# MATH 275: Calculus III

Lecture Notes by Angel V. Kumchev

# Contents

# Preface

These lecture notes originated from a set of class notes that I used to suplement my lectures for M408D *Sequences, Series, and Multivariable Calculus* at UT-Austin, and later for MATH 275 here, at TU. About 2006, the TU mathematics department set a goal to include in MATH 275 some topics that had previously been excluded "for lack of class time." I found that difficult to achieve within the constraints imposed by the organization of the standard calculus text that we were using at the time, yet quite manageable if I moved some material around and kept the applications more focused on geometry. Those changes were significant enough to diminish the value of the textbook as a primary reference for the course—it became a (rather expensive) source of homework assignments and pretty pictures. That motivated me to add some exercises and illustrations to the class notes and to turn them in a self-contained resource—these lecture notes.

I have not written these notes in vacuum, so there are several people whose help and suggestions I gratefully acknowledge. The earliest version of my UT-Austin class notes benefited greatly from numerous exchanges with Michael Tehranchi and Vrej Zarikian, and my good friend and colleague Alexei Kolesnikov helped me with the addition of some new material to the current version. My colleagues John Chollet and Tatyana Sorokina have used these notes in their own teaching and in the process have not only helped me find various typos, but also have provided me with insights regarding possible improvements—as have many of their and my own former students.

The present version of these notes differs from earlier versions in two main ways. First, I added a couple of new topics: the method of Lagrange multipliers (Lecture #12) and the general theory of change of variables in multiple integrals (Lecture #19). While those topics are not officially a part of the MATH 275 curriculum, I was convinced over time that they are too important not to be mentioned. The second major addition to this version is the inclusion of more visual illustrations and of several *Mathematica* tutorials. Among the latter, I want to mention specifically Appendix D, which is a version of the introduction to *Mathematica* that is part of the computer laboratories used in calculus courses at TU. To that end, I thank my colleagues Raouf Boules, Geoffrey Goodson, Ohoe Kim and Michael O'Leary for allowing me to use their work.

LECTURE 1

# Three-Dimensional Coordinate Systems

You are familiar with coordinate systems in the plane from single-variable calculus. The first coordinate system in the plane that you have seen is most likely the Cartesian system, which consists of two perpendicular axes. These axes are usually positioned so that one of them, called the $x$-axis, is horizontal and oriented from left to right and the other, called the $y$-axis, is vertical and oriented upward (see Figure 1.1). The intersection point of the two axes is called the origin and is commonly denoted by $O$.

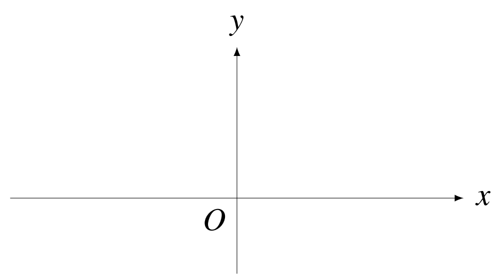



FIGURE 1.1. Two-dimensional Cartesian coordinate system

## 1.1. Cartesian coordinates in space

The *three-dimensional Cartesian coordinate system* consists of a point $O$ in space, called the *origin*, and of three mutually perpendicular axes through $O$, called the *coordinate axes* and labeled the *x-axis*, the *y-axis*, and the *z-axis*. The labeling of the axes is usually done according to the *right-hand rule*: if one places one's right hand in space so that the origin lies on one's palm, the $y$-axis points toward one's arm, and the $x$-axis points toward one's curled fingers, then the $z$-axis points toward one's thumb.

There are several standards for drawing two-dimensional models of three-dimensional objects. In particular, there are three somewhat standard ways to draw a three-dimensional coordinate system. These are shown on Figure 1.2. Model (c) is the one commonly used in manual sketches. Model (a) is a version of model (c) that is commonly used in textbooks and will be the one we shall use in these notes; unlike model (c), it is based on real mathematics that transforms three-dimensional coordinates to two-dimensional ones. Model (b) is the default model preferred by computer algebra systems such as *Mathematica*.

The coordinate axes determine several notable objects in space. First, we have the *coordinate planes*: these are the unique planes determined by the three different pairs of coordinate axes. The coordinate planes determined by the $x$- and $y$-axes, by the $x$- and $z$-axes, and by the $y$- and $z$-axes are called the *xy-plane*, the *xz-plane*, and the *yz-plane*, respectively. The three coordinate planes divide space into eight "equal" parts, called *octants*; the octants are similar to the four quadrants

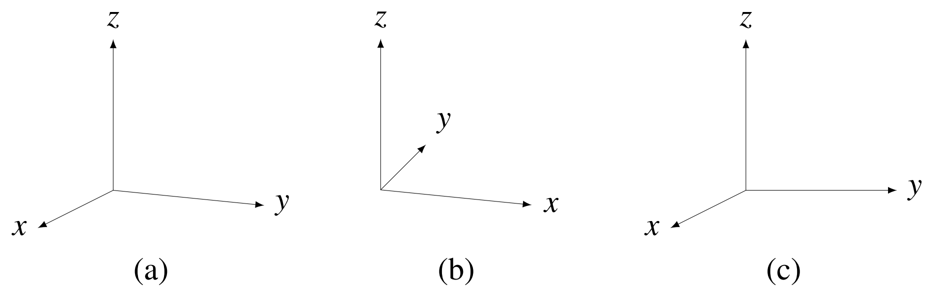



Figure 1.2. Three-dimensional Cartesian coordinate systems

that the two coordinate axes divide the $xy$-plane into. The octants are labeled I through VIII, so that octants I through IV lie above the respective quadrants in the $xy$-plane and octants V through VIII lie below those quadrants.

Next, we define the Cartesian coordinates of a point $P$ in space. We start by finding its orthogonal projections onto the $yz$-, the $xz$-, and the $xy$-planes. Let us denote those points by $P_x$, $P_y$, and $P_z$, respectively (see Figure 1.3). Let $a$ be the directed distance from $P_x$ to $P$: that is,

$$a = \begin{cases} |PP_x| & \text{if the direction from } P_x \text{ to } P \text{ is the same as that of the } x\text{-axis,} \\ -|PP_x| & \text{if the direction from } P_x \text{ to } P \text{ is opposite to that of the } x\text{-axis.} \end{cases}$$

Further, let $b$ be the directed distance from $P_y$ to $P$, and let $c$ be the directed distance from $P_z$ to $P$. The *three-dimensional Cartesian coordinates* of $P$ are $(a, b, c)$ and we write $P(a, b, c)$.

Example 1.1. Plot the points $A(4, 3, 0)$, $B(2, -2, 2)$, and $C(2, 2, -2)$.

Example 1.2. Determine the sets of points described by the equations: $z = 1$; $x = -2$; and $y = 3, x = -1$.

Answer. The equation $z = 1$ represents a plane parallel to the $xy$-plane. The equation $x = -2$ represents a plane parallel to the $yz$-plane. The equations $y = 3, x = -1$ represent a vertical line perpendicular to the $xy$-plane and passing through the point $(-1, 3, 0)$. □

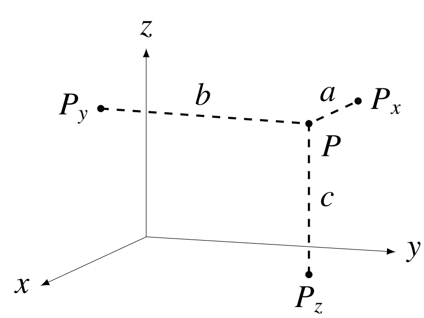



Figure 1.3. Cartesian coordinates of a point in space

## 1.2. Distance in space

Let $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ be two points in space. Then the distance between $P_1$ and $P_2$ is given by

$$|P_1P_2| = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}. \tag{1.1}$$

EXAMPLE 1.3. The distance between the points $A(4, 3, 0)$ and $B(2, -2, 2)$ is

$$|AB| = \sqrt{(4-2)^2 + (3+2)^2 + (0-2)^2} = \sqrt{33}.$$

□

EXAMPLE 1.4. Find the equation of the sphere of radius 3 centered at $C(1, -2, 0)$.

SOLUTION. The sphere of radius 3 centered at $C(1, -2, 0)$ is the set of points $P(x, y, z)$ such that

$$|PC| = \sqrt{(x-1)^2 + (y+2)^2 + (z-0)^2} = 3,$$

that is, the points $P(x, y, z)$, whose coordinates satisfy the equation

$$(x-1)^2 + (y+2)^2 + z^2 = 9.$$ □

By generalizing this example, we obtain the following proposition.

PROPOSITION 1.1. *The sphere with center $C(a, b, c)$ and radius $r > 0$ has an equation*

$$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2. \tag{1.2}$$

*Conversely, the set of points $(x, y, z)$ satisfying (1.2) is a sphere in space with center $(a, b, c)$ and radius $r$.*

EXAMPLE 1.5. The equation

$$x^2 + y^2 + z^2 - 4x + 6y + z = 7$$

describes a sphere. Find its center and radius.

SOLUTION. We can rewrite the above equation as

$$x^2 - 4x + 4 + y^2 + 6y + 9 + z^2 + z + \tfrac{1}{4} = 7 + 4 + 9 + \tfrac{1}{4}$$
$$(x-2)^2 + (y+3)^2 + (z + \tfrac{1}{2})^2 = \tfrac{81}{4}.$$

The last equation is of the form (1.2) with $a = 2$, $b = -3$, $c = -\frac{1}{2}$, and $r = \frac{9}{2}$. Hence, the given equation describes a sphere with center $(2, -3, -\frac{1}{2})$ and radius $\frac{9}{2}$. □

## 1.3. Cylindrical coordinates

You should be familiar with the polar coordinate system in the plane from single-variable calculus (see also Appendix C). In the remainder of this lecture, we define two sets of coordinates in space that extend the polar coordinates in the plane to the three-dimensional setting. We start with the so-called "cylindrical coordinates". The basic idea of the cylindrical coordinate system is to replace the Cartesian coordinates $x, y$ in the $xy$-plane by the corresponding polar coordinates and to keep the Cartesian $z$-coordinate.

Let $P$ be a point in space with Cartesian coordinates $(x, y, z)$. The *cylindrical coordinates* of $P$ are the triple $(r, \theta, z)$, where $(r, \theta)$ are the polar coordinates of the point $P_0(x, y)$ in the $xy$-plane (see Figure 1.4). In other words, we pass between Cartesian and cylindrical coordinates in space

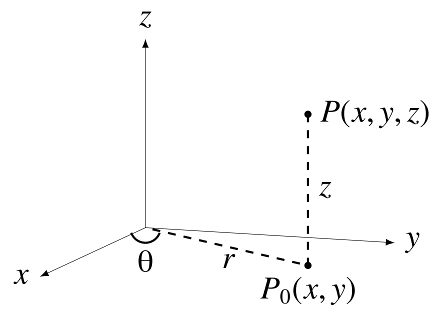



FIGURE 1.4. Cylindrical coordinates of a point in space

by passing between Cartesian and polar coordinates in the $xy$-plane and keeping the $z$-coordinate unchanged. The formulas for conversion from cylindrical to Cartesian coordinates are

$$x = r\cos\theta, \qquad y = r\sin\theta, \qquad z = z, \tag{1.3}$$

and those for conversion from Cartesian to cylindrical coordinates are

$$r = \sqrt{x^2 + y^2}, \qquad \theta = \arctan(y/x), \qquad z = z. \tag{1.4}$$

The formula $\theta = \arctan(y/x)$ in (1.4) comes with a caveat: unless the point $(x, y)$ lies in the first quadrant of the $xy$-plane, this formula yields a reference angle which needs to be adjusted to obtain the actual value of $\theta$. (This is similar to the conversion from Cartesian to polar coordinates in the plane: see §C.2 in Appendix C).

EXAMPLE 1.6. Find the Cartesian coordinates of the points $P(2, \pi, e)$ and $Q(4, \arctan(-2), 2.7)$.

SOLUTION. The Cartesian coordinates of $P$ are $(x, y, e)$, where

$$x = 2\cos\pi = -2, \qquad y = 2\sin\pi = 0.$$

That is, $P(-2, 0, e)$. The Cartesian coordinates of $Q$ are $(x, y, 2.7)$, where

$$x = 4\cos(\arctan(-2)) = \frac{4}{\sqrt{5}}, \qquad y = 4\sin(\arctan(-2)) = -\frac{8}{\sqrt{5}}.$$

That is, $Q(\frac{4}{\sqrt{5}}, -\frac{8}{\sqrt{5}}, 2.7)$. □

EXAMPLE 1.7. Find the cylindrical coordinates of the points $P(3, 4, 5)$ and $Q(-2, 2, -1)$.

SOLUTION. The cylindrical coordinates of $P$ are $(r, \theta, 5)$, where

$$r = \sqrt{3^2 + 4^2} = 5, \qquad \tan\theta = \frac{4}{3},$$

and $\theta$ lies in the first quadrant. Hence, $\theta = \arctan(\frac{4}{3})$ and the cylindrical coordinates of $P$ are $(5, \arctan(\frac{4}{3}), 5)$. The cylindrical coordinates of $Q$ are $(r, \theta, -1)$, where

$$r = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}, \qquad \tan\theta = -1,$$

and $\theta$ lies in the second quadrant. Hence, $\theta = \frac{3\pi}{4}$ and the cylindrical coordinates of $P$ are $(2\sqrt{2}, \frac{3\pi}{4}, -1)$. □

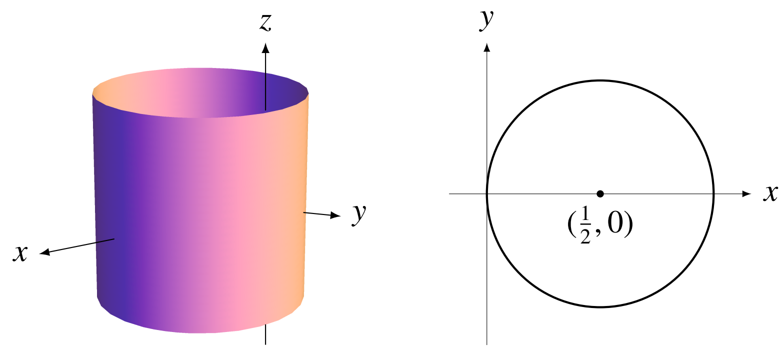



Figure 1.5. The cylinder $x^2 + y^2 = x$ and its cross-section

Example 1.8. Find the Cartesian equation of the surface $r = \cos\theta$. What kind of surface is this?

