. Show that the moment of inertia of the box about a line through its center and parallel to the side of length a is $M(b^2 + c^2)/12$.

57.[R] A right solid circular cone has altitude h, radius a, constant density, and mass M.

(a) Why is its moment of inertia about its axis less that Ma^2?

(b) Show that its moment of inertia about its axis is $3Ma^2/10$.

58.[R] Let P_0 be a fixed point in a solid of mass M. Show that for all choices of three mutually perpendicular lines that meet at P_0 the sum of the moments of inertia of the solid about the lines is the same.

59.[C] [An exercise showing that the center of mass does not depend on the choice of coordinates.]

17.S Chapter Summary

This chapter generalizes the notion of a definite integral over an interval to integrals over plane sets, surfaces, and solids. These definitions are almost the same, the integral of $f(P)$ over a set being the limit of sums of the form $\sum f(P_i)\ \Delta A_i$, $\sum f(P_i)\ \Delta S_i$, or $\sum f(P_i)\ \Delta V_i$ for integrals over plane sets, surfaces, or solids, respectively.

If $f(P)$ denotes the density at P, then in each case, the integrals give the total mass.

The average value concept extends easily to functions of several variables. For instance, if $f(P)$ is defined on some plane region R, its average value over R is defined as

$$\frac{1}{area(R)} \int\limits_R f(P)\ dA.$$

Sometimes these "multiple integrals" (also known as "double" or "triple" integrals) can be calculated by repeated integrations over intervals, that is, as "iterated integrals." This requires a description of the region in an appropriate coordinate system and replaces dA or dV by an expression based on the area or volume of a small patch swept out by small changes in the coordinates, as recorded in Table 17.S.1.

Coordinate System	Substitution
Rectangular (2-d)	$dA = dx\ dy$
Rectangular (3-3)	$dV = dx\ dy\ dz$
Polar	$dA = r\ dr\ d\theta$
Cylindrical	$dV = r\ dr\ d\theta\ dz$
Cylindrical (surface)	$dS = r\ d\theta\ dz$
Spherical	$dV = \rho^2 \sin(\phi)\ d\phi\ d\rho\ d\theta$
Spherical (surface)	$dS = \rho^2 \sin(\phi)\ d\phi\ d\theta$

Table 17.S.1:

An integral over a surface S, $\int_S f(P)\ dS$, can often be replaced by an integral over the projection of S onto a plane R, replacing dS by $dA \cos(\gamma)$, where γ is the angle between a normal to S and a normal to R.

To remember these formulas, be able to draw Figures 17.S.1 and 17.S.2, including the labels.

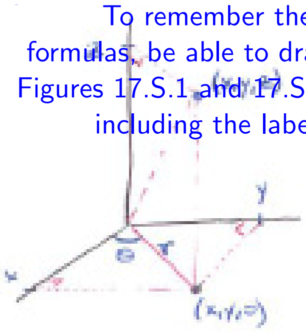


Figure 17.S.1:

Figure 17.S.2:

EXERCISES for 17.S *Key:* R–routine, M–moderate, C–challenging

1.[R] The temperature at the point (x, y) at time t is $T(x, y, t) = e^{-tx} \sin(x + 3y)$. Let $f(t)$ be the average temperature in the rectangle $0 \le x \le \pi$, $0 \le y \le \pi/2$ at time t. Find df/dt.

2.[R] Let f be a function such that $f(-x, y) = -f(x, y)$.

Key Facts

Formula	Significance
$\int_R 1\ dA$	Area of R
$\int_R 1\ dV$	Volume of R
$\frac{\int_R f(P)\ dA}{\text{Area of } R}$ or $\frac{\int_R f(P)\ dV}{\text{Volume of } R}$	Average value of f over R
$\int_R \sigma(P)\ dA$ or $\int_R \delta(P)\ dV$	Total mass of R, M (σ and δ denote density)
$\int_R y\sigma(P)\ dA$, $\int_R x\sigma(P)\ dA$	Moments, M_x and M_y about x and y axes, respectively. (A moment can be computed around any line in the plane.)
$\int_R f(P)\sigma(P)\ dA$, $\int_R f(P)\sigma(P)\ dV$ where $f(P)$ is the square of the distance from P to some fixed line L	Moment of inertia around L for planar and solid regions, respectively.
$\int_R x^2\sigma(P)\ dA$, $\int_R y^2\sigma(P)\ dA$	Second moments, M_{xx} and M_{yy} about x and y axes, respectively.
$\left(\frac{M_y}{M}, \frac{M_x}{M}\right)$	Center of mass, $(\bar{x}, \bar{y})$
$\int_R z\delta(P)\ dV$	Moment M_{xy}
$\int_R y\delta(P)\ dV$	Moment M_{xz}
$\int_R x\delta(P)\ dV$	Moment M_{yz}
$\left(\frac{M_{yz}}{M}, \frac{M_{xz}}{M}, \frac{M_{xy}}{M}\right)$	Center of mass of solid, $(\bar{x}, \bar{y}, \bar{z})$

Table 17.S.2:

Relations Between Rectangular Coordinates and Spherical or Cylindrical Coordinates

$x = \rho\sin(\phi)\cos(\theta)$	$x = r\cos(\theta)$
$y = \rho\sin(\phi)\sin(\theta)$	$y = r\sin(\theta)$
$z = \rho\cos(\phi)$	$z = z$

Table 17.S.3:

(a) Give some examples of such functions.

(b) For what type regions R in the xy plane is $\int_R f(x,y)\,dA$ certainly equal to 0?

3.[R] Find $\int_R (2x^3y^2+7)\,dA$ where R is the square with vertices $(1,1)$, $(-1,1)$, $(-1,-1)$, and $(1,-1)$. Do this with as little work as possible.

4.[R] Let $f(x,y)$ be a continuous function. Define $g(x)$ to be $\int_R f(P)\,dA$, where R is the rectangle with vertices $(3,0)$, $(3,5)$, $(x,0)$, and $(x,5)$, $x>3$. Express dg/dx as a suitable integral.

5.[R] Let R be a plane lamina in the shape of the region bounded by the graph of the equation $r=2a\sin(\theta)$ $(a>0)$. If the variable density of the lamina is given by $\sigma(r,\theta)=\sin(\theta)$, find the center of mass R.

In Exercises 6 to 9 find the moment of inertia of a homogeneous lamina of mass M of the given shape, around the given line.

6.[R] A disk of radius a, about the line perpendicular to it through its center.

7.[R] A disk of radius a, about a line perpendicular to it through a point on the circumference.

8.[R] A disk of radius a, about a diameter.

9.[R] A disk of radius a, about a tangent.

10.[C] Let $\mathcal{S}$ be the sphere of radius a and center at the origin. The integral $\int_{\mathcal{S}}(xz+y^2)dS$ can be done with little effort.

(a) Why is $\int_{\mathcal{S}} xz dS = 0$?

(b) Why is $\int_{\mathcal{S}} x^2 dS = \int_{\mathcal{S}} y^2 dS = \int_{\mathcal{S}} z^2 dS$?

(c) Why is $\int_{\mathcal{S}} y^2 dS = \int_{\mathcal{S}} (a^2/3) dS$?

(d) Show that $\int_{\mathcal{S}}(xz+y^2)dS = 4\pi a^2/3$.

11.[C] Let $f(P)$ and $g(P)$ be continuous functions defined on the plane region R.

(a) Show that

$$\left(\int_R f(P)g(P)\,dA\right)^2 \le \left(\int_R f(P)^2\,dA\right)\left(\int_R g(P)^2\,dA\right).$$

Hint: Review the proof of the Cauchy-Schwarz inequality presented in the CIE on Average Speed and Class Size on page 684.

(b) Show that if equality occurs in the inequality in (a), then f is a constant times g.

12.[C] (Courtesy of G. D. Chakerian.) A solid region $\mathcal{S}$ is bounded below by the $x-y$ plane, above by the surface $z = f(P)$, and the sides by the surface of a cylinder, as shown in Figure 17.S.3.

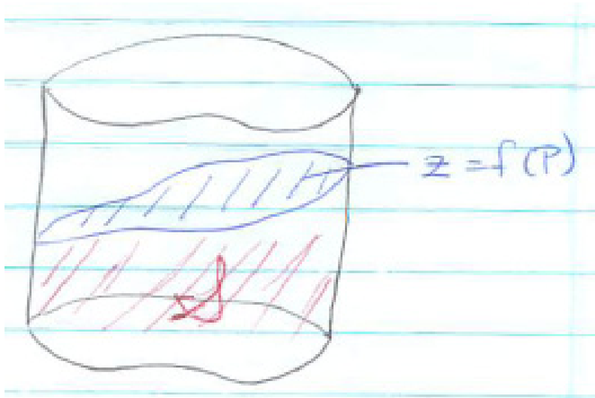


Figure 17.S.3:

The volume of $\mathcal{S}$ is V. If V is fixed, show that the top surface that minimizes the height of the centroid of $\mathcal{S}$ is a horizontal plane. NOTE: Water in a glass illustrates this, for nature minimizes the height of the centroid of the water. HINT: See Exercise 11.

Exercises 13 to 19 explore the average distance for all points on a curve or in a region. Recall that the distance from a point to a curve is the shortest distance from the point to the curve.

13.[M] Find the average distance from points in a disk of radius a to the center of the disk.

(a) Set up the pertinent definite integral in rectangular coordinates.

(b) Set it up in polar coordinates.

(c) Evaluate the easier integral in (a) and (b).

14.[M] Find the average distance from points in a square of side a to the center of the square.

(a) Set up the pertinent definite integral in rectangular coordinates.

(b) Set it up in polar coordinates.

(c) Evaluate the easier integral in (a) and (b).

15.[M] Find the average distance from points in a ball of radius a to the center of the ball.

(a) Set up the pertinent definite integral in rectangular coordinates.

(b) Set it up in spherical coordinates.

(c) Evaluate the easier integral in (a) and (b).

16.[M] Find the average distance from points in a cube of side a to the center of the cube.

(a) Set up the pertinent definite integral in rectangular coordinates.

(b) Set it up in polar coordinates.

(c) Evaluate the easier integral in (a) and (b).

17.[M] Find the average distance from points in a square of side a to the border of the square.

(a) Set up the pertinent definite integral in rectangular coordinates.

(b) Set it up in polar coordinates.

(c) Evaluate the easier integral in (a) and (b).

18.[M] Find the average distance from the points in a disk of radius a to the circular border.

(a) Before doing any calculations, decide whether the average distance is greater than $a/2$ or less than $a/2$. Explain how you made this decision.

(b) Carry out the calculation using a convenient coordinate system.

19.[C] Let A and B be two points in the xy-plane. A curve (in the xy-plane) consists of all points P such that the sum of the distances from P to A and P to B is constant, say $2a$. Consider the distance from P to A as a function of arclength on the curve. Find the average of that distance.

Calculus is Everywhere # 22

Solving the Wave Equation

In the *The Wave in a Rope* Calculus is Everywhere in the previous chapter we encountered the partial differential equation

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 y}{\partial t^2}. \tag{C.22.1}$$

Now we will solve this equation to find y as a function of x and t. First, we solve some simpler equations, which will help us solve (C.22.1).

EXAMPLE 6 Let $u(x, y)$ satisfy the equation $\partial u/\partial x = 0$. Find the form of $u(x, y)$.
SOLUTION Since $\partial u/\partial x$ is 0, $u(x, y)$, for a fixed value of y, is constant. Thus, $u(x, y)$ depends only on y, and can be written in the form $h(y)$ for some function h of a single variable.

On the other hand, any function $u(x, y)$ that can be written in the form $h(y)$ has the property that $\partial u/\partial x = 0$ is any function that can be written as a function of y alone. ⋄

EXAMPLE 7 Let $u(x, y)$ satisfy

$$\frac{\partial^2 u}{\partial x \partial y} = 0. \tag{C.22.2}$$

Find the form of $u(x, y)$.
SOLUTION We know that

$$\frac{\partial\left(\frac{\partial u}{\partial y}\right)}{\partial x} = \frac{\partial^2 u}{\partial x \partial y} = 0.$$

By Example 6,

$$\frac{\partial u}{\partial y} = h(y) \qquad \text{for some function } h(y).$$

By the Fundamental Theorem of Calculus, for any number b,

$$u(x, b) - u(x, 0) = \int_0^b \frac{\partial u}{\partial y} dy = \int_0^b h(y) dy.$$

Let H be an antiderivative of h. Then

$$u(x,b) - u(x,0) = H(b) - H(0).$$

Replacing b by y shows that

$$u(x,y) = u(x,0) + H(y) - H(0).$$

That tells us that $u(x,y)$ can be expressed as the sum of a function of x and a function of y,

$$u(x,y) = f(x) + g(y). \tag{C.22.3}$$

◇

We will solve the wave equation (C.22.1) by using a suitable change of variables that transforms that equation into the one solved in Example 7.

The new variables are

$$p = x + ct \qquad \text{and} \qquad q = x - ct.$$

$x = \frac{1}{2}(p+q)$ and $t = \frac{1}{2c}(p-q)$.

One could solve these equations and express x and t as functions of p and q. We will apply the chain rule, where y is a function of p and q and p and q are functions of x and t, as indicated in Figure C.22.1. Thus $y(x,t) = u(p,q)$.

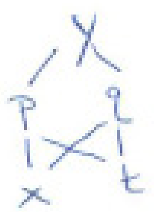


Figure C.22.1:

Keeping in mind that

$$\frac{\partial p}{\partial x} = 1, \qquad \frac{\partial p}{\partial t} = c, \qquad \frac{\partial q}{\partial x} = 1, \qquad \text{and } \frac{\partial q}{\partial t} = -c,$$

we have

$$\frac{\partial y}{\partial x} = \frac{\partial u}{\partial p}\frac{\partial p}{\partial x} + \frac{\partial u}{\partial q}\frac{\partial q}{\partial x} = \frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}.$$

Then

$$\begin{aligned}\frac{\partial^2 y}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right)\\ &= \frac{\partial}{\partial p}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right)\frac{\partial p}{\partial x} + \frac{\partial}{\partial q}\left(\frac{\partial u}{\partial p} + \frac{\partial u}{\partial q}\right)\frac{\partial q}{\partial x}\\ &= \left(\frac{\partial^2 u}{\partial p^2} + \frac{\partial^2 u}{\partial p \partial q}\right)\cdot 1 + \left(\frac{\partial^2 u}{\partial q \partial p} + \frac{\partial^2 u}{\partial^2 q}\right)\cdot 1.\end{aligned}$$

Thus

$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 u}{\partial p^2} + 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}. \tag{C.22.4}$$

A similar calculation shows that

$$\frac{\partial^2 y}{\partial t^2} = c^2\left(\frac{\partial^2 u}{\partial p^2} - 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}\right). \tag{C.22.5}$$

Substituting (C.22.4) and (C.22.5) in (C.22.1) leads to

$$\frac{\partial^2 u}{\partial p^2} + 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}. = \frac{1}{c^2}\left(c^2\right)\left(\frac{\partial^2 u}{\partial p^2} - 2\frac{\partial^2 u}{\partial p \partial q} + \frac{\partial^2 u}{\partial q^2}\right),$$

which reduces to

$$4\frac{\partial^2 u}{\partial p \partial q} = 0.$$

By Example 7, there are function $f(p)$ and $g(q)$ such that

$$y(x,t) = u(p,q) = f(p) + g(q).$$

or

$$y(x,t) = f(x+ct) + g(x-ct). \qquad \text{(C.22.6)}$$

The expression (C.22.6) is the most general solution of the wave equation (C.22.1).

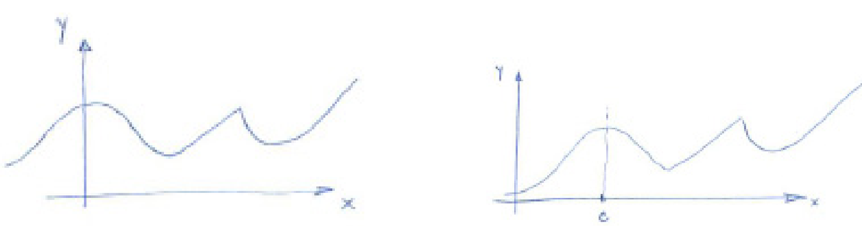
What does a solution (C.22.6) look like? What does the constant c tell us? To answer these questions, consider just

$$y(x,t) = g(x-ct). \qquad \text{(C.22.7)}$$

Here t represents time. For each value of t, $y(x,t) = g(x-ct)$ is simply a function of x and we can graph it in the xy plane. For $t = 0$, (C.22.7) becomes

$$y(x,0) = g(x).$$

That is just the graph of $y = g(x)$, whatever g is, as shown in Figure C.22.2(a).



(a) (b)

Figure C.22.2: (a) $t = 0$, (b) $t = 1$.

Now consider $y(x,t)$ when $t = 1$, which we may think of as "one unit of time later." Then

$$y = y(x,1) = g(x - c \cdot 1) = g(x-c).$$

The value of $y(x, 1)$ is the same as the value of g at $x - c$, c units to the left of x. So the graph at $t = 1$ is the graph of f in Figure C.22.2(a) shifted to the *right* c units, as in Figure C.22.2(b).

As t increases, the initial "wave" shown in Figure C.22.2(a) moves further to the right at the constant speed, c. Thus c tells us the velocity of the moving wave. That fact will play a role in Maxwell's prediction that electro-magnetic waves travel at the speed of light, as we will see in the Calculus is Everywhere at the end of Chapter 18.

EXERCISES

1.[R] Which functions $u(x, y)$ have both $\partial u/\partial x$ and $\partial u/\partial y$ equal to 0 for all x and y?

2.[R] Let $u(x, y)$ satisfy the equation $\partial^2 u/\partial x^2 = 0$. Find the form of $u(x, y)$.

3.[R] Show that any function of the form (C.22.3) satisfies equation (C.22.2).

4.[R] Verify that any function of the form (C.22.6) satisfies the wave equation.

5.[M] We interpreted $y(x, t) = g(x - ct)$ as the description of a wave moving with speed c to the right. Interpret the equation $y(x, t) = f(x + ct)$.

6.[M] Let k be a positive constant.

(a) What are the solutions to the equation

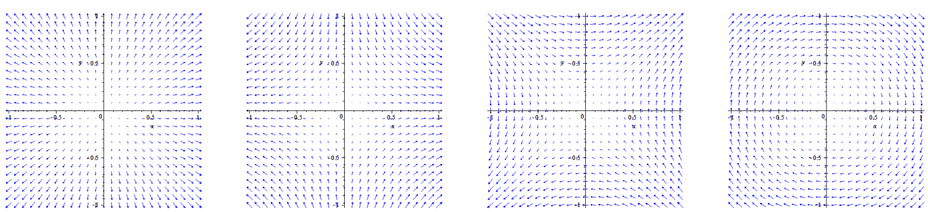
$$\frac{\partial^2 y}{\partial x^2} = k\frac{\partial^2 y}{\partial t^2}?$$

(b) What is the speed of the "waves"?

Chapter 18

The Theorems of Green, Stokes, and Gauss

Imagine a fluid or gas moving through space or on a plane. Its density may vary from point to point. Also its velocity vector may vary from point to point. Figure 18.0.1 shows four typical situations. The diagrams shows flows in the plane because it's easier to sketch and show the vectors there than in space.



(a) (b) (c) (d)

Figure 18.0.1: Four typical vector fields in the plane.

The plots in Figure 18.0.1 resemble the slope fields of Section 3.6 but now, instead of short segments, we have vectors, which may be short or long. Two questions that come to mind when looking at these vector fields:

- For a fixed region of the plane (or in space), is the amount of fluid in the region increasing or decreasing or not changing?

- At a given point, does the field create a tendency for the fluid to rotate? In other words, if we put a little propeller in the fluid would it turn? If so, in which direction, and how fast?

This chapter provides techniques for answering these questions which arise in several areas, such as fluid flow, electromagnetism, thermodynamics, and gravity. These techniques will apply more generally, to a general vector field. Applications come from magnetics as well as fluid flow.

Throughout we assume that all partial derivatives of the first and second orders exist and are continuous.

18.1 Conservative Vector Fields

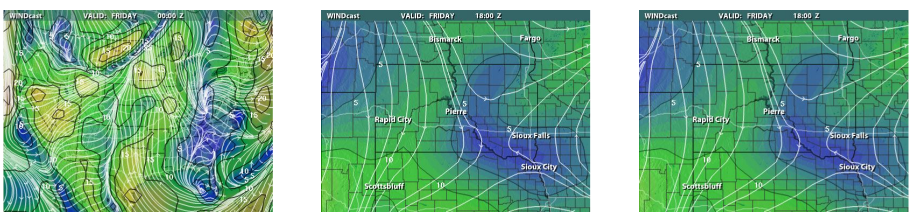
In Section 15.3 we defined integrals of the form

$$\int_C (P\ dx + Q\ dy + R\ dz). \tag{18.1.1}$$

where P, Q, and R are scalar functions of x, y, and z and C is a curve in space. Similarly, in the xy-plane, for scalar functions of x and y, P and Q, we have

$$\int_C (P\ dx + Q\ dy).$$

Instead of three scalar fields, P, Q, and R, we could think of a single vector function $\mathbf{F}(x,y,z) = P(x,y,z)\mathbf{i} + Q(x,y,z)\mathbf{j} + R(x,y,z)\mathbf{k}$. Such a function is called a **vector field**, in contrast to a scalar field. It's hard to draw a vector field defined in space. However, it's easy to sketch one defined only on a plane. Figure 18.1.1 shows three wind maps, showing the direction and speed of the winds for (a) the entire United States, (b) near Pierre, SD and (c) near Tallahassee, FL on April 24, 2009.



(a) (b) (c)

Figure 18.1.1: Wind maps showing (a) a source and (b) a saddle. Obtained from `www.intellicast.com/National/Wind/Windcast.aspx` on April 23, 2009. [Another idea for these sample plots is to use maps from Hurricane Katrina.]

Introducing the formal vector $d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$, we may rewrite (18.1.1) as

$$\int_C \mathbf{F}\cdot\ d\mathbf{r}.$$

The vector notation is compact, is the same in the plane and in space, and emphasizes the idea of a vector field. However, the clumsy notations

$$\int_C (P\,dx + Q\,dy + R\,dz) \qquad \text{and} \qquad \int_C (P(x,y,z)\,dx + Q(x,y,z)\,dy + R(x,y,z)\,dz)$$

do have two uses: to prove theorems and to carry out calculations.

Conservative Vector Fields

Recall the definition of a conservative vector field from Section 15.3.

> **DEFINITION** *(Conservative Field)* A vector field **F** defined in some planar or spatial region is called **conservative** if
>
> $$\int_{C_1} \mathbf{F}\cdot\ dr = \int_{C_2} \mathbf{F}\cdot\ dr$$
>
> whenever C_1 and C_2 are any two simple curves in the region with the same initial and terminal points.

An equivalent definition of a conservative vector field **F** is that for any simple closed curve C in the region $\oint_C \mathbf{F}\cdot\ d\mathbf{r} = 0$, as Theorem 18.1.1 implies. A closed curve is a curve that begins and ends at the same point, forming a loop. It is simple if it passes through no point — other than its start and finish points — more than once. A curve that starts at one point and ends at a different point is simple if it passes through no point more than once. Figure 18.1.2 shows some curves that are simple and some that are not.

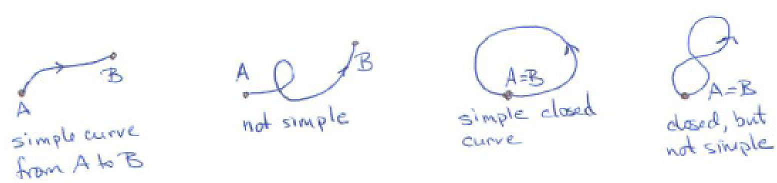


Figure 18.1.2:

Theorem 18.1.1. *A vector field* **F** *is conservative if and only if* $\oint_C \mathbf{F}\cdot\ dr = 0$ *for every simple closed curve in the region where* **F** *is defined.*

Proof

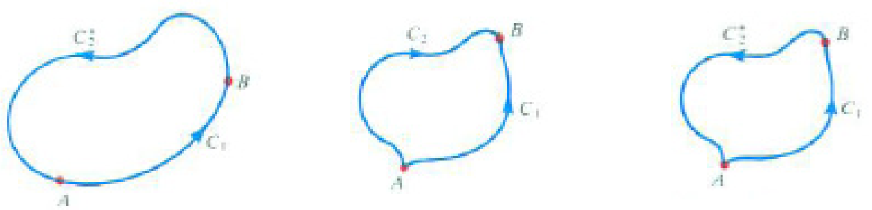
Assume that **F** is a conservative and let C be simple closed curve that starts and ends at the point A. Pick a point B on the curve and break C into two curves: C_1 from A to B and C_2^* from B to A, as indicated in Figure 18.1.3(a).

Let C_2 be the curve C_2^* traversed in the opposite direction, from A to B. Then, since **F** is conservative,

Note the sign change.

$$\oint_C \mathbf{F}\cdot\ d\mathbf{r} \;=\; \int_{C_1} \mathbf{F}\cdot\ d\mathbf{r} + \int_{C_2^*} \mathbf{F}\ d\mathbf{r} = \int_{C_1} \mathbf{F}\cdot\ d\mathbf{r} - \int_{C_2} \mathbf{F}\cdot\ d\mathbf{r} = 0.$$

On the other hand, assume that $\mathbf{F}$ has the property that $\oint_C \mathbf{F}\cdot\, d\mathbf{r} = 0$ for any simple closed curve C in the region. Let C_1 and C_2 be two simple curves in the region, starting at A and ending at B. Let $-C_2$ be C_2 taken in the reverse direction. (See Figures 18.1.3(b) and (c).) Then C_1 followed by $-C_2$ is a closed curve C from A back to A. Thus



(a) (b) (c)

Figure 18.1.3:

$$0 = \oint_C \mathbf{F}\cdot\, dr \;=\; \int_{C_1} \mathbf{F}\cdot\, dr + \int_{-C_2} \mathbf{F}\cdot\, dr = \int_{C_1} \mathbf{F}\cdot\, dr - \int_{C_2} \mathbf{F}\cdot\, dr.$$

Consequently,

$$\int_{C_1} \mathbf{F}\cdot\, dr = \int_{C_2} \mathbf{F}\cdot\, dr.$$

This concludes both directions of the argument. •

In this proof we tacitly assumed that C_1 and C_2 overlap only at their endpoints, A and B. Exercise 26 treats the case when the curves intersect elsewhere also.

Every Gradient Field is Conservative

Whether a particular vector field is conservative is important in the study of gravity, electro-magnetism, and thermodynamics. In the rest of this section we describe ways to determine whether a vector field $\mathbf{F}$ is conservative.

The first method that may come to mind is to evaluate $\oint \mathbf{F}\cdot d\mathbf{r}$ for every simple closed curve and see if it is always 0. If you find a case where it is not 0, then $\mathbf{F}$ is not conservative. Otherwise you face the task of evaluating a never-ending list of integrals checking to see if you always get 0. That is a most impractical test. Later in this section partial derivatives will be used to obtain a much simpler test. The first test involves gradients.

Gradient Fields Are Conservative

The fundamental theorem of calculus asserts that $\int_a^b f'(x)\ dx = f(b) - f(a)$. The next theorem asserts that $\int_C \nabla f \cdot d\mathbf{r} = f(B) - f(A)$, where f is a function of two or three variables and C is a curve from A to B. Because of its resemblance to the fundamental theorem of calculus, Theorem 18.1.2 is sometimes called the **fundamental theorem of vector fields**.

Any vector field that is the gradient of a scalar field turns out to be conservative. That is the substance of Theorem 18.1.2, which says, "The circulation of a gradient field of a scalar function f along a curve is the difference in values of f at the end points."

Theorem 18.1.2. *Let f be a scalar field defined in some region in the plane or in space. Then the gradient field $\mathbf{F} = \nabla f$ is conservative. In fact, for any points A and B in the region,*

$$\int_C \nabla f \cdot\ d\mathbf{r} = f(B) - f(A).$$

Proof

For simplicity take the planar case. Let C be given by the parameterization $\mathbf{r} = \mathbf{G}(t)$ for t in $[a, b]$. Let $\mathbf{G}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$. Then,

$$\int_C \nabla f \cdot d\mathbf{r} = \int_C \left(\frac{\partial f}{\partial x}\ dx + \frac{\partial f}{\partial y}\ dy \right) = \int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \right)\ dt.$$

The integrand $(\partial f/\partial x)(dx/dt) + (\partial f/\partial y)(dy/dt)$ is reminiscent of the chain rule in Section 16.3. If we introduce the function H defined by

$$H(t) = f(x(t), y(t)),$$

then the chain rule asserts that

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

Thus

$$\int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \right)\ dt = \int_a^b \frac{dH}{dt}\ dt = H(b) - H(a)$$

by the fundamental theorem of calculus. But

$$H(b) = f(x(b), y(b)) = f(B)$$

and

$$H(a) = f(x(a), y(a)) = f(A).$$

Consequently,

$$\int_C \nabla f \cdot d\mathbf{r} = f(B) - f(A), \tag{18.1.2}$$

and the theorem is proved. •

In differential form Theorem 18.1.2 reads

If f is defined as the xy-plane, and C starts at A and ends at B,

$$\int_C \left(\frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy \right) = f(B) - f(A) \tag{18.1.3}$$

If f is defined in space, then,

$$\int_C \left(\frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy + \frac{\partial f}{\partial z}\, dz \right) = f(B) - f(A). \tag{18.1.4}$$

Note that one vector equation (18.1.2) covers both cases (18.1.3) and (18.1.4). This illustrates an advantage of vector notation.

It is a much more pleasant task to evaluate $f(B) - f(A)$ than to compute a line integral.

EXAMPLE 1 Let $f(x, y, z) = \frac{1}{\sqrt{x^2+y^2+z^2}}$, which is defined everywhere except at the origin. (a) Find the gradient field $\mathbf{F} = \nabla f$, (b) Compute $\int_C \mathbf{F} \cdot\, d\mathbf{r}$ where C is any curve from $(1, 2, 2)$ to $(3, 4, 0)$.

SOLUTION (a) Straightforward computations show that

$$\frac{\partial f}{\partial x} = \frac{-x}{(x^2 + y^2 + z^2)^{3/2}}, \frac{\partial f}{\partial y} = \frac{-y}{(x^2 + y^2 + z^2)^{3/2}}, \frac{\partial f}{\partial z} = \frac{-z}{(x^2 + y^2 + z^2)^{3/2}}.$$

So

$$\nabla f = \frac{-z\mathbf{i} - y\mathbf{j} - z\mathbf{k}}{(x^2 + y^2 + z^2)^{3/2}}. \tag{18.1.5}$$

If we let $\mathbf{r}(x, y, z) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $r = \|\mathbf{r}\|$, and $\widehat{\mathbf{r}} = \mathbf{r}/r$, then (18.1.5) can be written more simply as

$$\mathbf{F} = \nabla f = \frac{-\mathbf{r}}{r^3} = \frac{-\widehat{\mathbf{r}}}{r^2}.$$

(b) For any curve C from $(1,2,2)$ to $(3,4,0)$,

$$\begin{aligned}\int_C \nabla f \cdot \, d\mathbf{r} &= f(3,4,0) - f(1,2,2) = \frac{1}{\sqrt{3^2+4^2+0^2}} - \frac{1}{\sqrt{1^2+2^2+2^2}} \\ &= \frac{1}{5} - \frac{1}{3} = -\frac{2}{15}.\end{aligned}$$

◇

For a constant k, positive or negative, any vector field, $\mathbf{F} = k\widehat{\mathbf{r}}/r^2$, is called an **inverse square central field**. They play an important role in the study of gravity and electromagnetism.

In Example 1 $\|\nabla f\| = \frac{\|-\mathbf{r}\|}{r^3} = \frac{r}{r^3} = \frac{1}{r^2}$ and $f(x,y,z) = \frac{1}{r}$. In the study of gravity, ∇f measures gravitational attraction, and f measures "potential."

EXAMPLE 2 Evaluate $\oint_C (y\ dx + x\ dy)$ around a closed curve C taken counterclockwise.

SOLUTION In Section 15.3 it was shown that if the area enclosed by a curve C is A, then $\oint_C x\ dy = A$ and $\oint_C y\ dx = -A$. Thus,

$$\oint_C (y\ dx + x\ dy) = -A + A = 0.$$

A second solution uses Theorem 18.1.2. Note that

$$\nabla(xy) = \frac{\partial(xy)}{\partial x}\mathbf{i} + \frac{\partial(xy)}{\partial y}\mathbf{j} = y\mathbf{i} + x\mathbf{j},$$

that is, the gradient of xy is $y\mathbf{i} + x\mathbf{j}$.

Hence, byTheorem 18.1.2, if the endpoints of C are A and B

$$\oint_C (y\ dx + x\ dy) = \oint_C \nabla(xy) \cdot d\mathbf{r} = xy|_A^B.$$

Because C is a closed curve, $A = B$ and so the integral is 0. ◇

A differential form $P(x,y,z)\ dx + Q(x,y,z)\ dy + R(x,y,z)\ dz$ is called **exact** if there is a scalar function f such that $P(x,y,z) = \partial f/\partial x$, $Q(x,y,z) = \partial f/\partial y$, and $R(x,y,z) = \partial f/\partial z$. In that case, the expression takes the form

$$\frac{\partial f}{\partial x}\ dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z}\ dz.$$

That is the same thing as saying that the vector field $\mathbf{F} = P(x,y,z)\mathbf{i} + Q(x,y,z)\mathbf{j} + R(x,y,z)\mathbf{k}$ is a gradient field: $\mathbf{F} = \nabla f$.

If F is Conservative Must It Be a Gradient Field?

The proof of the next theorem is similar to the proof of the second part of the Fundamental Theorem of Calculus. We suggest you review that proof (page 536) before reading the following proof.

FTC II states that every continuous function has an antiderivative.

The question may come to mind, *"If* **F** *is conservative, is it necessarily the gradient of some scalar function?"* The answer is "yes." That is the substance of the next theorem, but first we need to introduce some terminology about regions.

A region $\mathcal{R}$ in the plane is **open** if for each point P in $\mathcal{R}$ there is a disk with center at P that lies entirely in $\mathcal{R}$. For instance, a square *without its edges* is open. However, a square *with its edges is not open.*

An open region in space is defined similarly, with "disk" replaced by "ball."

An open region $\mathcal{R}$ is **arcwise-connected** if any two points in it can be joined by a curve that lies completely in $\mathcal{R}$. In other words, it consists of just one piece.

Theorem 18.1.3. *Let* $\mathbf{F}$ *be a conservative vector field defined in some arcwise-connected region* $\mathcal{R}$ *in the plane (or in space). Then there is a scalar function* f *defined in that region such that* $\mathbf{F} = \nabla f$.

Proof

Consider the case when $\mathbf{F}$ is planar, $\mathbf{F} = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$. (The case where $\mathbf{F}$ is defined in space is similar.) Define a scalar function f as follows. Let (a,b) be a fixed point in $\mathcal{R}$ and (x,y) be any point in $\mathcal{R}$. Select a curve C in $\mathcal{R}$ that starts at (a,b) and ends at (x,y).

Define $f(x,y)$ to be $\int_C \mathbf{F}\cdot\, d\mathbf{r}$. Since $\mathbf{F}$ is conservative, the number $f(x,y)$ depends only on the point (x,y) and not on the choice of C. (See Figure 18.1.4.)

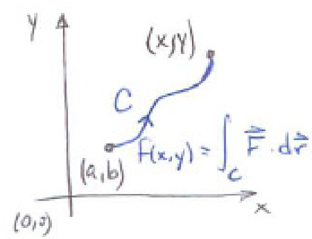


Figure 18.1.4:

All that remains is to show that $\nabla f = \mathbf{F}$; that is, $\partial f/\partial x = P$ and $\partial f/\partial y = Q$. We will go through the details for the first case, $\partial f/\partial x = P$. The reasoning for the other partial derivative is similar.

Let (x_0,y_0) be an arbitrary point in $\mathcal{R}$ and consider the difference quotient whose limit is $\partial f/\partial x(x_0,y_0)$, namely,

$$\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h},$$

for h small enough so that (x_0+h,y_0) is also in the region.

Let C_1 be any curve from (a,b) to (x_0,y_0) and let C_2 be the straight path from (x_0,y_0) to (x_0+h,y_0). (See Figure 18.1.5.) Let C be the curve from $(0,0)$ to the point(x_0+h,y_0) formed by taking C_1 first and continuing on C_2. Then

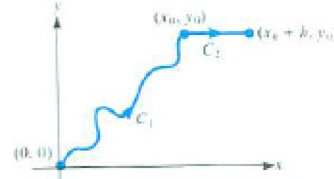


Figure 18.1.5:

$$f(x_0,y_0) = \int\limits_{C_1} \mathbf{F}\cdot d\mathbf{r},$$

and

$$f(x_0 + h, y_0) = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_{C_1} \mathbf{F} \cdot d\mathbf{r} + \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

Thus

$$\frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} = \frac{\int_{C_2} \mathbf{F} \cdot d\mathbf{r}}{h} = \frac{\int_{C_2} (P(x, y)\ dx + Q(x, y)\ dy)}{h}.$$

On C_2, y is constant, $y = y_0$; hence $dy = 0$. Thus $\int_{C_2} Q(x, y)\ dy = 0$. Also,

$$\int_{C_2} P(x, y)\ dx = \int_x^{x+h} P(x, y)\ dx.$$

See Section 6.3 for the MVT for Definite Integrals

By the Mean-Value Theorem for definite integrals, there is a number x^* between x and $x + h$ such that

$$\int_x^{x+h} P(x, y)\ dx = P(x^*, y_0)h.$$

Hence

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) &= \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \\ &= \lim_{h \to 0} \frac{1}{h} \int_{x_0}^{x_0+h} P(x, y_0)\ dx = \lim_{h \to 0} P(x^*, y_0) = P(x_0, y_0). \end{aligned}$$

Consequently,

$$\frac{\partial f}{\partial x}(x_0, y_0) = P(x_0, y_0),$$

as was to be shown.

In a similar manner, we can show that

$$\frac{\partial f}{\partial y}(x_0, y_0) = Q(x_0, y_0).$$

•

For a vector field $\mathbf{F}$ defined throughout some region in the plane (or space) the following three properties are therefore equivalent: Figure 18.1.6 tells us that any one of the three properties, (1), (2), or (3), describes a conservative field. We used property (3) as the definition.

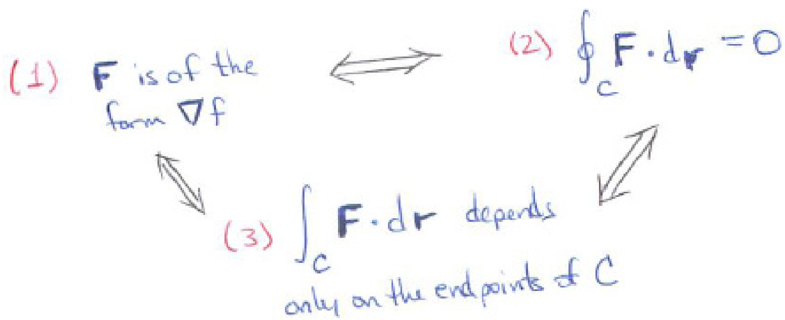


Figure 18.1.6: Double-headed arrows (⇔) mean "if and only if" or "is equivalent to." (Single-headed arrows (⇒) mean "implies.")