Solution. We have

$$r = \cos\theta \iff r^2 = r\cos\theta \iff x^2 + y^2 = x.$$

The last equation is the equation of a circle of radius $\frac{1}{2}$ centered at the point $(\frac{1}{2}, 0)$ in the $xy$-plane:

$$x^2 + y^2 = x \iff x^2 - x + \tfrac{1}{4} + y^2 = \tfrac{1}{4} \iff (x - \tfrac{1}{2})^2 + y^2 = \tfrac{1}{4}.$$

Here, however, we use the equation $x^2 + y^2 = x$ to describe a surface in space, so we view it as an equation in $x, y, z$. Since $z$ does not appear in the equation, if the $x$- and $y$-coordinates of a point satisfy the equation $x^2 + y^2 = x$, then the point is on the surface. In other words, if $a^2 + b^2 = a$, then all points $(a, b, z)$, whose $x$-coordinate is $a$ and $y$-coordinate is $b$, are on the surface. We also know that all the points $(x, y, 0)$, with $x^2 + y^2 = x$, are on the surface: these are the points on the circle in the $xy$-plane with the same equation. Putting these two observations together, we conclude that the given surface is the cylinder that consists of all the vertical lines passing through the points of the circle $x^2 + y^2 = x$ in the $xy$-plane (see Figure 1.5). □

## 1.4. Spherical coordinates

We now introduce the so-called "spherical coordinates" in space. In order to define this set of coordinates, we choose a special point $O$, the *origin*, a plane that contains $O$, and a pair of mutually perpendicular axes passing through $O$—one that lies in the chosen plane and another that is orthogonal to the chosen plane. Since we want to relate the spherical coordinates of a point to its Cartesian coordinates, it is common to choose $O$ to be the origin $O(0, 0, 0)$ of the Cartesian coordinate system, the special plane to be the $xy$-plane, and the two axes to be the $x$-axis and the $z$-axis. With these choices, the *spherical coordinates* of a point $P(x, y, z)$ in space are $(\rho, \theta, \phi)$, where $\rho = |OP|$, $\theta$ is the angular polar coordinate of the point $P_0(x, y)$ in the $xy$-plane, and $\phi$ is the angle between $OP$ and the positive direction of the $z$-axis (see Figure 1.6); it is common to restrict $\phi$ to the range $0 \le \phi \le \pi$.

The angles $\theta$ and $\phi$ are often called the *longitude* and the *latitude* of $P$, because with the right choice of the spherical coordinate system, the angles $\theta$ and $90^\circ - \phi$ for a point on the surface of the Earth are its geographic longitude and latitude.

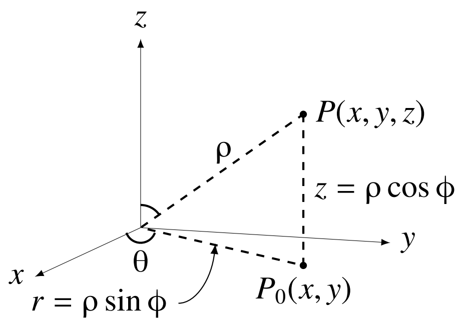



FIGURE 1.6. Spherical coordinates of a point in space

Consider a point $P(x, y, z)$ in space and let $P_0(x, y, 0)$. Also, let $(\rho, \theta, \phi)$ be the spherical coordinates of $P$. From the right triangle $\triangle OPP_0$, we find that

$$z = \pm|P_0P| = |OP|\cos\phi = \rho\cos\phi, \qquad |OP_0| = |OP|\sin\phi = \rho\sin\phi.$$

On the other hand, using polar coordinates in the $xy$-plane, we find that

$$x = |OP_0|\cos\theta = \rho\sin\phi\cos\theta, \qquad y = |OP_0|\sin\theta = \rho\sin\phi\sin\theta.$$

Thus, the formulas for conversion from spherical to Cartesian coordinates are

$$x = \rho\sin\phi\cos\theta, \qquad y = \rho\sin\phi\sin\theta, \qquad z = \rho\cos\phi. \tag{1.5}$$

The formulas for conversion from Cartesian to spherical coordinates are

$$\rho = \sqrt{x^2 + y^2 + z^2}, \qquad \theta = \arctan(y/x), \qquad \phi = \arccos(z/\rho), \tag{1.6}$$

where the formula for $\theta$ comes with some strings attached (the same as in the cases of polar and cylindrical coordinates).

EXAMPLE 1.9. Find the Cartesian coordinates of the point $P(2, \pi/3, 3\pi/4)$.

SOLUTION. The Cartesian coordinates of $P$ are

$$x = 2\sin(3\pi/4)\cos(\pi/3) = \frac{1}{\sqrt{2}}, \quad y = 2\sin(3\pi/4)\sin(\pi/3) = \frac{\sqrt{3}}{\sqrt{2}}, \quad z = 2\cos(3\pi/4) = -\sqrt{2}.$$

□

EXAMPLE 1.10. Find the spherical coordinates of the points $P(1, -1, -\sqrt{6})$ and $Q(-2, 2, -1)$.

SOLUTION. The spherical coordinates of $P$ are

$$\rho = \sqrt{1 + 1 + 6} = \sqrt{8}, \qquad \phi = \arccos\left(-\sqrt{6}/\sqrt{8}\right) = \arccos\left(-\sqrt{3}/2\right) = \frac{5\pi}{6},$$

and the solution $\theta$ of

$$\tan\theta = -1, \qquad 3\pi/2 < \theta < 2\pi.$$

Hence, $\theta = 7\pi/4$ and the spherical coordinates of $P$ are $(\sqrt{8}, 7\pi/4, 5\pi/6)$.

The spherical coordinates of $Q$ are

$$\rho = \sqrt{4 + 4 + 1} = 3, \qquad \phi = \arccos(-1/3),$$

and the solution $\theta$ of

$$\tan\theta = -1, \qquad \pi/2 < \theta < \pi.$$

Hence, $\theta = 3\pi/4$ and the spherical coordinates of $Q$ are $(3, 3\pi/4, \arccos(-1/3))$. □

EXAMPLE 1.11. Find the Cartesian equation of the surface $\rho^2(\sin^2\phi - 2\cos^2\phi) = 1$.

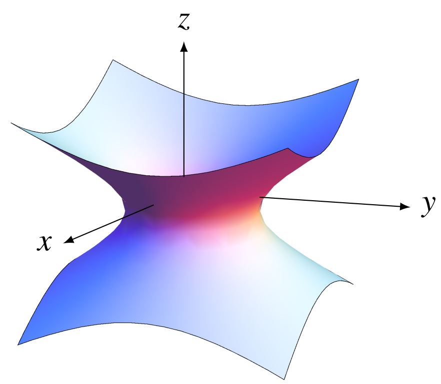
SOLUTION. Since

$$x^2 + y^2 = \rho^2 \sin^2\phi \cos^2\theta + \rho^2 \sin^2\phi \sin^2\theta = \rho^2 \sin^2\phi,$$

we have

$$\rho^2 \sin^2\phi - 2\rho^2 \cos^2\phi = 1 \quad \Longleftrightarrow \quad x^2 + y^2 - 2z^2 = 1.$$

The last equation represents a surface known as "one-sheet hyperboloid"; it is shown on Figure 1.7. □



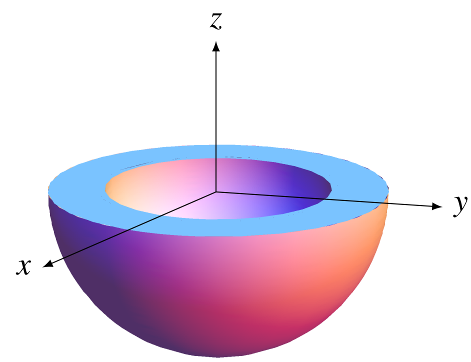




FIGURE 1.7. The hyperboloid $x^2 + y^2 - 2z^2 = 1$

FIGURE 1.8. The region $2 \le \rho \le 3$, $\pi/2 \le \phi \le \pi$

EXAMPLE 1.12. Describe the solid defined by the inequalities

$$2 \le \rho \le 3, \qquad \pi/2 \le \phi \le \pi.$$

SOLUTION. The second inequality restricts the points of the solid to the lower half-space. The inequality $\rho \le 3$ describes the points in space that are within distance of 3 from the origin, that is, the interior of the sphere $x^2 + y^2 + z^2 = 9$. Similarly the inequality $\rho \ge 2$ describes the exterior of the sphere $x^2 + y^2 + z^2 = 4$. Thus, the given solid is obtained from the lower half of the interior of the sphere $x^2 + y^2 + z^2 = 9$ by removing from it the interior of the smaller sphere $x^2 + y^2 + z^2 = 4$ (see Figure 1.8). □

EXAMPLE 1.13. Find the cylindrical and spherical equations of the sphere $x^2 + y^2 + z^2 + 4z = 0$.

SOLUTION. By (1.4), $x^2 + y^2 = r^2$, so the cylindrical form of the given equation is

$$x^2 + y^2 + z^2 + 4z = 0 \quad \Longleftrightarrow \quad r^2 + z^2 + 4z = 0.$$

By (1.6), $x^2 + y^2 + z^2 = \rho^2$, so the spherical form of the given equation is

$$\rho^2 + 4\rho\cos\phi = 0 \quad \Longleftrightarrow \quad \rho = -4\cos\phi.$$ □

## Exercises

1.1. Which of the points $P(1, 3, 2)$, $Q(0, -2, -1)$, and $R(5, 3, -3)$: is closest to the $xy$-plane; lies in the $xz$-plane; lies in the $yz$-plane; is farthest from the origin?

1.2. Describe the set in $\mathbb{R}^3$ represented by the equation $x + 2y = 2$.

1.3. Find the side lengths of the triangle with vertices $A(0, 6, 2)$, $B(3, 4, 1)$ and $C(1, 3, 4)$. Is this a special kind of a triangle: equilateral, isosceles, right?

1.4. Find an equation of the sphere that passes through $P(4, 2, 1)$ and has center $C(1, 1, 1)$.

1.5. Show that the equation $x^2 + y^2 + z^2 = x + 2y + 4z$ represents a sphere and find its center and radius.

Describe verbally the set in $\mathbb{R}^3$ represented by the given equation or inequality.

1.6. $x = -3$

1.7. $z \le 2$

1.8. $-1 \le y \le 2$

1.9. $(x - 1)^2 + y^2 + z^2 \ge 4$

1.10. $4 < x^2 + y^2 + z^2 < 9$

1.11. $xy = 0$

1.12. Write inequalities to describe the half-space consisting of all points to the right of the plane $y = -2$.

1.13. Write inequalities to describe the interior of the upper hemisphere of radius 2 centered at the origin.

Find the Cartesian coordinates of the point with the given cylindrical coordinates.

1.14. $(2, 0, 3)$ 1.15. $(3, \pi/4, -3)$ 1.16. $(3, 7\pi/6, 2)$ 1.17. $(3, 3\pi/2, -1)$

Find the cylindrical coordinates of the point with the given Cartesian coordinates.

1.18. $(1, -1, 4)$ 1.19. $(-2, 2, 2)$ 1.20. $(\sqrt{3}, 1, -2)$ 1.21. $(3, 3, 2)$

Find the Cartesian coordinates of the point with the given spherical coordinates.

1.22. $(1, 0, 0)$ 1.23. $(3, \pi/4, 3\pi/4)$ 1.24. $(3, 7\pi/6, \pi/2)$ 1.25. $(2, \pi/2, \pi/4)$

Find the spherical coordinates of the point with the given Cartesian coordinates.

1.26. $(1, 1, 2)$ 1.27. $(-1, 1, -\sqrt{2})$ 1.28. $(\sqrt{3}, 1, 0)$ 1.29. $(\sqrt{2}, 1, -1)$

Find the cylindrical and spherical equations of the surface represented by the given Cartesian equation.

1.30. $x^2 + y^2 = 2z$ 1.31. $x^2 + 2x + y^2 + z^2 = 1$ 1.32. $z = \sqrt{x^2 + y^2}$

Find the Cartesian equation of the surface represented by the given cylindrical or spherical equation. If possible, describe the surface verbally.

1.33. $\rho \cos \phi = -2$

1.34. $r = 2 \sin \theta$

1.35. $r^2 + z^2 = 4$

1.36. $\rho = 2 \sin \phi \sin \theta$

1.37. $\rho \sin^2 \phi - 2 \cos \phi = 0$

1.38. $r \cos \theta - 2r \sin \theta = 3$

Describe verbally the set in $\mathbb{R}^3$ represented by the given conditions.

1.39. $\theta = \pi/4$

1.40. $\rho = 2$

1.41. $r = 1,\ 0 \le z \le 1$

1.42. $2 \le \rho \le 4,\ 0 \le \phi \le \pi/2$

1.43. $\rho \le 1,\ -\pi/2 \le \theta \le \pi/2$

1.44. $\rho \sin \phi \le 1,\ \pi/2 \le \phi \le \pi$

LECTURE 2

# Vectors

## 2.1. Geometric vectors

A *vector* is a quantity (such as velocity, acceleration, force, impulse, etc.) that has both *magnitude* and *direction*. In mathematics, we consider two types of vectors: geometric and algebraic. A *geometric vector* is a directed line segment, represented visually as an arrow: the length of the arrow represents the magnitude of the vector and the arrow points in the direction of the vector. We shall denote vectors by small boldface letters: $\mathbf{a}, \mathbf{b}, \dots, \mathbf{z}$. If $\mathbf{v}$ is a vector represented by an arrow with initial point $A$ and terminal point $B$, we may also write $\mathbf{v} = \overrightarrow{AB}$; this is consistent with an alternative notation for vectors: $\vec{a}, \vec{b}, \dots, \vec{z}$.

We say that two vectors $\mathbf{u}$ and $\mathbf{v}$ are *equal*, and write $\mathbf{u} = \mathbf{v}$, if $\mathbf{u}$ and $\mathbf{v}$ have equal magnitudes (lengths) and the same direction. The vector of zero length is called the *zero vector* and is denoted $\mathbf{0}$; it is the only vector with no specific direction.

## 2.2. Vector addition

Given two vectors $\mathbf{u} = \overrightarrow{AB}$ and $\mathbf{v} = \overrightarrow{BC}$ (represented so that the terminal point of the first vector is the initial point of the second), we define *their sum* $\mathbf{u} + \mathbf{v}$ by (see Figure 2.1(a))

$$\mathbf{u} + \mathbf{v} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

When $\mathbf{u} = \overrightarrow{AB}$ and $\mathbf{v} = \overrightarrow{CD}$ are two vectors in general position, with $C \neq B$, we first represent $\mathbf{v}$ by an equal arrow $\overrightarrow{BD'}$ and then apply the above definition: $\mathbf{u} + \mathbf{v} = \overrightarrow{AD'}$, as shown on Figure 2.1(b).

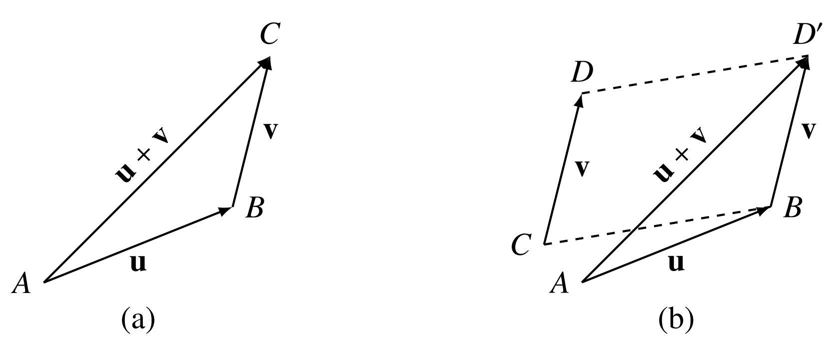



FIGURE 2.1. Vector addition

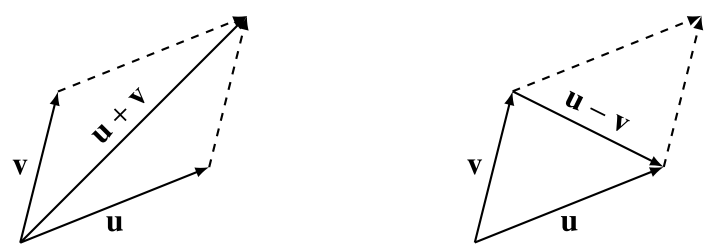
### 2.3. Multiplication by scalars

Let $\mathbf{v}$ be a vector and let $c$ be a real number (called also a *scalar*). The *scalar multiple* $c\mathbf{v}$ is the vector whose length is $|c|$ times the length of $\mathbf{v}$ and whose direction is the same as the direction of $\mathbf{v}$ when $c > 0$, or opposite to the direction of $\mathbf{v}$ when $c < 0$. If $c = 0$ or $\mathbf{v} = \mathbf{0}$, then $c\mathbf{v} = \mathbf{0}$.

The vector $(-1)\mathbf{v}$ has the same length as $\mathbf{v}$ but the opposite direction. We call this vector the *negative* (or *opposite*) of $\mathbf{v}$ and denote it by $-\mathbf{v}$. For two vectors $\mathbf{u}$ and $\mathbf{v}$, we define *their difference* $\mathbf{u} - \mathbf{v}$ by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v}).$$

We visualize the sum and the difference of two vectors $\mathbf{u}$ and $\mathbf{v}$ by drawing them so that they share the same initial point. We then consider the parallelogram they define. The sum $\mathbf{u} + \mathbf{v}$ is represented by the diagonal of the parallelogram that passes through the common initial point of $\mathbf{u}$ and $\mathbf{v}$ and the difference $\mathbf{u} - \mathbf{v}$ is represented by the other diagonal (see Figure 2.2).




(a) (b)

FIGURE 2.2. The sum and the difference of $\mathbf{u}$ and $\mathbf{v}$

### 2.4. Coordinate representation of vectors. Algebraic vectors

Often, it is convenient to introduce a coordinate system and to work with vectors algebraically instead of geometrically. If we represent a (geometric) vector $\mathbf{a}$ by the arrow $\overrightarrow{OA}$, where $O$ is the origin of a Cartesian coordinate system, the terminal point $A$ of $\mathbf{a}$ has coordinates $(a_1, a_2)$ or $(a_1, a_2, a_3)$, depending on the dimension. We call these coordinates the *coordinates* or the *components* of the *algebraic vector* $\mathbf{a}$ and write

$$\mathbf{a} = \langle a_1, a_2 \rangle \qquad \text{or} \qquad \mathbf{a} = \langle a_1, a_2, a_3 \rangle.$$

We will use the same algebraic vector to represent all the geometric vectors equal to $\overrightarrow{OA}$. For example, if $A(1, -1)$, $B(-1, 3)$, and $C(0, 2)$, then

$$\overrightarrow{BC} = \overrightarrow{OA} = \langle 1, -1 \rangle.$$

Thus, for any two points $A(a_1, a_2, a_3)$ and $B(b_1, b_2, b_3)$, the algebraic vector $\mathbf{v}$ representing the arrow $\overrightarrow{AB}$ is

$$\mathbf{v} = \langle b_1 - a_1, b_2 - a_2, b_3 - a_3 \rangle.$$

From now on, we shall state all the formulas for three-dimensional vectors. One usually obtains the respective two-dimensional definition or result simply by dropping the last component of the vectors; any exceptions to this rule will be mentioned explicitly when they occur.

## 2.5. Operations with algebraic vectors

Note that when we write a vector $\mathbf{a}$ algebraically, the distance formula provides a convenient tool for computing the length (also called the *norm*) $\|\mathbf{a}\|$ of that vector: if $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$, then

$$\|\mathbf{a}\| = \left(a_1^2 + a_2^2 + a_3^2\right)^{1/2}. \tag{2.1}$$

Furthermore, if $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$ are vectors and $c$ is a scalar, we have the following coordinate formulas for $\mathbf{a} + \mathbf{b}$, $\mathbf{a} - \mathbf{b}$, and $c\mathbf{a}$:

$$\mathbf{a} + \mathbf{b} = \langle a_1 + b_1, a_2 + b_2, a_3 + b_3 \rangle, \tag{2.2}$$

$$\mathbf{a} - \mathbf{b} = \langle a_1 - b_1, a_2 - b_2, a_3 - b_3 \rangle, \tag{2.3}$$

$$c\mathbf{a} = \langle ca_1, ca_2, ca_3 \rangle. \tag{2.4}$$

In other words, we add vectors, subtract vectors, and multiply vectors by scalars by simply performing the respective operations componentwise.

Example 2.1. Consider the vectors $\mathbf{a} = \langle 1, -2, 3 \rangle$ and $\mathbf{b} = \langle 0, 1, -2 \rangle$. Then

$$\begin{aligned} 2\mathbf{a} + \mathbf{b} &= \langle 2, -4, 6 \rangle + \langle 0, 1, -2 \rangle = \langle 2, -3, 4 \rangle, \\ \|2\mathbf{a} + \mathbf{b}\| &= \sqrt{2^2 + (-3)^2 + 4^2} = \sqrt{29}. \end{aligned}$$

The algebraic operations with vectors have several properties that resemble familiar properties of the algebraic operations with numbers.

Theorem 2.1. *Let* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *denote vectors and let* $c, c_1, c_2$ *denote scalars. Then:*

i) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
ii) $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
iii) $\mathbf{a} + \mathbf{0} = \mathbf{a}$
iv) $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$
v) $c(\mathbf{a} + \mathbf{b}) = c\mathbf{a} + c\mathbf{b}$
vi) $(c_1 + c_2)\mathbf{a} = c_1\mathbf{a} + c_2\mathbf{a}$
vii) $(c_1c_2)\mathbf{a} = c_1(c_2\mathbf{a})$
viii) $1\mathbf{a} = \mathbf{a}$
ix) $\|\mathbf{a}\| \geq 0$*, with* $\|\mathbf{a}\| = 0$ *if and only if* $\mathbf{a} = \mathbf{0}$
x) $\|c\mathbf{a}\| = |c| \cdot \|\mathbf{a}\|$
xi) $\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|$ *(the triangle inequality)*

Proof. The proofs of parts i)–viii) are all similar: they use (2.2)–(2.4) to deduce i)–viii) from the respective properties of numbers. For example, to prove the distributive law vi), we write $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and argue as follows:

$$\begin{aligned} (c_1 + c_2)\mathbf{a} &= \langle (c_1 + c_2)a_1, (c_1 + c_2)a_2, (c_1 + c_2)a_3 \rangle && \text{by (2.4)} \\ &= \langle c_1a_1 + c_2a_1, c_1a_2 + c_2a_2, c_1a_3 + c_2a_3 \rangle && \text{by the distributive law for numbers} \\ &= \langle c_1a_1, c_1a_2, c_1a_3 \rangle + \langle c_2a_1, c_2a_2, c_2a_3 \rangle && \text{by (2.2)} \\ &= c_1\mathbf{a} + c_2\mathbf{a} && \text{by (2.3).} \end{aligned}$$

The proofs of ix) and x) are also easy. For instance,

$$\begin{aligned}\|c\mathbf{a}\| &= \left((ca_1)^2 + (ca_2)^2 + (ca_3)^2\right)^{1/2} \\ &= \left(c^2\left(a_1^2 + a_2^2 + a_3^2\right)\right)^{1/2} = \sqrt{c^2}\left(a_1^2 + a_2^2 + a_3^2\right)^{1/2} = |c| \cdot \|\mathbf{a}\|.\end{aligned}$$

The proof of xi) is slightly more involved, so we leave it to the exercises. □

## 2.6. Unit vectors and the standard basis

The simplest nonzero vectors are:

$$\mathbf{e}_1 = \langle 1, 0, 0\rangle, \quad \mathbf{e}_2 = \langle 0, 1, 0\rangle, \quad \mathbf{e}_3 = \langle 0, 0, 1\rangle.$$

These are called the *standard basis vectors*. Sometimes, especially in physics and engineering, the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ are denoted by $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$, respectively, but we shall stick with the above notation, because it is more common in mathematics (it is also more consistent with generalizations to higher dimensions).

Because of (2.2)–(2.4), it is always possible to express a given vector as a combination of the three standard basis vectors. For example,

$$\langle 1, 2, 4\rangle = \langle 1, 0, 0\rangle + \langle 0, 2, 0\rangle + \langle 0, 0, 4\rangle = \langle 1, 0, 0\rangle + 2\langle 0, 1, 0\rangle + 4\langle 0, 0, 1\rangle = \mathbf{e}_1 + 2\mathbf{e}_2 + 4\mathbf{e}_3.$$

More generally, if $\mathbf{a} = \langle a_1, a_2, a_3\rangle$, we have

$$\mathbf{a} = \langle a_1, 0, 0\rangle + \langle 0, a_2, 0\rangle + \langle 0, 0, a_3\rangle = a_1\mathbf{e}_1 + a_2\mathbf{e}_2 + a_3\mathbf{e}_3.$$

Notice that $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ have length 1. In general, vectors of length 1 are called *unit vectors*. We remark that if $\mathbf{a}$ is any nonzero vector (i.e., $\mathbf{a} \neq \mathbf{0}$), then there is a unique unit vector that points in the same direction as $\mathbf{a}$—namely, the vector

$$\mathbf{u} = \frac{1}{\|\mathbf{a}\|}\mathbf{a} = \frac{\mathbf{a}}{\|\mathbf{a}\|}.$$

Indeed, if $c = \|\mathbf{a}\|^{-1}$, then $\mathbf{a}$ and $\mathbf{u} = c\mathbf{a}$ point in the same direction (since $c > 0$) and

$$\|\mathbf{u}\| = \|c\mathbf{a}\| = |c| \cdot \|\mathbf{a}\| = \|\mathbf{a}\|^{-1} \cdot \|\mathbf{a}\| = 1.$$

This simple computation is known as *normalization*, and $\mathbf{u}$ is called the *normalization* of $\mathbf{a}$.

Example 2.2. Consider the vector $\mathbf{a} = \langle 2, 1, 2\rangle$. Its length is

$$\|\mathbf{a}\| = \sqrt{2^2 + 1^2 + 2^2} = 3,$$

so the normalization of $\mathbf{a}$ is the vector

$$\tfrac{1}{3}\mathbf{a} = \left\langle \tfrac{2}{3}, \tfrac{1}{3}, \tfrac{2}{3}\right\rangle.$$

## 2.7. The dot product

We can now add two vectors and we can multiply a vector by a scalar. Is there a way to multiply two vectors? The first way that comes to mind is to multiply the vectors componentwise:

$$\mathbf{ab} = \langle a_1, a_2, a_3\rangle\langle b_1, b_2, b_3\rangle = \langle a_1b_1, a_2b_2, a_3b_3\rangle,$$

but this way of multiplying two vectors turns out to be of little practical use. Instead, through the remainder of this lecture, we shall talk about two meaningful ways to multiply vectors: the dot product and the cross product.

Definition. Given vectors $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, we define their *dot product* $\mathbf{a} \cdot \mathbf{b}$ to be

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + a_3 b_3. \tag{2.5}$$

Notice that the dot product of two vectors is a scalar, not a vector. Sometimes, the dot product is called also the *inner product* or the *scalar product of* **a** *and* **b**. Because of this, you should avoid the temptation to refer to the scalar multiple $c\mathbf{a}$ as the "scalar product of $c$ and **a**".