Almost A Test For Being Conservative

Figure 18.1.6 describes three ways of deciding whether a vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ is conservative. Now we give a simple way to tell that it is *not* conservative. The method is simpler than finding a particular line integral $\int_C \mathbf{F}\cdot\, d\mathbf{r}$ that is not 0.

Remember that we have assumed that all of the functions we encounter in this chapter have continuous first and second partial derivatives.

The test depends on the fact that the two orders in which are may compute a second-order mixed partial derivative give the same result. (We used this fact in Section 16.8 in a thermodynamics context.)

Consider an expression of the form $P\ dx + Q\ dy + R\ dz$ (or equivalently a vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$). If the form is exact, then $\mathbf{F}$ is a gradient and there is a scalar function f such that

$$\frac{\partial f}{\partial x} = P, \qquad \frac{\partial f}{\partial y} = Q, \qquad \frac{\partial f}{\partial z} = R.$$

Since

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

we have

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

Similarly we find

$$\frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y} \qquad \text{and} \qquad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}.$$

To summarize,

If the vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ is conservative, then

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0, \qquad \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} = 0, \qquad \frac{\partial R}{\partial x} - \frac{\partial P}{\partial z} = 0. \tag{18.1.6}$$

If at least one of these three equations (18.1.6) doesn't hold, then $P\ dx + Q\ dy + R\ dz$ is *not* exact (and $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ is not conservative).

EXAMPLE 3 Show that $\cos(y)\ dx + \sin(xy)\ dy + \ln(1+x)\ dz$ is *not* exact.
SOLUTION Checking whether the first equation in (18.1.6) holds we compute

$$\frac{\partial(\sin(xy))}{\partial x} - \frac{\partial(\cos(y))}{\partial y},$$

which equals

$$y\cos(xy) + \sin(y),$$

which is not 0. There's no need to check the remaining two equations in (18.1.6). The expression $\sin(xy)\ dx + \cos(y)\ dy + \ln(1+x)\ dz$ is not exact. (Equivalently, the vector field $\sin(xy)\mathbf{i} + \cos(y)\mathbf{j} + \ln(1+x)\mathbf{k}$ is not a gradient field, hence not conservative.) ◇

Notice that we completed Example 3 without doing any integration.

We can restate the three equations (18.1.6) as a single vector equation, by introducing a 3 by 3 formal determinant

$$\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{pmatrix} \tag{18.1.7}$$

Expanding this as though the nine entries were numbers, we get

$$\mathbf{i}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) - \mathbf{j}\left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right) + \mathbf{k}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right). \tag{18.1.8}$$

If the three scalar equations in (18.1.6) hold, then (18.1.8) is the $\mathbf{0}$-vector. In view of the importance of the vector (18.1.8), it is given a name.

> **DEFINITION** *(Curl of a Vector Field)* The **curl** of the vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ is the vector field given by the formula (18.1.7) or (18.1.8). It is denoted $\mathbf{curl}\,\mathbf{F}$.

The formal determinant (18.1.7) is like the one for the cross product of two vectors. For this reason, it is also denoted $\nabla \times \mathbf{F}$ (read as "del cross F"). That's a lot easier to write than (18.1.8), which refers to the components. Once again we see the advantage of vector notation.

The definition also applies to a vector field $\mathbf{F} = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ in the plane. Writing $\mathbf{F}$ as $P(x, y)\mathbf{i} + Q(x, y)\mathbf{j} + 0\mathbf{k}$ and observing that $\partial Q/\partial z = 0$ and $\partial P/\partial z = 0$, we find that

$$\nabla \times \mathbf{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}.$$

EXAMPLE 4 Compute the curl of $\mathbf{F} = xyz\mathbf{i} + x^2\mathbf{j} - xy\mathbf{k}$.
SOLUTION The curl of $\mathbf{F}$ is given by

$$\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xyz & x^2 & -xy, \end{pmatrix}$$

which is short for

$$\begin{aligned} &\left(\frac{\partial}{\partial y}(-xy) - \frac{\partial}{\partial z}(x^2)\right)\mathbf{i} - \left(\frac{\partial}{\partial x}(-xy) - \frac{\partial}{\partial z}(xyz)\right)\mathbf{j} + \left(\frac{\partial}{\partial x}(x^2) - \frac{\partial}{\partial y}(xyz)\right)\mathbf{k} \\ &= (-x - 0)\mathbf{i} - (-y - xy)\mathbf{j} + (2x - xz)\mathbf{k} \\ &= -x\mathbf{i} + (y + xy)\mathbf{j} + (2x - xz)\mathbf{k}. \end{aligned}$$

◇

If any case, in view of (18.1.6), for vector fields in space or in the xy-plane we have this theorem.

Theorem 18.1.4. *If* $\mathbf{F}$ *is a conservative vector field, then* $\nabla \times \mathbf{F} = \mathbf{0}$.

You may wonder why the vector field **curl F** obtained from the vector field **F** is called the "curl of **F**." Here we came upon the concept purely mathematically, but, as you will see in Section 18.6 it has a physical significance: If **F** describes a fluid flow, the curl of **F** describes the tendency of the fluid to rotate and form whirlpools — in short, to "curl."

The Converse of Theorem 18.1.4 Isn't True

Warning: The converse of Theorem 18.1.4 is false.

It would be delightful if the converse of Theorem 18.1.4 were true. Unfortunately, it is not. There are vector fields **F** whose curls are **0** that are not conservative. Example 5 provides one such **F** in the xy-plane. Its curl is **0** but it is not conservative, that is, $\nabla \times \mathbf{F} = \mathbf{0}$ and there is a closed curve C with $\oint_C \mathbf{F} \cdot d\mathbf{r}$ not zero.

EXAMPLE 5 Let $\mathbf{F} = \frac{-y\mathbf{i}}{x^2+y^2} + \frac{x\mathbf{j}}{x^2+y^2}$. Show that (a) $\nabla \times \mathbf{F} = \mathbf{0}$, but (b) $\mathbf{F}$ is not conservative.

SOLUTION (a) We must compute

$$\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{-y}{x^2+y^2} & \frac{x}{x^2+y^2} & 0 \end{pmatrix}$$

which equals

$$\left(\frac{\partial(0)}{\partial y} - \frac{\partial}{\partial z}\left(\frac{x}{x^2+y^2}\right)\right)\mathbf{i} - \left(\frac{\partial(0)}{\partial x} - \frac{\partial}{\partial z}\left(\frac{-y}{x^2+y^2}\right)\right)\mathbf{j} + \left(\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) - \frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right)\right)\mathbf{k}.$$

The $\mathbf{i}$ and $\mathbf{j}$ components are clearly 0, and a direct computation shows that the $\mathbf{k}$ component is

$$\frac{y^2-x^2}{(x^2+y^2)^2} - \frac{y^2-x^2}{(x^2+y^2)^2} = 0.$$

Thus the curl of $\mathbf{F}$ is $\mathbf{0}$.

(b) To show that $\mathbf{F}$ is *not* conservative, it suffices to exhibit a closed curve C such that $\oint_C \mathbf{F} \cdot d\mathbf{r}$ is not 0. One such choice for C is the unit circle parameterized counterclockwise by

$$x = \cos(\theta), \qquad y = \sin(\theta), \qquad 0 \le \theta \le 2\pi.$$

On this curve $x^2 + y^2 = 1$. Figure 18.1.7 shows a few values of $\mathbf{F}$ at points on C. Clearly $\int_C \mathbf{F} \cdot d\mathbf{r}$, which measures circulation, is positive, not 0. However, if you have any doubt, here is the computation of $\int_C \mathbf{F} \cdot d\mathbf{r}$:

Recall that, on C, $x^2 + y^2 = 1$.

$$\begin{aligned} \oint_C \mathbf{F} \cdot d\mathbf{r} &= \oint_C \left(\frac{-y\,dx}{x^2+y^2} + \frac{x\,dy}{x^2+y^2}\right) \\ &= \int_0^{2\pi} (-\sin\theta\ d(\cos\theta) + \cos\theta\ d(\sin\theta)) \\ &= \int_0^{2\pi} (\sin^2\theta + \cos^2\theta)\ d\theta = \int_0^{2\pi} d\theta = 2\pi. \end{aligned}$$

This establishes (b), $\mathbf{F}$ is not conservative. ◇

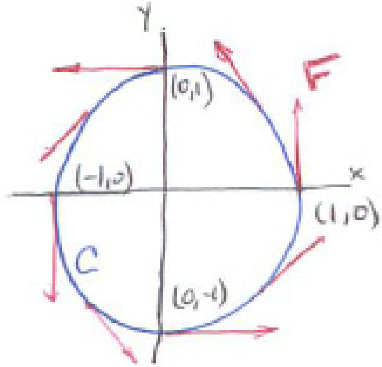


Figure 18.1.7:

The curl of **F** being **0** is not enough to assure us that a vector field **F** is conservative. An extra condition must be satisfied by **F**. This condition concerns the domain of **F**. This extra assumption will be developed for planar fields in Section 18.2 and for spatial fields **F** in Section 18.6. Then we will have a simple test for determining whether a vector field is conservative.

Summary

We showed that a vector field being conservative is equivalent to its being the gradient of a scalar field. Then we defined the curl of a vector field. If a field is denoted **F**, the curl of **F** is a new vector field denoted **curl F** or $\nabla \times \mathbf{F}$. If **F** is conservative, then $\nabla \times \mathbf{F}$ is **0**. However, if the curl of **F** is **0**, it does not follow that **F** is conservative. An extra assumption (on the domain of **F**) must be added. That assumption will be described in the next section.

EXERCISES for Section 18.1 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 to 4 answer "True" or "False" and explain.

1.[R] "If **F** is conservative, then $\nabla \times \mathbf{F} = \mathbf{0}$."

2.[R] "If $\nabla \times \mathbf{F} = \mathbf{0}$, then **F** is conservative."

3.[R] "If **F** is a gradient field, then $\nabla \times \mathbf{F} = \mathbf{0}$."

4.[R] "If $\nabla \times \mathbf{F} = \mathbf{0}$, then **F** is a gradient field."

5.[R] Using information in this section, describe various ways of showing a vector field **F** is *not* conservative.

6.[R] Using information in this section, describe various ways of showing a vector field **F** *is* conservative.

7.[R] Decide if each of the following sets is open, closed, neither open nor closed, or both open and closed.

(a) unit disk with its boundary

(b) unit disk without any of its boundary points

(c) the x-axis

(d) the entire xy-plane

(e) the xy-plane with the x-axis removed

(f) a square with all four of its edges (and corners)

(g) a square with all four of its edges but with its corners removed

(h) a square with none of its edges (and corners)

8.[R] In Example 1 we computed a certain line integral by using the fact that the vector field $(-x\mathbf{i} - y\mathbf{j} - z\mathbf{k})/(x^2 + y^2 + z^2)^{3/2}$ is a gradient field. Compute that integral directly, without using the information that the field is a gradient.

9.[R] Let $\mathbf{F} = y\cos(x)\mathbf{i} + (\sin(x) + 2y)\mathbf{j}$.

(a) Show that $\mathbf{curl}\,\mathbf{F}$ is $\mathbf{0}$ and **F** is defined in an arcwise-connected region of the plane.

(b) Construct a "potential function" f whose gradient is **F**.

10.[R] Let $f(x,y,z) = e^{3x}\ln(z+y^2)$. Compute $\int_C \nabla f \cdot d\mathbf{r}$, where C is the straight path from $(1,1,1)$ to $(4,3,1)$.

11.[R] We obtained the first of the three equations in (18.1.6). Derive the other two.

12.[R] Find the curl of $\mathbf{F}(x,y,z) = e^{x^2}yz\mathbf{i} + x^3\cos^2 3y\mathbf{j} + (1+x^6)\mathbf{k}$.

13.[R] Find the curl of $\mathbf{F}(x,y) = \tan^2(3x)\mathbf{i} + e^{3x}\ln(1+x^2)\mathbf{j}$.

14.[R] Using theorems of this section, explain why the curl of a gradient is $\mathbf{0}$, that is, $\mathbf{curl}(\nabla f) = \mathbf{0}$ ($\nabla \times \nabla f = \mathbf{0}$) for a scalar function $f(x,y,z)$. HINT: No computations are needed.

15.[R] By a computation using components, show that for the scalar function $f(x,y,z)$, $\mathbf{curl}\,\nabla f = \mathbf{0}$.

16.[R] Let $f(x,y) = \cos(x+y)$. Evaluate $\int_C \nabla f \cdot \, d\mathbf{r}$, where C is the curve that lies on the parabola $y = x^2$ and goes from $(0,0)$ to $(2,4)$.

17.[R] In Example 5 we computed $\oint_C \mathbf{F} \cdot d\mathbf{r}$, where $\mathbf{F} = \frac{-y\mathbf{i}+x\mathbf{j}}{x^2+y^2}$ and C is the unit circle with center at the origin. Compute the integral when C is the circle of radius 5 with center at the origin.

18.[M] In Example 5 we computed $\oint_C \mathbf{F} \cdot d\mathbf{r}$ where $\mathbf{F} = \frac{-y\mathbf{i}+x\mathbf{j}}{x^2+y^2}$ and C is the unit circle with center at the origin.

(a) *Without doing any new computations*, evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ where C is the square path with vertices $(1,0)$, $(2,0)$, $(2,1)$, $(1,1)$, $(1,0)$.

(b) Evaluate the integral in (a) by a direct computation, breaking the integral into four integrals, one over each edge.

19.[M] If $\mathbf{F}$ and $\mathbf{G}$ are conservative, is $\mathbf{F} + \mathbf{G}$?

20.[M] By a direct computation, show that $\mathbf{curl}(f\mathbf{F}) = \nabla f \times \mathbf{F} + f\,\mathbf{curl}\,\mathbf{F}$.

21.[M] By a direct computation, show that $\mathbf{curl}(\mathbf{F} \times \mathbf{G}) = (\mathbf{G} \cdot \nabla)\mathbf{F} - (\mathbf{F} \cdot \nabla)\mathbf{G} + \mathbf{F}(\nabla \cdot \mathbf{G}) - \mathbf{G}(\nabla \cdot \mathbf{F})$. Each of the first two terms has a form not seen before now in this text. Here is how to interpret them when $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ and $\mathbf{G} = G_1\mathbf{i} + G_2\mathbf{j} + G_3\mathbf{k}$:

$$(\mathbf{G} \cdot \nabla)\mathbf{F} = G_1\frac{\partial F_1}{\partial x} + G_2\frac{\partial F_2}{\partial y} + G_3\frac{\partial F_3}{\partial z}.$$

22.[M] If $\mathbf{F}$ and $\mathbf{G}$ are conservative, is $\mathbf{F} \times \mathbf{G}$?

23.[M] Explain why the curl of a gradient field is the zero vector, that is, $\nabla \times \nabla f = \mathbf{0}$.

24.[M] Assume that $\mathbf{F}(x, y)$ is conservative. Let C_1 be the straight path from $(0, 0, 0)$ to $(1, 0, 0)$, C_2 the straight path from $(1, 0, 0)$ to $(1, 1, 1)$. If $\int_{C_1} \mathbf{F}\ d\mathbf{r} = 3$ and $\int_{C_2} \mathbf{F}\ d\mathbf{r} = 4$, what can be said about $\int_C \mathbf{F}\ d\mathbf{r}$, where C is the straight path from $(0, 0, 0)$ to $(1, 1, 1)$?

25.[M] Let $\mathbf{F}(x, y)$ be a field that can be written in the form

$$\mathbf{F}(x, y) = g(\sqrt{x^2 + y^2})\frac{x\mathbf{i} + y\mathbf{j}}{\sqrt{x^2 + y^2}}$$

where g is a scalar function. If we denote $x\mathbf{i}+y\mathbf{j}$ as $\mathbf{r}$, then $\mathbf{F}(x, y) = g(r)\widehat{\mathbf{r}}$, where $r = \|\mathbf{r}\|$ and $\widehat{\mathbf{r}} = \|\mathbf{r}\|/r$. Show that $\oint_C \mathbf{F}\cdot\ d\mathbf{r} = 0$, for any path $ABCDA$ of the form shown in Figure 18.1.8. (The path consists of two circular arcs and parts of two rays from the origin.)

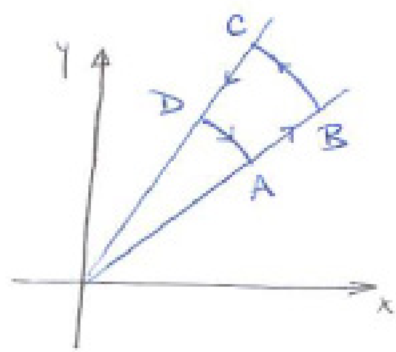


Figure 18.1.8:

26.[M] In Theorem 18.1.1 we proved that $\partial f/\partial x = P$. Prove that $\partial f/\partial y = Q$.

27.[C] In view of the previous exercise, we may expect $\mathbf{F}(x, y) = g(\sqrt{x^2 + y^2})\frac{x\mathbf{i}+y\mathbf{j}}{\sqrt{x^2+y^2}}$ to be conservative. Show that it is by showing that $\mathbf{F}$ is the gradient of $G(x, y) =$

$H(\sqrt{x^2+y^2})$, where H is an antiderivative of g, that is, $H' = g$.

28.[C] The domain of a vector field $\mathbf{F}$ is all of the xy-plane. Assume that there are two points A and B such that $\int_C \mathbf{F}\, d\mathbf{r}$ is the same for all curves C from A to B. Deduce that $\mathbf{F}$ is conservative.

29.[C] A gas at temperature T_0 and pressure P_0 is brought to the temperature $T_1 > T_0$ and pressure $P_1 > P_0$. The work done in this process is given by the line integral in the TP- plane

$$\int_C \left(\frac{RT\, dP}{P} - R\, dT \right),$$

where R is a constant and C is the curve that records the various combinations of T and P during the process. Evaluate this integral over the following paths, shown in Figure 18.1.9.

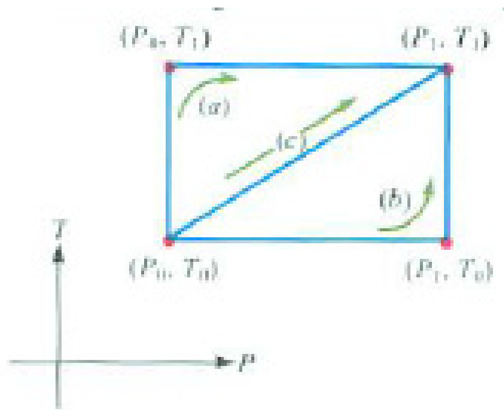


Figure 18.1.9:

(a) The pressure is kept constant at P_0 while the temperature is raised from T_0 to T_1; then the temperature is kept constant at T_1 while the pressure is raised from P_0 to P_1.

(b) The temperature is kept constant at T_0 while the pressure is raised from P_0 to P_1; then the temperature is raised from T_0 to T_1 while the pressure is kept constant at P_1.

(c) Both pressure and temperature are raised simultaneously in such a way that the path from (P_0, T_0) to (P_1, T_1) is straight.

Because the integrals are path dependent, the differential expression $RT\, dP/P - R\, dT$ defines a thermodynamic quantity that depends on the process, not just on the state. Vectorially speaking, the vector field $(RT/P)\mathbf{i} - R\mathbf{j}$ is not conservative.

30.[C] Assume that $\mathbf{F}(x, y)$ is defined throughout the xy-plane and that $\oint_C \mathbf{F}(x, y) \cdot d\mathbf{r} = 0$ for every closed curve that can fit inside a disk of diameter 0.01. Show that $\mathbf{F}$ is conservative.

31.[C] This exercise completes the proof of Theorem 18.1.1 in the case when C_1 and C_2 overlap outside of their endpoints A and B. In that case; introduce a third simple curve from A to B that overlaps C_1 and C_2 only at A and B. Then an argument similar to that in the proof of Theorem 18.1.1 can dispose of this case.

32.[C] We proved that $\lim \frac{\int_{x_0}^{x_0+h} P(x,y_0)\ dx}{h}$ equals $P(x_0, y_0)$, by using the Mean Value Theorem for definite integrals. Find a different proof of this result that uses a part of the Fundamental Theorem of Calculus.

18.2 Green's Theorem and Circulation

In this section we discuss a theorem that relates an integral of a vector field over a closed curve C in a plane to an integral of a related scalar function over the region $\mathcal{R}$ whose boundary is C. We will also see what this means in terms of the circulation of a vector field.

Statement of Green's Theorem

We begin by stating Green's Theorem and explaining each term in it. Then we will see several applications of the theorem. Its proof is at the end of the next section.

There are two analogs of Green's Theorem in space; they are discussed in Sections 18.5 and 18.6.

Green's Theorem

Let C be a simple, closed counterclockwise curve in the xy-plane, bounding a region $\mathcal{R}$. Let P and Q be scalar functions defined at least on an open set containing $\mathcal{R}$. Assume P and Q have continuous first partial derivatives. Then

$$\oint_C (P\ dx + Q\ dy) = \int_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ dA.$$

Recall, from Section 18.1, that a curve is closed when it starts and ends at the same point. It's simple when it does not intersect itself (except at its start and end). These restrictions on C ensure that it is the boundary of a region $\mathcal{R}$ in the xy-plane.

Since P and Q are independent of each other, Green's Theorem really consists of two theorems:

$$\int_C P\ dx = -\int_R \frac{\partial P}{\partial y}\ dA \qquad \text{and} \qquad \oint_C Q\ dy = \int_R \frac{\partial Q}{\partial x}\ dA. \qquad (18.2.1)$$

EXAMPLE 1 In Section 15.3 we showed that if the counterclockwise curve C bounds a region $\mathcal{R}$, then $\oint_C y\ dx$ is the negative of the area of $\mathcal{R}$. Obtain this result with the aid of Green's Theorem.

SOLUTION Let $P(x,y) = y$, and $Q(x,y) = 0$. Then Green's Theorem says that

$$\oint_C y\ dx = -\int_R \frac{\partial y}{\partial y}\ dA.$$

Since $\partial y/\partial y = 1$, it follows that $\oint y\ dx$ is $-\int_{\mathcal{R}} 1\ dA$, the negative of the area of $\mathcal{R}$. ◇

Green's Theorem and Circulation

What does Green's Theorem say about a vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$? First of all, $\oint_C (P\ dx + Q\ dy)$ now becomes simply $\oint_C \mathbf{F} \cdot\ d\mathbf{r}$.

The right hand side of Green's Theorem looks a bit like the curl of a vector field in the plane. To be specific, we compute the curl of $\mathbf{F}$:

$$\begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ P(x,y) & Q(x,y) & 0 \end{pmatrix} = 0\mathbf{i} - 0\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}$$

Thus the curl of $\mathbf{F}$ equals the vector function

$$\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}. \tag{18.2.2}$$

To obtain the (scalar) integrand on the right-hand side of (18.2.2), we "dot (18.2.2) with $\mathbf{k}$,"

$$\left(\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}\right) \cdot \mathbf{k} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Green's Theorem Expressed in Terms of Circulation

We can now express Green's Theorem using vectors. In particular, circulation around a closed curve can be expressed in terms of a double integral of the curl over a region.

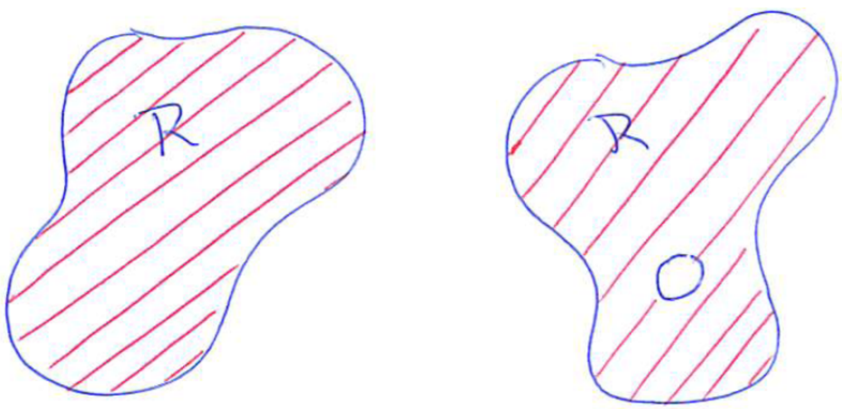
If the counterclockwise closed curve C bounds the region $\mathcal{R}$, then

$$\oint\limits_C \mathbf{F} \cdot\ d\mathbf{r} = \int\limits_{\mathcal{R}} (\nabla \times \mathbf{F}) \cdot \mathbf{k}\ dA.$$

Recall that if $\mathbf{F}$ describes the flow of a fluid in the xy-plane, then $\oint_C \mathbf{F} \cdot d\mathbf{r}$ represents its circulation, or tendency to form whirlpools. This theorem tells us that the magnitude of the curl of $\mathbf{F}$ represents the tendency of the fluid to rotate. If the curl of $\mathbf{F}$ is $\mathbf{0}$ everywhere, then $\mathbf{F}$ is called **irrotational** — there is no rotational tendency.

This form of Green's theorem provides an easy way to show that a vector field $\mathbf{F}$ is conservative. It uses the idea of a simply-connected region. Informally "a simply-connected region in the xy-plane comes in one piece and has no holes." More precisely, an arcwise-connected region $\mathcal{R}$ in the plane or in space is **simply-connected** if each closed curve in $\mathcal{R}$ can be shrunk gradually to a point while remaining in $\mathcal{R}$.

Figure 18.2.1 shows two regions in the plane. The one on the left is simply-connected, while the one on the right is not simply connected. For instance, the



(a) (b)

Figure 18.2.1: Regions in the plane that are (a) simply connected and (b) not simply connected.

xy-plane is simply connected. So is the xy-plane without its positive x-axis. However, the xy-plane, without the origin is *not* simply connected, because a circular path around the origin cannot be shrunk to a point while staying within the region.

If the origin is removed from xyz-space, what is left *is* simply connected. However, if we remove the z-axis, what is left is *not* simply connected.

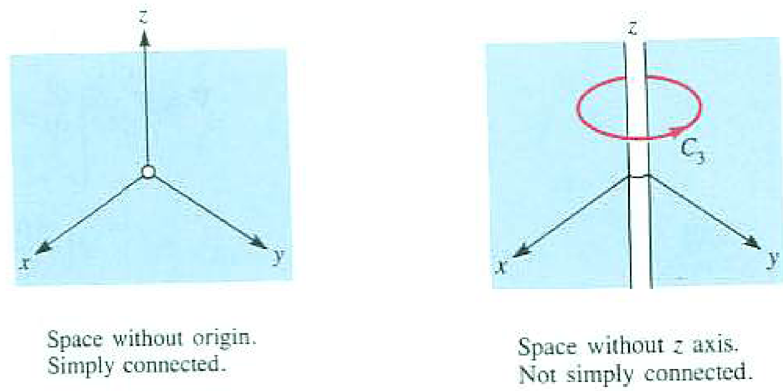
Figure 18.2.2(b) shows a curve that cannot be shrunk to a point while avoiding the z-axis.

Now we can state an easy way to tell whether a vector field is conservative.

Theorem. *If a vector field* $\mathbf{F}$ *is defined in a simply-connected region in the* xy*-plane and* $\nabla \times \mathbf{F} = \mathbf{0}$ *throughout that region, then* $\mathbf{F}$ *is conservative.*

Proof

Let C be any simple closed curve in the region and $\mathcal{R}$ the region it bounds.



(a) (b)

Figure 18.2.2: (a) xyz-space with the origin removed is simply connected. (b) xyz-space with the z-axis removed is not simply connected.

We wish to prove that the circulation of $\mathbf{F}$ around C is $\mathbf{0}$. We have

$$\oint_C \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{R}} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{k}\; dA.$$

Since $\mathbf{curl}\,\mathbf{F}$ is $\mathbf{0}$ throughout $\mathcal{R}$, it follows that $\oint_C \mathbf{F}\cdot\, d\mathbf{r} = 0$. •

In Example 5 in Section 18.1, there is a vector field whose curl is $\mathbf{0}$ but is not conservative. In view of the theorem just proved, its domain must not be simply connected. Indeed, the domain of the vector field in that example is the xy-plane with the origin deleted.

EXAMPLE 2 Let $\mathbf{F}(x, y, z) = e^x y\mathbf{i} + (e^x + 2y)\mathbf{j}$.

1. Show that $\mathbf{F}$ is conservative.

2. Exhibit a scalar function f whose gradient is $\mathbf{F}$.

SOLUTION

1. A straightforward calculation shows that $\nabla\times\mathbf{F} = \mathbf{0}$. Since $\mathbf{F}$ is defined throughout the xy-plane, a simply-connected region, Theorem 18.2 tells us that $\mathbf{F}$ is conservative.

2. By Section 18.1, we know that there is a scalar function f such that $\nabla f = \mathbf{F}$. There are several ways to find f. We show one of these methods here. Additional approaches are pursued in Exercises 7 and 8.

The approach chosen here follows the construction in the proof of Theorem 18.1.3. For a point (a,b), define $f(a,b)$ to equal $\int_C \mathbf{F}\cdot d\mathbf{r}$, where C is any curve from $(0,0)$ to (a,b). Any curve with the prescribed endpoints will do. For simplicity, choose C to be the curve that goes from $(0,0)$ to (a,b) in a straight line. (See Figure 18.2.3.) When a is not zero, we can use x as a parameter and write this segment as: $x=t$, $y=(b/a)t$ for $0\le t\le a$. (If $a=0$, we would use y as a parameter.) Then

$$\begin{aligned} f(a,b) &= \int_C (e^x y\,dx + (e^x+2y)\,dy) = \int_0^a \left(e^t\frac{b}{a}t\,dt + \left(e^t + 2\frac{b}{a}t\right)\frac{b}{a}\,dt\right) \\ &= \frac{b}{a}\int_0^a \left(te^t + e^t + 2\frac{b}{a}t\right)\,dt = \frac{b}{a}\left((t-1)e^t + e^t + \frac{b}{a}t^2\right)\Big|_0^a \\ &= \frac{b}{a}\left(te^t + \frac{b}{a}t^2\right)\Big|_0^a = be^a + b^2. \end{aligned}$$

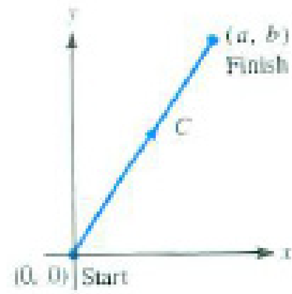


Figure 18.2.3:

Since $f(a,b)=be^a+b^2$, we see that $f(x,y)=ye^x+y^2$ is the desired function. One could check this by showing that the gradient of f is indeed $e^x y\mathbf{i} + (e^x+2y)\mathbf{j}$. Other suitable potential functions f are $e^x y + y^+ k$ for any constant k.

$ye^x + y^2 + k$ for any constant k, also would be a potential.

◇

The next example uses the *cancellation principle*, which is based on the fact that the sum of two line integrals in opposite direction on a curve is zero. This idea is used here to develop the two-curve version of Green's Theorem and then several more times before the end of this chapter.

Green's Theorem — The Two-Curve Case

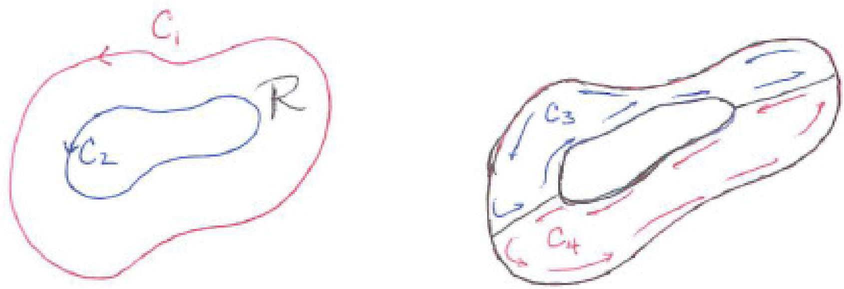
EXAMPLE 3 Figure 18.2.4(a) shows two closed counterclockwise curves C_1, and C_2 that enclose a ring-shaped region $\mathcal{R}$ in which $\nabla\times\mathbf{F}$ is $\mathbf{0}$. Show that the circulation of $\mathbf{F}$ over C_1 equals the circulation of $\mathbf{F}$ over C_2.

SOLUTION Cut $\mathcal{R}$ into two regions, each bounded by a simple curve, to which we can apply Theorem 18.2. Let C_3 bound one of the regions and C_4 bound the other, with the usual counterclockwise orientation. On the cuts, C_3 and C_4 go in opposite directions. On the outer curve C_3 and C_4 have the same orientation as C_1. On the inner curve they are the opposite orientation of C_2. (See Figure 18.1.2(b).) Thus

$$\int_{C_3} \mathbf{F}\cdot\ dr + \int_{C_4} \mathbf{F}\cdot\ dr = \int_{C_1} \mathbf{F}\cdot\ dr - \int_{C_2} \mathbf{F}\cdot\ dr. \tag{18.2.3}$$

By Theorem 18.2 each integral on the left side of (18.2.3) is 0. Thus

$$\int_{C_1} \mathbf{F}\cdot d\mathbf{r} = \int_{C_2} \mathbf{F}\cdot d\mathbf{r} \tag{18.2.4}$$



(a) (b)

Figure 18.2.4:

◇

Example 3 justifies the "two-curve" variation of Green's Theorem:

Two-Curve Version of Green's Theorem

Assume two nonoverlapping curves C_1 and C_2 lie in a region where $\mathbf{curl}\,\mathbf{F}$ is $\mathbf{0}$ and form the border of a ring. Then, if C_1 and C_2 both have the same orientation,

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

This theorem tells us "as you move a closed curve within a region of zero-curl, you don't change the circulation." The next Example illustrates this point.

EXAMPLE 4 Let $\mathbf{F} = \frac{-y\mathbf{i}+x\mathbf{j}}{x^2+y^2}$ and C be the closed counterclockwise curve bounding the square whose vertices are $(-2,-2)$, $(2,-2)$, $(2,2)$, and $(-2,2)$. Evaluate the circulation of $\mathbf{F}$ around C as easily as possible.

SOLUTION This vector field appeared in Example 5 of Section 18.1. Since its curl is $\mathbf{0}$, at all points except the origin, where $\mathbf{F}$ is not defined, we may use the two-curve version of Green's Theorem. Thus $\oint_C \mathbf{F} \cdot d\mathbf{r}$ equals the circulation of $\mathbf{F}$ over the unit circle in Example 5, hence equals 2π.

This is a lot easier than integrating $\mathbf{F}$ directly over each of the four edges of the square. ◇

How to Draw $\nabla \times \mathbf{F}$

For the planar vector field $\mathbf{F}$, its curl, $\nabla \times \mathbf{F}$, is of the form $z(x,y)\mathbf{k}$. If $z(x,y)$ is positive, the curl points directly up from the page. Indicate this by the symbol $\odot$, which suggests the point of an arrow or the nose of a rocket. If $z(x, y)$ is negative, the curl points down from the page. To show this, use the symbol $\oplus$, which suggests the feathers of an arrow or the fins of a rocket. Figure 18.2.5 illustrates their use.

This is standard notation in physics.

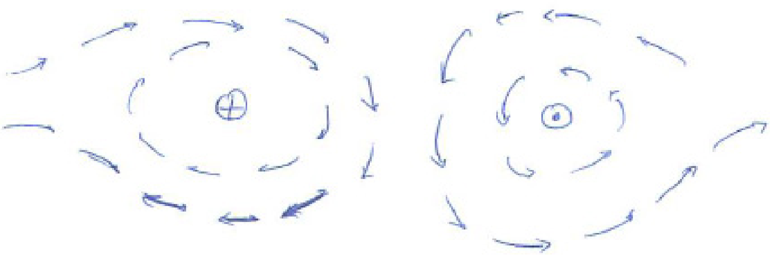


Figure 18.2.5:

Summary

We first expressed Green's theorem in terms of scalar functions

$$\oint_C (P\ dx + Q\ dy) = \int_{\mathcal{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ dA.$$

We then translated it into a statement about the circulation of a vector field;

$$\oint_C \mathbf{F}\cdot\ d\mathbf{r} = \int_{\mathcal{R}} (\nabla \times \mathbf{F}) \cdot \mathbf{k}\ dA.$$

In this theorem the closed curve C is oriented counterclockwise.

With the aid of this theorem we were able to show the following important result:

If the curl of $\mathbf{F}$ is $\mathbf{0}$ *and if the domain of* $\mathbf{F}$ *is simply connected*, then $\mathbf{F}$ is conservative.

Also, in a region in which $\nabla \times \mathbf{F} = \mathbf{0}$, the value of $\oint_C \mathbf{F}\cdot\ d\mathbf{r}$ does not change as you gradually change C to other curves in the region.

EXERCISES for Section 18.2 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 through 4 verify Green's Theorem for the given functions P and Q and curve C.

1.[R] $P = xy$, $Q = y^2$ and C is the border of the square whose vertices are $(0,0)$, $(1,0)$, $(1,1)$ and $(0,1)$.

2.[R] $P = x^2$, $Q = 0$ and C is the boundary of the unit circle with center $(0,0)$.

3.[R] $P = e^y$, $Q = e^x$ and C is the triangle with vertices $(0,0)$, $(1,0)$, and $(0,1)$.

4.[R] $P = \sin(y)$, $Q = 0$ and C is the boundary of the portion of the unit disk with center $(0,0)$ in the first quadrant.

5.[R] Figure 18.2.6 shows a vector field for a fluid flow $\mathbf{F}$. At the indicated points A, B, C, and D tell when the curl of $\mathbf{F}$ is pointed up, down or is $\mathbf{0}$. (Use the $\odot$ and $\oplus$ notation.) HINT: When the fingers of your right hand copy the direction of the flow, your thumb points in the direction of the curl, up or down.

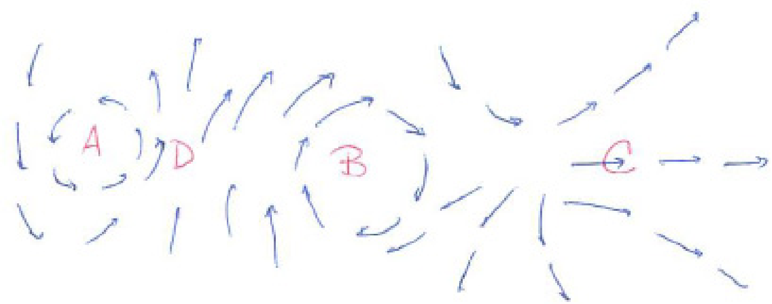


Figure 18.2.6:

6.[R] Assume that $\mathbf{F}$ describes a fluid flow. Let P be a point in the domain of $\mathbf{F}$ and C a small circular path around P.

(a) If the curl of $\mathbf{F}$ points upward, in what direction is the fluid tending to turn near P, clockwise or counterclockwise?

(b) If C is oriented clockwise, would $\oint_C \mathbf{F} \cdot d\mathbf{r}$ to be positive or negative?

7.[R] In Example 2 we constructed a function f by using a straight path from $(0,0)$ to (a,b). Instead, construct f by using a path that consists of two line segments, the first from $(0,0)$ to $(a,0)$, and the second, from $(a,0)$ to (a,b).

8.[R] In Example 2 we constructed a function f by using a straight path from $(0,0)$ to (a,b). Instead, construct f by using a path that consists of two line segments, the first from $(0,0)$ to $(0,b)$, and the second from $(0,b)$ to (a,b).

9.[R] Another way to construct a potential function f for a vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$

is to work directly with the requirement that $\nabla f = \mathbf{F}$. That is, with the equattions

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

(a) Integrate $\frac{\partial f}{\partial x} = e^x y$ with respect to x to conclude that $f(x,y) = e^x y + C(y)$. Note that the "constant of integration" can be any function of y, which we call $C(y)$. (Why?)