Example 2.3. The dot product of $\mathbf{a} = \langle 4, 2, -1 \rangle$ and $\mathbf{b} = \langle -2, 2, 1 \rangle$ is

$$\mathbf{a} \cdot \mathbf{b} = (4)(-2) + (2)(2) + (-1)(1) = -5.$$

The dot product of $\mathbf{u} = 2\mathbf{e}_1 + 3\mathbf{e}_2 + \mathbf{e}_3$ and $\mathbf{v} = \mathbf{e}_1 + 4\mathbf{e}_2$ is

$$\mathbf{u} \cdot \mathbf{v} = (2)(1) + (3)(4) + (1)(0) = 14.$$

The next theorem summarizes some important algebraic properties of the dot product.

Theorem 2.2. *Let* **a**, **b**, *and* **c** *denote vectors and let* $c_1, c_2$ *denote scalars. Then:*

i) $\mathbf{a} \cdot \mathbf{a} = \|\mathbf{a}\|^2$
ii) $\mathbf{a} \cdot \mathbf{0} = 0$
iii) $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
iv) $(c_1\mathbf{a}) \cdot (c_2\mathbf{b}) = c_1 c_2 (\mathbf{a} \cdot \mathbf{b})$
v) $(\mathbf{a} + \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot \mathbf{c} + \mathbf{b} \cdot \mathbf{c}$

Proof. These follow from the definition of the dot product and from the properties of the real numbers. For example, if $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3 \rangle$, we have

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= a_1 b_1 + a_2 b_2 + a_3 b_3 && \text{by (2.5)} \\ &= b_1 a_1 + b_2 a_2 + b_3 a_3 && \text{by the commutative law for real numbers} \\ &= \mathbf{b} \cdot \mathbf{a} && \text{by (2.5).} \end{aligned}$$

That is, iii) holds. □

We said earlier that our definition of the dot product is motivated by applications. The next theorem relates the dot product of two vectors to their geometric representations.

Theorem 2.3. *If* **a** *and* **b** *are two vectors, then*

$$\mathbf{a} \cdot \mathbf{b} = \|\mathbf{a}\| \, \|\mathbf{b}\| \cos\theta,$$

*where* $\theta$ *is the angle between* **a** *and* **b**.

Corollary 2.4. *If* **a** *and* **b** *are two nonzero vectors and* $\theta$ *is the angle between them, then*

$$\cos\theta = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \, \|\mathbf{b}\|}.$$

Example 2.4. Find the angle between $\mathbf{a} = \langle 1, 2, -1 \rangle$ and $\mathbf{b} = \langle 0, 2, 1 \rangle$.

Solution. The angle $\theta$ between **a** and **b** satisfies

$$\cos\theta = \frac{(1)(0) + (2)(2) + (-1)(1)}{\sqrt{1^2 + 2^2 + (-1)^2}\sqrt{0^2 + 2^2 + 1^2}} = \frac{3}{\sqrt{6}\sqrt{5}} = \frac{3}{\sqrt{30}},$$

so $\theta = \arccos\left(\frac{3}{\sqrt{30}}\right) = 0.9911\ldots.$ □

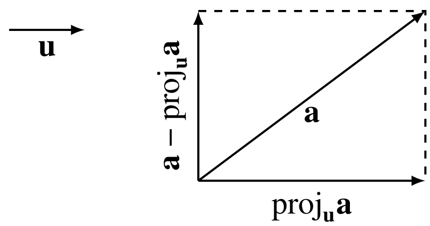



FIGURE 2.3. The projection of a vector onto a unit vector

We say that two vectors $\mathbf{a}$ and $\mathbf{b}$ are *orthogonal* (or *perpendicular*) if $\mathbf{a}\cdot\mathbf{b} = 0$. By the corollary, $\mathbf{a}\cdot\mathbf{b} = 0$ implies that the angle between $\mathbf{a}$ and $\mathbf{b}$ is 90°, so this definition makes perfect sense.

Note that if $\mathbf{a} = \langle a_1, a_2, a_3\rangle$, then we can use the dot products of $\mathbf{a}$ and the standard basis vectors $\mathbf{e}_j$ to extract the components of $\mathbf{a}$:

$$a_1 = \mathbf{a}\cdot\mathbf{e}_1, \qquad a_2 = \mathbf{a}\cdot\mathbf{e}_2, \qquad a_3 = \mathbf{a}\cdot\mathbf{e}_3.$$

More generally, suppose that $\mathbf{u}$ is a unit vector. Then $\mathbf{a}\cdot\mathbf{u}$ is a scalar and

$$\mathrm{proj}_{\mathbf{u}}\mathbf{a} = (\mathbf{a}\cdot\mathbf{u})\mathbf{u}$$

is a vector that is parallel to $\mathbf{u}$. This vector is called the *(orthogonal) projection of* $\mathbf{a}$ *onto* $\mathbf{u}$. It has the property that the difference between it and the original vector $\mathbf{a}$ is perpendicular to $\mathbf{u}$:

$$\left(\mathbf{a} - \mathrm{proj}_{\mathbf{u}}\mathbf{a}\right)\cdot\mathbf{u} = 0.$$

The two-dimensional case of this is illustrated on Figure 2.3. Note that the size of the dot product $\mathbf{a}\cdot\mathbf{u}$ equals the length of the projection; the sign of that dot product indicates whether the projection and $\mathbf{u}$ have the same direction (positive sign) or opposite directions (negative sign).

## 2.8. The cross product

The dot product of two vectors is a scalar. In this section, we define another product of two vectors: their "cross product", which is a vector. Since our discussion requires basic familiarity with $2\times 2$ and $3\times 3$ determinants, a quick look at Appendix A may be helpful before proceeding further.

DEFINITION. Let $\mathbf{a} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3\rangle$ be two three-dimensional vectors. Then their *cross product* $\mathbf{a}\times\mathbf{b}$ is the vector

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\mathbf{e}_3.$$

REMARKS. There are a couple of things we should point out:

1. The cross product is defined only for three-dimensional vectors. There is no analog of this operation for two-dimensional vectors.
2. The defining formula may be confusing at first. On the other hand, following the exact labeling is important because of the properties of the cross product (see Theorem 2.5

below). Another representation of the cross product that is much easier to remember is

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}. \tag{2.6}$$

This expression is not exactly a determinant in the sense described in Appendix A (since its first row consists of vectors instead of numbers), but if we expand this "determinant" as if we were dealing with the determinant of a regular $3\times 3$ matrix, we get exactly the cross product of $\mathbf{a}$ and $\mathbf{b}$.

The following theorem summarizes the algebraic properties of the cross product.

THEOREM 2.5. *Let* $\mathbf{a}$, $\mathbf{b}$, *and* $\mathbf{c}$ *denote three-dimensional vectors and let* $c, c_1, c_2$ *denote scalars. Then*

i) $\mathbf{b}\times\mathbf{a} = -(\mathbf{a}\times\mathbf{b})$
ii) $(c_1\mathbf{a})\times(c_2\mathbf{b}) = c_1c_2(\mathbf{a}\times\mathbf{b})$
iii) $\mathbf{a}\times(\mathbf{b}+\mathbf{c}) = \mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$
iv) $(\mathbf{a}+\mathbf{b})\times\mathbf{c} = \mathbf{a}\times\mathbf{c}+\mathbf{b}\times\mathbf{c}$
v) $\mathbf{a}\times\mathbf{a} = \mathbf{0}$ *and* $\mathbf{a}\times\mathbf{0} = \mathbf{0}$
vi) $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{a}\times\mathbf{b})\times\mathbf{c}$ *may fail.*

EXAMPLE 2.5. We have

$$\mathbf{e}_1\times\mathbf{e}_2 = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\mathbf{e}_3 = \mathbf{e}_3,$$

and similarly,

$$\mathbf{e}_2\times\mathbf{e}_3 = \mathbf{e}_1, \qquad \mathbf{e}_3\times\mathbf{e}_1 = \mathbf{e}_2.$$

By property i), we also have

$$\mathbf{e}_2\times\mathbf{e}_1 = -\mathbf{e}_3, \qquad \mathbf{e}_1\times\mathbf{e}_3 = -\mathbf{e}_2, \qquad \mathbf{e}_3\times\mathbf{e}_2 = -\mathbf{e}_1.$$

Note that

$$\mathbf{e}_1\times(\mathbf{e}_2\times\mathbf{e}_1) = \mathbf{e}_1\times(-\mathbf{e}_3) = -(\mathbf{e}_1\times\mathbf{e}_3) = \mathbf{e}_2$$

and

$$(\mathbf{e}_1\times\mathbf{e}_2)\times\mathbf{e}_1 = \mathbf{e}_3\times\mathbf{e}_1 = \mathbf{e}_2,$$

so $\mathbf{e}_1\times(\mathbf{e}_2\times\mathbf{e}_1) = (\mathbf{e}_1\times\mathbf{e}_2)\times\mathbf{e}_1$. On the other hand,

$$\mathbf{e}_1\times(\mathbf{e}_1\times\mathbf{e}_2) = \mathbf{e}_1\times\mathbf{e}_3 = -\mathbf{e}_2,$$

whereas by property v),

$$(\mathbf{e}_1\times\mathbf{e}_1)\times\mathbf{e}_2 = \mathbf{0}\times\mathbf{e}_2 = \mathbf{0}.$$

Thus, $\mathbf{e}_1\times(\mathbf{e}_1\times\mathbf{e}_2) \neq (\mathbf{e}_1\times\mathbf{e}_1)\times\mathbf{e}_2$; this illustrates property vi) above. □

EXAMPLE 2.6. Compute the cross product of $\mathbf{a} = \langle 3, -2, 1\rangle$ and $\mathbf{b} = \langle 1, -1, 1\rangle$.

SOLUTION. By (2.6),

$$\begin{aligned}\mathbf{a}\times\mathbf{b} &= \begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ 3 & -2 & 1\\ 1 & -1 & 1\end{vmatrix}\\ &= \begin{vmatrix}-2 & 1\\ -1 & 1\end{vmatrix}\mathbf{e}_1 - \begin{vmatrix}3 & 1\\ 1 & 1\end{vmatrix}\mathbf{e}_2 + \begin{vmatrix}3 & -2\\ 1 & -1\end{vmatrix}\mathbf{e}_3\\ &= -\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3 = \langle -1, -2, -1\rangle\,.\end{aligned}$$

□

SECOND SOLUTION. We can also argue by using the algebraic properties in Theorem 2.5 and the values of the cross products of $\mathbf{e}_1, \mathbf{e}_2$, and $\mathbf{e}_3$:

$$\begin{aligned}&(3\mathbf{e}_1 - 2\mathbf{e}_2 + \mathbf{e}_3)\times(\mathbf{e}_1 - \mathbf{e}_2 + \mathbf{e}_3)\\ &= 3\mathbf{e}_1\times\mathbf{e}_1 - 2\mathbf{e}_2\times\mathbf{e}_1 + \mathbf{e}_3\times\mathbf{e}_1 - 3\mathbf{e}_1\times\mathbf{e}_2 + 2\mathbf{e}_2\times\mathbf{e}_2 - \mathbf{e}_3\times\mathbf{e}_2 + 3\mathbf{e}_1\times\mathbf{e}_3 - 2\mathbf{e}_2\times\mathbf{e}_3 + \mathbf{e}_3\times\mathbf{e}_3\\ &= \mathbf{0} + 2\mathbf{e}_3 + \mathbf{e}_2 - 3\mathbf{e}_3 + \mathbf{0} + \mathbf{e}_1 - 3\mathbf{e}_2 - 2\mathbf{e}_1 + \mathbf{0} = -\mathbf{e}_1 - 2\mathbf{e}_2 - \mathbf{e}_3.\end{aligned}$$

□

Finally, we discuss the geometric meaning of the cross product.

THEOREM 2.6. *Let* $\mathbf{a} = \langle a_1, a_2, a_3\rangle$ *and* $\mathbf{b} = \langle b_1, b_2, b_3\rangle$ *be three-dimensional vectors. Then:*

i) $\mathbf{a}\times\mathbf{b}$ *is perpendicular to both* $\mathbf{a}$ *and* $\mathbf{b}$*. The direction of* $\mathbf{a}\times\mathbf{b}$ *is determined by the right-hand rule.*
ii) $\|\mathbf{a}\times\mathbf{b}\| = \|\mathbf{a}\|\,\|\mathbf{b}\|\,|\sin\theta|$*, where* $\theta$ *is the angle between* $\mathbf{a}$ *and* $\mathbf{b}$*. That is, the length of* $\mathbf{a}\times\mathbf{b}$ *is equal to the area of the parallelogram determined by* $\mathbf{a}$ *and* $\mathbf{b}$*.*
iii) $\mathbf{a}\times\mathbf{b} = \mathbf{0}$ *if and only if* $\mathbf{a}$ *and* $\mathbf{b}$ *are parallel.*
iv) *The volume* $V$ *of the parallelepiped determined by three vectors* $\mathbf{a}, \mathbf{b}$*, and* $\mathbf{c}$ *which are not* coplanar *(that is, they do not lie in one plane) is given by the formula*

$$V = \left|\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})\right|.$$

*Furthermore, three vectors* $\mathbf{a}, \mathbf{b}$*, and* $\mathbf{c}$ *are coplanar if and only if* $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = 0$.

EXAMPLE 2.7. Find two unit vectors perpendicular to $\mathbf{a} = \langle 1, 2, 3\rangle$ and $\mathbf{b} = \langle 2, 1, 1\rangle$.

SOLUTION. We know from part i) of Theorem 2.6 that $\mathbf{a}\times\mathbf{b}$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. We find that

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix}\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3\\ 1 & 2 & 3\\ 2 & 1 & 1\end{vmatrix} = \begin{vmatrix}2 & 3\\ 1 & 1\end{vmatrix}\mathbf{e}_1 - \begin{vmatrix}1 & 3\\ 2 & 1\end{vmatrix}\mathbf{e}_2 + \begin{vmatrix}1 & 2\\ 2 & 1\end{vmatrix}\mathbf{e}_3 = -\mathbf{e}_1 + 5\mathbf{e}_2 - 3\mathbf{e}_3.$$

Hence,

$$\|\mathbf{a}\times\mathbf{b}\| = \sqrt{(-1)^2 + 5^2 + (-3)^2} = \sqrt{35},$$

and two unit vectors orthogonal to $\mathbf{a}$ and $\mathbf{b}$ are the normalization of $\mathbf{a}\times\mathbf{b}$ and its opposite:

$$\mathbf{u} = \frac{\mathbf{a}\times\mathbf{b}}{\|\mathbf{a}\times\mathbf{b}\|} = \left\langle \frac{-1}{\sqrt{35}}, \frac{5}{\sqrt{35}}, \frac{-3}{\sqrt{35}}\right\rangle \qquad \text{and} \qquad -\mathbf{u} = \left\langle \frac{1}{\sqrt{35}}, \frac{-5}{\sqrt{35}}, \frac{3}{\sqrt{35}}\right\rangle.$$

□

EXAMPLE 2.8. Find the area of the triangle with vertices $P(1,2,3)$, $Q(-1,3,2)$, and $R(3,-1,2)$.