(b) Next, differentiate the result found in (a) with respect to y. This gives two formulas for $\frac{\partial f}{\partial y}$: $e^x + C'(y)$ and $e^x + 2y$. Use this fact to explain why $C'(y) = 2y$.

(c) Solve the equation for C found in (b).

(d) Combine the results of (a) and (c) to obtain the general form for a potential function for this vector field.

In Exercises 10 through 13

(a) check that $\mathbf{F}$ is conservative in the given domain, that is $\nabla \times \mathbf{F} = \mathbf{0}$, and the domain of $\mathbf{F}$ is simply connected

(b) construct f such that $\nabla f = \mathbf{F}$, using integrals on curves

(c) construct f such that $\nabla f = \mathbf{F}$, using antiderivatives, as in Exercise 9.

10.[R] $\mathbf{F} = 3x^2y\ vi + x^3\mathbf{j}$, domain the xy-plane

11.[R] $\mathbf{F} = y\cos(xy)\ vi + (x\cos(xy) + 2y)\mathbf{j}$, domain the xy-plane

12.[R] $\mathbf{F} = (ye^{xy} + 1/x)\mathbf{i} + xe^{xy}\mathbf{j}$, domain all xy with $x > 0$

13.[R] $\mathbf{F} = \frac{2y\ln(x)}{x}\mathbf{i} + (\ln(x))^2\mathbf{j}$, domain all xy with $x > 0$

14.[R] Verify Green's Theorem when $\mathbf{F}(xy) = x\mathbf{i} + y\mathbf{j}$ and $\mathcal{R}$ is the disk of radius a and center at the origin.

15.[R] In Example 1 we used Green's Theorem to show that $\oint_C y\ dx$ is the negative of the area that C encloses. Use Green's Theorem to show that $\oint_C x\ dy$ equals that area. (We obtained this result in Section 15.3 without Green's Theorem.)

16.[R] Let A be a plane region with boundary C a simple closed curve swept out counterclockwise. Use Green's theorem to show that the area of A equals

$$\frac{1}{2}\oint(-y\ dx + x\ dy).$$

17.[R] Use Exercise 16 to find the area of the region bounded by the line $y = x$ and the curve

$$\begin{cases} x &= t^6 + t^4 \\ y &= t^3 + t \end{cases} \qquad \text{for } t \text{ in } [0, 1].$$

18.[R] Assume that $\mathbf{curl}\,\mathbf{F}$ at $(0, 0)$ is -3. Let C sweep out the boundary of a circle of radius a, center at $(0, 0)$. When a is small, estimate the circulation $\int_C \mathbf{F} \cdot d\mathbf{r}$.

19.[R] Which of these fields are conservative:

(a) $x\mathbf{i} - y\mathbf{j}$

(b) $\frac{x\mathbf{i} - y\mathbf{j}}{x^2 + y^2}$

(c) $3\mathbf{i} + 4\mathbf{j}$

(d) $(6xy - y^3)\mathbf{i} + (4y + 3x^2 - 3xy^2)\mathbf{j}$

(e) $\frac{y\mathbf{i} - x\mathbf{j}}{1 + x^2y^2}$

(f) $\frac{x\mathbf{i} + y\mathbf{j}}{x^2 + y^2}$

20.[R] Figure 18.2.7 shows a fluid flow $\mathbf{F}$. All the vectors are parallel, but their magnitudes increase from bottom to top. A small simple curve C is placed in the flow.

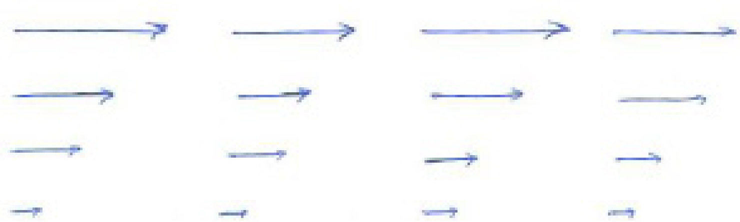


Figure 18.2.7:

(a) Is the circulation around C positive, negative, or 0? Justify your opinion.

(b) Assume that a wheel with small blades is free to rotate around its axis, which is perpendicular to the page. When it is inserted into this flow, which way would it turn, or would it not turn at all? (Don't just say, "It would get wet.")

21.[R] Let $\mathbf{F}(x, y) = y^2\mathbf{i}$.

(a) Sketch the field.

(b) Without computing it, predict when $(\nabla \times \mathbf{F}) \cdot \mathbf{k}$ is positive, negative or zero.

(c) Compute $(\nabla \times \mathbf{F}) \cdot \mathbf{k}$.

(d) What would happen if you dipped a wheel with small blades free to rotate around its axis, which is perpendicular to the page, into this flow.

22.[R] Check that the curl of the vector field in Example 2 is $\mathbf{0}$, as asserted.

23.[R] Explain in words, without explicit calculations, why the circulation of the field $f(r)\widehat{\mathbf{r}}$ around the curve $PQRSP$ in Figure 18.2.8 is zero. As usual, f is a scalar function, $r = ||\mathbf{r}||$, and $\widehat{r} = \mathbf{r}/r$.

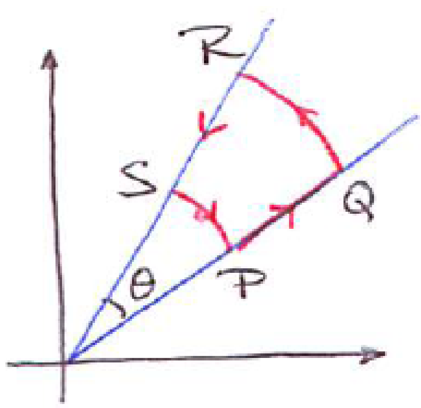


Figure 18.2.8: ARTIST: Please color the four sides of the closed curve.

In Exercises 24 to 27 let $\mathbf{F}$ be a vector field defined everywhere in the plane except a the point P shown in Figure 18.2.9. Assume that $\nabla \times \mathbf{F} = \mathbf{0}$ and that $\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = 5$.

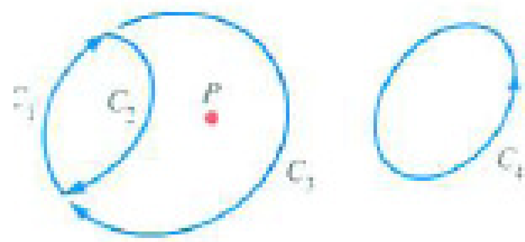


Figure 18.2.9:

24.[R] What, if anything, can be said about $\int_{C_2} \mathbf{F} \cdot d\mathbf{r}$?

25.[R] What, if anything, can be said about $\int_{C_3} \mathbf{F} \cdot d\mathbf{r}$?

26.[R] What, if anything, can be said about $\int_{C_4} \mathbf{F} \cdot d\mathbf{r}$?

27.[R] What, if anything, can be said about $\int_{C} \mathbf{F} \cdot d\mathbf{r}$, where C is the curve formed by C_1 followed by C_3?

In Exercises 28 to 31 show that the vector field is conservative and then construct a scalar function of which it is the gradient. Use the method in Example 2.

28.[R] $2xy\mathbf{i} + x^2\mathbf{j}$

29.[R] $\sin(y)\mathbf{i} + (x\cos(y) + 3)\mathbf{j}$

30.[R] $(y+1)\mathbf{i} + (x+1)\mathbf{j}$

31.[R] $3y\sin^2(xy)\cos(xy)\mathbf{i} + (1 + 3x\sin^2(xy)\cos(xy))\mathbf{j}$

32.[R] Show that

(a) $3x^2y\ dx + x^3\ dy$ is exact.

(b) $3xy\ dx + x^2\ dy$ is not exact.

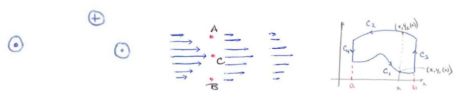
33.[R] Show that $(x\ dx + y\ dy)/(x^2+y^2)$ is exact and exhibit a function f such that df equals the given expression. (That is, find f such that $\nabla f \cdot d\mathbf{r}$ agrees with the given differential form.)

34.[R] Let $\mathbf{F} = \widehat{\mathbf{r}}/\|\mathbf{r}\|$ in the xy plane and let C be the circle of radius a and center $(0,0)$.

(a) Evaluate $\oint_C \mathbf{F} \cdot \mathbf{n}\ ds$ without using Green's theorem.

(b) Let C now be the circle of radius 3 and center $(4, 0)$. Evaluate $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$, doing as little work as possible.

35.[R] Figure 18.2.10(a) shows the direction of a vector field at three points. Draw a vector field compatible with these values. (No zero-vectors, please.)



(a) (b) (c)

Figure 18.2.10:

36.[R] Consider the vector field in Figure 18.2.10(b). Will a paddle wheel turn at A? At B? At C? If so, in which direction?

37.[R] Use Exercise 16 to obtain the formula for area in polar coordinates:

$$\text{Area} = \frac{1}{2} \int_{\alpha}^{\beta} r^2 \, d\theta.$$

Hint: Assume C is given parametrically as $x = r(\theta)\cos(\theta)$, $y = r(\theta)\sin(\theta)$, for $\alpha \le \theta \le \beta$.

38.[M] A curve is given parametrically by $x = t(1 - t^2)$, $y = t^2(1 - t^3)$, for t in $[0, 1]$.

(a) Sketch the points corresponding to $t = 0$, 0.2, 0.4, 0.6, 0.8, and 1.0, and use them to sketch the curve.

(b) Let $\mathcal{R}$ be the region enclosed by the curve. What difficulty arises when you try to compute the area of $\mathcal{R}$ by a definite integral involving vertical or horizontal cross sections?

(c) Use Exercise 16 to find the area of $\mathcal{R}$.

39.[M] Repeat Exercise 38 for $x = \sin(\pi t)$ and $y = t - t^2$, for t in $[0, 1]$. In (a), let $t = 0$, 1/4, 1/2, 3/4, and 1.

40.[C] Assume that you know that Green's Theorem is true when $\mathcal{R}$ is a triangle and C its boundary.

(a) Deduce that it therefore holds for quadrilaterals.

(b) Deduce that it holds for polygons.

41.[C] Assume that $\nabla \times \mathbf{F} = \mathbf{0}$ in the region $\mathcal{R}$ bounded by an exterior curve C_1 and two interior curves C_2 and C_3, as in Figure 18.2.11. Show that $\int_{C_1} \mathbf{F} \cdot \, d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot \, d\mathbf{r} + \int_{C_3} \mathbf{F} \cdot \, d\mathbf{r}$.

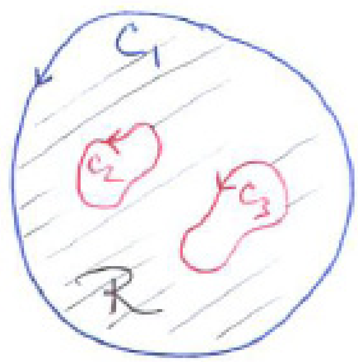


Figure 18.2.11:

42.[C] We proved that $\int_{\mathcal{R}} \frac{\partial Q}{\partial y} \, dA = \int_C Q \, dy$ in a special case. Prove it in this more general case, in which we assume less about the region $\mathcal{R}$. Assume that $\mathcal{R}$ has the description $a \leq x \leq b$, $y_1(x) \leq y \leq y_2(x)$. Figure 18.2.10(c) shows such a region, which need not be convex. The curved path C breaks up into four paths, two of which are straight (or may be empty), as in Figure 18.2.10(c).

43.[C] We proved the second part of (18.2.1), namely that $\oint_C Q \, dy = \int_R \partial Q/\partial x \, dA$. Prove the first part, $\oint_C P \, dx = -\int_R \partial p/\partial y \, dA$.

18.3 Green's Theorem, Flux, and Divergence

In the previous section we introduced Green's Theorem and applied it to discover a theorem about circulation and curl. That concerned the line integral of $\mathbf{F} \cdot \mathbf{T}$, the tangential component of $\mathbf{F}$, since $\mathbf{F} \cdot\ d\mathbf{r}$ is short for $(\mathbf{F} \cdot \mathbf{T})\ ds$. Now we will translate Green's Theorem into a theorem about the line integral of $\mathbf{F} \cdot \mathbf{n}$, the normal component of $\mathbf{F}$, $\oint \mathbf{F} \cdot \mathbf{n}\ ds$. Thus Green's Theorem will provide information about the flow of the vector field $\mathbf{F}$ across a closed curve C (see Section 15.4).

Green's Theorem Expressed in Terms of Flux

Let $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$ and C be a counterclockwise closed curve. (We use M and N now, to avoid confusion with P and Q needed later.) At a point on a closed curve the **unit exterior normal vector** (or **unit outward normal vector**) $\mathbf{n}$ is perpendicular to the curve and points outward from the region enclosed by the curve. To compute $\mathbf{F} \cdot \mathbf{n}$ in terms of M and N, we first express $\mathbf{n}$ in terms of $\mathbf{i}$ and $\mathbf{j}$.

The vector

$$\mathbf{T} = \frac{dx}{ds}\mathbf{i} + \frac{dy}{ds}\mathbf{j}$$

is tangent to the curve, has length 1, and points in the direction in which the curve is swept out. A typical $\mathbf{T}$ and $\mathbf{n}$ are shown in Figure 18.3.1. As Figure 18.3.1 shows, the exterior unit normal $\mathbf{n}$ has its x component equal to the y component of $\mathbf{T}$ and its y component equal to the negative of the x component of $\mathbf{T}$. Thus

$$\mathbf{n} = \frac{dy}{ds}\mathbf{i} - \frac{dx}{ds}\mathbf{j}.$$

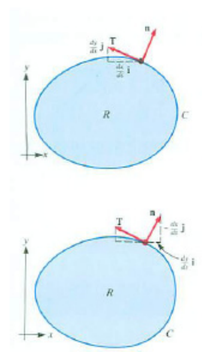


Figure 18.3.1:

Consequently, if $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$, then

$$\begin{aligned}\oint_C \mathbf{F} \cdot \mathbf{n} &= \oint_C (M\mathbf{i} + N\mathbf{j}) \cdot \left(\frac{dy}{ds}\mathbf{i} - \frac{dx}{ds}\mathbf{j}\right)\ ds = \oint_C \left(M\frac{dy}{ds} - N\frac{dx}{ds}\right)\ ds \\ &= \oint_C (M\ dy - N\ dx) = \oint_C (-N\ dx + M\ dy). \qquad (18.3.1)\end{aligned}$$

In (18.3.1), $-N$ plays the role of P and M plays the role of Q in Green's Theorem. Since Green's Theorem states that

$$\oint_C (P\ dx + Q\ dy) = \int_R \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\ dA$$

we have

$$\oint_C (-N\ dx + M\ dy) = \int_R \left(\frac{\partial M}{\partial x} - \frac{\partial(-N)}{\partial y}\right)\ dA$$

or simply, if $\mathbf{F} = M\mathbf{i} + N\mathbf{j}$, then

$$\oint_C \mathbf{F} \cdot \mathbf{n}\, ds = \int_R \left(\frac{\partial M}{\partial x} + \frac{\partial N}{\partial y} \right) dA.$$

In our customary "P and Q" notations, we have

Green's Theorem Expressed in Terms of Flux

If $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$, then

$$\oint_C \mathbf{F} \cdot \mathbf{n}\, ds = \int_{\mathcal{R}} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) dA$$

where C is the boundary of $\mathcal{R}$.

The expression

$$\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y},$$

the sum of two partial derivatives, is call the **divergence** of $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$. It is written $\operatorname{div} \mathbf{F}$ or $\nabla \cdot \mathbf{F}$. The latter notation is suggested by the "symbolic" dot product

$$\left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} \right) \cdot (P\mathbf{i} + Q\mathbf{j}) = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}.$$

It is pronounced "del dot eff". Theorem 18.3 is called "the divergence theorem in the plane." It can be written as

Divergence Theorem in the Plane

$$\oint_C \mathbf{F} \cdot \mathbf{n}\, ds = \int_{\mathcal{R}} \operatorname{div} \mathbf{F}\, dA$$

where C is the boundary of $\mathcal{R}$.

EXAMPLE 1 Compute the divergence of (a) $\mathbf{F} = e^{xy}\mathbf{i} + \arctan(3x)\mathbf{j}$ and (b) $\mathbf{F} = -x^2\mathbf{i} + 2xy\mathbf{j}$.

SOLUTION

(a) $\frac{\partial}{\partial x} e^{xy} + \frac{\partial}{\partial y} \arctan(3x) = ye^{xy} + 0 = ye^{xy}$

(b) $\frac{\partial}{\partial x}(-x^2) + \frac{\partial}{\partial y}(2xy) = -2x + 2x = 0.$

◇

The double integral of the divergence of $\mathbf{F}$ over a region describes the amount of flow across the border of that region. It tells how rapidly the fluid is leaving (diverging) or entering the region (converging). Hence the name "divergence".

In the next section we will be using the divergence of a vector field defined in space, $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$, where P, Q and R are functions of x, y, and z. It is defined as the sum of three partial derivatives

$$\nabla \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

It will play a role in measuring flux across a surface.

EXAMPLE 2 Verify that $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$ equals $\int_R \nabla \cdot \mathbf{F}\, dA$, when $\mathbf{F}(x, y) = x\mathbf{i} + y\mathbf{j}$, R is the disk of radius a and center at the origin and C is the boundary curve of R.

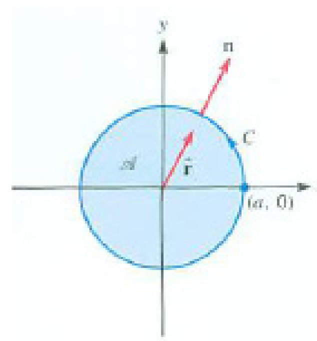


Figure 18.3.2:

SOLUTION First we compute $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$, where C is the circle bounding $\mathcal{R}$. (See Figure 18.3.2.)

Since C is a circle centered at $(0, 0)$, the unit exterior normal $\mathbf{n}$ is $\widehat{\mathbf{r}}$:

$$\mathbf{n} = \widehat{\mathbf{r}} = \frac{x\mathbf{i} + y\mathbf{j}}{\|x\mathbf{i} + y\mathbf{j}\|} = \frac{x\mathbf{i} + y\mathbf{j}}{a}.$$

Thus, remembering that $\oint_C ds$ is just the arclength of C,

$$\begin{aligned} \oint_C \mathbf{F} \cdot \mathbf{n}\, ds &= \oint_C (x\mathbf{i} + y\mathbf{j}) \cdot \left(\frac{x\mathbf{i} + y\mathbf{j}}{a} \right) ds = \oint_C \frac{x^2 + y^2}{a}\, ds \\ &= \oint_C \frac{a^2}{a}\, ds = a \oint_C ds = a(2\pi a) = 2\pi a^2. \end{aligned} \qquad (18.3.2)$$

Next we compute $\int_{\mathcal{R}} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) dA$. Since $P = x$ and $Q = y$, $\partial P/\partial x + \partial Q/\partial y = 1 + 1 = 2$. Then

$$\int_{\mathcal{R}} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} \right) dA = \int_{\mathcal{R}} 2\, dA,$$

which is twice the area of the disk $\mathcal{R}$, hence $2\pi a^2$. This agrees with (18.3.2). ◇

As the next example shows, a double integral can provide a way to compute the flux: $\oint \mathbf{F} \cdot \mathbf{n}\, ds$.

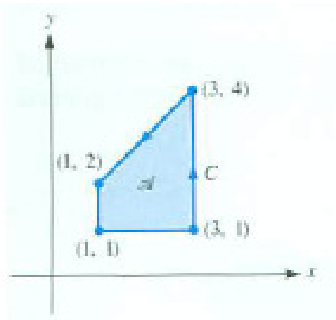


Figure 18.3.3:

EXAMPLE 3 Let $\mathbf{F} = x^2\mathbf{i} + xy\mathbf{j}$. Evaluate $\oint \mathbf{F}\cdot\mathbf{n}\ ds$ over the curve that bounds the quadrilateral with vertices $(1,1)$, $(3,1)$, $(3,4)$, and $(1,2)$ shown in Figure 18.3.3.

SOLUTION The line integral could be evaluated directly, but would require parameterizing each of the four edges of C. With Green's Theorem we can instead evaluate an integral over a single plane region.

Let $\mathcal{R}$ be the region that C bounds. By Green's theorem

$$\begin{aligned}\oint_C \mathbf{F}\cdot\mathbf{n}\ ds &= \int_{\mathcal{R}} \nabla\cdot\mathbf{F}\ dA = \int_{\mathcal{R}} \left(\frac{\partial(x^2)}{\partial x} + \frac{\partial(xy)}{\partial y}\right)\ dA \\ &= \int_{\mathcal{R}} (2x+x)\ dA = \int_{\mathcal{R}} 3x\ dA.\end{aligned}$$

See Exercise 15.

Then

$$\int_{\mathcal{R}} 3x\ dA = \int_1^3 \int_1^{y(x)} 3x\ dy\ dx,$$

where $y(x)$ is determined by the equation of the line that provides the top edge of $\mathcal{R}$. We easily find that the line through $(1,2)$ and $(3,4)$ has the equation $y = x+1$. Therefore,

$$\int_{\mathcal{R}} 3x\ dA = \int_1^3 \int_1^{x+1} 3x\ dy\ dx.$$

The inner integration gives

$$\int_1^{x+1} 3x\ dy = 3xy|_{y=1}^{y=x+1} = 3x(x+1) - 3x = 3x^2.$$

The second integration gives

$$\int_1^3 3x^2\ dx = x^3|_1^3 = 27 - 1 = 26$$

◇

A Local View of div F

We have presented a "global" view of $\operatorname{div}\mathbf{F}$, integrating it over a region $\mathcal{R}$ to get the total divergence across the boundary of $\mathcal{R}$. But there is another way of viewing $\operatorname{div}\mathbf{F}$, "locally." This approach makes uses an extension of the Permanence Principle of Section 2.5 to the plane and to space.

Let $P = (a, b)$ be a point in the plane and $\mathbf{F}$ a vector field describing fluid flow. Choose a very small region $\mathcal{R}$ around P, and let C be its boundary. (See Figure 18.3.4.) Then the net flow out of $\mathcal{R}$ is

$$\oint_C \mathbf{F}\cdot\mathbf{n}\, ds.$$

By Green's theorem, the net flow is also

$$\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\, dA.$$

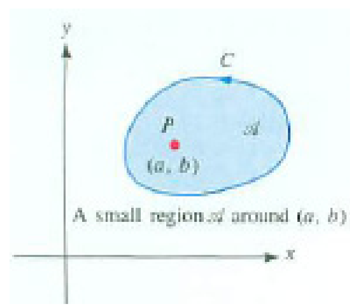


Figure 18.3.4:

Now, since $\operatorname{div}\mathbf{F}$ is continuous and $\mathcal{R}$ is small, $\operatorname{div}\mathbf{F}$ is almost constant throughout $\mathcal{R}$, staying close to the divergence of $\mathbf{F}$ at (a, b). Thus

$$\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\, dA \approx \operatorname{div}\mathbf{F}(a, b)\ \text{Area}(\mathcal{R}).$$

or, equivalently,

$$\frac{\text{Net flow out of } R}{\text{Area of } \mathcal{R}} \approx \operatorname{div}\mathbf{F}(a, b). \tag{18.3.3}$$

This means that

$$\operatorname{div}\mathbf{F} \text{ at } P$$

is a measure of the rate at which fluid tends to leave a small region around P. Hence another reason for the name "divergence." If $\operatorname{div}\mathbf{F}$ is positive, fluid near P tends to get less dense (diverge). If $\operatorname{div}\mathbf{F}$ is negative, fluid near P tends to accumulate (converge).

Moreover, (18.3.3) suggests a different definition of the divergence $\operatorname{div}\mathbf{F}$ at (a, b), namely

Diameter is defined in Section 17.1.

Local Definition of $\operatorname{div}\mathbf{F}(a, b)$

$$\operatorname{div}\mathbf{F}(a, b) = \lim_{\text{Diameter of } \mathcal{R}\to 0} \frac{\oint_C \mathbf{F}\cdot\mathbf{n}\, ds}{\text{Area of } \mathcal{R}}$$

where $\mathcal{R}$ is a region enclosing (a, b) whose boundary C is a simple closed curve.

This definition appeals to our physical intuition. We began by defining div $\mathbf{F}$ mathematically, as $\partial P/\partial x + \partial Q/\partial y$. We now see its physical meaning, which is independent of any coordinate system. This coordinate-free definition is the basis for Section 18.9.

EXAMPLE 4 Estimate the flux of $\mathbf{F}$ across a small circle C of radius a if div $\mathbf{F}$ at the center of the circle is 3.

SOLUTION The flux of $\mathbf{F}$ across C is $\oint_C \mathbf{F}\cdot\mathbf{n}\ ds$, which equals $\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\ dA$, where $\mathcal{R}$ is the disk that C bounds. Since div $\mathbf{F}$ is continuous, it changes little in a small enough disk, and we treat it as almost constant. Then $\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\ dA$ is approximately (3)(Area of $\mathcal{R}$) $= 3(\pi a^2) = 3\pi a^2$. ◇

Proof of Green's Theorem

As Steve Whitaker of the chemical engineering department at the University of California at Davis has observed, "The concepts that one must understand to *prove* a theorem are frequently the concepts one must understand to *apply* the theorem." So read the proof slowly at least twice. It is not here just to show that Green's theorem is true. After all, it has been around for over 150 years, and no one has said it is false. Studying a proof strengthens one's understanding of the fundamentals.

In this proof we use the concepts of a double integral, an iterated integral, a line integral, and the fundamental theorem of calculus. So the proof provides a quick review of four basic ideas.

We prove that $\oint_{\mathcal{R}} Q\ dy = \int_{\mathcal{R}} \frac{\partial Q}{\partial x}\ dA$. The proof that $\oint_C P\ dx = -\int \frac{\partial P}{\partial y}\ dA$ is similar.

To avoid getting involved in distracting details we assume that $\mathcal{R}$ is **strictly convex**: It has no dents and its border has no straight line segments. The basic ideas of the proof show up clearly in this special case. Thus $\mathcal{R}$ has the description $a \le x \le b$, $y_1(x) \le y \le y_2(x)$, as shown in Figure 18.3.5. We will express both $\int_{\mathcal{R}} \frac{\partial Q}{\partial y}\ dA$ and $\int_C Q\ dy$ as definite integrals over the interval $[a, b]$.

First, we have

$$\int_{\mathcal{R}} \frac{\partial Q}{\partial y}\ dA = \int_a^b \int_{y_1(x)}^{y_2(x)} \frac{\partial Q}{\partial y}\ dy\ dx.$$

By the Fundamental Theorem of Calculus,

$$\int_{y_1(x)}^{y_2(x)} \frac{\partial Q}{\partial y}\ dy = Q(x, y_2(x)) - Q(x, y_1(x)).$$

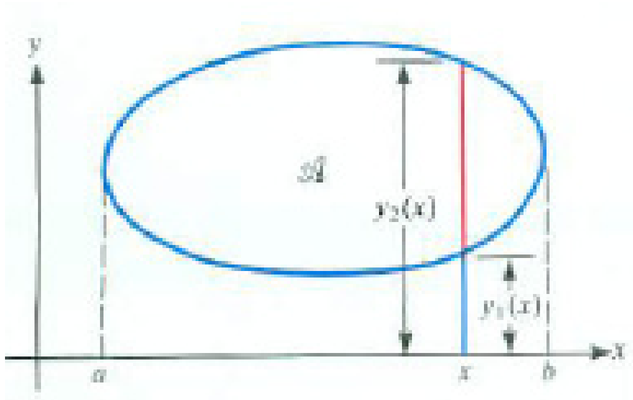


Figure 18.3.5: ARTIST: Please change $\mathcal{A}$ with $\mathcal{R}$.

Hence

$$\int_{\mathcal{R}} \frac{\partial Q}{\partial y}\, dA = \int_a^b \left(Q(x,y_2(x)) - Q(x,y_1(x))\right)\, dx. \tag{18.3.4}$$

Next, to express $\int_C -Q\, dx$ as an integral over $[a.b]$, break the closed path C into two successive paths, one along the bottom part of $\mathcal{R}$, described by $y = y_1(x)$, the other along the top part of $\mathcal{R}$, described by $y = y_2(x)$. Denote the bottom path C_1 and the top path C_2. (See Figure 18.3.6.)

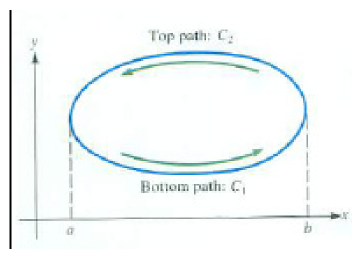


Figure 18.3.6:

Then

$$\oint_C (-Q)\, dx = \int_{C_1} (-Q)\, dx + \int_{C_2} (-Q)\, dx. \tag{18.3.5}$$

But

$$\int_{C_1} (-Q)\, dx = \int_{C_1} \left(-Q(x,y_1(x))\right)\, dx = \int_a^b \left(-Q(x,y_1(x))\right)\, dx,$$

and

$$\int_{C_2} (-Q)\, dx = \int_{C_2} \left(-Q(x,y_2(x))\right)\, dx = \int_b^a \left(-Q(x,y_2(x))\right)\, dx = \int_a^b Q(x,y_2(x))\, dx.$$

Thus by (18.3.5),

$$\begin{aligned}\oint_C (-Q)\, dx &= \int_a^b -Q(x,y_1(x))\, dx + \int_a^b Q(x,y_2(x))\, dx \\ &= \int_a^b \left(Q(x,y_2(x)) - Q(x,y_1(x))\right)\, dx.\end{aligned}$$

This is also the right side of (18.3.4) and concludes the proof.

Summary

We introduced the "divergence" of a vector field $\mathbf{F} = P\mathbf{i} + Q\mathbf{j}$, namely the scalar field $\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$ denoted $\operatorname{div} \mathbf{F}$ or $\nabla \cdot \mathbf{F}$.

We translated Green's Theorem into a theorem about the flux of a vector field in the xy-plane. In symbols, the divergence theorem in the plane says that

$$\oint\limits_{C} \mathbf{F} \cdot \mathbf{n} \, ds = \int\limits_{\mathcal{R}} \operatorname{div} \mathbf{F} \, dA.$$

"The integral of the normal component of $\mathbf{F}$ around a simple closed curve equals the integral of the divergence of $\mathbf{F}$ over the region which the curve bounds."

From this it follows that

$$\operatorname{div} \mathbf{F}(P) = \lim_{\text{diameter of } \mathcal{R} \to 0} \frac{\oint_C \mathbf{F} \cdot \mathbf{n} \, ds}{\text{Area of } \mathcal{R}} = \lim_{\text{diameter of } \mathcal{R} \to 0} \frac{\text{Flux across } C}{\text{Area of } \mathcal{R}}$$

where C is the boundary of the region $\mathcal{R}$, which contains P.

We concluded with a proof of Green's theorem, that provides a review of several basic concepts.

EXERCISES for Section 18.3 *Key:* R–routine, M–moderate, C–challenging

1.[R] State the divergence form of Green's Theorem in symbols.

2.[R] State the divergence form of Green's Theorem in words, using no symbols to denote the vector fields, etc.

In Exercises 3 to 6 compute the divergence of the given vector fields.

3.[R] $\mathbf{F} = x^3y\mathbf{i} + x^2y^3\mathbf{j}$

4.[R] $\mathbf{F} = \arctan(3xy)\mathbf{i} + (e^{y/x})\mathbf{j}$

5.[R] $\mathbf{F} = \ln(x+y)\mathbf{i} + xy(\arcsin y)^2\mathbf{j}$

6.[R] $\mathbf{F} = y\sqrt{1+x^2}\mathbf{i} + \ln((x+1)^3(\sin(y))^{3/5}e^{x+y})\mathbf{j}$

In Exercises 7 to 10 compute $\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\, dA$ and $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ and check that they are equal.

7.[R] $\mathbf{F} = 3x\mathbf{i} + 2y\mathbf{j}$, and $\mathcal{R}$ is the disk of radius 1 with center $(0,0)$.

8.[R] $\mathbf{F} = 5y^3\mathbf{i} - 6x^2\mathbf{j}$, and $\mathcal{R}$ is the disk of radius 2 with center $(0,0)$.

9.[R] $\mathbf{F} = xy\mathbf{i} + x^2y\mathbf{j}$, and $\mathcal{R}$ is the square with vertices $(0,0)$, $(a,0)$ (a,b) and $(0,b)$, where a, $b > 0$.

10.[R] $\mathbf{F} = \cos(x+y)\mathbf{i} + \sin(x+y)\mathbf{j}$, and $\mathcal{R}$ is the triangle with vertices $(0,0)$, $(a,0)$ and (a,b), where a, $b > 0$.

In Exercises 11 to 14 use Green's Theorem expressed in terms of divergence to evaluate $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ for the given $\mathbf{F}$, where C is the boundary of the given region R.

11.[R] $\mathbf{F} = e^x \sin y\mathbf{i} + e^{2x}\cos(y)\mathbf{j}$, and R is the rectangle with vertices $(0,0)$, $(1,0)$, $(0,\pi/2)$, and $(1,\pi/2)$.

12.[R] $\mathbf{F} = y\tan(x)\mathbf{i} + y^2\mathbf{j}$, and R is the square with vertices $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$.

13.[R] $\mathbf{F} = 2x^3y\mathbf{i} - 3x^2y^2\mathbf{j}$, and R is the triangle with vertices $(0,1)$, $(3,4)$, and $(2,7)$.

14.[R] $\mathbf{F} = \frac{-\mathbf{i}}{xy^2} + \frac{\mathbf{j}}{x^2y}$, and R is the triangle with vertices $(1,1)$, $(2,2)$, and $(1,2)$. HINT: Write $\mathbf{F}$ with a common denominator.

15.[R] In Example 3 we found $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ by computing a double integral. Instead, evaluate the integral $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ directly.

16.[R] Let $\mathbf{F}(x,y) = \mathbf{i}$, a constant field.

(a) Evaluate directly the flux of $\mathbf{F}$ around the triangular path, $(0,0)$ to $(1,0)$, to $(0,1)$ back to $(0,0)$.

(b) Use the divergence of $\mathbf{F}$ to evaluate the flux in (a).

17.[R] Let a be a "small number" and $\mathcal{R}$ be the square with vertices (a, a), $(-a, a)$, $(-a, -a)$, and $(a, -a)$, and C its boundary. If the divergence of $\mathbf{F}$ at the origin is 3, estimate $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$.

18.[R] Assume $\|\mathbf{F}(P)\| \leq 4$ for all points P on a curve of length L that bounds a region $\mathcal{R}$ of area A. What can be said about the integral $\int_{\mathcal{R}} \nabla \cdot \mathbf{F}\, dA$?

19.[R] Verify the divergence form of Green's Theorem for $\mathbf{F} = 3x\mathbf{i} + 4y\mathbf{j}$ and C the square whose vertices are $(2, 0)$, $(5, 0)$, $(5, 3)$, and $(2, 3)$.

A vector field $\mathbf{F}$ is said to be **divergence free** when $\nabla \cdot \mathbf{F} = 0$ at every point in the field.

20.[R] Figure 18.3.7 shows four vector fields. Two are divergence-free and two are not. Decide which two are not, copy them onto a sheet of drawing paper, and sketch a closed curve C for which $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$ is not 0.

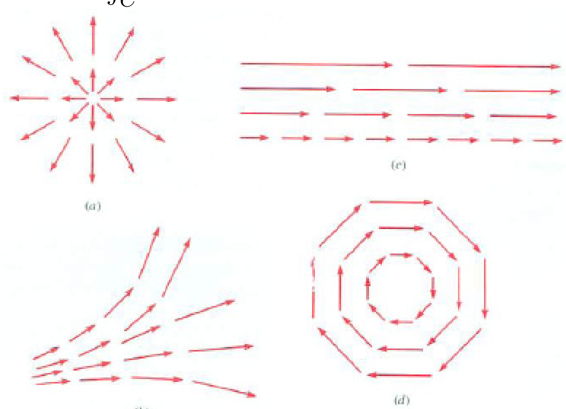


Figure 18.3.7:

21.[R] For a vector field $\mathbf{F}$,

(a) Is the curl of the gradient of $\mathbf{F}$ always $\mathbf{0}$?

(b) Is the divergence of the gradient of $\mathbf{F}$ always 0?

(c) Is the divergence of the curl of $\mathbf{F}$ always 0?

(d) Is the gradient of the divergence of $\mathbf{F}$ always $\mathbf{0}$?

22.[R] Figure 18.3.8 describes the flow $\mathbf{F}$ of a fluid. Decide whether $\nabla \cdot \mathbf{F}$ is positive, negative, or zero at each of the points A, B, and C.

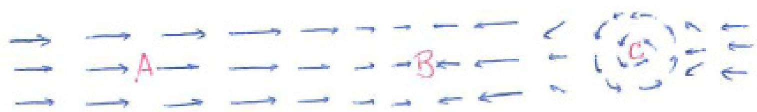


Figure 18.3.8:

23.[R] If $\operatorname{div}\mathbf{F}$ at $(0.1, 0.1)$ is 3 estimate $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$, where C is the curve around the square whose vertices are $(0,0)$, $(0.2,0)$, $(0.2,0.2)$, $(0,0.2)$.

24.[M] Find the area of the region bounded by the line $y = x$ and the curve

$$\begin{cases} x &= t^6 + t^4 \\ y &= t^3 + t \end{cases}$$

for t in $[0,1]$. HINT: Use Green's Theorem.

25.[M] Let f be a scalar function. Let $\mathcal{R}$ be a convex region and C its boundary taken counterclockwise. Show that

$$\int\limits_{\mathcal{R}} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\right)\, dA = \oint\limits_{C} \left(\frac{\partial f}{\partial x}\, dy - \frac{\partial f}{\partial x}\, dx\right).$$

26.[M] Let $\mathbf{F}$ be the vector field whose formula in polar coordinates is $\mathbf{F}(r,\theta) = r^n\widehat{\mathbf{r}}$, where $\mathbf{r} = x\mathbf{i}+y\mathbf{j}$, $r = \|\mathbf{r}\|$, and $\widehat{\mathbf{r}} = \mathbf{r}/r$. Show that the divergence of $\mathbf{F}$ is $(n+1)r^{n-1}$. HINT: First express $\mathbf{F}$ in rectangular coordinates. NOTE: See also Exercise 46 in Section 18.8.

27.[M] A region with a hole is bounded by two oriented curves C_1 and C_2, as in Figure 18.3.9. which shows typical exterior-pointing unit normal vectors.

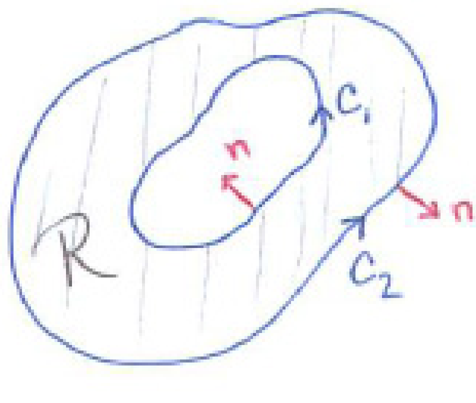


Figure 18.3.9:

Find an equation expressing $\int_{\mathcal{R}} \nabla \cdot \mathbf{F}\, dA$ in terms of $\oint_{C_1} \mathbf{F} \cdot \mathbf{n}\, ds$ and $\oint_{C_2} \mathbf{F} \cdot \mathbf{n}\, ds$. HINT: Break R into two regions that have no holes, as in Exercises 34 and 35.

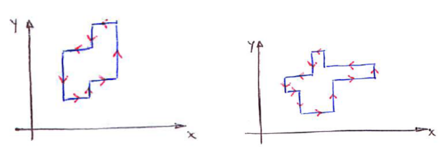
28.[M] The region R is bounded by the curves C_1 and C_2, as in Figure 18.3.10.

Figure 18.3.10:

(a) Show that $\oint_{C_1} \mathbf{F} \cdot \mathbf{n}\, ds - \int_{C_2} \mathbf{F} \cdot \mathbf{n}\, ds = \int_R (\nabla \cdot \mathbf{F})\, dA$.

(b) If $\nabla \cdot \mathbf{F} = 0$ in $\mathcal{R}$, show that $\int_{C_1} \mathbf{F} \cdot \mathbf{n}\, ds = \int_{C_2} \mathbf{F} \cdot \mathbf{n}\, ds$.

29.[M] Let $\mathbf{F}$ be a vector field in the xy-plane whose flux across any rectangle is 0. Show that its flux across the curves in Figure 18.3.11(a) and (b) is also 0.



(a) (b)

Figure 18.3.11:

30.[M] Assume that the circulation of $\mathbf{F}$ along every circle in the xy-plane is 0. Must $\mathbf{F}$ be conservative?

31.[C] The field $\mathbf{F}$ is defined throughout the xy-plane. If the flux of $\mathbf{F}$ across every circle is 0, must the flux of $\mathbf{F}$ across every square be 0? Explain.

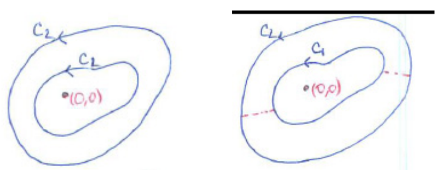
32.[C] Let $\mathbf{F}(x, y)$ describe a fluid flow. Assume $\nabla \cdot \mathbf{F}$ is never 0 in a certain region

R. Show that none of the stream lines in the region closes up to form a loop within $\mathcal{R}$. HINT: At each point P on a stream line, $\mathbf{F}(P)$ is tangent to that streamline.

33.[C] Let $\mathcal{R}$ be a region in the xy-plane bounded by the closed curve C. Let $f(x,y)$ be defined on the plane. Show that

$$\int_{\mathcal{R}} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x^2}\right)\, dA = \oint_C D_{\mathbf{n}}(f)\, ds.$$

34.[C] Assume that $\mathbf{F}$ is defined everywhere in the xy-plane except at the origin and that the divergence of $\mathbf{F}$ is identically 0. Let C_1 and C_2 be two counterclockwise simple curves circling the origin. C_1 lies within the region within C_2. Show that $\oint_{C_1} \mathbf{F}\cdot\mathbf{n}\, ds = \int_{C_2} \mathbf{F}\cdot\mathbf{n}\, ds$. (See Figure 18.3.12(a).)



(a) (b)

Figure 18.3.12:

HINT: Draw the dashed lines in Figure 18.3.12(b) to cut the region between C_1 and C_2 into two regions.

35.[C] (This continues Exercise 34.) Assume that $\mathbf{F}$ is defined everywhere in the xy-plane except at the origin and that the divergence of $\mathbf{F}$ is identically 0. Let C_1 and C_2 be two counterclockwise simple curves circling the origin. They may intersect. Show that $\oint_{C_1} \mathbf{F}\cdot\mathbf{n}\, ds = \oint_{C_2} \mathbf{F}\cdot\mathbf{n}\, ds$. The message from this Exercise is this: if the divergence of $\mathbf{F}$ is 0, you are permitted to replace an integral over a complicated curve by an integral over a simpler curve.

36.[C]

(a) Draw enough vectors for the field $\mathbf{F}(x,y) = (x\mathbf{i} + y\mathbf{j})/(x^2+y^2)$ to show what it looks like.

(b) Compute $\nabla\cdot\mathbf{F}$.

(c) Does your sketch in (a) agree with what you found for $\nabla\cdot\mathbf{F}$. in (b)? (If not, redraw the vector field.)

18.4 Central Fields and Steradians

Central fields are a special but important type of vector field that appear in the study of gravity and the attraction or repulsion of electric charges. These fields radiate from a point mass or point charge. Physicists invented these fields in order to avoid the mystery of "action at a distance." One particle acts on another directly, through the vector field it creates. This comforts students of gravitation and electromagnetism by glossing over the riddle of how an object can act upon another without any intervening object such as a rope or spring.

Figure 18.4.1:

Central Fields

A **central field** is a continuous vector field defined everywhere in the plane (or in space) except, perhaps, at a point $\mathcal{O}$, with these two properties:

1. Each vector points towards (or away from) $\mathcal{O}$.

2. The magnitudes of all vectors at a given distance from $\mathcal{O}$ are equal.

$\mathcal{O}$ is call the center, or pole, of the field. A central field is also called "radially symmetric." There are various ways to think of a central vector field. For such a field in the plane, all the vectors at points on a circle with center $\mathcal{O}$ are perpendicular to the circle and have the same length, as shown in Figures 18.4.1 and 18.4.2.

The same holds for central vector fields in space, with "circle" replaced by "sphere."

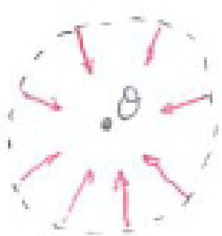


Figure 18.4.2:

The formula for a central vector field has a particularly simple form. Let the field be $\mathbf{F}$ and P any point other than $\mathcal{O}$. Denote the vector $\overrightarrow{OP}$ by $\mathbf{r}$ and its magnitude by r and $\mathbf{r}/r$ by $\widehat{\mathbf{r}}$. Then there is a scalar function f, defined for all positive numbers, such that

$$\mathbf{F}(P) = f(r)\widehat{\mathbf{r}}.$$

The magnitude of $\mathbf{F}(P)$ is $\|f(r)\|$. If $f(r)$ is positive, $\mathbf{F}(P)$ points away from $\mathcal{O}$. If $f(r)$ is negative, $\mathbf{F}(P)$ points toward $\mathcal{O}$.

To conclude this introduction to central fields we point out that a central field is a vector-valued function of more than one variable. Because the point P with coordinates (x, y, z) is also associated with the vector $\mathbf{r} = \overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ we may denote $\mathbf{F}(P)$ as $\mathbf{F}(x, y, z)$ or $\mathbf{F}(\mathbf{r})$.

Central Vector Fields in the Plane

Using polar coordinates with pole placed at the point $\mathcal{O}$, we may express a central field in the form

$$\mathbf{F}(\mathbf{r}) = f(r)\widehat{\mathbf{r}},$$

where $r = \|\mathbf{r}\|$ and $\widehat{\mathbf{r}} = \mathbf{r}/r$. The magnitude of $\mathbf{F}(\mathbf{r})$ is $|f(r)|$.

We already met such a field in Section 18.1 in the study of line integrals. In that case, $f(r) = 1/r$; the "field varied as the inverse first power." When, in Section 15.4, we encountered the line integral for the normal component of this field along a curve we found that it gives the number of radians the curve subtends.

See page 1239.

The vector field $\mathbf{F}(\mathbf{r}) = (1/r)\widehat{\mathbf{r}}$ can also be written as

$$\mathbf{F}(\mathbf{r}) = \frac{\mathbf{r}}{r^2}. \tag{18.4.1}$$

When glancing too quickly at (18.4.1), you might think its magnitude is inversely proportional to the square of r. However, the magnitude of the vector $\mathbf{r}$ in the numerator is r; the magnitude of $\mathbf{r}/r^2$ is $r/r^2 = 1/r$, the reciprocal of the first power of r.

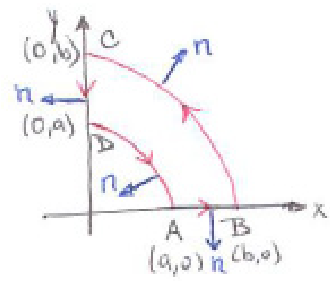


Figure 18.4.3:

EXAMPLE 1 Evaluate the flux $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$ for the central field $\mathbf{F}(x,y) = f(r)\widehat{\mathbf{r}}$, where $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$, over the closed curve shown in Figure 18.4.3. We have $a < b$ and the path goes from $A = (a,0)$ to $B = (b,0)$ to $C = (0,b)$, to $D = (0,a)$ and ends at $A = (a,0)$.

SOLUTION On the paths from A to B and from C to D the exterior normal, $\mathbf{n}$, is perpendicular to $\mathbf{F}$, so $\mathbf{F}\cdot\mathbf{n} = 0$, and these integrands contribute nothing to the integral. On BC, $\mathbf{F}$ equals $f(b)\widehat{\mathbf{r}}$. There $\widehat{\mathbf{r}} = \mathbf{n}$, so $\mathbf{F}\cdot\mathbf{n} = f(b)$ since $\mathbf{r}\cdot\mathbf{n} = 1$. Note that the length of arc BC is $(2\pi b)/4 = \pi b/2$. Thus

$$\int_B^C \mathbf{F}\cdot\mathbf{n}\,ds = \int_B^C f(b)\,ds = f(b)\int_B^C ds == \frac{\pi b}{2}f(b)$$

On the arc DC, $\widehat{\mathbf{r}} = -\mathbf{n}$. A similar calculation shows that

$$\int_D^C \mathbf{F}\cdot\mathbf{n}\,ds = -\frac{\pi}{2}af(a).$$

Hence

$$\oint_C \mathbf{F}\cdot\mathbf{n}\,ds \;=\; 0 + \frac{\pi}{2}bf(b) + 0 - \frac{\pi}{2}af(a) = \frac{\pi}{2}(bf(b) - af(a)).$$

◇

In order for a central field $f(r)\widehat{\mathbf{r}}$ to have zero flux around all paths of the special type shown in Figure 18.4.3, we must have

$$f(b)b - f(a)a = 0,$$

for all positive a and b. In particular,

$$f(b)b - f(1)1 = 0 \qquad \text{or} \qquad f(b) = \frac{f(1)}{b}.$$

Thus $f(r)$ must be inversely proportional to r and there is a constant c such that

$$f(r) = \frac{c}{r}.$$

If $f(r)$ is not of the form c/r, the vector field $\mathbf{F}(x,y) = f(r)\widehat{\mathbf{r}}$ does not have zero flux across these paths. In Exercise 5 you may compute the divergence of $(c/r)\widehat{\mathbf{r}}$ and show that it is zero.

> The only central vector fields with center at the origin in the plane with zero divergence are these whose magnitude is inversely proportional to the distance from the origin.

We underline "in the plane," because in space the only central fields with zero flux across closed surfaces have a magnitude inversely proportional to the square of the distance to the pole, as we will see in a moment.

Knowing that the central field $\mathbf{F} = \widehat{\mathbf{r}}/r$ has zero divergence enables us to evaluate easily line integrals of the form $\oint_C \frac{\widehat{\mathbf{r}}\cdot\mathbf{n}}{r}\, ds$, as the next example shows.

EXAMPLE 2 Let $\mathbf{F}(\mathbf{r}) = \widehat{\mathbf{r}}/r$. Evaluate $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ where C is the counterclockwise circle of radius 1 and center $(2,0)$, as shown in Figure 18.4.4.

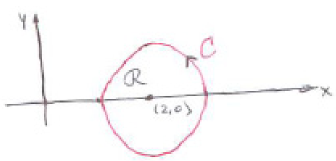


Figure 18.4.4:

SOLUTION Exercise 5 shows that the field $\mathbf{F}$ has 0-divergence throughout C and the region R that C bounds. By Green's Theorem, the integral also equals the integral of the divergence over R:

$$\oint_C \mathbf{F}\cdot\mathbf{n}\, ds = \int_R \nabla\cdot\mathbf{F}\, dA. \tag{18.4.2}$$

Since the divergence of $\mathbf{F}$ is 0 throughout R, the integral on the right side of (18.4.2) is 0. Therefore $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds = 0$. ◇

The next example involves a curve that surrounds a point where the vector field $\mathbf{F} = \widehat{\mathbf{r}}/r$ is not defined.

EXAMPLE 3 Let C be a simple closed curve enclosing the origin. Evaluate $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$, where $\mathbf{F} = \widehat{\mathbf{r}}/r$.

SOLUTION Figure 18.4.5 shows C and a small circle D centered at the origin and situated in the region that C bounds. Without a formula describing C, we could not compute $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$ directly. However, since the divergence of $\mathbf{F}$ is 0 throughout the region bounded by C and D, we have, by the Two-Curve Case of Green's Theorem,

See page 1239 in Section 15.4.

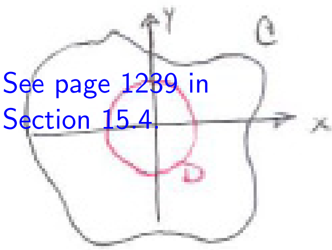


Figure 18.4.5:

$$\oint_C \mathbf{F}\cdot\mathbf{n}\,ds = \oint_D \mathbf{F}\cdot\mathbf{n}\,ds. \tag{18.4.3}$$

The integral on the right-hand side of (18.4.3) is easy to compute directly. To do so, let the radius of D be a. Then for points P on D, $\mathbf{F}(P) = \widehat{\mathbf{r}}/a$. Now, $\widehat{\mathbf{r}}$ and $\mathbf{n}$ are the same unit vector. So $\widehat{\mathbf{r}}\cdot\mathbf{n} = 1$. Thus

$$\oint_D \mathbf{F}\cdot\mathbf{n}\,ds = \oint_D \frac{\widehat{\mathbf{r}}\cdot\mathbf{n}}{a}\,ds = \int_D \frac{1}{a}\,ds = \frac{1}{a}2\pi a = 2\pi.$$

Hence $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds = 2\pi$. ◇

Central Fields in Space

A central field in space with center at the origin has the form $\mathbf{F}(x,y,z) = \mathbf{F}(\mathbf{r}) = f(r)\widehat{\mathbf{r}}$ We show that if the flux of $\mathbf{F}$ over any surface bounding certain special regions is zero then $f(r)$ must be inversely proportional to the square of r.

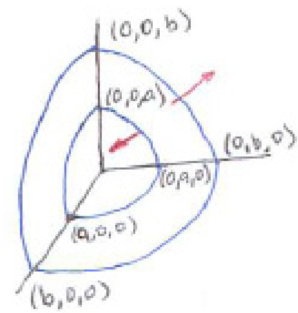


Figure 18.4.6:

Consider the surface S shown in Figure 18.4.6. It consists of an octant of two concentric spheres, one of radius a, the other of radius b, $a < b$, together with the flat surfaces on the coordinate planes. Let $\mathcal{R}$ be the region bounded by the surface S. On its three flat sides $\mathbf{F}$ is perpendicular to the exterior normal. On the outer sphere $\mathbf{F}(x,y,z)\cdot\mathbf{n} = f(b)$. On the inner sphere $\mathbf{F}(x,y,z)\cdot\mathbf{n} = -f(a)$. Thus

Surface area of a sphere of radius r is $4\pi r^2$.

$$\oint_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,dS = f(b)(\frac{1}{8})(4\pi b^2) - f(a)(\frac{1}{8})(4\pi a^2) = \frac{\pi}{2}(f(b)b^2 - f(a)a^2).$$

Since this is to be 0 for all positive a and b, it follows that there is a constant c, such that

Compare with Example 1.

$$f(r) = \frac{c}{r^2}.$$

The magnitude must be proportional to the "inverse square."

The following fact is justified in Exercise 28:

> The only central vector field with center at the origin in the plane with zero divergence are these whose magnitude is inversely proportional to the distance from the origin.

A Geometric Application

See Sections 18.7 and 18.9.

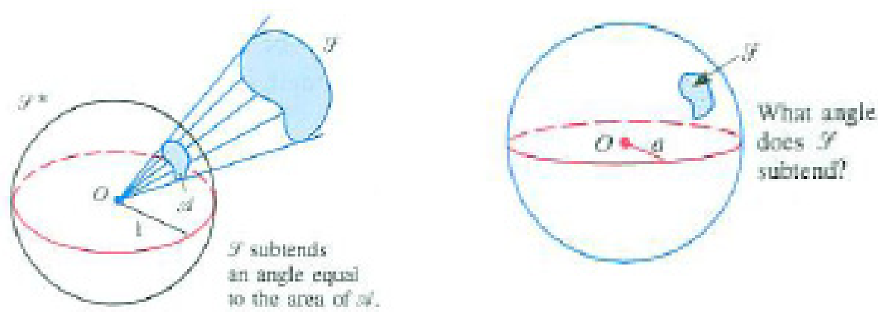
As we will see later in this chapter an "inverse square" central field is at the heart of gravitational theory and electrostatics. Now we show how it is used in geometry, a result we will apply in both areas.

In Section 15.4 we showed how radian measure could be expressed in terms of the line integral $\int_C (\widehat{\mathbf{r}}/r) \cdot \mathbf{n}\, ds$, that is, in terms of the central field whose magnitude is inversely proportional to the *first power* of the distance from the center. That was based on circular arcs in a plane. Now we move up one dimension and consider patches on surfaces of spheres, which will help us measure solid angles.

Let $\mathcal{O}$ be a point and $\mathcal{S}$ a surface such that each ray from $\mathcal{O}$ meets $\mathcal{S}$ in at most one point. Let $\mathcal{S}^*$ be the unit sphere with center at $\mathcal{O}$. The rays from $\mathcal{O}$ that meet $\mathcal{S}$ intersect $\mathcal{S}^*$ in a set that we call $\mathcal{R}$, as shown in Figure 18.4.7(a). Let the area of $\mathcal{R}$ be A. The solid angle subtended by $\mathcal{S}$ at $\mathcal{O}$ is said to have a measure of A steradians

Steradians comes from *stereo*, the Greek word for space, and *radians*.

For instance, a closed surface $\mathcal{S}$ that encloses $\mathcal{O}$ subtends a solid angle of 4π steradians, because the area of the unit sphere is 4π.



(a) (b)

Figure 18.4.7:

EXAMPLE 4 Let $\mathcal{S}$ be part of the surface of a sphere of radius a, $\mathcal{S}_a$, whose center is $\mathcal{O}$. Find the angle subtended by $\mathcal{S}$ at $\mathcal{O}$. (See Figure 18.4.7(b).)

SOLUTION The entire sphere $\mathcal{S}_a$ subtends an angle of 4π steradians because it has an area $4\pi a^2$. We therefore have the proportion

$$\frac{\text{Angle } \mathcal{S} \text{ subtends}}{\text{Angle } \mathcal{S}_a \text{ subtends}} = \frac{\text{Area of } \mathcal{S}}{\text{Area of } \mathcal{S}_a},$$

or

$$\frac{\text{Angle } \mathcal{S} \text{ subtends}}{4\pi} = \frac{\text{Area of } \mathcal{S}}{4\pi a^2}.$$

Hence

$$\text{Angle } \mathcal{S} \text{ subtends} = \frac{\text{Area of } \mathcal{S}}{a^2} \quad \text{steradians.}$$

◇

EXAMPLE 5 Let $\mathcal{S}$ be a surface such that each ray from the point $\mathcal{O}$ meets $\mathcal{S}$ in at most one point. Find an integral that represents in steradians the solid angle that $\mathcal{S}$ subtends at $\mathcal{O}$.

SOLUTION Consider a very small patch of $\mathcal{S}$. Call it $d\mathcal{S}$ and let its area be dA. If we can estimate the angle that this patch subtends at $\mathcal{O}$, then we will have the local approximation that will tell us what integral represents the total solid angle subtended by $\mathcal{S}$.

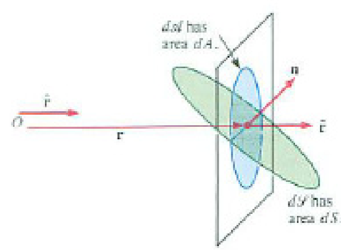


Figure 18.4.8:

Let $\mathbf{n}$ be a unit normal at a point in the patch, which we regard as essentially flat, as in Figure 18.4.8. Let $d\mathcal{A}$ be the projection of the patch $d\mathcal{S}$ on a plane perpendicular to $\mathbf{r}$, as shown in Figure 18.4.8. The area of $d\mathcal{A}$ is approximately dA, where

$$dA = \widehat{\mathbf{r}} \cdot \mathbf{n} \, dS.$$

Now, $d\mathcal{S}$ and $d\mathcal{A}$ subtend approximately the same solid angle, which according to Example 4 is about

$$\frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{\|\mathbf{r}\|^2} \, dS \quad \text{steradians.}$$

Consequently $\mathcal{S}$ subtends a solid angle of

$$\int\limits_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{\|\mathbf{r}\|^2} \, dS \quad \text{steradians.}$$

◇

The following special case will be used in Section 18.5.

Let $\mathcal{O}$ be a point in the region bounded by the closed surface $\mathcal{S}$. Assume each ray from $\mathcal{O}$ meets $\mathcal{S}$ in exactly one point, and let $\mathbf{r}$ denote the position vector from $\mathcal{O}$ to that point. Then

$$\int_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r^2}\, dS = 4\pi. \tag{18.4.4}$$

Incidentally, (18.4.4) is easy to establish when $\mathcal{S}$ is a sphere of radius a and center at the origin. In that case $\widehat{\mathbf{r}} = \mathbf{n}$, so $\widehat{\mathbf{r}} \cdot \mathbf{n} = 1$. Also, $r = a$. Then (18.4.4) becomes $\int_{\mathcal{S}} (1/a^2)\, dS = (1/a^2) 4\pi a^2 = 4\pi$. However, it is not obvious that (18.4.4) holds far more generally, for instance when $\mathcal{S}$ is a sphere and the origin is *not* its center, or when $\mathcal{S}$ is not a sphere.

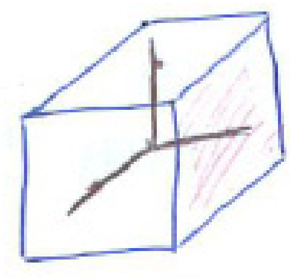


Figure 18.4.9:

EXAMPLE 6 Let $\mathcal{S}$ be the cube of side 2 bounded by the six planes $x = \pm 1$, $y = \pm 1$, $z = \pm 1$, shown in Figure 18.4.9. Find $\oint_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r^2}\, dS$, where $\mathcal{S}$ is one of the six faces of the cube.

SOLUTION Each of the six faces subtends the same solid angle at the origin. Since the entire surface subtends 4π steradians, each face subtends $4\pi/6 = 2\pi/3$ steradians. Then the flux over each face is

$$\int_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r^2}\, dS = \frac{2\pi}{3}.$$

◇

Figure 18.4.10:

In physics books you will see the integral $\int_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r^2}\, dS$ written using other notations, including:

$$\int_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot \mathbf{n}}{r^3}\, dS, \qquad \int_{\mathcal{S}} \frac{\widehat{\mathbf{r}} \cdot d\mathbf{S}}{r^2}, \qquad \int_{\mathcal{S}} \frac{\mathbf{r} \cdot d\mathbf{S}}{r^3}, \qquad \int_{\mathcal{S}} \frac{\cos(\mathbf{r}, \mathbf{n})}{r^2}\, dS.$$

The symbol $d\mathbf{S}$ is short for $\mathbf{n}\, dS$, and calls to mind Figure 18.4.10, which shows a small patch on the surface, together with an exterior normal unit vector.

Recall that $\cos(\mathbf{r}, \mathbf{n})$ denotes the cosine of the angle between $\mathbf{r}$ and $\mathbf{n}$; see also Section 14.2.

Summary

We investigated central vector fields. In the plane the only divergence-free central fields are of the form $(c/r)\widehat{\mathbf{r}}$ where c is a constant, "an inverse first power." In space the only incompressible central fields are of the form $(c/r^2)\widehat{\mathbf{r}}$, "an inverse second power." The field $\widehat{\mathbf{r}}/r^2$ can be used to express the size of a solid angle of a surface $\mathcal{S}$ in steradians as an integral: $\int_{\mathcal{S}} \widehat{\mathbf{r}} \cdot \mathbf{n}/r^2 \; dS$. In particular, if $\mathcal{S}$ encloses the center of the field, then $\int_{\mathcal{S}} \widehat{\mathbf{r}} \cdot \mathbf{n}/r^2 \; dS = 4\pi$.

Incompressible vector fields have divergence zero, and are discussed again in Section 18.6.

EXERCISES for Section 18.4 *Key:* R–routine, M–moderate, C–challenging

1.[R] Define a central field in words, using no symbols.

2.[R] Define a central field with center at $\mathcal{O}$, in symbols.

3.[R] Give an example of a central field in the plane that

(a) does not have zero divergence,

(b) that does have zero divergence.

4.[R] Give an example of a central field in space that

(a) that is not divergence-free,

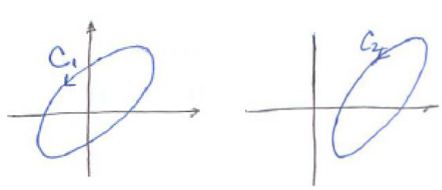
(b) that is divergence-free.

5.[R] Let $\mathbf{F}(x,y)$ be an inverse-first-power central field in the plane $\mathbf{F}(x,y) = (c/r)\widehat{\mathbf{r}}$, where $\mathbf{r} = x\mathbf{i}+y\mathbf{j}$. Compute the divergence of $\mathbf{F}$. HINT: First write $\mathbf{F}(x,y)$ as $\frac{cx\mathbf{i}+cy\mathbf{j}}{x^2+y^2}$.

6.[R] Show that the curl of a central vector field in the plane is $\mathbf{0}$.

7.[R] Show that the curl of a central vector field in space is $\mathbf{0}$.

8.[R] Let $\mathbf{F}(\mathbf{r}) = \widehat{\mathbf{r}}/r$. Evaluate $\oint_C \mathbf{F}\cdot\mathbf{n}\, ds$ as simply as you can for the two ellipses in Figure 18.4.11.



(a) (b)

Figure 18.4.11:

9.[R] Figure 18.4.12 shows a cube of side 2 with one corner at the origin.

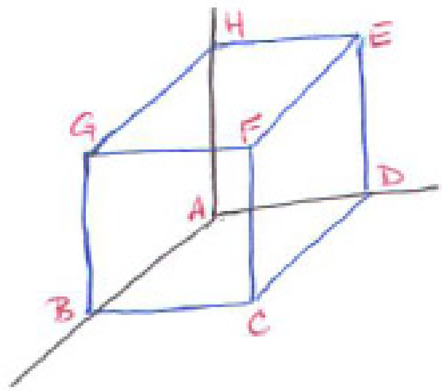


Figure 18.4.12:

Evaluate as easily as you can the integral of the function $\widehat{\mathbf{r}} \cdot \mathbf{n}/r^2$ over

(a) the square $EFGH$,

(b) the square $ABCD$,

(c) the entire surface of the cube.

10.[R] Let $\mathbf{F}(\mathbf{r}) = \widehat{\mathbf{r}}/r^3$. Evaluate the flux of $\mathbf{F}$ over the sphere of radius 2 and center at the origin.

11.[R] A pyramid is made of four congruent equilateral triangles. Find the steradians subtended by one face at the centroid of the pyramid. (No integration is necessary.)

12.[R] How many steradians does one face of a cube subtend at

(a) One of the four vertices not on that face?

(b) The center of the cube? NOTE: No integration is necessary.

13.[M] In Example 2 the integral $\oint_C \widehat{\mathbf{r}} \cdot \mathbf{n}/r \; ds$ turned out to be 0. How would you explain this in terms of subtended angles?

14.[M] Let $\mathbf{F}$ and $\mathbf{G}$ be central vector fields in the plane with different centers.

(a) Show that the vector field $\mathbf{F} + \mathbf{G}$ is not a central field.

(b) Show that the divergence of $\mathbf{F} + \mathbf{G}$ is 0.

15.[M] In Example 6, we evaluated a surface integral by interpreting it in terms of the size of a subtended solid angle. Evaluate the integral directly, without that knowledge.

16.[M] Let $\mathcal{S}$ be the triangle whose vertices are $(1,0,0)$, $(0,1,0)$, $(0,0,1)$. Evaluate $\int_{\mathcal{S}} \frac{\widehat{\mathbf{r}}\cdot\mathbf{n}}{r^2}\, dS$ by using steradians.

17.[M] Evaluate the integral in Exercise 16 directly.

18.[M] Let $\mathbf{F}(x,y,z) = \frac{x\mathbf{i}+y\mathbf{j}+0\mathbf{k}}{x^2+y^2}$ be a vector field in space.

(a) What is the domain of $\mathbf{F}$?

(b) Sketch $\mathbf{F}(1,1,0)$ and $\mathbf{F}(1,1,2)$ with tails at the given points.

(c) Show $\mathbf{F}$ is not a central field.

(d) Show its divergence is 0.

Exercises 19 to 26 are related.

19.[M] Let $\mathbf{F}$ be a planar central field. Show that $\nabla\times\mathbf{F}$ is $\mathbf{0}$. Hint: $\mathbf{F}(x,y) = \frac{g(\sqrt{x^2+y^2}(x\mathbf{i}+y\mathbf{j}))}{\sqrt{x^2+y^2}}$ for some scalar function g.

20.[M] (This continues Exercise 19.) Show that $\mathbf{F}$ is a gradient field; to be specific, $\mathbf{F} = \nabla g(\sqrt{x^2+y^2})$.

21.[C] Carry out the computation to show that the *only* central fields in space that have zero divergence have the form $\mathbf{F}(\mathbf{r}) = c\widehat{\mathbf{r}}/r^2$, if the origin of the coordinates is at the center of the field.

22.[C] If we worked in four-dimensional space instead of the two-dimensional plane or three-dimensional space, which central fields do you think would have zero divergence? Carry out the calculation to confirm your conjecture.

23.[C] Let $\mathbf{F} = \widehat{\mathbf{r}}/r^2$ and S be the surface of the lopsided pyramid with square base, whose vertices are $(0,0,0)$, $(1,1,0)$, $(0,1,0)$, $(0,1,1)$, $(1,1,1)$.

(a) Sketch the pyramid.

(b) What is the integral of $\mathbf{F}\cdot\mathbf{n}$ over the square base?

(c) What is the integral of $\mathbf{F}\cdot\mathbf{n}$ over each of the remaining four faces?

(d) Evaluate $\oint_S \mathbf{F} \cdot \mathbf{n}\, dS$.

24.[C] Let C be the circle $x^2 + y^2 = 4$ in the xy-plane. For each point Q in the disk bounded by C consider the central field with center Q, $\mathbf{F}(P) = \overrightarrow{PQ}/\|PQ\|^2$. Its magnitude is inversely proportional to the first power of the distance P is from Q. For each point Q consider the flux of $\mathbf{F}$ across C.

(a) Evaluate directly the flux when Q is the origin $(0, 0)$.

(b) If Q is not the origin, evaluate the flux of $\mathbf{F}$.

(c) Evaluate the flux when Q lies on C.

25.[C] Let $\mathbf{F}$ be the central field in the plane, with center at $(1, 0)$ and with magnitude inversely proportional to the first power of the distance to $(1, 0)$: $\mathbf{F}(x, y) = \frac{(x-1)\mathbf{i}+y\mathbf{j}}{\|(x-1)\mathbf{i}+y\mathbf{j}\|^2}$. Let C be the circle of radius 2 and center at $(0, 0)$.

(a) By thinking in terms of subtended angle, evaluate the flux $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds$.

(b) Evaluate the flux by carrying out the integration.

26.[C] This exercise gives a geometric way to see why a central force is conservative. Let $\mathbf{F}(x, y) = f(r)\widehat{\mathbf{r}}$. Figure 18.4.13 shows $\mathbf{F}(x, y)$ and a short vector $\overrightarrow{d\mathbf{r}}$ and two circles.

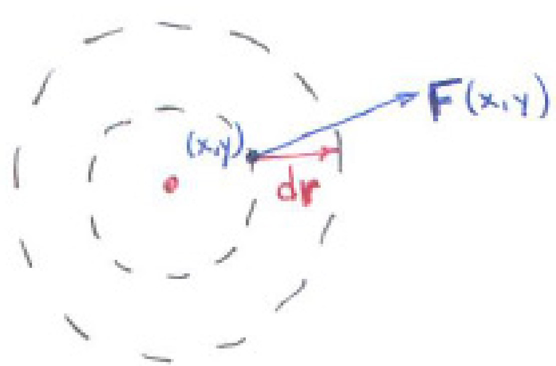


Figure 18.4.13:

(a) Why is $\mathbf{F}(x, y) \cdot\ d\mathbf{r}$ approximately $f(r)\ dr$, where dr is the difference in the radii of the two circles?

(b) Let C be a curve from A to B, where $A = (a, \alpha)$ and $B = (b, \beta)$ in polar coordinates. Why is $\int_C \mathbf{F} \cdot \, d\mathbf{r} = \int_a^b f(r) \, dr$?

(c) Why is $\mathbf{F}$ conservative?

Skill Drill

27.[R] Show that the derivative of $\frac{1}{3}\tan^3(x) - \tan(x) + x$ is $\tan^4(x)$.

28.[R] Use integration by parts to show that

$$\int \tan^n(x) \, dx = \frac{\tan^{n-1}(x)}{n-1} - \int \tan^{n-2}(x) \, dx.$$

29.[R] Entry 16 in the Table of Antiderivatives in the front cover of this book is:

$$\int \frac{dx}{x(ax+b} = \frac{1}{b} \ln \left| \frac{x}{ax+b} \right|.$$

(a) Use a partial fraction expansion to evaluate the antiderivative.

(b) Use differentiation to check that this formula is correct.

30.[R] Repeat Exercise 29 for entry 17 in the Table of Antiderivatives:

$$\int \frac{dx}{x(ax+b)} = \frac{1}{b} \ln \left| \frac{x}{ax+b} \right|.$$

31.[R] Show that $x \arccos(x) - \sqrt{1-x^2}$ is an integral of $\arccos(x)$.

32.[R] Find $\int \arctan(x)$.

33.[R]

(a) Find $\int xe^{ax} \, dx$.

(b) Use integration by parts to show that

$$\int x^m e^{ax}\,dx = \frac{x^m e^{ax}}{a} - \frac{m}{a}\int x^{m-1} e^{ax}\,dx.$$

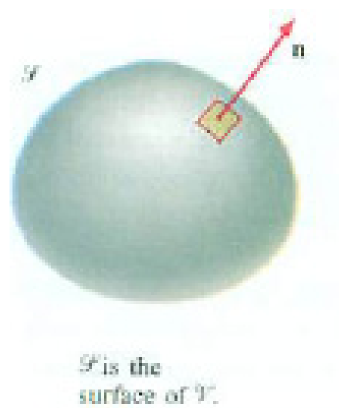
(c) Verify the equation in (b) by differentiating the right hand side.

18.5 The Divergence Theorem in Space (Gauss' Theorem)

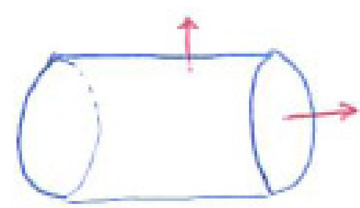
In Sections 18.2 and 18.3 we developed Green's theorem and applied it in two forms for a vector field $\mathbf{F}$ in the plane. One form concerned the line integral of the tangential component of $\mathbf{F}$, $\oint_C \mathbf{F}\cdot\mathbf{T}\,ds$, also written as $\oint_C \mathbf{F}\cdot d\mathbf{r}$. The other concerned the integral of the normal component of $\mathbf{F}$, $\oint_C \mathbf{F}\cdot\mathbf{n}\,ds$. In this section we develop the **Divergence Theorem**, an extension of the second form from the plane to space. The extension of the first form to space is the subject of Section 18.6. In Section 18.7 the Divergence Theorem will be applied to electro-magnetism.

The Divergence (or Gauss's) Theorem

Consider a region $\mathcal{R}$ in space bounded by a surface $\mathcal{S}$. For instance, $\mathcal{R}$ may be a ball and $\mathcal{S}$ its surface. This is a case encountered in the elementary theory of electro-magnetism. In another case, $\mathcal{R}$ is a right circular cylinder and $\mathcal{S}$ is its surface, which consists of two disks and its curved side. See Figure 18.5.1(a). Both figures show typical unit exterior normals, perpendicular to the surface.



(a)



(b)

Figure 18.5.1:

The Divergence Theorem relates an integral over the surface to an integral over the region it bounds.

Theorem. *Divergence Theorem —One-Surface Case. Let $\mathcal{V}$ be the region in space bounded by the surface $\mathcal{S}$. Let* $\mathbf{n}$ *denote the exterior unit normal of $\mathcal{V}$*

along the boundary $\mathcal{S}$. Then

$$\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\ ds = \int_{\mathcal{V}} \nabla\cdot\mathbf{F}\ dV$$

for any vector field $\mathbf{F}$ *defined on* $\mathcal{V}$.

State the Theorem aloud.

In words: "The integral of the normal component of $\mathbf{F}$ over a surface equals the integral of the divergence of $\mathbf{F}$ over the region the surface bounds."

The integral $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\ dS$ is called the **flux** of the field $\mathbf{F}$ across the surface $\mathcal{S}$.

If $\mathbf{F} = P\mathbf{i}+Q\mathbf{j}+R\mathbf{k}$ and $\cos(\alpha)$, $\cos(\beta)$, and $\cos(\gamma)$ are the direction cosines of the exterior normal, then the Divergence Theorem reads

$$\int_{\mathcal{S}} (P\mathbf{i}+Q\mathbf{j}+R\mathbf{k})\cdot(\cos(\alpha)\mathbf{i}+\cos(\beta)\mathbf{j}+\cos(\gamma)\mathbf{k})\ dS = \int_{\mathcal{V}} \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\ dV.$$

Evaluating the dot product puts the Divergence Theorem in the form

Direction cosines are defined in Section 14.4.

$$\int_{\mathcal{S}} (P\cos(\alpha)+Q\cos(\beta)+R\cos(\gamma))\ dS = \int_{\mathcal{V}} \left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\ dV.$$

When the Divergence Theorem is expressed in this form, we see that it amounts to three scalar theorems:

$$\int_{\mathcal{S}} P\cos(\alpha)\ dS = \int_{\mathcal{V}} \frac{\partial P}{\partial x}\ dV, \quad \int_{\mathcal{S}} Q\cos(\beta)\ dS = \int_{\mathcal{V}} \frac{\partial Q}{\partial y}\ dV, \quad \text{and} \quad \int_{\mathcal{S}} R\cos(\gamma)\ dS = \int_{\mathcal{V}} \frac{\partial R}{\partial z}\ dV. \tag{18.5.1}$$

As is to be expected, establishing these three equations proves the Divergence Theorem. We delay the proof to the end of this section, after we have shown how the Divergence Theorem is applied.

You could have guessed the result in this Example by thinking in terms of the solid angle and steradians. Why?

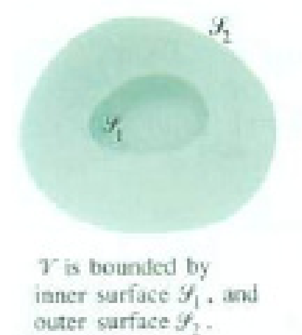


Figure 18.5.2:

Two-Surface Version of the Divergence Theorem

The Divergence Theorem also holds if the solid region has several holes like a piece of Swiss cheese. In this case, the boundary consists of several separate closed surfaces. The most important case is when there is just one hole and hence an inner surface $\mathcal{S}_1$ and an outer surface $\mathcal{S}_2$ as shown in Figure 18.5.2.