SOLUTION. The area of $\triangle PQR$ is one half of the area of the parallelogram determined by the vectors $\overrightarrow{PQ} = \langle -2, 1, -1\rangle$ and $\overrightarrow{PR} = \langle 2, -3, -1\rangle$. We know from part ii) of Theorem 2.6 that the area of the parallelogram is the length of

$$\overrightarrow{PQ} \times \overrightarrow{PR} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ -2 & 1 & -1 \\ 2 & -3 & -1 \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ -3 & -1 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} -2 & -1 \\ 2 & -1 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} -2 & 1 \\ 2 & -3 \end{vmatrix} \mathbf{e}_3 = -4\mathbf{e}_1 - 4\mathbf{e}_2 + 4\mathbf{e}_3.$$

Therefore, the area of the triangle is $\frac{1}{2}\sqrt{16+16+16} = 2\sqrt{3}$. □

## 2.9. Vectors in *Mathematica*

**2.9.1. Representation.** We represent a vector in *Mathematica* as a list of numbers (the components of the vector). For example, to define a vector $\mathbf{a} = \langle a_1, a_2, a_3\rangle$, we write

```
a = {a1, a2, a3}
```

To refer to a particular component of a vector (or a list), we use double square brackets. For example, the second component of the vector **a** defined above can be accessed as `a[[2]]`. Although we are only interested in two- and three-dimensional vectors here, we should note that *Mathematica* can handle vectors of higher dimensions just as easily.

On the geometric side, *Mathematica* has a built-in command `Arrow`, which produces a geometric vector in two or three dimensions. In three dimensions, its basic syntax is

```
Arrow[{{x1, y1, z1}, {x2, y2, z2}}]
```

This results in an arrow with initial point $(x_1, y_1, z_1)$ and terminal point $(x_2, y_2, z_2)$. (See *Mathematica*'s help for various formatting options and more advanced usage.) To get a two-dimensional vector, we simply use two-dimensional points as arguments:

```
Arrow[{{x1, y1}, {x2, y2}}]
```

For technical reasons, the `Arrow` command produces a "graphics primitive" instead of an actual plot of the geometric vector. For example, if we try to execute the command

```
Arrow[{{1,2}, {4,-1}}]
```

*Mathematica* will simply repeat it in response. To see the result of the above command, we need to "wrap" it in a `Graphics` command like this:

```
Graphics[Arrow[{{1,2}, {4,-1}}]]
```

Similarly, to display a three-dimensional geometric vector, we use a `Graphics3D` command. Here is an example:

```
Graphics3D[Arrow[{{1,2,0}, {1,-1,2}}]]
```

**2.9.2. Vector addition and scalar multiplication.** Vector addition, vector subtraction and scalar multiplication are denoted in the usual way: if $c$ is a scalar and **a** and **b** are two vectors (of equal dimensions), then

| | |
|---|---|
| `a+b` | denotes the vector sum $\mathbf{a} + \mathbf{b}$ |
| `a-b` | denotes the vector difference $\mathbf{a} - \mathbf{b}$ |
| `c*a`, `c a` | denote the scalar multiple $c\mathbf{a}$ |

For example, let us define the vectors $\mathbf{a} = \langle 2, 7, -3 \rangle$ and $\mathbf{b} = \langle -3, 0, 5 \rangle$.

```
a = {2,7,-3}
b = {-3,0,5}
```

We can calculate $2\mathbf{a}$, $\mathbf{a} - \mathbf{b}$, and $\mathbf{a} + 3\mathbf{b}$ as follows:

```
2a
a-b
a+3*b

{4,14,-6}
{5,7,-8}
{-7,7,12}
```

**2.9.3. The dot and cross products.** *Mathematica* has built-in commands for calculating the dot and cross products: `Dot` and `Cross`, respectively. These commands also have infix operator forms. A period between vectors performs a dot product, and a × symbol between vectors performs a cross product. To enter the symbol ×, first hit the `Esc` key, type "cross", and then hit again the `Esc` key. As an example, let us use *Mathematica* to compute the dot and cross products of two generic three-dimensional vectors:

```
a = {a1, a2, a3}
b = {b1, b2, b3}
```

Here are the two versions of the dot product:

```
a.b
Dot[a,b]

a1 b1 + a2 b2 + a3 b3
a1 b1 + a2 b2 + a3 b3
```

And one of the versions of the cross product (see the *Mathematica* notebook file for the other version):

```
Cross[a,b]

{-a3 b2 + a2 b3, a3 b1 - a1 b3, -a2 b1 + a1 b2}
```

**2.9.4. Graphical representation revisited.** Using the `Arrow` command can be cumbersome when only basic drawing functionality is needed. In such situations, it is much simpler to use a user-defined command that "wraps around" `Arrow` and makes it more user-friendly (but cuts some of its functionality). The command `vector` defined below presents one such option.

```
vector[a_, b_, rgb_:{0,0,0}] := Which[
  Length[a] == Length[b] == 2,
   Graphics[{RGBColor[rgb[[1]],rgb[[2]],rgb[[3]]], Arrow[{a,b}]}],
  Length[a] == Length[b] == 3,
   Graphics3D[{RGBColor[rgb[[1]],rgb[[2]],rgb[[3]]], Arrow[{a,b}]}],
  True, Print["Bad dimensions!"]
 ]
```

The command `vector` can be called in two ways. In its more basic form,

```
vector[{x1, y1}, {x2, y2}]
vector[{x1, y1, z1}, {x2, y2, z2}]
```

the `vector` command returns a graphical representation of a two- or three-dimensional arrow with the given initial and terminal points (similarly to `Arrow`). For example, try

```
vector[{1,2}, {-1,0}]
vector[{1,1,0}, {1,2,-1}]
```

When the dimensions of the two inputs do not match, or if they match but are neither 2 nor 3, the `vector` command prints out the error message "`Bad dimensions!`":

```
vector[{1,2}, {1,2,-1}]

Bad dimensions!
```

The `vector` command allows the user to include a third, optional argument in the form of a pre-defined color or a three-dimensional vector to be interpreted as an RGB color-specification. For example, the following four commands will render their outputs in red, red, blue, and teal, respectively:

```
vector[{1,2}, {-1,0}, Red]
vector[{1,2}, {-1,0}, {1,0,0}]
vector[{1,2}, {-1,0}, {0,0,1}]
vector[{1,2}, {-1,0}, {0,0.5,0.5}]
```

## Exercises

2.1. Let $O(0,0)$, $A(1,2)$, $B(3,0)$, and $M(-2,1)$ be points in $\mathbb{R}^2$. Draw the geometric vectors $\mathbf{v} = \overrightarrow{AB}$ and $\mathbf{w} = \overrightarrow{MO}$ and use them to draw the following vectors $\mathbf{v} + 2\mathbf{w}$, $\mathbf{w} - \mathbf{v}$, and $\mathbf{w} + \frac{1}{2}\mathbf{v}$.

Find $\|\mathbf{b}\|$, $\mathbf{a} - 3\mathbf{b}$, $\|\mathbf{a} + \frac{1}{2}\mathbf{b}\|$, and $2\mathbf{a} + \mathbf{b} - 3\mathbf{e}_2$ for the given vectors $\mathbf{a}$ and $\mathbf{b}$.

2.2. $\mathbf{a} = \langle 1, -2 \rangle$, $\mathbf{b} = \langle 2, 0 \rangle$

2.3. $\mathbf{a} = 3\mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b} = -\mathbf{e}_1 + \mathbf{e}_2$

2.4. $\mathbf{a} = \langle 2, 0, -1 \rangle$, $\mathbf{b} = \langle -1, 1, 1 \rangle$

2.5. $\mathbf{a} = \mathbf{e}_2 + \frac{1}{2}\mathbf{e}_3$, $\mathbf{b} = -2\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$

2.6. Find a vector $\mathbf{b}$ that has the same direction as $\mathbf{a} = \langle 1, -1 \rangle$ and length $\|\mathbf{b}\| = 3$.

Find $\mathbf{a} \cdot \mathbf{b}$ for the given vectors $\mathbf{a}$ and $\mathbf{b}$.

2.7. $\mathbf{a} = \langle 3, -2, 1 \rangle$, $\mathbf{b} = \langle 1, 0, 2 \rangle$

2.8. $\mathbf{a} = \langle 2, 0, -1 \rangle$, $\mathbf{b} = \langle 2, 1, 4 \rangle$

2.9. $\mathbf{a} = 3\mathbf{e}_1 + 2\mathbf{e}_2 + 4\mathbf{e}_3$, $\mathbf{b} = -\mathbf{e}_1 + 2\mathbf{e}_2 + \frac{1}{2}\mathbf{e}_3$

2.10. $\|\mathbf{a}\| = 2$, $\|\mathbf{b}\| = 3$, $\sphericalangle(\mathbf{a}, \mathbf{b}) = 60°$

Find the angle between the given vectors.

2.11. $\mathbf{a} = \langle 2, 0, -1 \rangle$, $\mathbf{b} = \langle 2, 1, 4 \rangle$

2.12. $\mathbf{a} = \langle 2, 0 \rangle$, $\mathbf{b} = \left\langle 1, \sqrt{3} \right\rangle$

2.13. $\mathbf{a} = 3\mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b} = -\mathbf{e}_1 + \mathbf{e}_3$

2.14. $\mathbf{a} = \mathbf{e}_2 + \mathbf{e}_3$, $\mathbf{b} = \mathbf{e}_1 + 2\mathbf{e}_2 - 3\mathbf{e}_3$

2.15. Find the three angles of the triangle with vertices $A(1,0,0)$, $B(2,1,0)$, $C(0,1,2)$.

Determine whether the given vectors are parallel, perpendicular, or neither.

2.16. $\mathbf{a} = \langle 2, 1, -1 \rangle$, $\mathbf{b} = \left\langle \frac{1}{2}, 1, 2 \right\rangle$

2.17. $\mathbf{a} = \langle 1, -2, 2 \rangle$, $\mathbf{b} = \langle -2, 4, -4 \rangle$

2.18. $\mathbf{a} = \langle 1, 2, 3 \rangle$, $\mathbf{b} = \left\langle 1, \sqrt{2}, \sqrt{3} \right\rangle$

2.19. Find the projection of $\mathbf{a} = \langle 1, -1, 4 \rangle$ onto $\mathbf{u} = \left\langle \frac{1}{3}, \frac{2}{3}, -\frac{2}{3} \right\rangle$.

2.20. Find the projection of $\mathbf{a} = \langle 2, -1, 2 \rangle$ onto the direction of $\mathbf{b} = \langle 2, -1, 0 \rangle$.

Find $\mathbf{a} \times \mathbf{b}$ for the given vectors $\mathbf{a}$ and $\mathbf{b}$.

2.21. $\mathbf{a} = \langle 2, -3, 1 \rangle$, $\mathbf{b} = \langle 1, 1, 2 \rangle$

2.22. $\mathbf{a} = 3\mathbf{e}_1 + 2\mathbf{e}_2 + 2\mathbf{e}_3$, $\mathbf{b} = -\mathbf{e}_1 + 2\mathbf{e}_2 + \mathbf{e}_3$

2.23. $\mathbf{a} = \mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b} = 3\mathbf{e}_2 + \mathbf{e}_3$

2.24. $\mathbf{a} = \langle 1, -1, 2 \rangle$, $\mathbf{b} = \langle 1, 0, -3 \rangle$

Let $\mathbf{a} = \langle -1, 2, -4 \rangle$, $\mathbf{b} = \langle 7, 3, -4 \rangle$, and $\mathbf{c} = \langle -2, 1, 0 \rangle$. Evaluate the given expression or explain why it is undefined.

2.25. $\mathbf{a} \cdot (2\mathbf{b} + \mathbf{c})$

2.26. $\|\mathbf{b} \times (2\mathbf{a})\|$

2.27. $\mathbf{b} \times (\mathbf{c} \times \mathbf{a})$

2.28. $\mathbf{c} \cdot (-2\mathbf{a} \times \mathbf{b})$

2.29. $\|2\mathbf{c}\| \, (\mathbf{a} \times (\mathbf{a} + 3\mathbf{b}))$

2.30. $\mathbf{a} \times (\mathbf{c} \cdot \mathbf{b})$

2.31. Find a unit vector $\mathbf{u}$ orthogonal to both $\mathbf{x} = \langle -2, 3, -6 \rangle$ and $\mathbf{y} = \langle 2, 2, -1 \rangle$.

(a) by using the cross product;

(b) without using the cross product.

2.32. Find the area of the triangle with vertices $A(3, 2, 1)$, $B(2, 4, 5)$, and $C(3, 1, 4)$.

2.33. Find the volume of the parallelepiped with adjacent edges $OA$, $OB$, and $OC$, where $A(2, 3, 0)$, $B(3, -1, 2)$, and $C(1, 0, -4)$.

2.34. Use the method from the second solution of Example 2.6 to find $\mathbf{r} \times \mathbf{s}$, where

$$\mathbf{r} = \cos\phi\cos\theta\mathbf{e}_1 + \cos\phi\sin\theta\mathbf{e}_2 - \sin\phi\mathbf{e}_3, \quad \mathbf{s} = -\sin\phi\sin\theta\mathbf{e}_1 + \sin\phi\cos\theta\mathbf{e}_2.$$

2.35. Use the method from the second solution of Example 2.6 to find $\mathbf{u} \times \mathbf{v}$, where

$$\mathbf{u} = \mathbf{e}_1 + a\cos\phi\mathbf{e}_2 + a\sin\phi\mathbf{e}_3, \quad \mathbf{v} = -b\sin\phi\mathbf{e}_2 + b\cos\phi\mathbf{e}_3.$$

2.36. Let $\mathbf{r} = \langle x, y, z \rangle$ and $\mathbf{a} = \langle a, b, c \rangle$. Describe the set of points $P(x, y, z)$ such that $\|\mathbf{r} - \mathbf{a}\| = R$.