Theorem (Divergence Theorem — Two-Surface Case.)**.** *Let $\mathcal{V}$ be a region in space bounded by the surfaces $\mathcal{S}_1$ and $\mathcal{S}_2$. Let $\mathbf{n}^*$ denote the exterior normal along the boundary. Then*

$$\int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}^*\, dS + \int_{\mathcal{S}_2} \mathbf{F}\cdot\mathbf{n}^*\, dS = \int_{\mathcal{V}} \operatorname{div}\mathbf{F}\, dV$$

for any vector field defined on $\mathcal{V}$.

Compare with (18.2.4) in Exercise 3 in Section 18.2.

The importance of this form of the Divergence Theorem is that it allows us to conclude that the flux across each of the surfaces are the same provided these surfaces form the boundary of a solid where $\operatorname{div}\mathbf{F} = 0$.

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two closed surfaces that form the boundary of the region $\mathcal{V}$. Let $\mathbf{F}$ be a vector field defined on $\mathcal{V}$ such that the divergence of $\mathbf{F}$, $\nabla\cdot\mathbf{F}$, is 0 throughout $\mathcal{V}$. Then

$$\int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}\, dS = \int_{\mathcal{S}_2} \mathbf{F}\cdot\mathbf{n}\, dS \qquad (18.5.2)$$

The proof of this result closely parallels the derivation of (18.2.4) in Section 18.2.

The next example is a major application of (18.5.2), which enables us, if the divergence of $\mathbf{F}$ is 0, to replace the integral of $\mathbf{F}\cdot\mathbf{n}$ over a surface by an integral of $\mathbf{F}\cdot\mathbf{n}$ over a more convenient surface.

EXAMPLE 1 Let $\mathbf{F}(\mathbf{r}) = \widehat{\mathbf{r}}/r^2$, the inverse square vector field with center at the origin. Let $\mathcal{S}$ be a convex surface that encloses the origin. Find the flux of $\mathbf{F}$ over the surface, $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS$.

SOLUTION Select a sphere with center at the origin that does not intersect $\mathcal{S}$. This sphere should be very small in order to miss $\mathcal{S}$. Call this spherical surface $\mathcal{S}_1$ and its radius a. Then, by (18.5.2),

$$\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS = \int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}\, dS$$

But $\int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}\, dS$ is easy because the integrand $(\widehat{\mathbf{r}}/r^2)\cdot\mathbf{n}$ equals $\frac{\mathbf{r}\cdot\mathbf{n}}{r^2}$. Then, $\mathbf{r}\cdot\mathbf{n}$ is just 1. Thus:

$$\int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}\, dS = \int_{\mathcal{S}_1} \frac{1}{a^2}\, dS = \frac{1}{a^2}\int_{\mathcal{S}_1} dS = \frac{1}{a^2}4\pi a^2 = 4\pi.$$

◇

A **uniform** or **constant** vector field is a vector field where vectors at every point are all identical. Such fields are used in the next example.

EXAMPLE 2 Verify the Divergence Theorem for the constant field $\mathbf{F}(x, y, z) = 2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$ and the surface $\mathcal{S}$ of a cube whose sides have length 5 and is situated as shown in Figure 18.5.3.

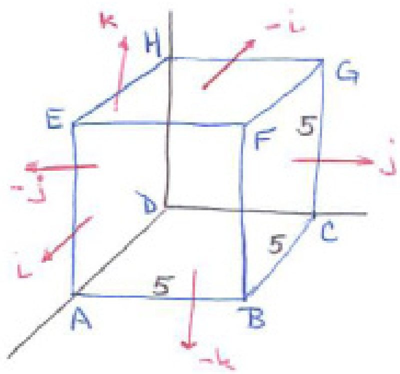


Figure 18.5.3:

SOLUTION To find $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS$ we consider the integral of $\mathbf{F}\cdot\mathbf{n}$ over each of the six faces.

On the bottom face, $ABCD$ the unit exterior normal is $-\mathbf{k}$. Thus

$$\mathbf{F}\cdot\mathbf{n} = (2\mathbf{i}+3\mathbf{j}+4\mathbf{k})\cdot(-\mathbf{k}) = -4.$$

So

$$\int\limits_{ABCD} \mathbf{F}\cdot\mathbf{n}\, dS = \int\limits_{ABCD} (-4)\, dS = -4 \int\limits_{ABCD} dS = (-4)(25) = -100.$$

The integral over the top face involves the exterior unit normal $\mathbf{k}$ instead of $-\mathbf{k}$. Then $\int_{EFGH} \mathbf{F}\cdot\mathbf{n}\, dS = 100$. The sum of these two integrals is 0. Similar computations show that the flux of $\mathbf{F}$ over the entire surface is 0.

The Divergence Theorem says that this flux equals $\int_{\mathcal{R}} \operatorname{div}\mathbf{F}\, dV$, where $\mathcal{R}$ is the solid cube. Now, $\operatorname{div}\mathbf{F} = \partial(2)/\partial x + \partial(3)/\partial y + \partial(4)/\partial z = 0+0+0 = 0$. So the integral of $\operatorname{div}\mathbf{F}$ over $\mathcal{R}$ is 0, verifying the divergence theorem. ◇

Why div **F** is Called the Divergence

Let $\mathbf{F}(x, y, z)$ be the vector field describing the **flow** for a gas. That is, $\mathbf{F}(x, y, z)$ is the product of the density of the gas at (x, y, z) and the velocity vector of the gas there.

The integral $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$ over a closed surface $\mathcal{S}$ represents the tendency of the gas to leave the region $\mathcal{R}$ that $\mathcal{S}$ bounds. If that integral is positive the gas is tending to escape or "diverge". If negative, the net effect is for the amount of gas in $\mathcal{R}$ to increase and become denser.

Let $\rho(x, y, z, t)$ be the density of the gas at time t at the point P, with units mass per unit volume. Then $\int_{\mathcal{R}} \rho\, dV$ is the total mass of gas in $\mathcal{R}$ at a given time. So the rate at which the mass in $\mathcal{R}$ changes is given by the derivative

$$\frac{d}{dt} \int_{\mathcal{R}} \rho\, dV.$$

If ρ is sufficiently well-behaved, mathematicians assure us that we may "differentiate past the integral sign." Then

$$\frac{d}{dt} \int_{\mathcal{R}} \rho\, dV = \int_{\mathcal{R}} \frac{\partial p}{\partial t}\, dV.$$

Therefore

$$\int_{\mathcal{R}} \frac{\partial p}{\partial t}\, dV = \int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$$

since both represent the rate at which gas accumulates in or escapes from $\mathcal{R}$. But, by the Divergence Theorem, $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS = \int_{\mathcal{R}} \nabla \cdot \mathbf{F}\, dV$, and so

$$\int_{\mathcal{R}} \nabla \cdot \mathbf{F}\, dV = \int_{\mathcal{R}} \frac{\partial p}{\partial t}\, dV$$

or,

$$\int_{\mathcal{R}} (\nabla \cdot \mathbf{F} - \frac{\partial p}{\partial t})\, dV = 0. \qquad (18.5.3)$$

From this is it possible to conclude that $\nabla \cdot \mathbf{F} - \frac{\partial p}{\partial t} = 0$?

Recall that the Zero-Integral Principle (see Section 6.3) says: If a continuous function f on an interval $[a, b]$ has the property that $\int_c^d f(x)\, dx = 0$ for every subinterval $[c, d]$ then $f(x) = 0$ on $[a, b]$. A natural extention of the Zero-Integral Principle (see Exercise 27) is:

Zero-Integral Principle in Space

Let $\mathcal{R}$ be a region in space, that is, a set of points in space that is bounded by a surface, and let f be a continuous function on $\mathcal{R}$. Assume that for every region $\mathcal{S}$ in $\mathcal{R}$, $\int_{\mathcal{S}} f(P)\, dS = 0$. Then $f(P) = 0$ for all P in $\mathcal{R}$.

Equation 18.5.3 holds not just for the solid $\mathcal{R}$ but for any solid region within $\mathcal{R}$. By the Zero-Integral Principle in Space, the integrand must be zero thoughout $\mathcal{R}$, and we conclude that

$$\nabla \cdot \mathbf{F} = \frac{\partial p}{\partial t}.$$

This equation tells us that $\operatorname{div} \mathbf{F}$ at a point P represents the rate gas is getting denser or lighter near P. That is why $\operatorname{div} \mathbf{F}$ is called the "divergence of $\mathbf{F}$". Where $\operatorname{div} \mathbf{F}$ is positive, the gas is dissipating. Where $\operatorname{div} \mathbf{F}$ is negative, the gas is collecting.

See Exercise 20 in Section 18.3.

For this reason a vector field for which the divergence is 0 is called **incompressible**. An incompressible is also called "divergence free".

We conclude this section with a proof of the Divergence Theorem.

Proof of the Divergence Theorem

We prove the theorem only for the special case that each line parallel to an axis meets the surface $\mathcal{S}$ in at most two points and $\mathcal{V}$ is convex. We prove the third equation in (18.5.1). The other two are established in the same way.

We wish to show that

$$\int_{\mathcal{V}} R\cos(\gamma)\ dS = \int_{\mathcal{V}} \frac{\partial R}{\partial z}\ dV. \tag{18.5.4}$$

Let $\mathcal{A}$ be the projection of $\mathcal{S}$ on the xy plane. Its description is

$$a \le x \le b, \qquad y_1(x) \le y \le y_2(x).$$

The description of $\mathcal{V}$ is then

$$a \le x \le b, \qquad y_1(x) \le y \le y_2(x), \qquad z_1(x,y) \le z \le z_2(x,y).$$

Then (see Figure 18.5.4)

$$\int_{\mathcal{V}} \frac{\partial R}{\partial z}\ dV = \int_a^b \int_{y_1(x)}^{y_2(x)} \int_{z_1(x,y)}^{z_2(x,y)} \frac{\partial R}{\partial z}\ dz\ dy\ dx. \tag{18.5.5}$$

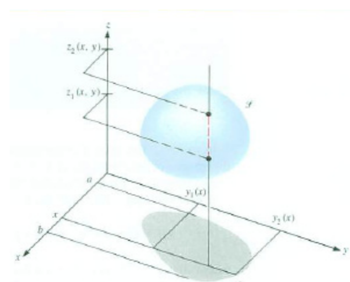


Figure 18.5.4:

The first integration gives

$$\int_{z_1(x,y)}^{z_2(x,y)} \frac{\partial R}{\partial z}\ dz = R(x,y,z_2) - R(x,y,z_1),$$

by the Fundamental Theorem of Calculus. We have, therefore,

$$\int_{\mathcal{V}} \frac{\partial R}{\partial z}\, dV = \int_a^b \int_{y_1(x)}^{y_2(x)} \left(R(x,y,z_2) - R(x,y,z_1)\right)\, dy\, dx,$$

hence

$$\int_{\mathcal{V}} \frac{\partial R}{\partial z}\, dV = \int_{\mathcal{A}} \left(R(x,y,z_2) - R(x,y,z_1)\right)\, dA.$$

This says that, essentially, on the "top half" of $\mathcal{V}$, where $0 < \gamma < \pi/2$, $dA = \cos(\gamma)\ dS$ is positive. And, on the bottom half of $\mathcal{V}$, where $\pi/2 < \gamma < \pi$, $dA = -\cos(\gamma)\ dS$. According to (17.7.1) in Section 17.7, the last integral equals

$$\int_{\mathcal{S}} R(x,y,z)\cos(\gamma)\ dS.$$

Thus

$$\int_{\mathcal{V}} \frac{\partial R}{\partial z}\, dV = \int_{\mathcal{S}} R\cos\gamma\ dS,$$

and (18.5.4) is established.

Similar arguments establish the other two equations in (18.5.1).

Summary

We stated the Divergence Theorem for a single surface and for two surfaces. They enable one to calculate the flux of a vector field $\mathbf{F}$ in terms of an integral of its divergence $\nabla \cdot \mathbf{F}$ over a region. This is especially useful for fields that are incompressible (divergence free). The most famous such field in space is the inverse-square vector field: $\widehat{\mathbf{r}}/r^2$. The flux across a surface of such a field depends on whether its center is inside or outside the surface. Specifically, if the center is at Q and the field is of the form $c\frac{\overrightarrow{QP}}{\|\overrightarrow{QP}\|^3}$, its flux across a surface not enclosing Q is 0. If it encloses Q, its flux is 4π. This is a consequence of the divergence theorem. It also can be explained geometrically, in terms of solid angles.

EXERCISES for Section 18.5 *Key:* R–routine, M–moderate, C–challenging

1.[R] State the Divergence Theorem in symbols.

2.[R] State the Divergence Theorem using only words, not using symbols, such as $\mathbf{F}$, $\nabla\cdot\mathbf{F}$, $\mathbf{n}$, $\mathcal{S}$, or $\mathcal{V}$.

3.[R] Explain why $\nabla\cdot\mathbf{F}$ at a point P can be expressed as a coordinate-free limit.

4.[R] What is the two-surface version of Gauss's theorem?

5.[R] Verify the divergence theorem for $\mathbf{F}(x,y,z) = x\mathbf{i} + y\mathbf{j} + 0\mathbf{k}$ and the surface $x^2 + y^2 + z^2 = 9$.

6.[R] Verify the divergence theorem for the field $\mathbf{F}(x,y,z) = x\mathbf{i}$ and the cube whose vertices are $(0,0,0)$, $(2,0,0)$, $(2,2,0)$, $(0,2,0)$, $(0,0,2)$, $(2,0,2)$, $(2,2,2)$, $(0,2,2)$.

7.[R] Verify the divergence theorem for $\mathbf{F} = 2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$ and the tetrahedron whose four vertices are $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(0,0,1)$.

8.[R] Verify the two-surface version of Gauss's theorem for $\mathbf{F}(x,y,z) = (x^2 + y^2 + z^2)(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$ and the surfaces are the spheres of radii 2 and 3 centered at the origin.

9.[R] Let $\mathbf{F} = 2x\mathbf{i} + 3y\mathbf{j} + (5z + 6x)\mathbf{k}$, and let $\mathbf{G} = (2x + 4z^2)\mathbf{i} + (3y + 5x)\mathbf{j} + 5z\mathbf{k}$. Show that

$$\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS = \int_{\mathcal{S}} \mathbf{G}\cdot\mathbf{n}\, dS,$$

where $\mathcal{S}$ is any surface bounding a region in space.

10.[R] Show that the divergence of $\widehat{\mathbf{r}}/r^2$ is 0. Hint: $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

In Exercises 11 to 18 use the Divergence Theorem.

11.[R] Let $\mathcal{V}$ be the solid region bounded by the xy plane and the paraboloid $z = 9 - x^2 - y^2$. Evaluate $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS$, where $\mathbf{F} = y^3\mathbf{i} + z^3\mathbf{j} + x^3\mathbf{k}$ and $\mathcal{S}$ is the boundary of $\mathcal{V}$.

12.[R] Evaluate $\int_{\mathcal{V}} \nabla\cdot\mathbf{F}\, dV$ for $\mathbf{F} = \sqrt{x^2 + y^2 + z^2}(x\mathbf{i} + y\mathbf{j} + z\mathbf{k})$ and $\mathcal{V}$ the ball of radius 2 and center at $(0,0,0)$.

In Exercises 13 and 14 find $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\, dS$ for the given $\mathbf{F}$ and $\mathcal{S}$.

13.[R] $\mathbf{F} = z\sqrt{x^2 + z^2}\mathbf{i} + (y+3)\mathbf{j} - x\sqrt{x^2 + z^2}\mathbf{k}$ and $\mathcal{S}$ is the boundary of the solid region between $z = x^2 + y^2$ and the plane $z = 4x$.

14.[R] $\mathbf{F} = x\mathbf{i} + (3y + z)\mathbf{j} + (4x + 2z)\mathbf{k}$ and $\mathcal{S}$ is the surface of the cube bounded by the planes $x = 1$, $x = 3$, $y = 2$, $y = 4$, $z = 3$ and $z = 5$.

15.[R] Evaluate $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$, where $\mathbf{F} = 4xz\mathbf{i} - y^2\mathbf{j} + yz\mathbf{k}$ and $\mathcal{S}$ is the surface of the cube bounded by the planes $x = 0$, $x = 1$, $y = 0$, $z = 0$, and $z = 1$, with the face corresponding to $x = 1$ removed.

16.[R] Evaluate $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$, where $\mathbf{F} = x\mathbf{i} + y\mathbf{j} + 2x\mathbf{k}$ and $\mathcal{S}$ is the boundary of the tetrahedron with vertices $(1, 2, 3)$, $(1, 0, 1)$ $(2, 1, 4)$, and $(1, 3, 5)$.

17.[R] Let $\mathcal{S}$ be a surface of area S that bounds a region $\mathcal{V}$ of volume V. Assume that $\|\mathbf{F}(P)\| \leq 5$ for all points P on the surface $\mathcal{S}$. What can be said about $\int_{\mathcal{V}} \nabla \cdot \mathbf{F}\, dV$?

18.[R] Evaluate $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$, where $\mathbf{F} = x^3\mathbf{i} + y^3\mathbf{j} + z^3\mathbf{k}$ and $\mathcal{S}$ is the sphere of radius a and center $(0, 0, 0)$.

In Exercises 19 to 22 evaluate $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$ for $\mathbf{F} = \widehat{\mathbf{r}}/r^2$ and the given surfaces, doing as little calculation as possible.

19.[R] $\mathcal{S}$ is the sphere of radius 2 and center $(5, 3, 1)$.

20.[R] $\mathcal{S}$ is the sphere of radius 3 and center $(1, 0, 1)$.

21.[R] $\mathcal{S}$ is the surface of the box bounded by the planes $x = -1$, $x = 2$, $y = 2$, $y = 3$, $z = -1$, and $z = 6$.

22.[R] $\mathcal{S}$ is the surface of the box bounded by the planes $x = -1$, $x = 2$, $y = -1$, $y = 3$, $z = -1$, and $z = 4$.

23.[M] Assume that the flux of $\mathbf{F}$ across every sphere is 0. Must the flux of $\mathbf{F}$ across the surface of every cube be 0 also?

24.[R] If $\mathbf{F}$ is always tangent to a given surface $\mathcal{S}$ what can be said about the integral of $\nabla \cdot \mathbf{F}$ over the region that $\mathcal{S}$ bounds?

25.[M] Let $\mathbf{F}(\mathbf{r}) = f(r)\widehat{\mathbf{r}}$ be a central vector field in space that has zero divergence. Show that $f(r)$ must have the form $f(r) = a/r^2$ for some constant a. HINT: Consider the flux of $\mathbf{F}$ across the closed surface in Figure 18.5.5.

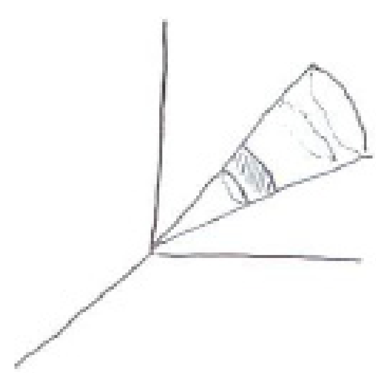


Figure 18.5.5:

26.[M] Let $\mathbf{F}$ be defined everywhere except at the origin and be divergence-free. Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two closed surfaces that enclose the origin. Explain why $\int_{\mathcal{S}_1} \mathbf{F}\cdot\mathbf{n}\; dS = \int_{\mathcal{S}_2} \mathbf{F}\cdot\mathbf{n}\; dS$. (The two surfaces may intersect.)

27.[M] Provide the details for the proof of the Zero-Integral Principle in Space. HINT: You need to consider the two cases when $f > 0$ and $f < 0$.

28.[M] Show that the flux of an inverse-square central field $c\widehat{\mathbf{r}}/r^2$ across any closed surface that bounds a region that does not contain the origin is zero.

29.[C]

(a) Show that the proof in the text of the Divergence Theorem applies to a tetrahedron. HINT: Choose your coordinate system carefully.

(b) Deduce that if the Divergence Theorem holds for a tetrahedron then it holds for any polyhedron. HINT: Each polyhedron can be cut into tetrahedra.

30.[C] In Exercise 25 you were asked to show generally that the only central fields with zero divergence are the inverse square fields. Show this, instead, by computing the divergence of $\mathbf{F}(x,y,z) = f(r)\hat{r}$, where $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$.

31.[C] Let $\mathbf{F}$ be defined everywhere in space except at the origin. Assume that

$$\lim_{\|\mathbf{r}\|\to\infty} \frac{\mathbf{F}(\mathbf{r})}{\|\mathbf{r}\|^2} = \mathbf{0}$$

and that $\mathbf{F}$ is defined everywhere except at the origin, and is divergence free. What can be said about $\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\; dS$, where $\mathcal{S}$ is the sphere of radius 2 centered at the origin?

We proved one-third of the Divergence Theorem. Exercises 32 and 33 concern the other two-thirds.

32.[C] Prove that

$$\int_{\mathcal{S}} Q\cos(\beta)\; dS = \int_{\mathcal{V}} \frac{\partial Q}{\partial y}\; dV.$$

33.[C] Prove that

$$\int_{\mathcal{S}} P\cos(\alpha)\; dS = \int_{\mathcal{V}} \frac{\partial P}{\partial x}\; dV.$$

34.[C] Let f be a scalar function $\mathbf{F}(x, y, z) = f(r)\widehat{\mathbf{r}}$, where $r = \|\mathbf{r}\}$ and $\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Show that if $\nabla \cdot \mathbf{F} = 0$, then $f(r) = c/r^2$ for some constant c.

18.6 Stokes' Theorem

In Section 18.1 we learned that Green's theorem in the xy-plane can be written as

$$\oint_C \mathbf{F} \cdot \, d\mathbf{r} = \int_{\mathcal{R}} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{k} \, dA,$$

where C is counterclockwise and C bounds the region $\mathcal{R}$. The general Stokes' Theorem introduced in this section extends this result to closed curves in space. It asserts that if the closed curve C bounds a surface $\mathcal{S}$, then

$$\oint_C \mathbf{F} \cdot \, d\mathbf{r} = \int_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n} \, dS.$$

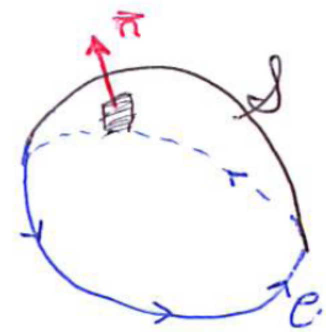


Figure 18.6.1:

As usual, the vector $\mathbf{n}$ is a unit normal to the surface. There are two such normals at each point on the surface. In a moment we describe how to decide which unit normal vector to use. The choice depends on the orientation of C.

In words, Stokes' theorem reads, "The circulation of a vector field around a closed curve is equal to the integral of the normal component of the curl of the field over any surface that the curve bounds."

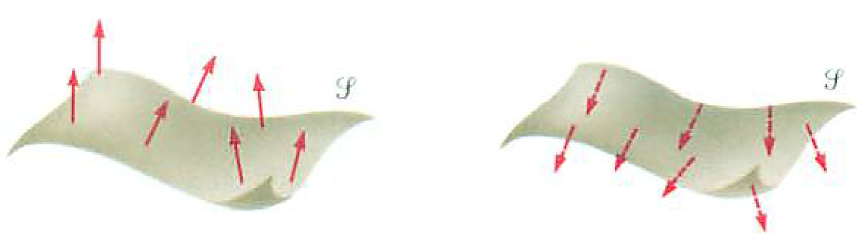
Stokes' published his theorem in 1854 (without proof, for it appeared as a question on a Cambridge University examination). By 1870 it was in common use. It is the most recent of the three major theorems discussed in this chapter, for Green published his theorem in 1828 and Gauss published the divergence theorem in 1839.

Figure 18.6.2:

Choosing the Normal n

In order to state Stokes' theorem precisely, we must describe what kind of surface $\mathcal{S}$ is permitted and which of the two possible normals $\mathbf{n}$ to choose.

The typical surfaces $\mathcal{S}$ that we consider have the property that it is possible to assign, at each point on $\mathcal{S}$, a unit normal $\mathbf{n}$ in a continuous manner. On the surface shown in Figure 18.6.2, there are two ways to do this. They are shown in Figure 18.6.3.



(a) (b)

Figure 18.6.3:

But, for the surface shown in Figure 18.6.4 (a Möbius band), it is impossible to make such a choice. If you start with choice (1) and move the normal continuously along the surface, by the time you return to the initial point on the surface at stage (9), you have the opposite normal. A surface for which a continuous choice *can* be made is called **orientable** or **two-sided**. Stokes' theorem holds for orientable surfaces, which include, for instance, any part of the surface of a convex body, such as a ball, cube or cylinder.

Right-hand rule for choosing **n**.

Consider an orientable surface $\mathcal{S}$, bounded by a parameterized curve C so that the curve is swept out in a definite direction. If the surface is part of a plane, we can simply use the right-hand rule to choose $\mathbf{n}$: The direction of $\mathbf{n}$ should match the thumb of the right hand if the fingers curl in the direction of C and the thumb and palm are perpendicular to the plane. If the surface is not flat, we still use the right-hand rule to choose a normal at points near C. The choice of one normal determines normals throughout the surface. Figure 18.6.5 illustrates the choice of $\mathbf{n}$. For instance, if C is counterclockwise in the xy-plane, this definition picks out the normal $\mathbf{k}$, not $-\mathbf{k}$.

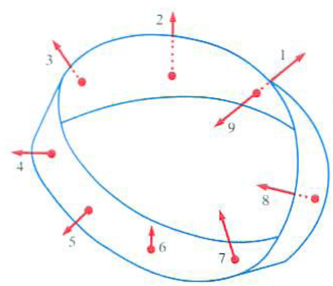
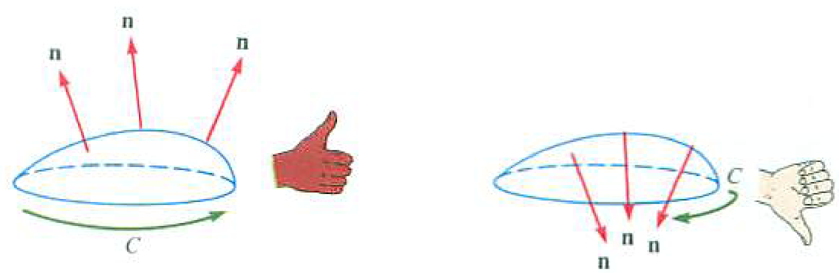


Figure 18.6.4: Follow the choices through all nine stages — there's trouble.



(a) (b)

Figure 18.6.5:

Theorem 18.6.1 (Stokes' theorem)**.** *Let $\mathcal{S}$ be an orientable surface bounded by the parameterized curve C. At each point of $\mathcal{S}$ let $\mathbf{n}$ be the unit normal chosen by the right-hand rule. Let $\mathbf{F}$ be a vector field defined on some region in space including $\mathcal{S}$. Then*

$$\oint_C \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{S}} (\nabla\times\mathbf{F})\cdot\mathbf{n}\, dS. \tag{18.6.1}$$

Some Applications of Stokes' Theorem

Stokes' theorem enables us to replace $\int_{\mathcal{S}}(\mathbf{curl\,F})\cdot\mathbf{n}\, dS$ by a similar integral over a surface that might be simpler than $\mathcal{S}$. That is the substance of the following special case of Stokes' theorem.

Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two surfaces bounded by the same curve C and oriented so that they yield the same orientation on C. Let $\mathbf{F}$ be a vector field defined on both $\mathcal{S}_1$ and $\mathcal{S}_2$. Then

$$\int_{\mathcal{S}_1} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n}\; dS = \int_{\mathcal{S}_2} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n}\; dS \qquad (18.6.2)$$

One way to evaluate some surface integrals is to choose a simpler surface.

The two integrals in (18.6.2) are equal since both equal $\oint_C \mathbf{F} \cdot\; d\mathbf{r}$.

EXAMPLE 1 Let $\mathbf{F} = xe^z\mathbf{i} + (x + xz)\mathbf{j} + 3e^z\mathbf{k}$ and let $\mathcal{S}$ be the top half of the sphere $x^2 + y^2 + z^2 = 1$. Find $\int_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n}\; dS$, where $\mathbf{n}$ is the outward normal. (See Figure 18.6.6.)

SOLUTION Let $\mathcal{S}^*$ be the flat base of the hemisphere. By (18.6.2),

$$\int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\; dS = \int_{\mathcal{S}^*} (\nabla \times \mathbf{F}) \cdot \mathbf{k}\; dS.$$

(On $\mathcal{S}^*$ note that $\mathbf{k}$, not $-\mathbf{k}$, is the correct normal to use.)

A straightforward calculation shows that

$$\nabla \times \mathbf{F} = -x\mathbf{i} + xe^z\mathbf{j} + (z + 1)\mathbf{k},$$

hence $(\nabla \times \mathbf{F}) \cdot \mathbf{k} = z + 1$. On $\mathcal{S}^*$, $z = 0$, so

$$\int_{\mathcal{S}^*} (\nabla \times \mathbf{F}) \cdot \mathbf{k}\; dS = \int_{\mathcal{S}^*} dS = \pi.$$

thus the original integral over $\mathcal{S}$ is also π. ◇

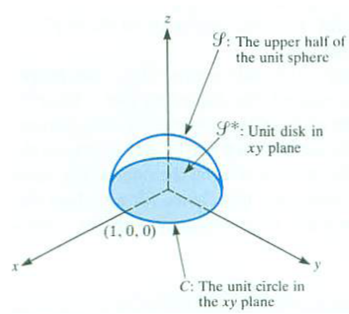


Figure 18.6.6: ARTIST: Add an arrow to indicate the unit circle in the plane is to be oriented counterclockwise. Also add "counterclockwise" to the text label for C.

Just as there are two-curve versions of Green's Theorem and of the Divergence Theorem, there is a two-curve version of Stokes' Theorem.

Stokes' Theorem for a Surface Bounded by Two Closed Curves

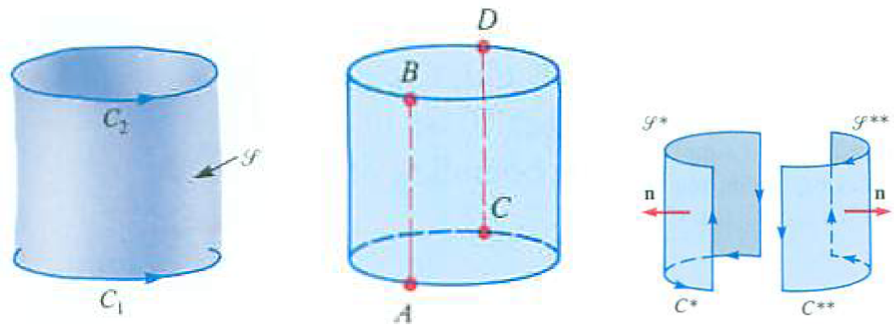
Let $\mathcal{S}$ be an orientable surface whose boundary consists of the two closed curves C_1 and C_2. Give C_1 an orientation. Orient $\mathcal{S}$ consistent with the the right-hand rule, as applied to C_1. Give C_2 the same orientation as C_1. (If C_2 is moved on $\mathcal{S}$ to C_1, the orientations will agree.) Then

$$\oint_{C_1} \mathbf{F} \cdot\; d\mathbf{r} - \oint_{C_2} \mathbf{F} \cdot\; d\mathbf{r} = \int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\; dS. \qquad (18.6.3)$$

Two-curve version of Stokes's Theorem

Proof

Figure 18.6.7(a) shows the typical situation.



(a) (b) (c)

Figure 18.6.7:

The cancellation principle was introduced in Section 18.2.

We will obtain (18.6.3) from Stokes's theorem with the aid of the cancellation principle. Introduce lines AB and CD on $\mathcal{S}$, cutting $\mathcal{S}$ into two surfaces, $\mathcal{S}^*$ and $\mathcal{S}^{**}$. (See Figure 18.6.7(c).) Now apply Stokes's theorem to $\mathcal{S}^*$ and $\mathcal{S}^{**}$. (See Figure 18.6.7(c).)

Let C^* be the curve that bounds $\mathcal{S}^*$, oriented so that where it overlaps C_1 it has the same orientation as C_1. Let C^{**} be the curve that bounds $\mathcal{S}^{**}$, again oriented to match C_1. (See Figure 18.6.7(c).)

By Stokes' theorem,

$$\oint_{C^*} \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{S}^*} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{n}\; dS \tag{18.6.4}$$

and

$$\oint_{C^{**}} \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{S}^{**}} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{n}\; dS. \tag{18.6.5}$$

Adding (18.6.4) and (18.6.5) and using the cancellation principle gives

$$\oint_{C_1} \mathbf{F}\cdot\, d\mathbf{r} - \oint_{C_2} \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{n}\; dS.$$

•

Recall, from Section 18.2, that $\mathbf{F}$ is irrotational when $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$.

In practice, it is most common to apply (18.6.3) when $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$. This is so important we state it explicitly:

Let $\mathbf{F}$ be a field such that $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$. Let C_1 and C_2 be two closed curves that together bound an orientable surface $\mathcal{S}$ on which $\mathbf{F}$ is defined. If C_1 and C_2 are similarly oriented, then

$$\oint_{C_1} \mathbf{F}\cdot\, d\mathbf{r} = \oint_{C_2} \mathbf{F}\cdot\, d\mathbf{r}. \qquad (18.6.6)$$

Equation (18.6.6) follows directly from (18.6.3) since $\int_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{n}\, dS = 0$.

EXAMPLE 2 Assume that $\mathbf{F}$ is irrotational and defined everywhere except on the z-axis. Given that $\oint_{C_1} \mathbf{F}\cdot\, d\mathbf{r} = 3$, find (a) $\oint_{C_2} \mathbf{F}\cdot\, d\mathbf{r}$ and (b) $\oint_{C_3} \mathbf{F}\cdot\, d\mathbf{r}$. (See Figure 18.6.8.)

SOLUTION (a) By (18.6.6), $\oint_{C_2} \mathbf{F}\cdot\, d\mathbf{r} = \oint_{C_1} \mathbf{F}\cdot\, d\mathbf{r} = 3$. (b) By Stokes' theorem, (18.6.1), $\oint_{C_3} \mathbf{F}\cdot\, d\mathbf{r} = 0$. ◇

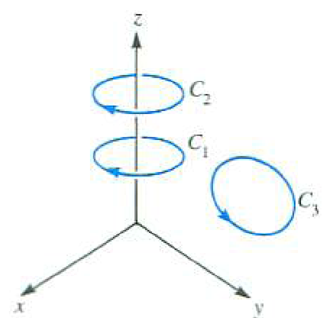


Figure 18.6.8:

Curl and Conservative Fields

In Section 18.1 we learned that if $\mathbf{F} = P\mathbf{i}+Q\mathbf{j}$ is defined on a simply connected region in the xy-plane and if $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$, then $\mathbf{F}$ is conservative. Now that we have Stokes' theorem, this result can be extended to a field $\mathbf{F} = P\mathbf{i}+Q\mathbf{j}+R\mathbf{k}$ defined on a simply connected region in space.

Theorem 18.6.2. *Let* $\mathbf{F}$ *be defined on a simply connected region in space. If* $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$*, then* $\mathbf{F}$ *is conservative.*

Proof

We provide only a sketch of the proof of this result. Let C be a simple closed curve situated in the simply connected region. To avoid topological complexities, we assume that it bounds an orientable surface $\mathcal{S}$. To show that $\oint_C \mathbf{F}\cdot d\mathbf{r} = 0$, we use the same short argument as in Section 18.2:

$$\oint_C \mathbf{F}\cdot\, d\mathbf{r} = \int_{\mathcal{S}} (\nabla\times\mathbf{F})\cdot\mathbf{n}\, dS = \int_{\mathcal{S}} 0\, dS = 0.$$

•

It follows from Theorem 18.6.2 that every central field $\mathbf{F}$ is conservative because a straightforward calculation shows that the curl of a central field is $\mathbf{0}$. (See Exercises 6 and 7 in Section 18.4.) Moreover, $\mathbf{F}$ is defined either throughout space or everywhere except at the center of the field.

Exercise 26 of Section 18.4 presents a purely geometric argument for why a central field is conservative.

In Sections 18.7 and 18.9 we will show how Stokes's theorem is applied in the theory of electromagnetism.

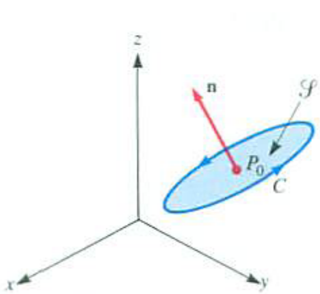


Figure 18.6.9:

Why Curl is Called Curl

Let **F** be a vector field describing the flow of a fluid, as in Section 18.1. Stokes's theorem will give a physical interpretation of **curl F**.

Consider a fixed point P_0 in space. Imagine a *small* circular disk $\mathcal{S}$ with center P_0. Let C be the boundary of $\mathcal{S}$ oriented in such a way that C and $\mathbf{n}$ fit the right-hand rule. (See Figure 18.6.9)

Now examine the two sides of the equation

$$\int_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n}\; dS = \oint_C \mathbf{F} \cdot \mathbf{T}\; ds. \qquad (18.6.7)$$

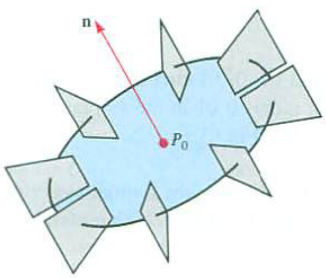


Figure 18.6.10:

The right side of (18.6.7) measures the tendency of the fluid to move along C (rather than, say, perpendicular to it.) Thus $\oint_C \mathbf{F} \cdot \mathbf{T}\; ds$ might be thought of as the "circulation" or "whirling tendency" of the fluid along C. For each tilt of the small disk $\mathcal{S}$ at P_0 — or, equivalently, each choice of unit normal vector $\mathbf{n}$ – the line integral $\oint_C \mathbf{F} \cdot \mathbf{T}\; ds$ measures a corresponding circulation. It records the tendency of a paddle wheel at P_0 with axis along $\mathbf{n}$ to rotate. (See Figure 18.6.10.)

Consider the left side of (18.6.7). If $\mathcal{S}$ is small, the integrand is almost constant and the integral is approximately

$$(\mathbf{curl}\,\mathbf{F})_{P_0} \cdot \mathbf{n} \cdot \text{Area of } \mathcal{S}, \qquad (18.6.8)$$

where $(\mathbf{curl}\,\mathbf{F})_{P_0}$ denotes the curl of **F** evaluated at P_0.

Keeping the center of $\mathcal{S}$ at P_0, vary the vector $\mathbf{n}$ by tilting the disk $\mathcal{S}$. For which choice of $\mathbf{n}$ will (18.6.8) be largest? Answer: For that $\mathbf{n}$ which has the same direction as the fixed vector $(\mathbf{curl}\,\mathbf{F})_{P_0}$. With that choice of $\mathbf{n}$, (18.6.8) becomes

$$\|(\mathbf{curl}\,\mathbf{F})_{P_0}\| \text{ Area of } \mathcal{S}\ .$$

The physical interpretation of curl

Thus a paddle wheel placed in the fluid at P_0 *rotates most quickly when its axis is in the direction of* **curl F** at P_0. The *magnitude* of **curl F** is a measure of *how fast* the paddle wheel can rotate when placed at P_0. Thus **curl F** records the direction and magnitude of maximum circulation at a given point. If **curl F** is **0**, there is no tendency of the fluid to rotate; that is why such vector fields are called irrotational.