2.37. Prove the *Cauchy–Schwarz inequality*: If $\mathbf{a}, \mathbf{b}$ are three-dimensional vectors, then

$$|\mathbf{a} \cdot \mathbf{b}| \leq \|\mathbf{a}\| \, \|\mathbf{b}\|.$$

2.38. Prove the *triangle inequality*: If $\mathbf{a}, \mathbf{b}$ are three-dimensional vectors, then

$$\|\mathbf{a} + \mathbf{b}\| \leq \|\mathbf{a}\| + \|\mathbf{b}\|.$$

2.39. Are the following statements true or false?

(a) If $\mathbf{a} \neq 0$ and $\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{c}$, then it follows that $\mathbf{b} = \mathbf{c}$.
(b) If $\mathbf{a} \neq 0$ and $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \mathbf{c}$, then it follows that $\mathbf{b} = \mathbf{c}$.
(c) If $\mathbf{a} \neq 0$, $\mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \mathbf{c}$, and $\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \mathbf{c}$, then it follows that $\mathbf{b} = \mathbf{c}$.

LECTURE 3

# Lines and Planes

## 3.1. Equations of a straight line in space

In our discussion of lines and planes, we shall use the notion of a position vector. A *position vector* is a vector whose initial point is at the origin. Thus, the components of a position vector are the same as the coordinates of its terminal point.

We shall specify a line in space using a point on the line and the direction of the line. If $\mathbf{r}_0$ is the position vector of a point $P_0(x_0, y_0, z_0)$ on the line and if $\mathbf{v}$ is a vector parallel to the line, then the terminal point of a position vector $\mathbf{r}$ is on the line if and only if

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{v} \qquad \text{for some } t \in \mathbb{R}. \tag{3.1}$$

We call (3.1) a *parametric vector equation* of the line $\ell$ through $P_0$ and parallel to $\mathbf{v}$. In this equation, we construct a generic point $P$ on $\ell$ in two steps: first we get on $\ell$ by going to the given point $P_0$; then we slide along $\ell$ by a multiple of the given vector $\mathbf{v}$ until we reach $P$ (see Figure 3.1). The value of the parameter $t$ indicates the distance along $\ell$ (measured in units equal to $\|\mathbf{v}\|$) that we need to move to reach $P$: for example, $t = 0$ corresponds to $P = P_0$ and $t = -2$ corresponds to a point that is at a distance $2\|\mathbf{v}\|$ from $P_0$ in the direction opposite to $\mathbf{v}$.

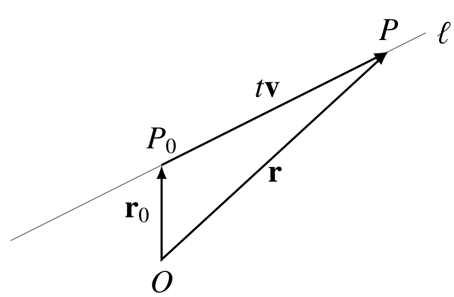



FIGURE 3.1. The parametric vector equation of a line

Writing $\mathbf{v}$ and $\mathbf{r}$ in component form, $\mathbf{v} = \langle a, b, c\rangle$ and $\mathbf{r} = \langle x, y, z\rangle$, we can express (3.1) in terms of the components of the vectors involved:

$$\langle x, y, z\rangle = \langle x_0, y_0, z_0\rangle + t\,\langle a, b, c\rangle = \langle x_0 + ta, y_0 + tb, z_0 + tc\rangle\,,$$

or equivalently:

$$x = x_0 + ta, \quad y = y_0 + tb, \quad z = z_0 + tc \qquad (t \in \mathbb{R}). \tag{3.2}$$

These are the *parametric scalar equations* of $\ell$. In general, the vector $\mathbf{v} = \langle a, b, c\rangle$ is determined up to multiplication by a nonzero scalar, that is, we can use any nonzero scalar multiple of $\mathbf{v}$ in place of $\mathbf{v}$. This will change the coefficients in equations (3.1) and (3.2), but not the underlying set of points. We shall refer to the coordinates $a, b, c$ of any vector $\mathbf{v}$ parallel to $\ell$ as *direction numbers* of $\ell$.

Another common way to specify a line $\ell$ in space are its symmetric equations. Given the scalar equations (3.2), with $abc \neq 0$, we can eliminate $t$ to obtain the equations

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}. \tag{3.3}$$

These are the *symmetric equations* of $\ell$. If one of $a, b$, or $c$ is zero—say, $b = 0$, we obtain the symmetric equations of $\ell$ by keeping that scalar equation and eliminating $t$ from the other two:

$$y = y_0, \quad \frac{x - x_0}{a} = \frac{z - z_0}{c}.$$

If two among $a, b, c$ are equal to zero, the corresponding scalar equations represent the symmetric equations of $\ell$ after discarding the third parametric equation: if $a = b = 0$, the symmetric equations of $\ell$ are

$$x = x_0, \quad y = y_0.$$

Finally, we can describe a line by the coordinates of any two distinct points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$ on the line. If $P_1$ and $P_2$ are such points, the vector $\overrightarrow{P_1P_2} = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1\rangle$ is parallel to the line, so we can write the symmetric equations (3.3) in the form

$$\frac{x - x_1}{x_2 - x_1} = \frac{y - y_1}{y_2 - y_1} = \frac{z - z_1}{z_2 - z_1}, \tag{3.4}$$

(or something like it, for other types of symmetric equations).

Example 3.1. Find the equations of the line $\ell$ through the points $P(1, 0, 3)$ and $Q(2, -1, 4)$.

Solution. To obtain the symmetric equations, we note that $\overrightarrow{PQ} = \langle 1, -1, 1\rangle$ is a vector parallel to $\ell$, so (3.3) can be written as

$$\frac{x - 1}{1} = \frac{y - 0}{-1} = \frac{z - 3}{1} \quad \Longleftrightarrow \quad x - 1 = -y = z - 3.$$

The parametric scalar equations (based again on $\overrightarrow{PQ}$) are

$$x = 1 + t, \quad y = 0 + t(-1) = -t, \quad z = 3 + t \qquad (t \in \mathbb{R}),$$

and the parametric vector equation is

$$\mathbf{r}(t) = \langle 1, 0, 3\rangle + t\,\langle 1, -1, 1\rangle = \langle 1 + t, -t, 3 + t\rangle \qquad (t \in \mathbb{R}).$$ □

Example 3.2. Find a parametric vector equation of the line $\ell$ passing through $P(-1, 2, -3)$ and parallel to the line $2(x + 1) = 4(y - 3) = z$.

Solution. We can write the equations of the second line as

$$\frac{x + 1}{\frac{1}{2}} = \frac{y - 3}{\frac{1}{4}} = z,$$

so $\mathbf{v} = \left\langle \frac{1}{2}, \frac{1}{4}, 1\right\rangle$ is a vector parallel to the second line. Thus, $\mathbf{v}$ is also parallel to $\ell$. That is, $\ell$ passes through $P$ and is parallel to $\mathbf{v}$. Its vector equation then is

$$\mathbf{r}(t) = \langle -1, 2, -3\rangle + t\left\langle \frac{1}{2}, \frac{1}{4}, 1\right\rangle = \left\langle -1 + \frac{1}{2}t, 2 + \frac{1}{4}t, -3 + t\right\rangle \qquad (t \in \mathbb{R}).$$ □

EXAMPLE 3.3. Find the point of intersection and the angle between the lines with the vector parametrizations

$$\mathbf{r}_1(t) = \langle 1 + t, -1 - t, -4 + 2t \rangle \qquad \text{and} \qquad \mathbf{r}_2(u) = \langle 1 - u, 1 + 3u, 2u \rangle\,.$$

SOLUTION. If it exists, the point of intersection of the given lines will have coordinates $(x, y, z)$ such that

$$x = 1 + t = 1 - u, \quad y = -1 - t = 1 + 3u, \quad z = -4 + 2t = 2u,$$

for some $t, u$. From the equations for $x$, we find that $t = -u$. Hence, the other two equations between $t$ and $u$ become

$$-1 + u = 1 + 3u, \quad -4 - 2u = 2u \qquad \Longrightarrow \qquad u = -1, \quad t = -u = 1.$$

Therefore, the point of intersection is $(2, -2, -2)$.

To find the angle between the lines, we find the angle $\theta$ between their parallel vectors: $\mathbf{v} = \langle 1, -1, 2 \rangle$ for the first line and $\mathbf{u} = \langle -1, 3, 2 \rangle$ for the second. Thus, by Corollary 2.4,

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \frac{-1 - 3 + 4}{\sqrt{6}\sqrt{14}} = 0 \qquad \Longrightarrow \qquad \theta = 90^\circ.$$

□

## 3.2. Equations of a plane

We shall specify a plane in space using a point on the plane and a direction perpendicular to the plane. If $\mathbf{r}_0$ is the position vector of a point $P_0(x_0, y_0, z_0)$ on the plane and $\mathbf{n} = \langle a, b, c \rangle$ is a vector perpendicular (also called *normal*) to the plane, then a generic point $P(x, y, z)$ is on the plane if and only if $\overrightarrow{P_0P}$ is perpendicular to $\mathbf{n}$ (see Fig. 3.2). Introducing the position vector $\mathbf{r}$ of $P$, we can express this condition as

$$\overrightarrow{P_0P} \cdot \mathbf{n} = 0 \quad \Longleftrightarrow \quad (\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n} = 0, \tag{3.5}$$

or equivalently,

$$\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0. \tag{3.6}$$

These are the *vector equations* of the given plane.

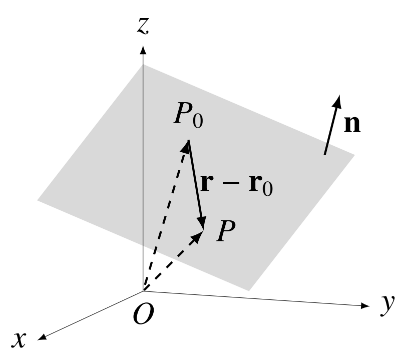



FIGURE 3.2. The vector equation of a plane

If we express the dot product in (3.5) in terms of the components of the vectors, we get

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0, \tag{3.7}$$

or the *scalar equation* of the given plane. Given the last equation, we can easily obtain the plane's *linear equation*:

$$ax + by + cz + d = 0, \tag{3.8}$$

simply by writing $d = -ax_0 - by_0 - cz_0$.

Example 3.4. Find the equation of the plane that contains $P(1, -2, 3)$ and is perpendicular to $\mathbf{e}_2 + 2\mathbf{e}_3$.

Solution. The scalar equation of this plane is

$$0(x - 1) + 1(y + 2) + 2(z - 3) = 0 \quad\Longleftrightarrow\quad y + 2z - 4 = 0. \qquad\square$$

Example 3.5. Find the equation of the plane $\pi$ that contains $P(3, -1, 5)$ and is parallel to the plane with equation $4x + 2y - 7z + 5 = 0$.

Solution. A normal vector for the given plane is $\mathbf{n} = \langle 4, 2, -7\rangle$. Hence, the scalar equation of $\pi$, which also has $\mathbf{n}$ as its normal vector, is

$$4(x - 3) + 2(y + 1) - 7(z - 5) = 0 \quad\Longleftrightarrow\quad 4x + 2y - 7z + 25 = 0. \qquad\square$$

Example 3.6. Find the equation of the plane that contains the points $P(1, 2, 3)$, $Q(-1, 3, 2)$, and $R(3, -1, 2)$.

Solution. The cross product $\overrightarrow{PQ} \times \overrightarrow{PR}$ is a normal vector for this plane. By Example 2.8,

$$\overrightarrow{PQ} \times \overrightarrow{PR} = -4\mathbf{e}_1 - 4\mathbf{e}_2 + 4\mathbf{e}_3,$$

so we can use the normal vector $\mathbf{n} = \langle 1, 1, -1\rangle$. We obtain the scalar equation

$$(x - 1) + (y - 2) - (z - 3) = 0 \quad\Longleftrightarrow\quad x + y - z = 0. \qquad\square$$

Example 3.7. Find the angle between the planes $2x - y + 3z = 5$ and $5x + 5y - z = 1$.

Solution. The angle between the two planes equals the angle between their normal vectors: $\mathbf{n}_1 = \langle 2, -1, 3\rangle$ and $\mathbf{n}_2 = \langle 5, 5, -1\rangle$, respectively. If $\theta$ is the angle between $\mathbf{n}_1$ and $\mathbf{n}_2$, we use the dot product to find

$$\cos\theta = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{\|\mathbf{n}_1\|\,\|\mathbf{n}_2\|} = \frac{10 - 5 - 3}{\sqrt{14}\,\sqrt{51}} = \frac{2}{\sqrt{714}}.$$

$\square$

Example 3.8. Find the line of intersection of the planes $x + y + z + 1 = 0$ and $x - y + z + 2 = 0$.

Solution. We shall compute the symmetric equations of the given line. First, by adding the two equations, we obtain the equation

$$2x + 2z + 3 = 0 \quad\Longleftrightarrow\quad x = -z - 3/2 \quad\Longleftrightarrow\quad x = \frac{z + 3/2}{-1}.$$

On the other hand, subtracting the two given equations, we get

$$2y - 1 = 0 \quad\Longleftrightarrow\quad y = 1/2.$$

Hence, the symmetric equations of the line of intersection are

$$x = \frac{z + 3/2}{-1}, \quad y = 1/2.$$

Note that these symmetric equations lead to the parametric scalar equations

$$x = t, \quad y = 1/2, \quad z = -3/2 - t \qquad (t \in \mathbb{R}). \qquad\square$$

Example 3.9. Find the distance from $P(3, -5, 2)$ to the plane $x - 2y + z = 5$.