A Vector Definition of Curl

In Section 18.1 **curl F** was defined in terms of the partial derivatives of the components of **F**. By Stokes' theorem, **curl F** is related to the circulation, $\oint_C \mathbf{F} \cdot\ d\mathbf{r}$. We exploit this relation to obtain a new view of **curl F**, free of coordinates.

Let P_0 be a point in space and let $\mathbf{n}$ be a unit vector. Consider a small disk $\mathcal{S}_{\mathbf{n}}(a)$, perpendicular to $\mathbf{n}$, whose center is P_0, and which has a radius a. Let $C_{\mathbf{n}}(a)$ be the boundary of $\mathcal{S}_{\mathbf{n}}(a)$, oriented to be compatible with the right-hand rule. Then

$$\int\limits_{\mathcal{S}_{\mathbf{n}}(a)} (\mathbf{curl}\,\mathbf{F}) \cdot \mathbf{n}\ dS = \oint\limits_{C_{\mathbf{n}}(a)} \mathbf{F} \cdot\ d\mathbf{r}.$$

As in our discussion of the physical meaning of curl, we see that

$$(\mathbf{curl}\,\mathbf{F})(P_0) \cdot \mathbf{n} \cdot \text{Area of } \mathcal{S}_{\mathbf{n}}(a) \approx \oint\limits_{C_{\mathbf{n}}(a)} \mathbf{F} \cdot d\mathbf{r},$$

or

$$(\mathbf{curl}\,\mathbf{F})(P_0) \cdot \mathbf{n} \approx \frac{\oint_{C_{\mathbf{n}}(a)} \mathbf{F} \cdot\ d\mathbf{r}}{\text{Area of } \mathcal{S}_{\mathbf{n}}(a)}.$$

Thus

$$(\mathbf{curl}\,\mathbf{F})(P_0) \cdot \mathbf{n} = \lim_{a \to 0} \frac{\oint_{C_{\mathbf{n}}(a)} \mathbf{F} \cdot\ d\mathbf{r}}{\text{Area of } \mathcal{S}_{\mathbf{n}}(a)}. \qquad (18.6.9)$$

Equation (18.6.9) gives meaning to the component of $(\mathbf{curl}\,\mathbf{F})(P_0)$ in any direction $\mathbf{n}$. So the magnitude and direction of **curl F** at P_0 can be described in terms of **F**, without looking at the components of **F**.

The magnitude of $(\mathbf{curl}\,\mathbf{F})_{P_0}$ is the maximum value of

$$\lim_{a \to 0} \frac{\oint_{C_{\mathbf{n}}(a)} \mathbf{F} \cdot\ d\mathbf{r}}{\text{Area of } \mathcal{S}_{\mathbf{n}}(a)}, \qquad (18.6.10)$$

for all unit vectors $\mathbf{n}$.
The direction of $(\mathbf{curl}\,\mathbf{F})_{P_0}$ is given by the vector $\mathbf{n}$ that maximizes the limit (18.6.10).

EXAMPLE 3 Let **F** be a vector field such that at the origin $\mathbf{curl}\,\mathbf{F} = 2\mathbf{i} + 4\mathbf{j} + 4\mathbf{k}$. Estimate $\oint_C \mathbf{F} \cdot\ d\mathbf{r}$ if C encloses a disk of radius 0.01 in the

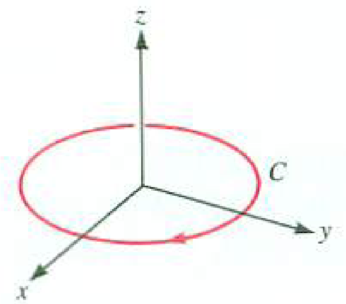


Figure 18.6.11:

xy-plane with center $(0,0,0)$. C is swept out clockwise. (See Figure 18.6.11.)
SOLUTION Let $\mathcal{S}$ be the disk whose border is C. Choose the normal to $\mathcal{S}$ that is consistent with the orientation of C and the right-hand rule. That choice is $-\mathbf{k}$. Thus

$$(\mathbf{curl\,F})\cdot(-\mathbf{k}) \approx \frac{\oint_C \mathbf{F}\cdot\, d\mathbf{r}}{\text{Area of } \mathcal{S}}.$$

The area of $\mathcal{S}$ is $\pi(0.01)^2$ and $\mathbf{curl\,F} = 2\mathbf{i}+3\mathbf{j}+4\mathbf{k}$. Thus

$$(2\mathbf{i}+3\mathbf{j}+4\mathbf{k})\cdot(-\mathbf{k}) \approx \frac{\oint_C \mathbf{F}\cdot\, d\mathbf{r}}{\pi(0.01)^2}.$$

From this it follows that

$$\oint_C \mathbf{F}\cdot\, d\mathbf{r} \approx -4\pi(0.01)^2.$$

◇

In a letter to the mathematician Tait written on November 7, 1870, Maxwell offered some names for $\nabla \times \mathbf{F}$:

> Here are some rough-hewn names. Will you like a good Divinity shape their ends properly so as to make them stick? ...
> The vector part $\nabla \times \mathbf{F}$ I would call the twist of the vector function. Here the word twist has nothing to do with a screw or helix. The word *turn* ... would be better than twist, for twist suggests a screw. Twirl is free from the screw motion and is sufficiently racy. Perhaps it is too dynamical for pure mathematicians, so for Cayley's sake I might say Curl (after the fashion of Scroll.)

His last suggestion, "curl," has stuck.

Proof of Stokes' Theorem

We include this proof because it reviews several basic ideas. The proof uses Green's theorem, the normal to a surface $z = f(x,y)$, and expressing an integral over a surface as an integral over its shadow on a plane. The approach is straightforward. As usual, we begin by expressing the theorem in terms of components. We will assume that the surface $\mathcal{S}$ meets each line parallel to an axis in at most one point. That permits us to project $\mathcal{S}$ onto each coordinate plane in an one-to-one fashion.

To begin we write $\mathbf{F}(x, y, z)$ as $P(x, y, z)\mathbf{i} + Q(x, y, z)\mathbf{j} + R(x, y, z)\mathbf{k}$, or, simply $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$. We will project $\mathcal{S}$ onto the xy-plane, so write the equation for $\mathcal{S}$ as $z - f(x, y) = 0$. A unit normal to $\mathcal{S}$ is

$$\mathbf{n} = \frac{-\frac{\partial f}{\partial x}\mathbf{i} - \frac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}}.$$

(Since the $\mathbf{k}$ component of $\mathbf{n}$ is positive, it is the correct normal, given by the right-hand rule.) Let C^* be the projection of C on the xy-plane, swept out counterclockwise.

See Exercise 9.

A straightforward computation shows that Stokes' theorem, expressed in components, reads

$$\int_C P\ dx + Q\ dy + R\ dz$$
$$= \int_{\mathcal{S}} \frac{\left(\frac{\partial R}{\partial x} - \frac{\partial Q}{\partial z}\right)\left(-\frac{\partial f}{\partial x}\right) - \left(\frac{\partial R}{\partial x} - \frac{\partial P}{\partial z}\right)\left(-\frac{\partial f}{\partial y}\right) + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)(1)}{\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}}\ dS.$$

As expected, this equation reduces to three equations, one for P, one for Q, and one for R.

We will establish the result for P, namely

$$\int_C P\ dx = \int_{\mathcal{S}} \frac{\frac{\partial P}{\partial z}(-\frac{\partial f}{\partial y}) - \frac{\partial P}{\partial y}(1)}{\sqrt{(\frac{\partial f}{\partial x})^2 + (\frac{\partial f}{\partial y})^2 + 1}}\ dS. \qquad (18.6.11)$$

To change the integral over $\mathcal{S}$ to an integral over its projection, $\mathcal{S}^*$, on the xy-plane, we replace dS by $\sqrt{(\partial f/\partial x)^2 + (\partial f/\partial y)^2 + 1}\ dA$. At the same time we project C onto a counterclockwise curve C^*. The square roots cancel leaving us with this equation in the xy-plane:

$$\int_{C*} P(x, y, f(x, y))\ dx = \int_R \left(-\frac{\partial P}{\partial z}\frac{\partial f}{\partial y} - \frac{\partial P}{\partial y}\right)\ dA. \qquad (18.6.12)$$

Finally, we apply Green's theorem to the left side of (18.6.12), and obtain:

$$\int_{C^*} P(x, y, f(x, y))\ dx = \int_{\mathcal{S}^*} -\frac{\partial P(x, y, f(x, y))}{\partial y}\ dA.$$

Be sure you understand each of the four steps in this proof, and why they are valid.

But
$$\frac{\partial P(x,y,f(x,y))}{\partial y} = \frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial f}{\partial y}. \tag{18.6.13}$$
Combining (18.6.12) and (18.6.13) completes the proof of (18.6.11).

In this proof we assumed that the surface $\mathcal{S}$ has a special form, meeting lines parallel to an axis just once. However, more general surfaces, such as the surface of a sphere or a polyhedron can be cut into pieces of the type treated in the proof. Exercise 48 shows why this observation then implies that Stokes' Theorem holds in these cases also.

Summary

Stokes' Theorem relates the circulation of a vector field over a closed curve C to the integral over a surface $\mathcal{S}$ that C bounds. The integrand over the surface is the component of the curl of the field perpendicular to the surface,

$$\int\limits_C \mathbf{F}\cdot\, d\mathbf{r} = \int\limits_{\mathcal{S}} (\mathbf{curl}\,\mathbf{F})\cdot\mathbf{n}\; dS.$$

The normal $\mathbf{n}$ is the normal vector to $\mathcal{S}$ given by the right-hand rule.

EXERCISES for Section 18.6 *Key:* R–routine, M–moderate, C–challenging

1.[C] We dealt only with the component P. What is the analog of (18.6.11) for Q? Prove it. HINT: The steps would parallel the steps used for P.

2.[R] State Stokes' Theorem (symbols permitted).

3.[R] State Stokes' Theorem in words (symbols not permitted).

4.[M] Explain why (18.6.5) holds if $\mathcal{S}_1$ and $\mathcal{S}_2$ together form the boundary surface $\mathcal{S}$ of a solid region R. Use the Divergence Theorem, not Stokes' Theorem.

5.[R] Let $F(r)$ be an antiderivative of $f(r)$. Show that $f(r)\widehat{\mathbf{r}}$ is the gradient of $F(r)$, hence is conservative. NOTE: $f(r)\frac{\mathbf{r}}{r} = f(r)\widehat{\mathbf{r}}$.

6.[M] Show that a central field $f(r)\widehat{\mathbf{r}}$ is conservative by showing that it is irrotational and defined on a simply connected region. HINT: Express $\widehat{\mathbf{r}}$ in terms of x, y and z. NOTE: See also Exercise 47.

7.[R]

(a) Use the fact that a gradient, ∇f, is conservative, to show that its curl is $\mathbf{0}$.

(b) Compute $\nabla \times \nabla f$ in terms of components to show that the curl of a gradient is $\mathbf{0}$.

8.[C] (See also Exercises 5 and 6.)

Sam: The only conservative fields in space that I know are the "inverse square central fields" with centers anywhere I please.

Jane: There are a lot more.

Sam: Oh?

Jane: Just start with any scalar function $f(x, y, z)$ with continuous partial derivation of the first and second orders. Then its gradient will be a conservative field.

Sam: O.K. But I bet there are still more.

Jane: No. I got them all.

Question: Who is right?

Exercises 9 to 14 concern the proof of Stokes' Theorem.

9.[C] Carry out the calculations in the proof that translated Stokes' Theorem into an equation involving the components P, Q, and R.

10.[C] Draw a picture of $\mathcal{S}$, $\mathcal{S}^*$, C and C^* that appear in the proof of Stokes' Theorem.

11.[C] Write the four steps involved in the proof of Stokes' Theorem, giving an explanation for each step.

12.[C] In the proof of Stokes' Theorem we used a normal $\mathbf{n}$. Show that it is the "correct" one, compatible with counterclockwise orientation of C^*.

13.[C]

(a) State Stokes' Theorem for $\int_C Q\ dy$.

(b) Prove Stokes' Theorem for $\int_C Q\ dy$.

(c) State Stokes' Theorem for $\int_C R\ dz$.

(d) Prove Stokes' Theorem for $\int_C R\ dz$.

14.[C] Draw a picture of $\mathcal{S}$, $\mathcal{S}^*$, C and C^* that appear in the proof.

Exercises 15 to 17 prepare you for Exercise 18.

15.[M] Assume that $\mathbf{G}$ is the curl of another vector field $\mathbf{F}$, $\mathbf{G} = \nabla \times \mathbf{F}$. Let $\mathcal{S}$ be a surface that bounds a solid region V. Let C be a closed curve on the surface $\mathcal{S}$ breaking $\mathcal{S}$ into two pieces $\mathcal{S}_1$ and $\mathcal{S}_2$.

16.[M] Using the Divergence Theorem, show that $\int_{\mathcal{S}} \mathbf{G} \cdot \mathbf{n}\ dS = 0$.

17.[M] Using Stokes' Theorem, show that $\int_{\mathcal{S}} \mathbf{G} \cdot \mathbf{n}\ dS = 0$. HINT: Break the integral into integrals over $\mathcal{S}_1$ and $\mathcal{S}_2$.

18.[R] Let $\mathbf{F} = e^{xy}\mathbf{i} + \tan(3yz)\mathbf{j} + 5z\mathbf{k}$ and $\mathcal{S}$ be the tetrahedron whose vertices are $(0,0,0)$, $(1,0,0)$, $(0,1,0)$, and $(0,0,1)$. Let $\mathcal{S}_1$ be the base of $\mathcal{S}$ in the xy-plane and $\mathcal{S}_2$ consist of the other three faces. Find $\int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\ dS$. HINT: think about the preceding two exercises.

19.[R] Assume that $\mathbf{F}$ is defined everywhere except on the z-axis and is irrotational. The curves C_1, C_2, C_3, and C_4 are as shown in Figure 18.6.12. What, if anything, can be said about

$$\oint_{C_1} \mathbf{F} \cdot\ d\mathbf{r}, \qquad \oint_{C_2} \mathbf{F} \cdot\ d\mathbf{r}, \qquad \oint_{C_3} \mathbf{F} \cdot\ d\mathbf{r}, \qquad \text{and} \qquad \oint_{C_4} \mathbf{F} \cdot\ d\mathbf{r}.$$

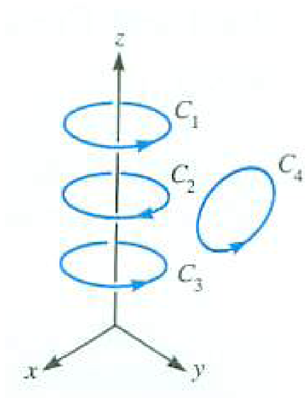


Figure 18.6.12:

In Exercises 20 to 23 verify Stokes' Theorem for the given $\mathbf{F}$ and surface $\mathcal{S}$.
20.[R] $\mathbf{F} = xy^2\mathbf{i} + y^3\mathbf{j} + y^2z\mathbf{k}$; $\mathcal{S}$ is the top half of the sphere $x^2 + y^2 + z^2 = 1$.

21.[R] $\mathbf{F} = y\mathbf{i} + xz\mathbf{j} + x^2\mathbf{k}$; $\mathcal{S}$ is the triangle with vertices $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$.

22.[R] $\mathbf{F} = y^5\mathbf{i} + x^3\mathbf{j} + z^4\mathbf{k}$; $\mathcal{S}$ is the portion of $z = x^2 + y^2$ below the plane $z = 1$.

23.[R] $\mathbf{F} = -y\mathbf{i} + x\mathbf{j} + z\mathbf{k}$, $\mathcal{S}$ is the portion of the cylinder $z = x^2$ inside the cylinder $x^2 + y^2 = 4$.

24.[R] Evaluate as simply as possible $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$, where $\mathbf{F}(x, y, z) = x\mathbf{i} - y\mathbf{j}$ and $\mathcal{S}$ is the surface of the cube bounded by the three coordinate planes and the planes $x = 1$, $y = 1$, $z = 1$, exclusive of the surface in the plane $x = 1$. (Let $\mathbf{n}$ be outward from the cube.)

25.[R] Using Stokes' Theorem, evaluate $\int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS$, where $\mathbf{F} = (x^2 + y - 4)\mathbf{i} + 3xy\mathbf{j} + (2xz + z^2)\mathbf{k}$, and $\mathcal{S}$ is the portion of the surface $z = 4 - (x^2 + y^2)$ above the xy plane. (Let $\mathbf{n}$ be the upward normal.)

In each of Exercises 26 to 29 use Stokes' Theorem to evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ for the given $\mathbf{F}$ and C. In each case assume that C is oriented counterclockwise when viewed from above.
26.[R] $\mathbf{F} = \sin(xy)\mathbf{i}$; C is the intersection of the plane $x + y + z = 1$ and the cylinder $x^2 + y^2 = 1$.

27.[R] $\mathbf{F} = e^x\mathbf{j}$; C is the triangle with vertices $(2, 0, 0)$, $(0, 3, 0)$ and $(0, 0, 4)$.

28.[R] $\mathbf{F} = xy\mathbf{k}$; C is the intersection of the plane $z = y$ with the cylinder $x^2 - 2x + y^2 = 0$.

29.[R] $\mathbf{F} = \cos(x + z)\mathbf{j}$; C is the boundary of the rectangle with vertices $(1, 0, 0)$, $(1, 1, 1)$, $(0, 1, 1)$, and $(0, 0, 0)$.

30.[R] Let $\mathcal{S}_1$ be the top half and $\mathcal{S}_2$ the bottom half of a sphere of radius a in space. Let $\mathbf{F}$ be a vector field defined on the sphere and let $\mathbf{n}$ denote an exterior normal to the sphere. What relation, if any, is there between $\int_{\mathcal{S}_1}(\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS$ and $\int_{\mathcal{S}_2}(\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS$?

31.[R] Let $\mathbf{F}$ be a vector field throughout space such that $\mathbf{F}(P)$ is perpendicular to the curve C at each point P on C, the boundary of a surface $\mathcal{S}$. What can one conclude about

$$\int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS?$$

32.[R] Let C_1 and C_2 be two closed curves in the xy-plane that encircle the origin and are similarly oriented, as in Figure 18.6.13.

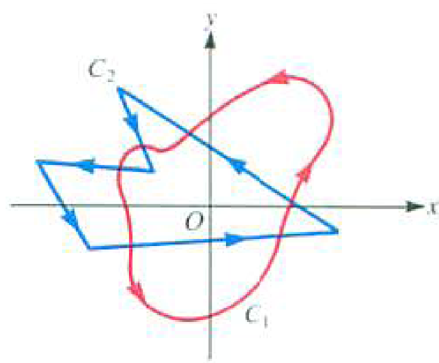


Figure 18.6.13:

Let $\mathbf{F}$ be a vector field defined throughout the plane except at the origin. Assume that $\nabla \times \mathbf{F} = \mathbf{0}$.

(a) Must $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$?

(b) What, it any, relation exists between $\oint_{C_1} \mathbf{F} \cdot d\mathbf{r}$ and $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r}$?

33.[R] Let $\mathbf{F}$ be defined everywhere in space except on the z-axis. Assume also that $\mathbf{F}$ is irrotational, $\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = 3$, and $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 5$. (See Figure 18.6.14.) What if, anything, can be said about

(a) $\oint_{C_3} \mathbf{F} \cdot d\mathbf{r}$,

(b) $\oint_{C_4} \mathbf{F} \cdot d\mathbf{r}$?

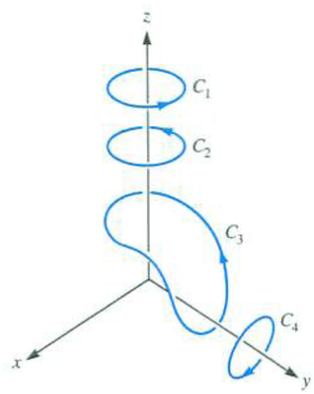


Figure 18.6.14:

34.[R] Which of the following sets are connected? simply connected?

(a) A circle ($x^2 + y^2 = 1$) in the xy-plane

(b) A disk ($x^2 + y^2 \leq 1$) in the xy-plane

(c) The xy-plane from which a circle is removed

(d) The xy-plane from which a disk is removed

(e) The xy-plane from which one point is removed

(f) xyz-space from which one point is removed

(g) xyz-space from which a sphere is removed

(h) xyz-space from which a ball is removed

(i) A solid torus (doughnut)

(j) xyz-space from which a solid torus is removed

(k) A coffee cup with one handle

(l) xyz-space from which a solid doughnut is removed

35.[R] Which central fields have curl $\mathbf{0}$?

36.[R] Let $\mathcal{V}$ be the solid bounded by $z = x + 2$, $x^2 + y^2 = 1$, and $z = 0$. Let $\mathcal{S}_1$ be the portion of the plane $z = x + 2$ that lies within the cylinder $x^2 + y^2 = 1$. Let C be the boundary of $\mathcal{S}_1$, with a counterclockwise orientation (as viewed from above). Let $\mathbf{F} = y\mathbf{i} + xz\mathbf{j} + (x + 2y)\mathbf{k}$. Use Stokes' Theorem for $\mathcal{S}_1$ to evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$.

37.[R] (See Exercise 36.) Let $\mathcal{S}_2$ be the curved surface of $\mathcal{V}$ together with the base of $\mathcal{V}$. Use Stokes' Theorem for $\mathcal{S}_2$ to evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$.

38.[R] Verify Stokes' theorem for the special case when $\mathbf{F}$ has the form ∇f, that is, is a gradient field.

39.[R] Let $\mathbf{F}$ be a vector field defined on the surface $\mathcal{S}$ of a convex solid. Show that $\int_{\mathcal{S}} (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS = 0$

(a) by the Divergence Theorem,

(b) by drawing a closed curve on C on $\mathcal{S}$ and using Stokes' Theorem on the two parts into which C divides $\mathcal{S}$.

40.[R] Evaluate $\oint_C \mathbf{F} \cdot d\mathbf{r}$ as simply as possible if $\mathbf{F}(x, y, z) = (-y\mathbf{i} + x\mathbf{j})/(x^2 + y^2)$ and C is the intersection of the plane $z = 2x + 2y$ and the paraboloid $z = 2x^2 + 3y^2$ oriented counterclockwise as viewed from above.

41.[R] Let $\mathbf{F}(x, y)$ be a vector field defined everywhere in the plane except at the origin. Assume that $\nabla \times \mathbf{F} = \mathbf{0}$. Let C_1 be the circle $x^2 + y^2 = 1$ counterclockwise; let C_2 be the circle $x^2 + y^2 = 4$ clockwise; let C_3 be the circle $(x - 2)^2 + y^2 = 1$ counterclockwise; let C_4 be the circle$(x - 1)^2 + y^2 = 9$ clockwise. Assuming that $\oint_{C_1} \mathbf{F} \cdot d\mathbf{r}$ is 5, evaluate

(a) $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r}$

(b) $\oint_{C_3} \mathbf{F} \cdot d\mathbf{r}$

(c) $\oint_{C_4} \mathbf{F} \cdot d\mathbf{r}$.

42.[M] Let $\mathbf{F}(x, y, z) = \mathbf{r}/\|\mathbf{r}\|^a$, where $\mathbf{r} = \mathbf{i} + u\mathbf{j} + z\mathbf{k}$ and a is a fixed real number.

(a) Show that $\nabla \times \mathbf{F} = \mathbf{0}$.

(b) Show that $\mathbf{F}$ is conservative.

(c) Exhibit a scalar function f such that $\mathbf{F} = \nabla f$.

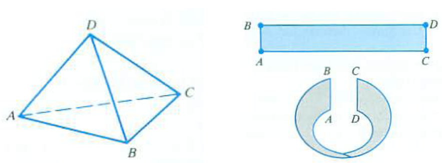
43.[M] Let $\mathbf{F}$ be defined throughout space and have continuous divergence and curl.

(a) For which $\mathbf{F}$ is $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS = 0$ for all spheres $\mathcal{S}$?

(b) For which $\mathbf{F}$ is $\oint_C \mathbf{F} \cdot\, d\mathbf{r} = 0$ for all circles C?

44.[M] Let C be the curve formed by the intersection of the plane $z = x$ and the paraboloid $z = x^2 + y^2$. Orient C to be counterclockwise when viewed from above. Evaluate $\oint_C (xyz\, dx + x^2\, dy + xz\, dz)$.

45.[M] Assume that Stokes' Theorem is true for triangles. Deduce that it then holds for the surface $\mathcal{S}$ in Figure 18.6.15(a), consisting of the three triangles DAB, DBC, DCA, and the curve $ABCA$.



(a) (b)

Figure 18.6.15:

46.[C] A Möbius band can be made by making a half-twist in a narrow rectangular strip, bringing the two ends together, and fastening them with glue or tape. See Figure 18.6.15(b).

(a) Make a Möbius band.

(b) Letting a pencil represent a normal $\mathbf{n}$ to the band, check that the Möbius band is not orientable.

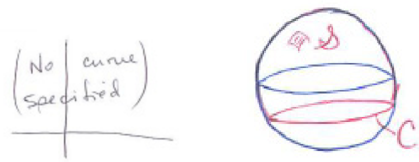
(c) If you form a band by first putting in a full twist (360°), is it orientable?

(d) What happens when you cut the bands in (a) and (c) down the middle? one third of the way from one edge to the other?

47.[C]

(a) Explain why the line integral of a central vector field $f(r)\widehat{\mathbf{r}}$ around the path in Figure 18.6.16(a) is 0.

(b) Deduce from (a) and the coordinate-free view of curl that the curl of a central field is **0**.



(a) (b)

Figure 18.6.16:

48.[C]

(a) The proof of Stokes' Theorem we gave would not apply to surfaces that are more complicated, such as the "top three fourths of a sphere," as shown in Figure 18.6.16(b). However, how could you cut $\mathcal{S}$ into pieces to each of which the proof applies? (Describe them in general terms, in words.)

(b) How could you use (a) to show that Stokes' Theorem holds for C and $\mathcal{S}$ in Figure 18.6.16(b)

49.[M] Sam has a different way to make the choice of **n**.

Sam: I think the book's way of choosing **n** is too complicated.

Jane: OK. How would you do it?

Sam: Glad you asked. First, I would choose a unit normal **n** at one point on the orientable surface.

Jane: That's a good start.

Sam: Then I choose unit normals in a continuous way everywhere on the surface starting at my initial choice.

Jane: And how would you finish?

Sam: My last step is to orient the boundary curve to be compatible with the right-hand rule.

Would this proposal work? If it does, would it agree with the approach in the text.

18.7 Connections Between the Electric Field and $\widehat{\mathbf{r}}/\|\mathbf{r}\|^2$

Even if you are not an engineer or physicist, as someone living in the 21st century you are surrounded by devices that depend on electricity. For that reason we now introduce one of the four equations that explain all of the phenomena of electricity and magnetism. Later in the chapter we will turn to the other three equations, all of which are expressed in terms of vector fields. The chapter concludes with a detailed description of how James Clerk Maxwell, using just these four equations, predicted that light is an electromagnetic phenomenon. Our explanation does not assume any prior knowledge of physics.

The Electric Field Due To a Single Charge

The starting point is some assumptions about the fundamental electrical charges, electrons and protons. An electron has a negative charge and a proton has a positive charge of equal absolute value. Two like charges exert a force of repulsion on each other; unlike charges attract each other.

Let C and P denote the location of charges q and q_0, respectively. Let $\mathbf{r}$ be the vector from C to P, as in Figure 18.7.1, so $r = \|\mathbf{r}\|$ is the distance between the two charges.

If both q and q_0 are protons or both are electrons, the force pushes the charges further apart. If one is a proton and the other is an electron, the force draws them closer. In both cases the magnitude of the force is inversely proportional to r^2, the square of the distance between the charges.

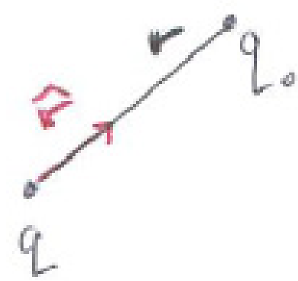


Figure 18.7.1: ARTIST: Please modify labeling to reflect that the charges are located at C and P with charges q and q_0, respectively.

Assume that q is positive, that is, is the charge of a proton. The magnitude of the force it exerts on charge q_0 is proportional to q and also proportional to q_0. It is also inversely proportional to r^2. So, for some constant k, the magnitude of the force is of the form

$$k\frac{q\, q_0}{r^2}.$$

It is directed along the vector $\mathbf{r}$. If q_0 is also positive, it is in the same direction as $\mathbf{r}$. If q_0 is negative, it is in the direction of $-\mathbf{r}$. We can summarize these observations in one vector equation

$$\mathbf{F} = k\frac{q\, q_0}{r^2}\widehat{\mathbf{r}} \tag{18.7.1}$$

where the constant k is positive.

For convenience in later calculations, k is replaced by $1/(4\pi\epsilon_0)$ The value of ϵ_0 depends on the units in which charge, distance, and force are measured. Then (18.7.1) is written

Read ϵ_0 as "epsilon zero" or "epsilon null."

$$\mathbf{F} = \frac{q\, q_0}{4\pi\epsilon_0 r^2}\widehat{\mathbf{r}}.$$

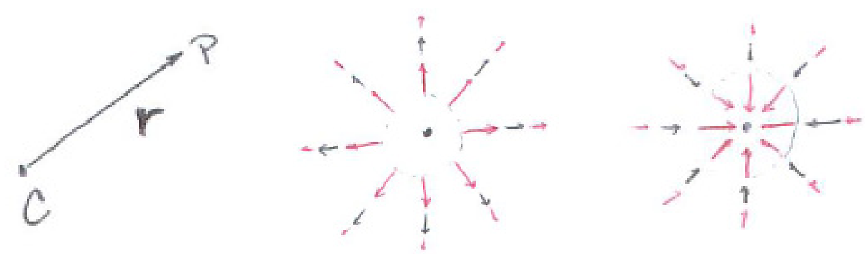
Physicists associate with a charge q a vector field. This field in turn exerts a force on other charges.

Consider a positive charge q at point C.

It "creates" a central inverse-square vector field $\mathbf{E}$ with center at C. It is defined everywhere except at C. Its value at a typical point P is

$$\mathbf{E}(P) = \frac{q\,\widehat{\mathbf{r}}}{4\pi\epsilon_0 r^2}$$

where $\overrightarrow{\mathbf{r}} = \overrightarrow{CP}$, as in Figure 18.7.2.



(a) (b) (c)

Figure 18.7.2:

The value of $\mathbf{E}$ depends only on q and the vector from C to P.

To find the force exerted by charge q on charge q_0 at P just multiply $\mathbf{E}$ by q_0, obtaining

$$\mathbf{F} = q_0\mathbf{E} \tag{18.7.2}$$

The field $\mathbf{E}$, which is a sheer invention, can be calculated in principle by putting a charge q_0 at P, observing the force $\mathbf{F}$ and then dividing $\mathbf{F}$ by q_0. The field $\mathbf{E}$ enables the charge q to "act at a distance" on other charges. It plays the role of a rubber band or a spring.

The Electric Field Due to a Distribution of Charge

Electrons and protons usually do not live in isolation. Instead, charge may be distributed on a line, a curve, a surface or in space.

Imagine a total charge Q occupying a region R in space. The density of the charge varies from point to point. Denote the density at P by $\delta(P)$. Like the density of mass it is defined as a limit as follows. Let $V(r)$ be a small ball of radius r and center at P. Then we have the definition

$$\delta(P) = \lim_{r\to 0^+} \frac{\text{charge in } V(r)}{\text{volume of } V(r)}.$$

The charge in $V(r)$ is approximately the volume of $V(r)$ times $\delta(P)$. We will be interested only in uniform charges, where the density is constant, with the fixed value δ. Thus the charge in a region of volume V is δV.

The field due to a uniform charge Q distributed in a region R is the sum of the fields due to the individual point charges in Q.

To describe that field we need the concept of the integral of a vector field. The definition is similar to the definition of the definite integral in Section 6.2. Let $\mathbf{F}(P)$ be a continuous vector field defined on some solid region R. Break R into regions R_1, R_2, $\dots$, R_n and choose a point P_i in R_i, $1 \le i \le n$. Let the volume of R_i be V_i. The sums $\sum_{i=1}^{n} \mathbf{F}(P_i)V_i$ have a limit as all R_i are chosen smaller and smaller. This limit, denoted $\int_R \mathbf{F}(P)\ dV$ is called the **integral of F over** R. Computationally, this integral can be computed componentwise. For example, if $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + P_3\mathbf{k}$ then $_R\mathbf{F}(P)$ then $\int_R \mathbf{F}\ dV = \int_R F_1\ dV\mathbf{i} + \int_R F_1\ dV\mathbf{j} + \int_R F_1\ dV\mathbf{k}$. Similar definitions hold for vector fields defined on surfaces or curves.

To estimate this field we partition R into small regions R_1, R_2, $\dots R_n$ and choose a point P_i in R_i, $i = 1, 2 \dots, n$. The volume of R_i is V_i. The charge in R_i is δV_i, where δ is the density of the charge. Figure 18.7.3 shows this contribution to the field at a point P.

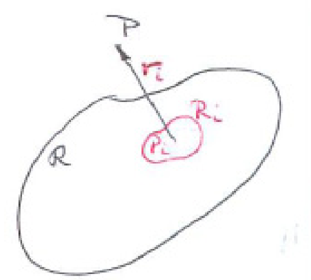


Figure 18.7.3:

Let $\mathbf{r}_i$ be the vector from P_i to P, and $r_i = |\mathbf{r}_i|$. Then the field due to the charge in this small patch R_i is approximately

$$\frac{\delta\ \widehat{\mathbf{r}}_i\ V_i}{4\pi\epsilon_0\ r_i^2}.$$

As an estimate of the field due to Q, we have the sum

$$\sum_{i=1}^{n} \frac{\delta\ \widehat{\mathbf{r}}_i\ V_i}{4\pi\epsilon_0 r_i^2}.$$

Taking limits as all the regions R_i are chosen smaller, we have

$$\mathbf{E}(P) = \text{Field at } P = \int_R \frac{\delta\ \widehat{\mathbf{r}}}{4\pi\epsilon_0 r^2}\ dV$$

Factoring out the constant $\delta/4\pi\epsilon_0$, we have

$$\mathbf{E}(P) = \frac{\delta}{4\pi\epsilon_0} \int_R \frac{\widehat{\mathbf{r}}}{r^2}\ dV$$

That is an integral over a solid region. If the charge is just on a surface S with uniform surface density σ, the field would be given by

$$\mathbf{E}(P) = \frac{\sigma}{4\pi\epsilon_0} \int_S \frac{\widehat{\mathbf{r}}}{r^2}\ dS.$$

If the charge lies on a line or a curve C, with uniform density λ, then

$$\mathbf{E}(P) = \frac{\lambda}{4\pi\epsilon_0} \int_C \frac{\widehat{\mathbf{r}}}{r^2}\, ds.$$

To illustrate the definition we compute one such field value directly. In Example 2 we solve the same problem much more simply.

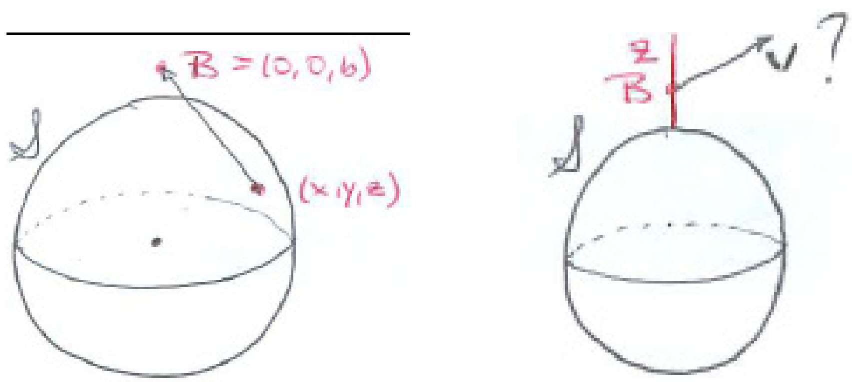
EXAMPLE 1 A charge Q is uniformly distributed on a sphere of radius a, $\mathcal{S}$. Find the electrostatic field $\mathbf{E}$ at a point B a distance $b > a$ from the center of the sphere.

SOLUTION We evaluate

$$\frac{\sigma}{4\pi\epsilon_0} \int_S \frac{\widehat{\mathbf{r}}}{r^2}\, dS. \tag{18.7.3}$$

Note that $\sigma = Q/4\pi a^2$, since the charge is uniform over an area of $4\pi a^2$.

Place a rectangular coordinate system with its origin at the center of the sphere and the z-axis on B, so that $B = (0, 0, b)$, as in Figure 18.7.4(a). Before we start to evaluate an integral, let us use the symmetry of the sphere



(a) (b)

Figure 18.7.4:

to predict something about the vector $\mathbf{E}(B)$. Could it look like the vector $\mathbf{v}$, which is not parallel to the z-axis, as in Figure 18.7.4(b)?

If you spin the sphere around the z-axis, the vector $\mathbf{v}$ would change. But the sphere is unchanged and so is the charge. So $\mathbf{E}(B)$ must be parallel to the z-axis. That means we know its x- and y-components are both 0. So we must find just its z-component, which is $\mathbf{E}(B) \cdot \mathbf{k}$.

Let (x, y, z) be a typical point on the sphere $\mathcal{S}$. Then

$$\mathbf{r} = (0\mathbf{i} + 0\mathbf{j} + b\mathbf{k}) - (x\mathbf{i} + y\mathbf{j} - z\mathbf{k}) = -x\mathbf{i} - y\mathbf{j} + (b - z)\mathbf{k}. \qquad (18.7.4)$$

So

$$\frac{\widehat{\mathbf{r}}}{r^2} = \frac{\mathbf{r}}{r^3} = \frac{-x\mathbf{i} - y\mathbf{j} + (b - z)\mathbf{k}}{(\sqrt{x^2 + y^2 + b^2 - 2bz + z^2})^3} = \frac{-x\mathbf{i} - y\mathbf{j} + (b - z)\mathbf{k}}{(a^2 + b^2 - 2bz)^{3/2}}. \qquad (18.7.5)$$

We need only the z-component of this,

$$\frac{b - z}{(a^2 + b^2 - 2bz)^{3/2}}.$$

The magnitude of $\mathbf{E}(B)$ is therefore

$$\frac{\sigma}{4\pi\epsilon_0} \int_S \frac{b - z}{(a^2 + b^2 - 2bz)^{3/2}}\, dS. \qquad (18.7.6)$$

We evaluate the integral in (18.7.6). To do this, introduce spherical coordinates in the standard position. We have $dS = a^2 \sin(\phi) d\phi\, d\theta$ and $z = a\cos(\phi)$. So (18.7.6) becomes

$$\int_0^{\pi} \int_0^{2\pi} \frac{(b - a\cos\phi)a^2 \sin\phi}{(a^2 + b^2 - 2ab\cos\phi)^{(3/2)}}\, d\theta\, d\phi;$$

which reduces, after the first integration with respect to θ, to

$$2\pi a^2 \int_0^{\pi} \frac{(b - a\cos\phi)\sin\phi\, d\phi}{(a^2 + b^2 - 2ab\cos\phi)^{3/2}} \qquad (18.7.7)$$

Let $u = \cos(\phi)$, hence $du = -\sin(\phi)\, d\phi$. This transforms (18.7.7) into

$$-2\pi a^2 \int_1^{-1} \frac{(b - au)\, du}{(a^2 + b^2 - 2abu)^{3/2}}. \qquad (18.7.8)$$

Then we make a second substitution, $v = a^2 + b^2 - 2abu$.
As you may check, this changes (18.7.8) into

$$\frac{2\pi a^2}{4ab^2} \int_{(b-a)^2}^{(b+a)^2} \frac{v + b^2 - a^2}{v^{3/2}}\, dv \qquad (18.7.9)$$

Write the integrand as the sum of $1/\sqrt{v}$ and $(b^2 - a^2)/v^{3/2}$, and use the Fundamental Theorem of Calculus, to show that (18.7.8) equals $4\pi a^2/b^2$.

Combining this with (18.7.9) shows that

$$\mathbf{E}(B) = \frac{\sigma}{4\pi\epsilon_0}\frac{4\pi a^2}{b^2}\mathbf{k} = \frac{Q}{4\pi\epsilon_0 b^2}\mathbf{k}.$$

◇

The result in this example, $Q/(4\pi\epsilon_0 b^2)\mathbf{k}$ is the same as if all the charge Q were at the center of the sphere. In other words, a uniform charge on a sphere acts on external particles as though the whole charge were placed at its center. This was discovered for the gravitational field by Newton and proved geometrically in his *Principia* of 1687.

Using Flux and Symmetry to Find E

We included Example 1 for two reasons. First, it reviews some integration techniques. Second, it will help you appreciate a much simpler way to find the field $\mathbf{E}$ due to a charge distribution.

Picture a charge Q distributed outside the region bound by a surface S, as in Figure 18.7.5.

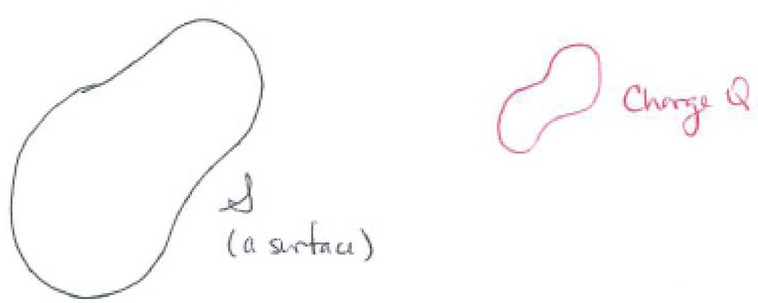


Figure 18.7.5:

The flux of $\mathbf{E}$ associated with a point charge q over a closed surface $\mathcal{S}$ is

$$\int_{\mathcal{S}} \mathbf{E}(P)\cdot\mathbf{n}\,dS = \int_{\mathcal{S}} \frac{\widehat{\mathbf{r}}\cdot\mathbf{n}}{4\pi\epsilon_0 r^2}\,dS = \frac{1}{4\pi\epsilon_0}\int_{\mathcal{S}} \frac{\mathbf{r}\cdot\mathbf{n}}{r^2}\,dS.$$

As we saw in Section 18.5 the integral is 4π when the charge is inside the solid bounded by the surface and 0 if the charge is outside. (See Exercise 28 in that section). Thus the total flux is q/ϵ_0 if the charge is inside and 0 if it is outside.

Consider a charge Q contained wholly within the region bounded by S. We will find the flux of a total charge Q distributed in a solid R inside a surface $\mathcal{S}$. (See Exercise 6 for the case when the charge is outside $\mathcal{S}$.)

Chop the solid R that the charge occupies into n small regions R_1, R_2, ..., R_n. In region R_i select a point P_i. Let the density of charge at P_i be $\delta(P_i)$. Thus the charge in R_i produces a flux of approximately $\delta(P_i)V_i/\epsilon_0$. Consequently

$$\sum_{i=1}^{n} \frac{\delta(P_i)V_i}{\epsilon_0}$$

estimates the flux produced by Q. Taking limits, we see that

$$\text{Flux across } S \text{ produced by } Q = \int_R \frac{\delta(P_i)}{\epsilon_0}\, dV$$

But $\int_R \delta(P_i)\, dV$ is the total charge Q. Thus we have

$$\text{Flux} = \frac{Q}{\epsilon_0}.$$

Thus we have one of the four fundamental equations of electrostatics:

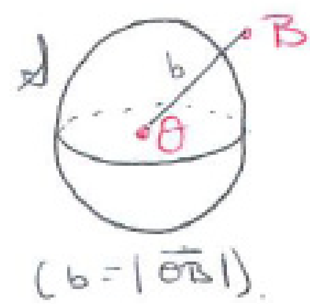


Figure 18.7.6:

Gauss' Law

The flux produced by a distribution of charge across a closed surface is the charge Q in the region bounded by the surface divided by ϵ_0.

The charge outside of S produces no flux across S. (More precisely, the negative flux across S cancels the positive flux.)

Let's illustrate the power of Gauss' Law by applying it to the case in Example 1.

EXAMPLE 2 A charge Q is distributed uniformly on a sphere of radius a. Find the electrostatic field $\mathbf{E}$ at a point B at a distance b from the center of a sphere of radius a, with $b > a$.

SOLUTION We don't need to introduce a coordinate system in Figure 18.7.6. By symmetry, the field at any point P outside the sphere is parallel to the vector $\overrightarrow{CP}$. Moreover, the magnitude of the field is the same for all points at a given distance from the origin C. Call this magnitude, $f(r)$, where r is the distance from C. We want to find $f(b)$.

To do this, imagine another sphere S^*, with center C and radius b, as in Figure 18.7.7.

The flux of $\mathbf{E}$ across S^* is $\int_{S^*} \mathbf{E} \cdot \mathbf{n}\, dS$.

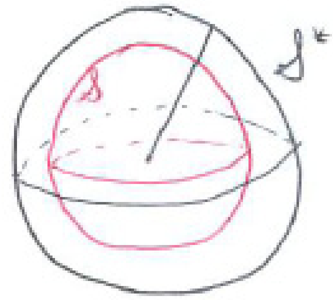


Figure 18.7.7:

But $\mathbf{E} \cdot \mathbf{n}$ is just $f(b)$ since $\mathbf{E}$ and $\mathbf{n}$ are parallel and $\mathbf{E}(P)$ has magnitude $f(b)$ for all points P on S^*. Thus $\int_{S^*} \mathbf{E} \cdot \mathbf{n}\, dS = \int_{S^*} f(b)\, dS = f(b) \int_{S^*}\, dS = f(b) 4\pi b^2$.

By Gauss' Law

$$\frac{Q}{\epsilon_0} = f(b)(4\pi b^2).$$

That tells us that

$$f(b) = \frac{Q}{4\pi\epsilon_0 b^2}.$$

This is the same result as in Example 1, but compare the work in each case. Symmetry and Gauss' Law provide an easy way to find the electrostatic field due to distribution of charge. ◇

The same approach shows that the field $\mathbf{E}$ produced by the spherical charge in Examples 1 and 2 inside the sphere is $\mathbf{0}$. Let $f(r)$ be the magnitude of $\mathbf{E}$ at a distance r from the center of the sphere. For $r > a$, $f(r) = Q/(4\pi\epsilon_0 r^2)$; for $0 < r < a$, $f(r) = 0$. The graph of f is shown in Figure 18.7.8.

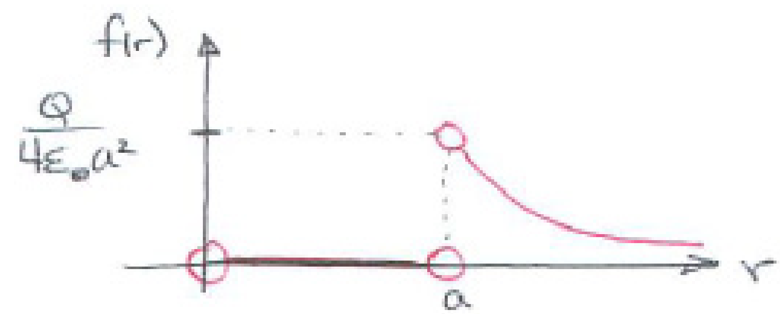


Figure 18.7.8:

If you are curious about $f(a)$ and $f(0)$, see Exercises 8 and 9.

Summary

The field due to a point charge q at a point C is given by the formula $\mathbf{E}(P) = \frac{1}{4\pi\epsilon_0}\frac{q\widehat{\mathbf{r}}}{r^2}$, where $\mathbf{r} = \overrightarrow{OP}$. This field produces a force $q_0\mathbf{E}(P)$ on a charge q_0 located at P.

The field due to a distribution of charge is obtained by an integration over a surface of solid region, depending where this charge is distributed.

We showed that a charge Q outside a surface produces a net flux of zero across the surface. However the flux produced by a charge within the surface is simply Q/ϵ_0. That is Gauss's Law.

We used Gauss's Law to find the field produced by a spherical distribution of charge.

EXERCISES for Section 18.7

Key: R–routine, M–moderate, C–challenging

1.[R] The charge q is positive and produces the electrostatic field $\mathbf{E}$. In what direction does $\mathbf{E}$ point at a charge q_0 that is (a) positive and (b) negative?

2.[R] Fill in the omitted details in the calculation in Exercise 1.

3.[R] Describe to a friend who knows no physics the field $\mathbf{E}$ produced by a point charge q.

4.[R] State Gauss's Law aloud several times.

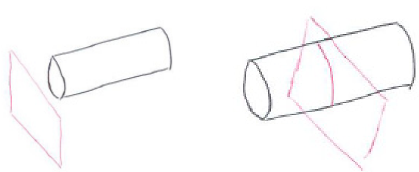
5.[R] Why do you think that the constant k was replaced by $1/4\pi\epsilon_0$. NOTE: Later we will see why it is convenient to have ϵ_0 in the denominator.

6.[R] Show that a charge Q distributed in a solid region R outside a closed surface $\mathcal{S}$ induces zero-flux across $\mathcal{S}$.

7.[R] A charge is distributed uniformly over an infinite plane. For any part of this surface of area A the charge is kA, where k is a constant. Find the field $\mathbf{E}$ due to the charge at any point P not in the plane.

(a) Use symmetry to say as much as you can about it. Be sure to discuss its direction.

(b) Show that the magnitude is constant by applying Gauss's Theorem to a cylinder whose axis is perpendicular to the plane and which does not intersect the plane.



(a) (b)

Figure 18.7.9:

(c) Find the magnitude of $\mathbf{E}$ by applying Gauss's Theorem to the cylinder in Figure 18.7.9(b). Let the area of the circular cross section be A and the area of its curved side be B.

8.[R] Find the field **E** of the charge in Example 1 at a point on the surface of the sphere. Why is Gauss's Law not applicable here? HINT: Let the point be $(0, 0, a)$.

9.[R] Find the field **E** of the charge in Example 1 at the center of the sphere. HINT: Use symmetry, don't integrate.

10.[R] Complete the graph in Figure 18.7.8. That is, fill in the function values corresponding to $r = 0$ and $r = a$.

11.[R] A charge is distributed uniformly along an infinite straight wire. The charge on a section of length l is kl. Find the field **E** due to this charge.

(a) Use symmetry to say as much as you can about the direction and magnitude of **E**.

(b) Find the magnitude by applying Gauss's Law to the cylinder of radius r and height h shown in Figure 18.7.10

(c) Find the force directly by an integral over the line, as in Example 1.

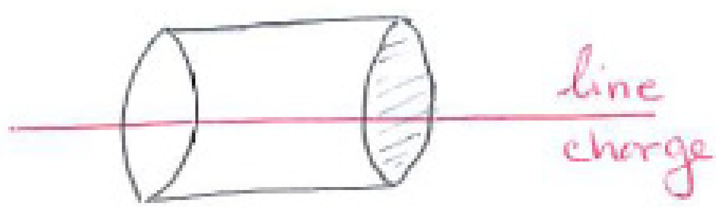
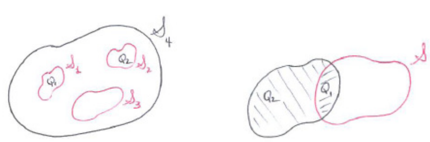


Figure 18.7.10:

12.[R] Figure 18.7.11(a) shows four surfaces. Inside S_1 is a total charge Q_1, and inside S_2 is a total charge Q_2. Find the total flux across each of the four surfaces.



(a) (b)

Figure 18.7.11:

13.[R] Imagine that there is a uniform distribution of charge Q throughout a ball of radius a. Use Gauss' Law to find the electrostatic field $\mathbf{E}$ produced by this charge

(a) at points outside the ball,

(b) at points inside the ball.

14.[R] Let $f(r)$ be the magnitude of the field in Exercise 13 at a distance r from the center of the ball. Graph $f(r)$ for $r \geq 0$.

15.[R] A charge Q lies partly inside a closed surface S and partly outside. Let Q_1 be the amount inside and Q_2 the amount outside, as in Figure 18.7.11(b). What is the flux across $\mathcal{S}$ of the charge Q?

16.[R] In Exercise 11 you found the field $\mathbf{E}$ due to a charge uniformly spread on an infinite line. If the charge density is λ, $\mathbf{E}$ at a point at a distance a from the line is $(\lambda/(2\pi a\epsilon_0))\,\mathbf{j}$.
Now assume that the line occupies only the right half of the x-axis, $[0, \infty)$.

(a) Using the result in Exercise 11, show that the $\mathbf{j}$-component of $\mathbf{E}(0, a)$ is $(\lambda/4\pi a\epsilon_0)\mathbf{j}$.

(b) By integrating over $[0, \infty)$, show that the $\mathbf{i}$-component of $\mathbf{E}$ at $(0, a)$ is $\lambda/(4\pi a\epsilon_0)\mathbf{i}$.

(c) What angle does $\mathbf{E}(0, a)$ make with the y-axis?

(d) Why is Gauss' Law of no use in determining the $\mathbf{i}$-component of $\mathbf{E}$ in this case.

17.[M] We showed that $\mathbf{E}(P) = \frac{\delta}{4\pi\epsilon_0}\int_R \frac{\widehat{\mathbf{r}}}{r^2}\,dV$ if the charge density is constant. Find the corresponding integral for $\mathbf{E}(P)$ when the charge density varies.

18.[C] In Example 1, we used an integral to find the electrostatic field outside a uniformly charged sphere. Carry out similar calculation to find the field inside the sphere. HINT: Is the square root of $(b-a)^2$ still $b-a$?

19.[C] Use the approach in Example 2 to find the electrostatic field *inside* a uniformly charged sphere.

20.[C] Graph the magnitude of the field in Example 1 as a function of the distance from the center of the sphere. This will need the results of Exercises 18 and 19.

21.[C] Find the field **E** in the Exercise 7 by integrating over the whole (infinite) plane. (Do not use Gauss's Theorem.)

18.8 Expressing Vector Functions in Other Coordinate Systems

We have expressed the gradient, divergence, and curl in terms of rectangular coordinates. However, students who apply vector analysis in engineering and physics courses will see functions expressed in polar, cylindrical, and spherical coordinates. This section shows how those expressions are found.

The Gradient in Polar Coordinates

Let $g(r,\theta)$ be a scalar function expressed in polar coordinates. Its gradient has the form $A(r,\theta)\widehat{\mathbf{r}} + B(r,\theta)\widehat{\theta}$, where $\widehat{\mathbf{r}}$ and $\widehat{\theta}$ are the unit vectors shown in Figure 18.8.1. The unit "radial vector" $\widehat{\mathbf{r}}$ points in the direction of increasing r. The unit "tangential vector" $\widehat{\theta}$ points in the direction determined by increasing θ. Note that $\widehat{\theta}$ is tangent to the circle through (r,θ) with center at the pole.

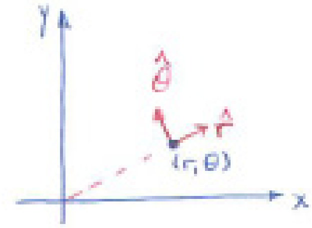


Figure 18.8.1:

Our goal is to find $A(r,\theta)$ and $B(r,\theta)$, which wedenote simply as A and B.

One might guess, in analogy with rectangular coordinates, that $A(r,\theta)$ would be $\partial g/\partial r$ and $B(r,\theta)$ would be $\partial g/\partial\theta$. That guess is part right and part wrong, for we will show that

$$\operatorname{grad} g = \frac{\partial g}{\partial r}\widehat{\mathbf{r}} + \frac{1}{r}\frac{\partial g}{\partial \theta}\widehat{\theta} \tag{18.8.1}$$

We reserve the use of ∇ for rectangular coordinates, and use grad in all other coordinate systems.

Note the appearance of $1/r$ in the $\widehat{\theta}$ component.

One way to obtain (18.8.1) is labor-intensive and not illuminating: express g, $\widehat{\mathbf{r}}$, and $\widehat{\theta}$ in terms of x, y, $\mathbf{i}$, $\mathbf{j}$ and use the formula for gradient in terms of rectangular coordinates, then translate back to polar coordinates. This approach, whose only virtue is that it offers good practice applying the chain rule for partial derivatives, is outlined in Exercises 17 and 18.

We will use a simpler way, that easily generalizes to the cylindrical and spherical coordinates. It exploits the connection between a gradient and directional derivative of g at a point P in the direction $\mathbf{u}$. In particular, it shows why the coefficient $1/r$ appears in (18.8.1).

Recall that if $\mathbf{u}$ is a unit vector, the directional derivative of g in the direction $\mathbf{u}$ is the dot product of $\operatorname{grad} g$ with $\mathbf{u}$:

$$D_{\mathbf{u}}g = \operatorname{grad} g \cdot \mathbf{u}.$$

In particular,

$$D_{\widehat{\mathbf{r}}}g = (A\widehat{\mathbf{r}} + B\widehat{\theta})\cdot\widehat{\mathbf{r}} = A$$

and

$$D_{\widehat{\theta}}g = (A\widehat{\mathbf{r}} + B\widehat{\theta}) \cdot \widehat{\theta} = B.$$

So all we need to do is find $D_{\widehat{\mathbf{r}}}g$ and $D_{\widehat{\theta}}g$.

First,

$$D_{\widehat{\mathbf{r}}}(g) \quad = \quad \lim_{\Delta r \to 0} \frac{g(r + \Delta r, \theta) - g(r, \theta)}{\Delta r} = \frac{\partial g}{\partial r}.$$

So $A(r, \theta) = \partial g/\partial r(r, \theta)$. That explains the expected part of (18.8.1).

Now we will see why B is *not* simply the partial derivation of g with respect to θ.

If we want to estimate a directional derivative at P of g in the direction $\mathbf{u}$ we pick a nearby point Q a distance Δs away in the direction of $\mathbf{u}$ and form the quotient

$$\frac{g(Q) - g(P)}{\Delta s} \tag{18.8.2}$$

Then we take the limit of (18.8.2) as $\Delta s \to 0$.

Now let $\mathbf{u}$ be $\widehat{\theta}$, and let's examine (18.8.2) in the case hwere $P = (r, \theta)$ and $Q = (r, \theta + \Delta\theta)$. The numerator in (18.8.2) is

$$g(r, \theta + \Delta\theta) - g(r, \theta).$$

We draw a picture to find Δs, as in Figure 18.8.2.

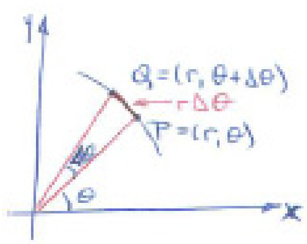


Figure 18.8.2:

The distance between P and Q is *not* $\Delta\theta$. Rather it is approximately $r\Delta\theta$ (when $\Delta\theta$ is small). That tells us that Δs in (18.8.2) is not $\Delta\theta$ but $r\Delta\theta$. Therefore

$$D_{\theta}g \quad = \quad \lim_{\Delta\theta \to 0} \frac{g(r, \theta + \Delta\theta) - g(r, \theta)}{r\Delta\theta} = \frac{1}{r} \lim_{\Delta\theta \to 0} \frac{g(r, \theta + \Delta\theta) - g(r, \theta)}{\Delta\theta} = \frac{1}{r}\frac{\partial g}{\partial\theta}.$$

Note $r\ \Delta\theta$ in the denominator.

That is why there is a $1/r$ in the formula (18.8.1) for the gradient of g. It occurs because a change $\Delta\theta$ in the parameter θ causes a point to move approximately the distance $r\Delta\theta$.

Divergence in Polar Coordinates

The divergence of $\mathbf{F}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ is simply $\partial P/\partial x + \partial Q/\partial y$. But what is the divergence of a vector field described in polar coordinates, $\mathbf{G}(r, \theta) = A(r, \theta)\widehat{\mathbf{r}} + B(r, \theta)\widehat{\theta}$. By now you are on guard, $\nabla \cdot \mathbf{G}$ is *not* the sum of $\partial A/\partial r$ and $\partial B/\partial\theta$).

To find $\nabla \cdot \mathbf{G}$, use the relation between $\nabla \cdot \mathbf{G}$ at $P = (r, \theta)$ and the flux across a small curve C that surrounds P.

$$\nabla \cdot \mathbf{G} = \lim_{\text{length of } C \to 0} \frac{\oint_C \mathbf{G} \cdot \mathbf{n}\ ds}{\text{Area within } C} \tag{18.8.3}$$

Note that (18.8.3) provides a coordinate-free description of divergence in the plane.

We are free to choose the small closed curve C to make it easy to estimate the flux across it. A curve C that corresponds to small changes Δr and $\Delta\theta$ is convenient is shown in Figure 18.8.3. We will use (18.8.3) to find the divergence at $P = (r, \theta)$. Now, P is not inside C; rather it is on C. However, since $\mathbf{G}$ is continuous, $\mathbf{G}(P)$ is the limit of values of $\mathbf{G}$ at points inside, so we may use (18.8.3).

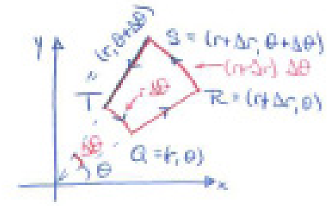


Figure 18.8.3: C is the curve $QRSTQ$

To estimate the flux across C, we estimate the flux across each of the four parts of the curve. Because these sections are short when Δr and $\Delta\theta$ are small, we may estimate the integral over each part by multiplying the value of the integrand at any point of the section (even at an end point) by the length of the section. As usual, $\widehat{\mathbf{n}}$ denotes an exterior unit vector perpendicular to C.

On QR and ST, $B\theta$ contributes to the flux (on RS and TQ it does not since $\mathbf{n} \cdot \theta$ is 0). On QR, θ is parallel to $\mathbf{n}$, as shown in Figure 18.8.4.

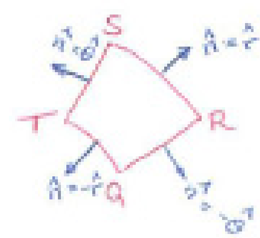


Figure 18.8.4:

However, on ST it points in the opposite direction, $\widehat{\theta} \cdot \widehat{\mathbf{n}}$ is -1. So, across ST, the flux contributed by $B\widehat{\theta}$ is approximately

$$(B\widehat{\theta} \cdot \widehat{\mathbf{n}})\Delta r = -B(r, \theta)\Delta r.$$

(We would get a better estimate by using $B(r + \frac{\Delta r}{2}, \theta)$ but $B(r, \theta)$ is good enough since B is continuous.)

On QR, $\widehat{\theta}$ and $\widehat{\mathbf{n}}$ point in almost the same direction, hence $\theta \cdot \widehat{\mathbf{n}}$ is close to 1 when $\Delta\theta$ is small. So on ST, $B\widehat{\theta}$ contributes approximately $B(r, \theta + \Delta\theta)\Delta r$ to the flux.

All told, the total contribution of $B\theta$ to the flux across C is

$$B(r, \theta + \Delta\theta)\Delta r - B(r, \theta)\Delta r \tag{18.8.4}$$

The contribution of $A\widehat{\mathbf{r}}$ to the flux is negligible on QR and ST because there $\widehat{\mathbf{r}}$ and $\widehat{\mathbf{n}}$ are perpendicular. On TQ, $\widehat{\mathbf{r}}$ and $\widehat{\mathbf{n}}$ point in almost directly opposite directions, hence $\widehat{\mathbf{r}} \cdot \widehat{\mathbf{n}}$ is near -1. The flux of $A\widehat{\mathbf{r}}$ there, is approximately

$$A(r, \theta)(\widehat{\mathbf{r}} \cdot \widehat{\mathbf{n}})r\Delta\theta = -A(r, \theta)r\Delta\theta. \tag{18.8.5}$$

On RS, which has radius $r + \Delta r$, $\widehat{\mathbf{r}}$ and $\widehat{\mathbf{n}}$ are almost identical, hence $\widehat{\mathbf{r}} \cdot \widehat{\mathbf{n}}$ is near 1. The contribution on RS, which has radius $r + \Delta r$ is approximately

$$A(r + \Delta r, \theta)(r + \Delta r)\Delta\theta. \tag{18.8.6}$$

Combining (18.8.4), (18.8.5) and (18.8.6), we see that the limit in (18.8.3) is the sum of two limits:

$$\lim_{\Delta r, \Delta\theta \to 0} \frac{A(r + \Delta r, \theta)(r + \Delta r)\Delta\theta - A(r, \theta)r\Delta\theta}{r\Delta r\Delta\theta} \tag{18.8.7}$$

The area within C is approximately, $r\Delta r\Delta\theta$.

and

$$\lim_{\Delta r,\Delta\theta\to 0}\frac{B(r,\theta+\Delta\theta)\Delta r - B(r,\theta)\Delta r}{r\Delta r\Delta\theta} \tag{18.8.8}$$

The first limit (18.8.7) equals

$$\lim_{\Delta r,\Delta\theta\to 0}\frac{1}{r}\frac{(r+\Delta r)A(r+\Delta r,\Delta\theta) - rA(r,\theta)}{\Delta r},$$

which is

$$\frac{1}{r}\frac{\partial(rA)}{\partial r}.$$

Note that r appears in the coefficient, $1/r$, and also in the function, rA, being differentiated.

The second limit (18.8.8) equals

$$\lim_{\Delta r,\Delta\theta\to 0}\frac{1}{r}\frac{B(r,\theta+\Delta\theta) - B(r,\theta)}{\Delta\theta},$$

hence is

$$\frac{1}{r}\frac{\partial B}{\partial\theta}.$$

Here r appears only once, in the coefficient.

Note the use of div, not $\nabla\cdot$.

All told, we have the desired divergence formula:

$$\operatorname{div}(A\widehat{\mathbf{r}} + B\theta) = \frac{1}{r}\frac{\partial(rA)}{\partial r} + \frac{1}{r}\frac{\partial B}{\partial\theta}. \tag{18.8.9}$$

Curl in the Plane

The curl of $\mathbf{F}(x,y) = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j} + 0\mathbf{k}$, a vector field in the plane, is given by the formula

$$\mathbf{curl}\,\mathbf{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}.$$

What is the formula for the curl when the field is described in polar coordinates: $\mathbf{G}(r,\theta) = A(r,\theta)\widehat{\mathbf{r}} + B(r,\theta)\widehat{\mathbf{n}}$? To find out we will reason as we did with divergence. This time we use

$$(\mathbf{curl}\,\mathbf{G})\cdot\widehat{\mathbf{n}} = \lim_{\text{length of } C\to 0}\frac{\oint_C \mathbf{G}\cdot\mathbf{k}\,ds}{\text{Area bounded by } C}.$$

where C is a closed curve around a fixed point in the (r, θ) plane, and the limit is taken as the length of C approaches 0. The curl is evaluated at a fixed point, which is on or within C.

See (18.6.9) on page 1591.

We compute the circulation of $\mathbf{G} = A\widehat{\mathbf{r}} + B\theta$ around the same curve used in the derivation of divergence in polar coordinates.

On TQ and RS, $A\widehat{\mathbf{r}}$, being perpendicular to the curve, contributes nothing to the circulation of $\mathbf{G}$ around C. On QR it contributes approximately

$$A(r, \theta)(\widehat{\mathbf{r}} \cdot \mathbf{T})\Delta r = A(r, \theta)\Delta r.$$

On ST, since there $\widehat{\mathbf{r}} \cdot \mathbf{T} = -1$, it contributes approximately

$$A(r, \theta + \Delta\theta)(\mathbf{r} \cdot \mathbf{T})\Delta r = -A(r, \theta + \Delta\theta)\Delta r.$$

A similar computation shows that $B\widehat{\theta}$ contributes to the total circulation approximately

$$B(r + \Delta r, \theta)(r + \Delta r)\Delta\theta - B(r, \theta)r\Delta\theta.$$

Therefore $(\nabla \times \mathbf{G} \cdot)\mathbf{k}$ in the sum of two limits:

$$\lim_{\Delta r, \Delta\theta \to 0} \frac{A(r, \theta)\Delta r - A(r, \theta + \Delta\theta)\Delta r}{r\Delta r\Delta\theta} = -\frac{1}{r}\frac{\partial A}{\partial \theta}$$

and

$$\lim_{\Delta r, \Delta\theta \to 0} \frac{B(r + \Delta r, \theta)(r + \Delta r)\Delta\theta - B(r, \theta)r\Delta\theta}{r\Delta r\Delta\theta} = \frac{1}{r}\frac{\partial(rB)}{\partial r}.$$

All told, we have

Note the use of **curl**, not $\nabla\times$.

$$\mathbf{curl}(A\widehat{\mathbf{r}} + B\theta) = \left(-\frac{1}{r}\frac{\partial A}{\partial \theta} + \frac{1}{r}\frac{\partial(rB)}{\partial r}\right)\mathbf{k}. \qquad (18.8.10)$$

EXAMPLE 1 Find the divergence and curl of $\mathbf{F} = r\theta^2\widehat{\mathbf{r}} + r^3\tan(\theta)\theta$.
SOLUTION The calculations are direct applications of (18.8.9) and (18.8.10). First, the divergence:

$$\begin{aligned} \operatorname{div}\mathbf{F} &= \frac{1}{r}\frac{\partial}{\partial r}\left(r \cdot r\theta^2\right) + \frac{1}{r}\frac{\partial}{\partial \theta}\left(r^3\tan(\theta)\right) \\ &= \frac{1}{r}\left(2r\theta^2\right) + \frac{1}{r}\left(r^3\sec^2(\theta)\right) = 2\theta^2 + r^2\sec^2(\theta). \end{aligned}$$

And, the curl:

$$\begin{aligned}\mathbf{curl}\,\mathbf{F} &= \left(-\frac{1}{r}\frac{\partial}{\partial\theta}\left(r\theta^2\right) + \frac{1}{r}\frac{\partial}{\partial r}\left(r\cdot r^3\tan(\theta)\right)\right)\mathbf{k} \\ &= \left(-\frac{1}{r}(2r\theta) + \frac{1}{r}\left(4r^3\tan(\theta)\right)\right)\mathbf{k} = \left(-2\theta + 4r^2\tan(\theta)\right)\mathbf{k}.\end{aligned}$$

◇

Cylindrical Coordinates

In cylindrical coordinates the gradient of $g(r,\theta,z)$ is

$$\text{grad}\, g = \frac{\partial g}{\partial r}\widehat{\mathbf{r}} + \frac{1}{r}\frac{\partial g}{\partial\theta}\widehat{\theta} + \frac{\partial g}{\partial z}\widehat{\mathbf{z}} \qquad (18.8.11)$$

Here $\widehat{\mathbf{z}}$ is the unit vector in the positive z direction, denoted $\mathbf{k}$ in Chapter 14. Note that (18.8.11) differs from (18.8.1) only by the extra term $(\partial g/\partial z)\widehat{\mathbf{z}}$. You can obtain (18.8.11) by computing directional derivatives of g along $\widehat{\mathbf{r}}$, $\widehat{\theta}$, and $\widehat{\mathbf{z}}$. The derivation is similar to the one that gave us the formula for the gradient of $g(r,\theta)$.

The divergence of $\mathbf{G}(r,\theta,z) = A\widehat{\mathbf{r}} + B\widehat{\theta} + C\widehat{\mathbf{z}}$ is given by the formula

$$\text{div}\,\mathbf{G} = \frac{1}{r}\frac{\partial(rA)}{\partial r} + \frac{\partial B}{\partial\theta} + \frac{\partial(rC)}{\partial z}. \qquad (18.8.12)$$

Note that the partial derivatives with respect to r and z are similar in that the factor r is present in both $\partial(rA)/\partial r$ and $\partial(rC)/\partial r$. You can obtain (18.8.12) by using the relation between $\nabla\cdot\mathbf{G}$ and the flux across the small surface determined by small changes Δr, $\Delta\theta$, and Δz.

The curl of $\mathbf{G} = A\widehat{\mathbf{r}} + B\theta + C\widehat{\mathbf{z}}$ is given by a formal determinant:

$$\mathbf{curl}\,\mathbf{G} = \frac{1}{r}\begin{vmatrix}\widehat{\mathbf{r}} & r\widehat{\theta} & \mathbf{k} \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial\theta} & \frac{\partial}{\partial z} \\ A & rB & C\end{vmatrix} \qquad (18.8.13)$$

To obtain this formula consider the circulation around three small closed curves lying in planes perpendicular to $\widehat{\mathbf{r}}$, $\widehat{\theta}$ and $\mathbf{k}$.

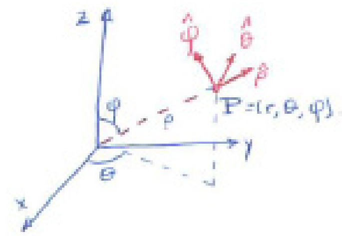


Figure 18.8.5:

Spherical Coordinates

In mathematics texts, spherical coordinates are denoted ρ, ϕ, θ. In physics and engineering a different notation is standard. There ρ is replaced by r, θ is the angle with z-axis, and ϕ plays the role of the mathematicians' θ, switching the roles of ϕ and θ. The formulas we state are in the mathematicians' notation.

The three basic unit vectors for spherical coordinates are denoted ρ, ϕ, θ. For instance, ρ points in the direction of increasing ρ. See Figure 18.8.5. Note that, at the point P, ϕ and θ are tangent to the sphere through P and center at the origin, while ρ is perpendicular to that sphere. Also, any two of ρ, ϕ, θ are perpendicular.

To obtain the formulas for $\nabla \cdot \mathbf{G}$ and $\nabla \times \mathbf{G}$, we would use the region corresponding to small changes $\Delta\rho$, $\Delta\phi$, and $\Delta\theta$, shown in Figure 18.8.6. That computation yields the following formulas:

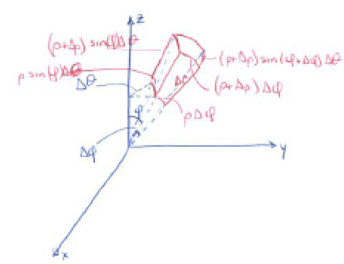


Figure 18.8.6:

If $g(\rho, \phi, \theta)$ is a scalar function,

$$\operatorname{grad} g = \frac{\partial g}{\partial \rho}\rho + \frac{1}{\rho}\frac{\partial g}{\partial \phi}\phi + \frac{1}{\rho\sin(\phi)}\frac{\partial g}{\partial \theta}\theta. \tag{18.8.14}$$

If $\mathbf{G}(\rho, \phi, \theta) = A\rho + B\phi + C\theta$

$$\operatorname{div} \mathbf{G} = \frac{1}{\rho^2}\frac{\partial(\rho^2 A)}{\partial r}) + \frac{1}{\rho\sin(\phi)}\frac{\partial(\sin(\phi)B)}{\partial \phi} + \frac{1}{\rho\sin(\phi)}\frac{\partial C}{\partial \theta}. \tag{18.8.15}$$

and

$$\begin{aligned}\mathbf{curl}\,\mathbf{G} \;&=\; \frac{1}{\rho}\left(\frac{1}{\sin(\phi)}\frac{\partial(\sin(\phi))C)}{\partial \phi} - \frac{1}{\rho\sin(\phi)}\frac{\partial B}{\partial \theta}\right)\rho \\ &\quad + \frac{1}{\rho}\left(\frac{1}{\sin(\phi)}\frac{\partial A}{\partial \theta} - \frac{\partial(\rho C)}{\partial \rho}\right)\phi + \frac{1}{\rho}\left(\frac{\partial(\rho B)}{\partial \rho} - \frac{\partial A}{\partial \phi}\right)\theta\end{aligned}$$

Each of these can be obtained by the method we used for polar coordinates. In each case, keep in mind that the change in ϕ or θ is not the same as the distance the corresponding point moves. However, a change in ρ is the same as the distance the corresponding point moves. For instance, the distance between (ρ, ϕ, θ) and $(\rho, \phi + \Delta\phi, \Delta\theta)$ is approximately $\rho\Delta\phi$ and the distance between (ρ, ϕ, θ) and $(\rho, \phi, \theta + \Delta\theta)$ is approximately $\rho\sin(\phi)\Delta\theta$.

An Application to Rotating Fluids

Consider a fluid rotating in a cylinder, for instance, in a centrifuge. If it rotates as a rigid body, then its velocity at a distance r from the axis of rotation has the form

$$\mathbf{G}(r,\theta) = cr\theta,$$

where c is a positive constant.

Then

$$\mathbf{curl}\,\mathbf{G} = \frac{1}{r}\frac{\partial(cr^2)}{\partial r}\mathbf{k} = 2c\mathbf{k}.$$

The curl is independent of r. That means that an imaginary paddle held with its axis held in a fixed position would rotate at the same rate no matter where it is placed.

Now consider the more general case with

$$\mathbf{G}(r,\theta) = cr^n\theta,$$

and n is an integer. Now

$$\mathbf{curl}\,\mathbf{G} = \frac{1}{r}\frac{\partial(cr^{n+1})}{\partial r}\mathbf{k} = c(n+1)r^{n-1}\mathbf{k}.$$

We just considered the case $n = 1$. If $n > 1$, the curl increases as r increases. The paddle wheel rotates faster if placed farther from the axis of rotation. The direction of rotation is the same as that of the fluid, counterclockwise.

Next consider the case $n = -2$. The speed of the fluid *decreases* as r increases. Now

$$\mathbf{curl}\,\mathbf{G} = c(-2+1)r^{-2-1}\mathbf{k} = -cr^{-3}\mathbf{k}.$$

The minus sign before the coefficient c tells us that the paddle wheel spins *clockwise* even though the fluid rotates counterclockwise. The farther the paddle wheel is from the axis, the slower it rotates.

Summary

We expressed gradient, divergence, and curl in several coordinate systems. Even though the basic unit vectors in each system may change direction from point to point, they remain perpendicular to each other. That simplified the computation of flux and circulation. The formulas are more complicated than those in rectangular coordinates because tha amount a parameter changes is not the same as the distance the corresponding point moves.

EXERCISES for Section 18.8 *Key:* R–routine, M–moderate, C–challenging

In Exercises 1 through 4 find and draw the gradient of the given functions of (r, θ) at $(2, \pi/4)$.

1.[R] r

2.[R] $r^2\theta$

3.[R] $e^{-r}\theta$

4.[R] $r^3\theta^2$

In Exercises 5 through 8 find the divergence of the given function

5.[R] $5\widehat{\mathbf{r}} + r^2\theta\widehat{\theta}$

6.[R] $r^3\theta\widehat{\mathbf{r}} + 3r\theta\widehat{\theta}$

7.[R] $r\widehat{\mathbf{r}} + r^3\widehat{\theta}$

8.[R] $r\sin(\theta)\widehat{\mathbf{r}} + r^2\cos(\theta)\widehat{\theta}$

In Exercises 9 through 12 compute the curl of the given function.