First solution. The distance from $P$ to the given plane is equal to the distance $|PQ|$, where $Q$ is the intersection point of the plane and the line through $P$ that is perpendicular to the plane. The normal vector to the plane $\mathbf{n} = \langle 1, -2, 1\rangle$ is parallel to the line $PQ$, so the symmetric equations of that line are

$$\frac{x-3}{1} = \frac{y+5}{-2} = \frac{z-2}{1}.$$

Combining these equations with the equation of the plane, we obtain a linear system

$$x - 3 = \frac{y+5}{-2}, \quad z - 2 = \frac{y+5}{-2}, \quad x - 2y + z = 5,$$

whose solution $(\frac{4}{3}, -\frac{5}{3}, \frac{1}{3})$ is the point $Q$. Therefore,

$$|PQ| = \left((-\tfrac{5}{3})^2 + (\tfrac{10}{3})^2 + (-\tfrac{5}{3})^2\right)^{1/2} = \tfrac{5}{3}\sqrt{6}. \qquad \square$$

We now give a second solution, which is somewhat trickier, but has the advantage that can be easily generalized (see the next example).

Second solution. Let $Q(u, v, w)$ be the same point as before. In this solution, we avoid the explicit computation of $u, v, w$.

We have $\overrightarrow{PQ} = \langle u - 3, v + 5, w - 2\rangle$. By the definition of $Q$, we know that $\overrightarrow{PQ}$ is normal to the given plane, so $\overrightarrow{PQ}$ is parallel to the normal vector $\mathbf{n} = \langle 1, -2, 1\rangle$. Hence, by the properties of the dot product,

$$\left|\mathbf{n} \cdot \overrightarrow{PQ}\right| = \|\mathbf{n}\|\,|PQ|\,\left|\cos \sphericalangle(\mathbf{n}, \overrightarrow{PQ})\right| = \sqrt{6}\,|PQ|,$$

since the angle between the two vectors is 0° or 180°. On the other hand, using the definition of the dot product, we have

$$\left|\mathbf{n} \cdot \overrightarrow{PQ}\right| = |(u-3) - 2(v+5) + (w-2)| = |u - 2v + w - 15| = |5 - 15| = 10.$$

Here, we have used that $u - 2v + w = 5$, because $Q$ lies on the given plane. Comparing the two expressions for $\left|\mathbf{n} \cdot \overrightarrow{PQ}\right|$, we conclude that $|PQ| = 10/\sqrt{6}$. $\square$

Example 3.10. Find the distance from $P(x_0, y_0, z_0)$ to the plane $ax + by + cz + d = 0$.

Solution*. We essentially repeat the second solution of the previous example. Let $Q(u, v, w)$ be the intersection point of the plane and the line through $P$ that is perpendicular to the plane. Then $\overrightarrow{PQ} = \langle u - x_0, v - y_0, w - z_0\rangle$. Since $\overrightarrow{PQ}$ is normal to the given plane, it is parallel to its normal vector $\mathbf{n} = \langle a, b, c\rangle$. Hence, by the properties of the dot product,

$$\left|\mathbf{n} \cdot \overrightarrow{PQ}\right| = \|\mathbf{n}\|\,|PQ|\,\left|\cos \sphericalangle(\mathbf{n}, \overrightarrow{PQ})\right| = \sqrt{a^2 + b^2 + c^2}\,|PQ|,$$

since the angle between the two vectors is 0° or 180°. On the other hand, using the definition of the dot product, we have

$$\begin{aligned}\left|\mathbf{n} \cdot \overrightarrow{PQ}\right| &= |a(u - x_0) + b(v - y_0) + c(w - z_0)| \\ &= |au + bv + cw - ax_0 - by_0 - cz_0| \\ &= |-d - ax_0 - by_0 - cz_0| = |ax_0 + by_0 + cz_0 + d|.\end{aligned}$$

Here, we have used that $au + bv + cw = -d$, because $Q$ lies on the given plane. Comparing the two expressions for $\left|\mathbf{n} \cdot \overrightarrow{PQ}\right|$, we conclude that the desired distance is

$$|PQ| = \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

□

## Exercises

3.1. Find the equations of the line that passes through the point $(1, -2, 3)$ and is perpendicular to the plane with equation $-2x + y - z + 4 = 0$.

3.2. Find the equations of the line that passes through the points $(0, -2, 3)$ and $(1, 1, 5)$.

3.3. Let $\ell_1$ be the line with equations $x = -1 + 2t$, $y = 3t$, $z = 2$ and let $\ell_2$ be the line through the points $(4, 3, 2)$ and $(6, -1, 4)$. Determine whether these two lines are parallel, intersecting, or skew. If they intersect, find their intersection point.

3.4. Find the intersection points of the line with equations $\frac{x-3}{2} = \frac{y+1}{5} = z$ and the coordinate planes.

3.5. Find the angle between the lines whose respective equations are

$$\frac{x+1}{-3} = \frac{y+2}{2} = \frac{z-2}{5} \quad \text{and} \quad \frac{x+1}{2} = \frac{y+2}{1} = \frac{z-2}{-2}.$$

3.6. Explain why the two lines from Exercise 3.5 intersect.

3.7. Find an equation of the plane that passes through the point $(1, -2, 3)$ and is perpendicular to the vector $\langle 3, 0, -2 \rangle$.

3.8. Find an equation of the plane that contains the line with equations $\frac{x-1}{2} = \frac{y+2}{3} = \frac{z-2}{4}$ and is parallel to the plane $3x - 2y = 5$.

3.9. Find an equation of the plane that contains the line with equations $\frac{x-1}{2} = \frac{y+2}{-1} = \frac{z}{4}$ and is perpendicular to the line $x = 1 - 2t,\ y = 4t,\ z = -3 + 2t$.

3.10. Find an equation of the plane that contains the line with equations $\frac{x-1}{2} = \frac{y+2}{3} = \frac{z}{4}$ and is perpendicular to the line $x = 1 - 2t,\ y = 3t,\ z = -3 + 2t$.

3.11. Find an equation of the plane through the points $(1, -1, 0)$, $(2, 0, 3)$, and $(-2, 2, 1)$.

3.12. Find an equation of the plane through the points $(1, 2, -1)$, $(-1, 0, 1)$, and $(2, 3, 0)$.

3.13. Find the equations of the intersection line of the planes $x + y - z = 2$ and $2x - 3y + z = 6$.

3.14. Find an equation of the plane that contains both the line $\ell_1$ with equations $x = 1 + 2t$, $y = -1 + 3t$, $z = -2t$ and the line $\ell_2$ with equations $x = -2 - 3t$, $y = t$, $z = 4 + 4t$.

3.15. Find an equation of the plane that contains both the line $\ell_1$ with equations $x = 1 - 2t$, $y = 2 + t$, $z = -1 - t$ and the line $\ell_2$ with equations $x = 1 + t$, $y = 2 - 3t$, $z = -1$.

Find the angle between the given planes.

3.16. $2x - y + 3z = 0$, $x + 2y + z = 4$

3.17. $x + y + z = 4$, $x + 2y - 3z = 6$

3.18. The planes $2x - 2y - z = 1$ and $6x - 6y - 3z = -5$ are parallel. Find the distance between them.

LECTURE 4

# Vector Functions

## 4.1. Definition

A *vector-valued function* (or a *vector function*) is a function whose domain is a set of real numbers and whose range is a set of vectors. In general, a vector function has the form

$$\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle = f_1(t)\mathbf{e}_1 + f_2(t)\mathbf{e}_2 + f_3(t)\mathbf{e}_3.$$

where $f_1$, $f_2$, and $f_3$ are functions in the regular sense (functions for which both the inputs and outputs are real numbers). The functions $f_1$, $f_2$, and $f_3$ are sometimes called the *component functions* of the vector function $\mathbf{f}$. When we want to distinguish that the values of a function $f$ are numbers (as opposed to vectors), we will refer to $f$ as a *scalar function*.

Example 4.1. As we know from Lecture #3, the parametric equation of the line in space passing through $(1, 0, -5)$ and parallel to the vector $\langle 1, 2, 1\rangle$ is, in fact, a vector function:

$$\mathbf{r}(t) = \langle 1 + t, 2t, -5 + t\rangle \qquad (t \in \mathbb{R}).$$

The line is the set of values of this function, **not** the graph (the graph is four-dimensional).

Here is another example of a vector function:

$$\mathbf{f}(t) = \left\langle \cos t, (t-5)^{-1}, \ln t\right\rangle ;$$

its domain is $(0, 5) \cup (5, \infty)$. □

## 4.2. Calculus of vector functions: Limits

Vector functions can have limits, derivatives, antiderivatives, and definite integrals—just like scalar functions. As we will see, all of these can be computed componentwise. Because of that, once one is proficient in the calculus of scalar-valued functions, one can easily master the calculus of vector functions. We begin with limits and continuity. The limit of a vector function is a vector.

Definition. If $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$ is defined near $a$, then the *limit of* $\mathbf{f}$ *as* $t \to a$ is defined by

$$\lim_{t\to a}\mathbf{f}(t) = \left\langle \lim_{t\to a} f_1(t), \lim_{t\to a} f_2(t), \lim_{t\to a} f_3(t)\right\rangle ,$$

provided that the limits of all three component functions exist.

Definition. A vector function $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$, defined near $a$, is *continuous at* $a$ if

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

Example 4.2. We have

$$\lim_{t\to 7}\left\langle \cos t, (t-5)^{-1}, \ln t\right\rangle = \left\langle \lim_{t\to 7}\cos t, \lim_{t\to 7}(t-5)^{-1}, \lim_{t\to 7}\ln t\right\rangle = \left\langle \cos 7, \tfrac{1}{2}, \ln 7\right\rangle .$$

## 4.3. Calculus of vector functions: Derivatives

Definition. If $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$ is continuous at $a$, then its *derivative* $\mathbf{f}'(a)$ *at* $a$ is defined by

$$\mathbf{f}'(a) = \lim_{t\to a} \frac{\mathbf{f}(t) - \mathbf{f}(a)}{t-a} = \lim_{h\to 0} \frac{\mathbf{f}(a+h) - \mathbf{f}(a)}{h},$$

provided that the limit on the right exists. If $\mathbf{f}'(a)$ exists, we say that $\mathbf{f}$ is *differentiable at* $a$.

Theorem 4.1. *A vector function* $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$ *is differentiable at* $a$ *if and only if all three component functions* $f_1$, $f_2$, $f_3$ *are. When* $\mathbf{f}'(a)$ *exists, we have*

$$\mathbf{f}'(a) = \left\langle f_1'(a), f_2'(a), f_3'(a)\right\rangle .$$

*In other words, the derivative of a vector function can be computed using componentwise differentiation.*

Proof. We have

$$\begin{aligned}
\mathbf{f}'(a) &= \lim_{t\to a} \frac{\mathbf{f}(t) - \mathbf{f}(a)}{t-a} \\
&= \lim_{t\to a} \frac{\langle f_1(t) - f_1(a), f_2(t) - f_2(a), f_3(t) - f_3(a)\rangle}{t-a} \\
&= \lim_{t\to a} \left\langle \frac{f_1(t) - f_1(a)}{t-a}, \frac{f_2(t) - f_2(a)}{t-a}, \frac{f_3(t) - f_3(a)}{t-a} \right\rangle \\
&= \left\langle f_1'(a), f_2'(a), f_3'(a)\right\rangle . \qquad \square
\end{aligned}$$

Example 4.3. If $\mathbf{f}(t) = \left\langle \cos t, (t-5)^{-1}, \ln t\right\rangle$, then

$$\mathbf{f}'(t) = \left\langle -\sin t, -(t-5)^{-2}, t^{-1}\right\rangle .$$

Theorem 4.2. *Let* $\mathbf{f}$ *and* $\mathbf{g}$ *be differentiable vector functions, let* $u$ *be a differentiable scalar function, and let* $c$ *be a scalar. Then the following formulas hold:*

i) $\left[\mathbf{f}(t) + \mathbf{g}(t)\right]' = \mathbf{f}'(t) + \mathbf{g}'(t)$;
ii) $\left[c\mathbf{f}(t)\right]' = c\mathbf{f}'(t)$;
iii) $\left[u(t)\mathbf{f}(t)\right]' = u(t)\mathbf{f}'(t) + u'(t)\mathbf{f}(t)$;
iv) $\left[\mathbf{f}(t)\cdot\mathbf{g}(t)\right]' = \mathbf{f}(t)\cdot\mathbf{g}'(t) + \mathbf{f}'(t)\cdot\mathbf{g}(t)$;
v) $\left[\mathbf{f}(t)\times\mathbf{g}(t)\right]' = \mathbf{f}(t)\times\mathbf{g}'(t) + \mathbf{f}'(t)\times\mathbf{g}(t)$;
vi) $\left[\mathbf{f}(u(t))\right]' = \mathbf{f}'(u(t))u'(t)$.

Proof. We shall prove iv). If $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$ and $\mathbf{g}(t) = \langle g_1(t), g_2(t), g_3(t)\rangle$, then

$$\mathbf{f}(t)\cdot\mathbf{g}(t) = f_1(t)g_1(t) + f_2(t)g_2(t) + f_3(t)g_3(t).$$

Hence, the left side of iv) equals

$$\begin{aligned}
\left[\mathbf{f}(t)\cdot\mathbf{g}(t)\right]' &= \left(f_1(t)g_1(t) + f_2(t)g_2(t) + f_3(t)g_3(t)\right)' \\
&= f_1'(t)g_1(t) + f_1(t)g_1'(t) + f_2'(t)g_2(t) + f_2(t)g_2'(t) + f_3'(t)g_3(t) + f_3(t)g_3'(t).
\end{aligned}$$

On the other hand, we have

$$\mathbf{f}'(t)\cdot\mathbf{g}(t) = \left\langle f_1'(t), f_2'(t), f_3'(t)\right\rangle \cdot \langle g_1(t), g_2(t), g_3(t)\rangle = f_1'(t)g_1(t) + f_2'(t)g_2(t) + f_3'(t)g_3(t),$$

$$\mathbf{f}(t)\cdot\mathbf{g}'(t) = \langle f_1(t), f_2(t), f_3(t)\rangle \cdot \left\langle g_1'(t), g_2'(t), g_3'(t)\right\rangle = f_1(t)g_1'(t) + f_2(t)g_2'(t) + f_3(t)g_3'(t),$$

so the right side of iv) equals

$$f_1'(t)g_1(t) + f_2'(t)g_2(t) + f_3'(t)g_3(t) + f_1(t)g_1'(t) + f_2(t)g_2'(t) + f_3(t)g_3'(t).$$

This establishes iv). □

## 4.4. Calculus of vector functions: Integrals

The definition of the definite integral of a continuous function via Riemann sums (see Appendix B) can also be extended to vector functions. Here, we will adopt an equivalent definition, which is a little less transparent but considerably easier to use.