9.[R] $r\widehat{\theta}$

10.[R] $r^3\theta\widehat{\mathbf{r}} + e^r\widehat{\theta}$

11.[R] $r\cos(\theta)\widehat{\mathbf{r}} + r\theta\widehat{\theta}$

12.[R] $1/r^3\widehat{\theta}$

13.[R] Find the directional derivative of $r^2\theta^3$ in the direction

(a) $\widehat{\mathbf{r}}$

(b) $\widehat{\theta}$

(c) $\mathbf{i}$

(d) $\mathbf{j}$

14.[R] What property of rectangular coordinates makes the formulas for gradient, divergence, and curl in those coordinates relatively simple?

15.[R] Estimate the flux of $r\theta vrhat = r^2\theta^3\widehat{\theta}$ around the circle of radius 0.01 with center at $(r, \theta) = (2, \pi/6)$.

16.[R] Estimate the circulation of the field in the preceding exercise around the same circle.

When translating between rectangular and polar coordinates, it may be necessary to express $\widehat{\mathbf{r}}$ and $\widehat{\theta}$ in terms of $\mathbf{i}$ and $\mathbf{j}$ and also $\mathbf{i}$ and $\mathbf{j}$ in terms of $\mathbf{r}$ and $\widehat{\theta}$. Exercise 17 and 18 concern this matter.

17.[R] Let (r,θ) be a point that has rectangular coordinates (x,y).

(a) Show that $\widehat{\mathbf{r}} = \cos(\theta)\mathbf{i} + \sin(\theta)\mathbf{j}$, which equals $x/\sqrt{x^2+y^2}\mathbf{i} + y/\sqrt{x^2+y^2}\mathbf{j} = \frac{x\mathbf{i}+y\mathbf{j}}{\sqrt{x^2+y^2}}$

(b) Show that $\theta = -\sin(\theta)\mathbf{i}+\cos(\theta)\mathbf{j}$, which equals $-y/\sqrt{x^2+y^2}\mathbf{i}+x/\sqrt{x^2+y^2}\mathbf{j}$.

(c) Draw a picture to accompany the calculations done in (a) and (b).

So we have $\widehat{\mathbf{r}}$ and θ in terms of $\mathbf{i}$ and $\mathbf{j}$:

$$\begin{cases} \widehat{\mathbf{r}} &= \frac{x\mathbf{i}+y\mathbf{j}}{\sqrt{x^2+y^2}} \\ \theta &= \frac{-y\mathbf{i}+x\mathbf{j}}{\sqrt{x^2+y^2}} \end{cases} \tag{18.8.16}$$

18.[R] Show that if (x,y) has polar coordinates (r,θ), then

$$\begin{cases} \mathbf{i} = \cos(\theta)\theta\widehat{\mathbf{r}} - \sin(\theta)\theta \\ \mathbf{j} = \sin(\theta)\theta\widehat{\mathbf{r}} + \cos(\theta)\theta \end{cases}$$

by solving the simultaneous equations (18.8.16) in the preceding exercise for $\mathbf{i}$ and $\mathbf{j}$.

In exercises 19 through 22

I. find the gradient of the given function, using the formula for gradient in rectangular coordinates,

II. find it by first expressing the function in polar coordinates and again for gradient in polar coordinates. (18.8.1),

show that the two results agree.

19.[R] x^2+y^2

20.[R] $\sqrt{x^2+y^2}$

21.[R] $3x+2y$

22.[R] $x/\sqrt{x^2+y^2}$

In Exercises 23 through 26

I. find the gradient of the given function, using its formula in polar coordinates, that is (18.8.1),

II. find it by first expressing the function in rectangular coordinates,

III. show that the two results agree.

23.[R] r^2
24.[R] $r^2\cos(\theta)$
25.[R] $r\sin(\theta)$
26.[R] e^r

In Exercise 27 and 28

I. find the divergence of the given vector field in rectangular coordinates,

II. find it by first expressing the function in polar coordinates and using (18.8.9),

III. show that the results agree.

27.[R] $x^2\mathbf{i} + y^2\mathbf{j}$

28.[R] $xy\mathbf{i}$

In Exercises 29 and 30

I. find the curl of the given vector field in rectangular coordinates,

II. find it by first expressing the function in polar coordinates and using (18.8.10),

III. show that the two results agree.

29.[R] $xy\mathbf{i} + x^2y^2\mathbf{j}$
30.[R] $(x/\sqrt{x^2+y^2})\mathbf{i}$

The next two exercises are useful in developing the formula for the gradient in cylindrical and spherical coordinates.
31.[R] Approximately how far is it from the points (r, θ, z) to

(a) $(r + \Delta r, \theta, z)$,

(b) $(r, \theta + \Delta\theta, z)$,

(c) $(r, \theta, z + \Delta z)$.

32.[R] Approximate the distance from the point (ρ, ϕ, θ) to

(a) $(\rho + \Delta\rho, \phi, \theta)$,

(b) $(\rho, \phi + \Delta\phi, \theta)$,

(c) $(\rho, \phi, \theta + \Delta\theta)$.

33.[M] Using the formulas for the gradient of $g(r, \phi, \theta)$, find the directional derivative of g in the direction

(a) $\widehat{\rho}$,

(b) $\widehat{\phi}$,

(c) $\widehat{\theta}$.

34.[M] Using the formulas for the gradient of $g(r, \theta, z)$, find the directional derivative of g in the direction

(a) $\widehat{\mathbf{r}}$,

(b) θ,

(c) $\mathbf{k}$.

35.[M] Without using the formula for the gradient, do Exercise 33.

36.[M] Without using the formula for the gradient, do Exercise 34.

37.[M] Using as few mathematical symbols as you can, state the formula for the divergence of a vector field given relative to $\widehat{\mathbf{r}}$ and θ.

38.[M] Using as few mathematical symbols as you can, state the formula for the curl of a vector field given relative to $\widehat{\mathbf{r}}$ and θ.

39.[M] In the formula for the divergence of $A\widehat{\mathbf{r}} + B\widehat{\theta}$, why do the terms rA and $1/r$ appear in $(1/r)(\partial(rA)/\partial r$ and rA? Explain in detail why $1/r$ appears.

40.[M] Obtain the formula for the gradient in cylindrical coordinates.

41.[M] Obtain the formula for curl in cylindrical coordinates.

42.[M] Obtain the formula for divergence in cylindrical coordinates.

43.[M] Obtain the formula for the gradient in spherical coordinates.

44.[M] Where did we use the fact that $\widehat{r}$ and $\widehat{\theta}$ are perpendicular when developing the expression for divergence in polar coordinates?

45.[M] Obtain the formula for the gradient of $g(r,\theta)$ in polar coordinates by starting with the formula for the gradient of $f(x,y)$ in rectangular coordinates. During the calculations you will have some happy moments as complicated expressions cancel and the identity $\cos^2(\theta)+\sin^2(\theta)=1$ simplifies expressions. (See Exercise 18.8.16.) Assume $g(r,\theta)=f(x,y)$, where $x=r\cos(\theta)$ and $y=r\sin(\theta)$. To express $\nabla f = \partial f/\partial x\mathbf{i}+\partial f/\partial y\mathbf{j}$ in terms of polar coordinates, it is necessary to express $\partial f/\partial x$, $\partial f/\partial y$, $\mathbf{i}$, and $\mathbf{j}$ in terms of partial derivative of $g(r,\theta)$ and $\widehat{\mathbf{r}}$ and θ.

(a) Show that $\partial r/\partial x=\cos(\theta)$, $\partial r/\partial y=\sin(\theta)$, $\partial\theta/\partial x=-(\sin(\theta))/r$, $\partial\theta/\partial y=(\cos\theta)/r$.

(b) Use the chain rule to express $\partial f/\partial x$ and $\partial f/\partial y$ in terms of partial derivatives of $g(r,\theta)$.

(c) Recalling the expression of $\mathbf{i}$ and $\mathbf{j}$ in terms of $\widehat{r}$ and $\widehat{\theta}$ in Exercise 18 obtain the gradient of $g(r,\theta)$ in polar coordinates.

46.[M] In Exercise 26 of Section 18.3, we found the divergence of $\mathbf{F}=r^n\widehat{\mathbf{r}}$ using rectangular coordinates. Find the divergence using polar coordinates formally. NOTE: The second way is much easier.

47.[M] In Exercise 6 of Section 18.6 we used rectangular coordinates to show that an irrotational planar central field is symmetric. Use the formula for curl in polar coordinates to obtain the same result. NOTE: This way is much easier.

48.[M] In Exercise 21 in Section 18.4 we used rectangular coordinates to show that an incompressible symmetric central field in the plane must have the form $\mathbf{F}(\mathbf{r})=(k/r)\widehat{\mathbf{r}}$. Obtain this result using the formula for divergence in polar coordinates.

18.9 Maxwell's Equations

At any point in space there is an electric field **E** and a magnetic field **B**. The electric field is due to charges (electrons and protons) whether stationary or moving. The magnetic field is due to moving charges.

To assure yourself that the magnetic field **B** is everywhere, hold up a pocket compass. The magnetic field, produced within the Earth, makes the needle point north.

All of the electrical phenomena and their applications can be explained by four equations, called **Maxwell's equations**. These equations allow **B** and **E** to vary in time. We state them for the simpler case when B and E are constant: $\partial\mathbf{B}/\partial t = \mathbf{0}$ and $\partial\mathbf{E}/\partial t = \mathbf{0}$. We met the first equation in the previous section. Here is the complete list

I. $\int_S \mathbf{E}\cdot\mathbf{n}\; dS = Q/\epsilon_0$, where S is a surface bounding a spatial region and Q is the change in that region. (Gauss's Law for Electricity)

II. $\oint_C \mathbf{E}\cdot d\mathbf{r} = 0$ for any closed curve C. (Faraday's Law of Induction)

III. $\int_S \mathbf{B}\cdot\mathbf{n}\; dS = 0$ for any surface S that bounds a spatial region. (Gauss's Law for Magnetism)

IV. $\oint_C \mathbf{B}\cdot d\mathbf{r} = \mu_0 \int_S \mathbf{J}\cdot\mathbf{n}\; dS$, where C bounds the surface $\mathcal{S}$ and $\mathbf{J}$ is the electric current flowing through $\mathcal{S}$. (Ampere's Law)

The constants ϵ_0 and μ_0 ("myoo zero") depend on the units used. They will be important in the CIE on Maxwell's Equations.

Each of the four statements about integrals can be translated into information about the behavior of **E** or **B** at each point.

In derivative or "local" form the four principles read:

I'. $\operatorname{div}\mathbf{E} = q/\epsilon_0$, where q is the charge density (Coulomb's Law)

II'. $\mathbf{curl}\,\mathbf{E} = \mathbf{0}$

III'. $\operatorname{div}\mathbf{B} = 0$

IV'. $\mathbf{curl}\,\mathbf{B} = \mu_0\mathbf{J}$

It turns out that $\frac{1}{\mu_0\epsilon_0}$ equals the square of the speed of light. Why that is justified is an astonishing story told in CIE 23.

Going Back and Forth Between "Local" and "Global."

Examples 1 and 2 show that Gauss's Law is equivalent to Coulomb's.

EXAMPLE 1 Obtain Gauss's Law for Electricity (I) from Coulomb's Law (I').
SOLUTION Let $\mathcal{V}$ be the solid region whose boundary is $\mathcal{S}$. Then

$$\begin{aligned} \int_{\mathcal{S}} \mathbf{E}\cdot\mathbf{n}\, dS &= \int_{\mathcal{V}} \nabla\cdot\mathbf{E}\, dV && \text{Divergence Theorem} \\ &= \int_{\mathcal{V}} \frac{q}{\epsilon_0}\, dV && \text{Coulomb's Law} \\ &= \frac{1}{\epsilon_0}\int_{\mathcal{V}} q\, dV = \frac{Q}{\epsilon_0}. \end{aligned}$$

Recall that the total charge in $\mathcal{V}$ is $Q = \int_{\mathcal{V}} q\, dV$. ◇

Does Gauss's law imply Coulomb's law? Example 2 shows that the answer is yes.

EXAMPLE 2 Deduce Coulomb's law (I') from Gauss's law for electricity (I).
SOLUTION Let $\mathcal{V}$ be any spatial region and let $\mathcal{S}$ be its surface. Let Q be the total charge in $\mathcal{V}$. Then

$$\begin{aligned} \frac{Q}{\epsilon_0} &= \int_{\mathcal{S}} \mathbf{E}\cdot\mathbf{n}\, dS && \text{Gauss's law} \\ &= \int_{\mathcal{V}} \nabla\cdot\mathbf{E}\, dV && \text{Divergence Theorem.} \end{aligned}$$

On the other hand,

$$Q = \int_{\mathcal{V}} q\, dV,$$

where q is the charge density. Thus

$$\int_{\mathcal{V}} \frac{q}{\epsilon_0}\, dV = \int_{\mathcal{V}} \nabla\cdot\mathbf{E}\, dV, \qquad \text{or} \qquad \int_{\mathcal{V}} \left(\frac{q}{\epsilon_0} - \nabla\cdot\mathbf{E}\right)\, dV = 0,$$

for all spatial regions. Since the integrand is assumed to be continuous, the "zero-integral principle" tells us that it must be identically 0. That is,

$$\frac{q}{\epsilon_0} - \nabla\cdot\mathbf{E} = 0,$$

which give us Coulomb's law. ◇

EXAMPLE 3 Show that II implies II'. That is, $\oint_C \mathbf{E}\cdot\, d\mathbf{r} = 0$ for closed curves implies $\mathbf{curl}\,\mathbf{E} = \mathbf{0}$.

SOLUTION By Stokes' theorem, for any orientable surface $\mathcal{S}$ bounded by a closed curve,

$$\int\limits_{\mathcal{S}} (\mathbf{curl}\,\mathbf{E}) \cdot \mathbf{n}\, dS = 0$$

The zero-integral principle implies that $(\mathbf{curl}\,\mathbf{E}) \cdot \mathbf{n} = 0$ at each point on the surface. Choosing $\mathcal{S}$ such that $\mathbf{n}$ is parallel to $\mathbf{curl}\,\mathbf{E}$ (if $\mathbf{curl}\,\mathbf{E}$ is not $\mathbf{0}$), implies that the magnitude of $\mathbf{curl}\,\mathbf{E}$ is 0, hence $\mathbf{curl}\,\mathbf{E}$ is $\mathbf{0}$. ◇

Maxwell, by studying the four equations, I', II', III', IV', deduced that electromagnetic waves travel at the speed of light, and therefore light is an electromagnetic phenomenon. In CIE 23 at the end of this chapter we show how he accomplished this, in one of the greatest creative insights in the history of science.

The exercises present the analogy of the four equations in integral form for the general case where $\mathbf{B}$ and $\mathbf{E}$ vary with time. It is here that $\mathbf{B}$ and $\mathbf{E}$ became tangled with each other; both appearing in the same equation. In this generality they are known as Maxwell's equations, in honor of James Clerk Maxwell (1831-1879), who put them in their final form in 1865.

Mathematics and Electricity

Benjamin Franklin, in his book *Experiments and Observations Made in Philadelphia*, published in 1751, made electricity into a science. For his accomplishments, he was elected a Foreign Associate of the French Academy of Sciences, an honor bestowed on no other American for over a century. In 1865, Maxwell completed the theory that Franklin had begun.

At the time that Newton Published his *Principia* on the gravitational field (1687), electricity and magnetism were the subjects of little scientific study. But the experiments of Franklin, Oersted, Henry, Ampère, Faraday, and others in the eighteenth and early nineteenth centuries gradually built up a mass of information subject to mathematical analysis. All the phenomena could be summarized in four equations, which in their final form appeared in Maxwell's *Treatise on Electricity and Magnetism*, published in 1873. For a fuller treatment, see *The Feynman Lectures on Physics*, vol. 2, Addison-Wesley, Reading, Mass., 1964.

Summary

We stated the four equations that describe electrostatic and magnetic fields that do not vary with time. Then we showed how to use the divergence theorem or Stokes' theorem to translate between their global and local forms. The exercises include the four equations in their general form, where $\mathbf{E}$ and $\mathbf{B}$ vary with time.

EXERCISES for Section 18.9 *Key:* R–routine, M–moderate, C–challenging

1.[R] Obtain II from II'.

2.[R] Obtain III' from III.

3.[R] Obtain III from III'.

4.[R] Obtain IV' from IV.

5.[R] Obtain IV from IV'.

In Exercises 6 to 9 use terms such as "circulation," "flux," "current," and "charge density" to express the given equation in words.

6.[R] I

7.[R] II

8.[R] III

9.[R] IV

10.[R] Which of the four laws tell us that an electric current produces a magnetic field?

11.[R] Which of the four laws tells us that a magnetic field produces an electric current?

In this section we assumed that the fields **E** and **B** do not vary in time, that is, $\partial\mathbf{E}/\partial t = \mathbf{0}$ and $\partial\mathbf{B}/\partial t = \mathbf{0}$. The general case, in empty space, where **E** and **B** depend on time, is also described by four equations, which we call $1, 2, 3, 4$. Numbers 1 and 3, do not involve time; they are similar to I' and III'.

1. $\nabla \cdot \mathbf{E} = q/\epsilon_0$
2. $\nabla \times \mathbf{E} = -\partial\mathbf{B}/\partial t$
3. $\nabla \cdot \mathbf{B} = 0$
4. $\nabla \times \mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\frac{d\mathbf{E}}{dt}$

(Here **J** is the current.)

12.[R] Which equation implies that a changing magnetic field creates an electric field?

13.[R] Which equation implies that a changing electrostatic field creates a magnetic field?

14.[R] Show that 2. is equivalent to

$$\oint_C \mathbf{E}\cdot\ dt = -\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot\mathbf{n}\ dS$$

Here, C bounds S. HINT: You may assume that $\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{B}\cdot\mathbf{n}\, dS$ equals $\int_{\mathcal{S}}(\partial B/\partial t)\cdot\mathbf{n}\ dS$.

15.[R] Show that 4. is equivalent to

$$\oint_C \mathbf{B}\cdot\ d\mathbf{r} = \mu_0\int_{\mathcal{S}} \mathbf{J}\cdot\mathbf{n}\ dS + \mu_0\epsilon_0\frac{\partial}{\partial t}\int_{\mathcal{S}} \mathbf{E}\cdot\mathbf{n}\ dS$$

(The circulation of **B** is related to the total current through the surface $\mathcal{S}$ that C bounds and to the rate at which the flux of **E** through $\mathcal{S}$ changes.)

18.S Chapter Summary

The first six sections developed three theorems: Green's Theorem, Gauss' Theorem (also called the Divergence Theorem), and Stokes' Theorem. The final four sections applied them to geometry and to physics and to expressing various functions in terms of non-rectangular coordinate systems. These four sections offer a way to deepen your understanding of the first six.

Name	Mathematical Expression	Physical Description
Green's Theorem	$\oint_C \mathbf{F}\cdot\mathbf{n}\,ds = \int_{\mathcal{R}} \nabla\cdot\mathbf{F}\,dA$	flux of $\mathbf{F}$ across C
	$\oint_C (-Qdx + Pdy) = \int_{\mathcal{R}} \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right)\,dA$	differential form
	$\oint_C \mathbf{F}\cdot\mathbf{T}\,ds = \oint_C \mathbf{F}\cdot d\mathbf{r} = \int_{\mathcal{R}} (\nabla\times\mathbf{F})\cdot\mathbf{k}\,dA$	circulation of $\mathbf{F}$ around C
	$\oint_C (Pdx + Qdy) = \int_{\mathbf{R}} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\,dA$	
Gauss' Theorem (Divergence Theorem)	$\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{N}\,dS = \int_R \nabla\cdot\mathbf{F}\,dV$	
Stokes' Theorem	$\oint_C \mathbf{F}\cdot\mathbf{T}\,ds = \int_{\mathcal{S}} (\nabla\times\mathbf{F})\cdot\mathbf{n}\,dS$ ($\mathcal{S}$ is a surface bounded by C with $\mathbf{n}$ compatible by orientation of C)	

Green's Theorem can be viewed as the planar version of either the Divergence Theorem or Stokes' Theorem.

Though div $\mathbf{F}$ and **curl** $\mathbf{F}$ were defined in terms of rectangular coordinates, they also have a meaning that is independent of any coordinates. For instance, if $\mathbf{F}$ is a vector field in space, the divergence of $\mathbf{F}$ at a point multiplied by the volume of a small region containing that point approximates the flux of $\mathbf{F}$ across the surface of that small region. More precisely,

$$\text{div}\,\mathbf{F} \text{ at } P \text{ equals the limit of } \frac{\int_{\mathcal{S}} \mathbf{F}\cdot\mathbf{n}\,ds}{\text{volume of } \mathcal{R}} \text{ as the diameter of } \mathcal{R} \text{ approaches } 0$$

The **curl** of $\mathbf{F}$ at P is a vector, so it's a bit harder to describe physically. Let $\mathbf{n}$ be a unit vector and C a small curve that lies in a plane through P, is perpendicular to $\mathbf{n}$, and surrounds P. Then the scalar component of **curl** $\mathbf{F}$ at P is the direction $\mathbf{n}$ multiplied by the area of the surface bounded by C gives the circulation of $\mathbf{F}$ along C.

A field whose curl is $\mathbf{0}$ is called irrotational. A field whose divergence is 0 is called incompressible (or divergence-free).

Of particular interest are conservative fields. A field $\mathbf{F}$ is conservative if its circulation on a curve depends only on the endpoints of the curve. If the domain of $\mathbf{F}$ is simply connected, $\mathbf{F}$ is conservative if and only if its curl is $\mathbf{0}$. A conservative field is expressible as the gradient of a scalar function.

Among the conservative fields are the symmetric central fields. If, in addition, they are divergence-free, they take a very special form that depends on the dimension of the problem.

Geometry	General Form of Divergence-Free Symmetric Central Fields	Description
$\mathbf{R}^2$ (plane)	$c\frac{\widehat{\mathbf{r}}}{r}$	inverse radial
$\mathbf{R}^3$ (space)	$c\frac{\widehat{\mathbf{r}}}{r^2}$	inverse square radial
$\mathbf{R}^n$	$c\frac{\widehat{\mathbf{r}}}{r^{n-1}}$	

In the case where $\mathbf{curl}\,\mathbf{F} = \mathbf{0}$ one can replace an integral $\int_A^B \mathbf{F} \cdot d\mathbf{r}$ by an integral over another curve joining A and B. This is most beneficial when the new line integral is easier to evaluate than the original one. Similarly, in a region where $\nabla \cdot \mathbf{F} = 0$ we can replace an integral $\int_{\mathcal{S}} \mathbf{F} \cdot \mathbf{n}\, dS$over the surface $\mathcal{S}$ with a more convenient integral over a different surface.

In applications in space the most important field is the inverse square central field, $\mathbf{F} = \frac{\widehat{\mathbf{r}}}{r^2}$. The flux of this field over a closed surface that does not enclose the origin is 0, but its flux over a surface that encloses the origin is 4π. If one thinks in terms of steradians, it is clear why the second integral is 4π: the flux of $\widehat{\mathbf{r}}/r^2$ also measures the solid angle subtended by a surface. Also, the first case becomes clear when one distinguishes the two parts of the surface where $\mathbf{n} \cdot \mathbf{r}$ is positive and where it is negative.

EXERCISES for 18.S *Key:* R–routine, M–moderate, C–challenging

1.[R] Match the vector fields given in mathematical symbols (a.-e.) with the written description (1.-5.)

a.	$\mathbf{F}(\mathbf{r})$	1.	an inverse cube central field
b.	$f(\mathbf{r})\widehat{\mathbf{r}}$	2.	a central field (center at origin)
c.	$f(r)\widehat{\mathbf{r}}$	3.	an arbitrary vector field
d.	$\widehat{\mathbf{r}}/r^2$	4.	a symmetric central field (center at origin)
e.	$\mathbf{r}/r^3$	5.	an inverse square central field

NOTE: There is not a one-to-one relation between the two columns.

2.[R] Use Green's theorem to evaluate $\oint_C (xy\ dx + e^x\ dy)$, where C is the curve that goes from $(0,0)$ to $(2,0)$ on the x-axis and returns from $(2,0)$ to $(0,0)$ on the parabola $y = 2x - x^2$.

3.[R] A curve C bounds a region $\mathcal{R}$ of area A.

(a) If $\oint_C \mathbf{F} \cdot d\mathbf{r} = -2$, estimate $\nabla \times \mathbf{F}$ at points in R.

(b) Would you use $\odot$ or $\oplus$ to indicate the curl?

4.[R] A curve C bounds a region $\mathcal{R}$ of area A.

(a) If $\oint_C \mathbf{F} \cdot \mathbf{n}\, ds = -2$, estimate $\nabla \cdot \mathbf{F}$ at points in $\mathcal{R}$.

(b) How did you decide whether $\nabla \cdot \mathbf{F}$ is positive or negative?

5.[R] A field $\mathbf{F}$ is called **uniform** if all its vectors are the same. Let $\mathbf{F}(x, y, z) = 3i$.

(a) Find the flux of $\mathbf{F}$ across each of the six faces of the cube in Figure 18.S.1 of side 3.

(b) Find the total flux of $\mathbf{F}$ across the surface of the box.

(c) Verify the divergence theorem for this $\mathbf{F}$.

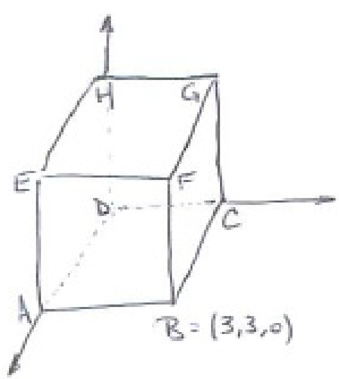


Figure 18.S.1:

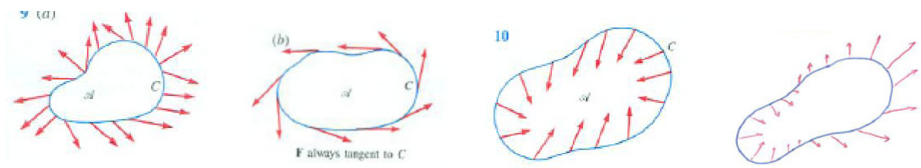
6.[R] Let $\mathbf{F}$ be the uniform field $\mathbf{F}(x, y, z) = 2\mathbf{i} + 3\mathbf{j} + 0\mathbf{k}$. Repeat Exercise 5 Carry out the preceding exercise for this field.

7.[R] See Exercise 8. Suppose you placed the point at which $\mathbf{E}$ is evaluated at $(a, 0, 0)$ instead of at $(0, 0, a)$.

(a) What integral in spherical coordinates arises?

(b) Would you like to evaluate it?

In Exercises 8 to 11, **F** is defined on the whole plane but indicated only at points on a curve C bounding a region $\mathcal{R}$. What can be said about $\int_{\mathcal{R}} \nabla \cdot \mathbf{F}\, dA$ in each case?



(a) (b) (c) (d)

Figure 18.S.2:

8.[R] See Figure 18.S.2(a).

9.[R] See Figure 18.S.2(b).

10.[R] See Figure 18.S.2(c).

11.[R] See Figure 18.S.2(d).

Exercises 12 to 15, **F** concern the same **F** as in Exercises 8 to 11. What can be said about $\int_{\mathcal{S}} \nabla \times \mathbf{F}\, dA$ in each case?

12.[R] See Figure 18.S.2(a).

13.[R] See Figure 18.S.2(b).

14.[R] See Figure 18.S.2(c).

15.[R] See Figure 18.S.2(d).

16.[R] Let C be the circle of radius 1 with center $(0,0)$.

(a) What does Green's theorem say about the line integral

$$\oint_C \left((x^2 - y^3)\, dx + (y^2 + x^3)\, dy\right)?$$

(b) Use Green's theorem to evaluate the integral in (a).

(c) Evaluate the integral in (a) directly.

17.[M] Let $\mathbf{F}(x,y) = (x+y)\mathbf{i} + x^2\mathbf{j}$ and let C be the counterclockwise path around the triangle whose vertices are $(0,0)$, $(1,1)$, and $(-1,1)$.

(a) Use the planar divergence theorem to evaluate $\int_C \mathbf{F} \cdot \mathbf{n} ds$, where $\mathbf{n}$ is the outward unit normal.

(b) Evaluate the line integral in (a) directly.

18.[M] Let b and c be positive numbers and $\mathcal{S}$ the "infinite rectangle" parallel to the xy-plane, consisting of the points (x, y, c) such that $0 \leq x \leq b$ and $b \geq 0$.

(a) If b were replaced by ∞, what is the solid angle $\mathcal{S}$ subtends at the origin? HINT: No integration is needed.

(b) Find the solid angle subtended by $\mathcal{S}$ when b is finite. HINT: See Exercise 93.

(c) Is the limit of your answer in (b) as $b \to \infty$ the same as your answer in (a)? HINT: It should be!

19.[M] Look back at the Fundamental Theorem of Calculus (Section 6.4), Green's Theorem (Section 18.2), the Divergence Theorem (Section 18.6), and Stokes' Theorem (Section 18.4). What single theme runs through all of them?

Calculus is Everywhere # 23
How Maxwell Did It

In a letter to his cousin, Charles Cay, dated January 5, 1965, Maxwell wrote:

> I have also a paper afloat containing an electromagnetic theory of light, which, till I am convinced to the contrary, I hold to be great guns. [Everitt, F., *James Clerk Maxwell: a force for physics*, Physics World, Dec 2006, `http://physicsworld.com/cws/article/print/26527`]

It indeed was "great guns," for out of his theory has come countless inventions, such as television, cell phones, and remote garage door openers. In a dazzling feat of imagination, Maxwell predicted that electrical phenomena create waves, that light is one such phenomenon, and that the waves travel at the speed of light, in a vacuum.

In this section we will see how those predictions came out of the four equations (I'), (II'), (III'), and (IV') in Section 18.9.

First, we take a closer look at the dimensions of the constants ε_0 and μ_0 that appear in (IV'),

$$\frac{1}{\mu_0\varepsilon_0}\nabla\times\mathbf{B}=\frac{\mathcal{J}}{\varepsilon_0}.$$

The constant ε_0 makes its appearance in the equation

$$\text{Force}=F=\frac{1}{4\pi\varepsilon_0}\frac{qq_0}{r^2}. \tag{C.23.1}$$

Since the force F is "mass times acceleration" its dimensions are

$$\text{mass}\cdot\frac{\text{length}}{\text{time}^2},$$

or, in symbols

$$m\frac{L}{T^2}.$$

The number 4π is a pure number, without any physical dimension.

The quantity qq_0 has the dimensions of "charge squared," q^2, and R^2 has dimensions L^2, where L denotes length.

Solving (C.23.1) for ε_0, we find the dimensions of ε_0. Since

$$\varepsilon_0=\frac{q^2}{4\pi Fr^2},$$

its dimensions are

$$\left(\frac{T^2}{mL}\right)\left(\frac{q^2}{L^2}\right) = \frac{T^2q^2}{mL^3}.$$

To figure out the dimensions of μ_0, we will use its appearance in calculating the force between two wires of length L each carrying a current I in the same direction and separated by a distance R. (Each generates a magnetic field that draws the other towards it.) The equation that describes that force is

$$\mu_0 = \frac{2\pi RF}{I^2L}.$$

Since R has the dimensions of length L and F has dimensions mL/T^2, the numerator has dimensions mL^2/T^2. The current I is "charge q per second," so I^2 has dimensions q^2/T^2. The dimension of the denominator is, therefore,

$$\frac{q^2L}{T^2}.$$

Hence μ_0 has the dimension

$$\frac{mL^2}{T^2} \cdot \frac{T^2}{q^2L} = \frac{mL}{q^2}.$$

The dimension of the product $\mu_0\varepsilon_0$ is therefore

$$\frac{mL}{q^2} \cdot \frac{T^2q^2}{mL^3} = \frac{T^2}{L^2}.$$

The dimension of $1/\mu_0\varepsilon_0$, the same as the square of speed. In short, $1/\sqrt{\mu_0\varepsilon_0}$ has the dimension of speed, "length divided by time."

Now we are ready to do the calculations leading to the prediction of waves traveling at the speed of light. We will use the equations (I'), (II'), (III'), and (IV'), as stated on page 1628, where the fields **B** and **E** vary with time. However, we assume there is no current, so $\mathcal{J} = \mathbf{0}$. We also assume that there is no charge q.

Recall the equation (IV')

$$\nabla \times \mathbf{B} = \mu_0\varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}.$$

Differentiating this equation with respect to time t we obtain

$$\frac{\partial}{\partial t}(\nabla \times \mathbf{B}) = \mu_0\varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{C.23.2}$$

As is easy to check, the operator $\frac{\partial}{\partial t}$ can be moved past the $\nabla\times$ to operate directly on **B**. Thus (C.23.2) becomes

$$\nabla \times \frac{\partial \mathbf{B}}{\partial t} = \mu_0\varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{C.23.3}$$

Recall the equation (II')

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}.$$

Taking the curl of both sides of this equation leads to

$$\nabla \left(-\nabla \times \mathbf{E}\right) = \nabla \times \frac{\partial \mathbf{B}}{\partial t}. \tag{C.23.4}$$

Combining (C.23.3) and (C.23.4) gives us an equation that involves **E** alone:

$$\nabla \times \left(-\nabla \times \mathbf{E}\right) = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{C.23.5}$$

An identity concerning "the curl of the curl," which tells us that

$$\nabla \times \left(\nabla \times \mathbf{E}\right) = \nabla \left(\nabla \cdot \mathbf{E}\right) - \left(\nabla \cdot \nabla\right) \mathbf{E}. \tag{C.23.6}$$

But $\nabla \cdot \mathbf{E} = 0$ is one of the four assumptions, namely (I), on the electromagnetic fields. By (C.23.5) and (C.23.6), we arrive at

$$\begin{aligned} \left(\nabla \cdot \nabla\right) \mathbf{E} &= \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \\ \text{or} \qquad \frac{\partial^2 \mathbf{E}}{\partial t^2} - \frac{1}{\mu_0 \varepsilon_0} \nabla^2 \mathbf{E} &= \mathbf{0}. \end{aligned} \tag{C.23.7}$$

The expression ∇^2 in (C.23.7) is short for

$$\begin{aligned} \nabla \cdot \nabla &= \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \cdot \left(\frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}\right) \\ &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \end{aligned} \tag{C.23.8}$$

In $(\nabla \cdot \nabla)\mathbf{E}$ we apply (C.23.8) to each of the three components of **E**. Thus $\nabla^2 \mathbf{E}$ is a vector. So is $\partial^2 \mathbf{E}/\partial t^2$ and (C.23.8) makes sense.

For the sake of simplicity, consider the case in which **E** has only an x-component, which depends only on x and t, $\mathbf{E}(x, y, z, t) = E(x, t)\mathbf{i}$, where E is a scalar function. Then (C.23.8) becomes

$$\frac{\partial^2}{\partial t^2} E(x,t)\mathbf{i} - \frac{1}{\mu_0 \varepsilon_0} \left(\frac{\partial^2 E}{\partial x^2} + \frac{\partial^2 E}{\partial y^2} + \frac{\partial^2 E}{\partial z^2}\right) \mathbf{i} = \mathbf{0},$$

from which it follows

$$\frac{\partial^2}{\partial t^2} E(x,t) - \frac{1}{\mu_0 \varepsilon_0} \frac{\partial^2 E}{\partial x^2} = 0. \tag{C.23.9}$$

Multiply (C.23.9) by $-\mu_0\varepsilon_0$ to obtain

$$\frac{\partial^2 E}{\partial x^2} - \mu_0\varepsilon_0 \frac{\partial^2 E}{\partial t^2} = 0.$$

This looks like the wave equation (see (16.3.11) on page 1298). The solutions are waves traveling with speed $1/\sqrt{\mu_0\varepsilon_0}$.

Maxwell then compares $\sqrt{\mu_0\varepsilon_0}$ with the velocity of light:

> In the following table, the principal results of direct observation of the velocity of light, are compared with the principal results of the comparison of electrical units $(1/\sqrt{\mu_0 v_0})$.

Velocity of light (meters per second)		Ratio of electrical units	
Fizeau	314,000,000	Weber	310,740,000
Sun's Parallax	308,000,000	Maxwell	288,000,000
Foucault	298,360,000	Thomson	282,000,000

Table C.23.1:

> It is magnificent that the velocity of light and the ratio of the units are quantities of the same order of magnitude. Neither of them can be said to be determined as yet with such a degree of accuracy as to enable us to assert that the one is greater or less than the other. It is to be hoped that, by further experiment, the relation between the magnitude of the two quantities may be more accurately determined.
>
> In the meantime our theory, which asserts that these two quantities are equal, and assigns a physical reason for this equality, is certainly not contradicted by the comparison of these results such as they are. [reference?]

On this basis Maxwell concluded that light is an "electromagnetic disturbance" and predicted the existence of other electromagnetic waves. In 1887, eight years after Maxwell's death, Heinrich Hertz produced the predicted waves, whose frequency placed them outside what the eye can see.

By 1890 experiments had confirmed Maxwell's conjecture. First of all, experiments gave the velocity of light as 299,766,000 meters per second and $\sqrt{1/\mu_0\varepsilon_0}$ as 299,550,000 meters per second.

Newton, in his *Principia* of 1687 related gravity on earth with gravity in the heavens. Benjamin Franklin, with his kite experiments showed that lightning was simply an electric phenomenon. From then through the early nineteenth century, Faraday, ???, ... showed that electricity and magnetism were inseparable. Then Maxwell joined them both to light. Einstein, in 1905(?), also by a mathematical argument, hypothesized that mass and energy were related, by his equation $E = mc^2$.

Calculus is Everywhere # 24
Heating and Cooling

Engineers who design a car radiator or a home air conditioner are interested in the distribution of temperature of a fin attached to a tube. We present one of the mathematical tools they use. Incidentally, the example shows how Green's Theorem is applied in practice.

A plane region $\mathcal{A}$ with boundary curve C is occupied by a sheet of metal. By various heating and cooling devices, the temperature along the border is held constant, independent of time. Assume that the temperature in $\mathcal{A}$ eventually stabilizes. This steady-state temperature at point P in $\mathcal{A}$ is denoted $T(P)$. What does that imply about the function $T(x, y)$?

First of all, heat tends to flow "from high to low temperatures," that is, in the direction of $-\nabla T$. According to Fourier's law, flow is proportional to the conductivity of the material k (a positive constant) and the magnitude of the gradient $\|\nabla T\|$. Thus

$$\oint_C (-k\nabla T) \cdot \mathbf{n} ds$$

measures the rate of heat loss across C.

Since the temperature in the metal is at a steady state, the heat in the region bounded by C remains constant. Thus

$$\oint_C (-k\nabla T) \cdot \mathbf{n} ds = 0.$$

Now, Green's theorem then tells us that

$$\int_{\mathcal{A}} \nabla \cdot (-k\nabla T) dA = 0$$

for any region $\mathcal{A}$ in the metal plate. Since $\nabla \cdot \nabla T$ is the Laplacian of T and k is not 0, we conclude that

$$\int_{\mathcal{A}} \left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right) dA = 0. \tag{C.24.1}$$

By the "zero integrals" theorem, the integrand must be 0 throughout $\mathcal{A}$,

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0.$$

This is an important step, since it reduces the study of the temperature distribution to solving a partial differential equation.

The expression

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2},$$

which is $\nabla \cdot \nabla T$, the divergence of the gradient of T, is called the **Laplacian** of T. If T is a function of x, y, and z, then its Laplacian has one more summand, $\partial^2 T/\partial z^2$. However, the vector notation remains the same, $\nabla \cdot \nabla T$. Even more compactly, it is often reduced to $\nabla^2 T$. Note that in spite of the vector notation, the Laplacian of a scalar field is again a scalar field. A function whose Laplacian is 0 is called "harmonic."

EXERCISES