Definition. If $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$ is continuous in $[a, b]$, then its *definite integral* $\int_a^b \mathbf{f}(t)\,dt$ is defined by

$$\int_a^b \mathbf{f}(t)\,dt = \left\langle \int_a^b f_1(t)\,dt, \int_a^b f_2(t)\,dt, \int_a^b f_3(t)\,dt \right\rangle.$$

Remark. Note that $\int_a^b \mathbf{f}(t)\,dt$ is a **vector** (not a scalar and not a vector function).

Example 4.4. If $\mathbf{f}(t) = \left\langle \cos t, (t-5)^{-1}, \ln t\right\rangle$, then

$$\begin{aligned}\int_1^2 \mathbf{f}(t)\,dt &= \left\langle \int_1^2 \cos t\,dt, \int_1^2 (t-5)^{-1}\,dt, \int_1^2 \ln t\,dt \right\rangle \\ &= \left\langle \big[\sin t\big]_1^2, \big[\ln|t-5|\big]_1^2, \big[t\ln t - t\big]_1^2 \right\rangle \\ &= \left\langle \sin 2 - \sin 1, \ln(\tfrac{3}{4}), 2\ln 2 - 1\right\rangle.\end{aligned}$$

Theorem 4.3. *Let* $\mathbf{f}$ *and* $\mathbf{g}$ *be continuous vector functions, let* $\mathbf{u}$ *be a vector, and let* $c$ *be a scalar. Then the following formulas hold:*

i) $\int_a^b \big[\mathbf{f}(t) + \mathbf{g}(t)\big]\,dt = \int_a^b \mathbf{f}(t)\,dt + \int_a^b \mathbf{g}(t)\,dt;$
ii) $\int_a^b \big[c\mathbf{f}(t)\big]\,dt = c\int_a^b \mathbf{f}(t)\,dt;$
iii) $\int_a^b \big[\mathbf{u}\cdot\mathbf{f}(t)\big]\,dt = \mathbf{u}\cdot\int_a^b \mathbf{f}(t)\,dt;$
iv) $\int_a^b \big[\mathbf{u}\times\mathbf{f}(t)\big]\,dt = \mathbf{u}\times\int_a^b \mathbf{f}(t)\,dt;$
v) $\int_a^b \mathbf{f}'(t)\,dt = \mathbf{f}(b) - \mathbf{f}(a).$

Proof. We shall prove iv). Both sides of the identity are three-dimensional vectors, so it suffices to show that their respective components are equal. If $\mathbf{u} = \langle u_1, u_2, u_3\rangle$ and $\mathbf{f}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle$, then

$$\mathbf{u}\times\mathbf{f}(t) = \langle u_2 f_3(t) - u_3 f_2(t), u_3 f_1(t) - u_1 f_3(t), u_1 f_2(t) - u_2 f_1(t)\rangle.$$

Thus, the first component of the left side of iv) is

$$\int_a^b \big(u_2 f_3(t) - u_3 f_2(t)\big)\,dt = u_2\int_a^b f_3(t)\,dt - u_3\int_a^b f_2(t)\,dt.$$

Since the right side of this identity is also the first component of the right side of iv), this shows that the first components of the two sides of iv) match. The proof that the remaining two components of the two sides of iv) match is similar. □

## 4.5. Vector functions in *Mathematica*

Defining and operating with vector functions in *Mathematica* is no different from defining and operating with scalar functions. For example, to define the function $\mathbf{f}(t) = \left\langle \cos t, (t-5)^{-1}, \ln t \right\rangle$ from Examples 4.2–4.4, we simply use the command

```
f[t_] := {Cos[t], 1/(t-5), Log[t]}
```

We can then use the commands `Limit`, `D`, `Integrate` and `NIntegrate` on the vector function `f[t]` just as we use them on scalar-valued functions:

```
 In[1]:= Limit[f[t], t->1]
Out[1]:= {Cos[1], -(1/4), 0}

 In[2]:= f'[t]
Out[2]:= {-Sin[t], -(1/(-5+t)^2), 1/t}

 In[3]:= D[f[2t], t]
Out[3]:= {-2Sin[2t], -(2/(-5+2t)^2), 1/t}

 In[4]:= NIntegrate[f[t], {t,1,4}]
Out[4]:= {-1.59827, -1.38629, 2.54518}

 In[5]:= Integrate[f[t], {t,0,1}]
Out[5]:= {Sin[1], -Log[5/4], -1}
```

### Exercises

4.1. Evaluate $\lim_{t\to a} \left\langle \frac{\ln(t+1)}{t}, \frac{\sin 2t}{t}, e^t \right\rangle$, where: (a) $a = 1$; (b) $a = 0$.

Find the derivative of the given vector function.

4.2. $\mathbf{f}(t) = \left\langle \ln\left(t^2+1\right), \arcsin 2t, e^t \right\rangle$

4.3. $\mathbf{f}(t) = \left\langle 5t^3 + 2t, e^{t^2} \sin t, e^{t^2} \cos t \right\rangle$

4.4. $\mathbf{f}(t) = t^3\mathbf{e}_1 - 3e^{2t}\mathbf{e}_3$

4.5. $\mathbf{f}(t) = \sqrt{t^2+1}\mathbf{e}_1 - \cos(2t)\mathbf{e}_2 + \sin(2t)\mathbf{e}_3$

Evaluate the given integral.

4.6. $\int_0^1 \left(\left(t^2+4\right)^{-1}\mathbf{e}_1 + t\left(t^2-2\right)^{-1}\mathbf{e}_3\right) dt$

4.7. $\int_{-1}^1 \left\langle 5t^3 + 2t, e^{2t}, \sin t \cos^4 t \right\rangle dt$

4.8. Find $\mathbf{f}(t)$, if $\mathbf{f}'(t) = 2t\mathbf{e}_1 + e^t\mathbf{e}_2 + (\ln t)\,\mathbf{e}_3$ and $\mathbf{f}(1) = \mathbf{e}_1 + \mathbf{e}_2$.

LECTURE 5

# Space Curves

Vector functions are a great tool for dealing with curves, both in the plane and in space. In this lecture, we use vector functions to give a unified treatment of plane curves and space curves.

## 5.1. Review of plane curves

Recall that a curve $\gamma$ in the plane is the set

$$\gamma = \Big\{(x, y) : x = x(t),\ y = y(t),\ t \in I\Big\},$$

where $I$ is some interval in $\mathbb{R}$ (finite or infinite; open, closed, or semiopen) and $x(t)$ and $y(t)$ are continuous functions of $t$. We refer to the pair of equations

$$x = x(t), \quad y = y(t) \qquad (t \in I) \tag{5.1}$$

as *parametric equations* or a *parametrization* of $\gamma$. A curve $\gamma$ has many parametrizations, and sometimes replacing one parametrization by another can change a particular property of the curve, even if the set of points stays the same. Thus, when we want to distinguish a particular parametrization, we talk about the *parametric curve* $\gamma$ with equations (5.1).

## 5.2. Definition of space curve

Next, we extend the above definitions to three dimensions.

Definition. If $x(t), y(t), z(t)$ are continuous functions defined on an interval $I$, then the set

$$\gamma = \Big\{(x, y, z) : x = x(t),\ y = y(t),\ z = z(t),\ t \in I\Big\}$$

is called a *space curve*. The equations

$$x = x(t), \quad y = y(t), \quad z = z(t) \qquad (t \in I) \tag{5.2}$$

are called *parametric equations* or a *parametrization* of $\gamma$; $t$ is the *parameter*.

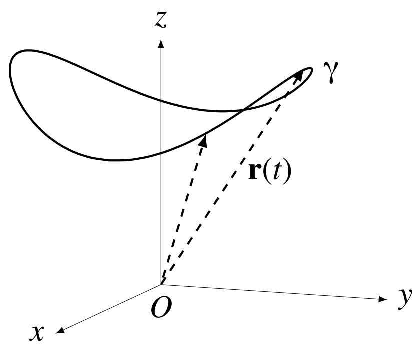
Alternatively, we can think of the curve $\gamma$ with equations (5.2) as the range of the vector function

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3 \qquad (t \in I). \tag{5.3}$$

We shall refer to (5.3) as the *parametric (vector) equation* of $\gamma$. Note that the plane curve (5.1) can be treated as a special case of a space curve given by the parametric equation

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (t \in I).$$

A good way to visualize the connection between a space curve $\gamma$ and its parametrization $\mathbf{r}(t)$ is to think of the vector $\mathbf{r}(t)$ as a position vector which moves as $t$ varies. The terminal point of $\mathbf{r}(t)$ then traces the points on the curve (see Figure 5.1(a), or execute the *Mathematica* code on Figure 5.1(b) for an interactive demonstration).




(a)

```
f[t_] := {t*Cos[t], t*Sin[t], t-1}
Manipulate[
  Show[{
    ParametricPlot3D[f[t], {t, 0, 2Pi}],
    Graphics3D[{Red, Arrow[{{0,0,0}, f[t]}]}]
  }],
  {t, 0, 2Pi}]
```

(b)

FIGURE 5.1. Vector functions and curves in space

EXAMPLE 5.1. Consider the space curve $\gamma$ with parametric equation

$$\mathbf{r}(t) = (\cos t)\mathbf{e}_1 + (\sin t)\mathbf{e}_2 + t\mathbf{e}_3 \qquad (0 \leq t < \infty).$$

The point $(\cos t, \sin t, 0)$ traces the unit circle $x^2 + y^2 = 1$ in the $xy$-plane in the counterclockwise direction, so the point $(\cos t, \sin t, t)$ on $\gamma$ moves in the positive direction of the $z$-axis along a spiral lying above that circle. This curve is known as the *helix*; it is shown on Figure 5.2. □

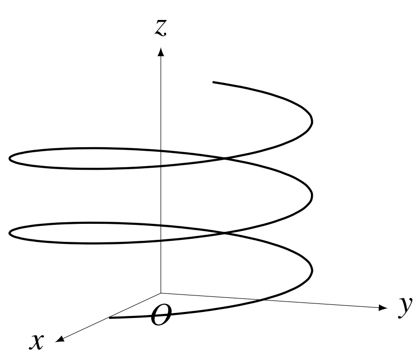



FIGURE 5.2. The helix

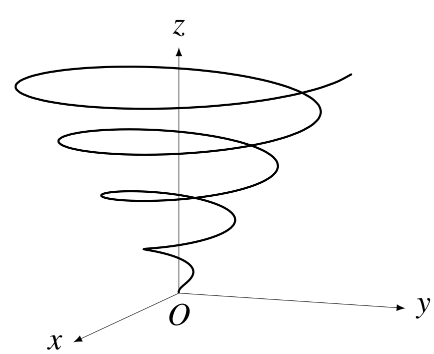



FIGURE 5.3. $\mathbf{r}(t) = \langle t\cos t, t\sin t, t\rangle$

EXAMPLE 5.2. Consider the space curve $\gamma$ with parametric equation

$$\mathbf{r}(t) = (t\cos t)\mathbf{e}_1 + (t\sin t)\mathbf{e}_2 + t\mathbf{e}_3 \qquad (0 \leq t < \infty).$$

The initial point of the curve is the origin. For any given $t$, we have

$$(t\cos t)^2 + (t\sin t)^2 = t^2,$$

so the $x$, $y$, and $z$-coordinates of a point on $\gamma$ always satisfy the equation $x^2 + y^2 = z^2$. Thus, $\gamma$ is an infinite spiral that starts at the origin and unwinds upward and outward; it is shown on Figure 5.3. □

EXAMPLE 5.3. Given two points $P_1(x_1, y_1, z_1)$ and $P_2(x_2, y_2, z_2)$, the space curve

$$\mathbf{r}(t) = (1 - t)\langle x_1, y_1, z_1\rangle + t\langle x_2, y_2, z_2\rangle \qquad (0 \leq t \leq 1)$$

is the line segment with endpoints $P_1$ and $P_2$. Indeed, we can express $\mathbf{r}(t)$ in the form

$$\mathbf{r}(t) = \langle x_1 + t(x_2 - x_1), y_1 + t(y_2 - y_1), z_1 + t(z_2 - z_1)\rangle ,$$

which is the parametric equation of the line through $P_1$ that is parallel to $\overrightarrow{P_1P_2}$. Thus, if the domain for $t$ was the whole real line $\mathbb{R}$, the curve would be the straight line through $P_1$ and $P_2$. Restricting the domain for $t$ to the closed interval $[0, 1]$, we obtain the piece of the line between the endpoints of $\mathbf{r}(0) = \langle x_1, y_1, z_1\rangle$ and $\mathbf{r}(1) = \langle x_2, y_2, z_2\rangle$; that is precisely the line segment $P_1P_2$. □

## 5.3. Smooth curves

At first glance, continuous curves appear to be a good model for the geometric objects that we think of as "curves": that is, two- or three-dimensional objects that are essentially one-dimensional. However, it turns out that there are some rather peculiar continuous parametric curves. For example, in 1890 the Italian mathematician Giuseppe Peano stunned the mathematical community with his construction of a continuous plane curve that fills the entire unit square (see §5.7); there are also examples of continuous space curves that fill an entire cube. In order to rule out such pathological curves, we shall restrict our attention to a special class of parametric curves called "piecewise smooth".

Definition. A curve $\gamma$ given by a vector function $\mathbf{r}(t)$, $t \in I$, is called *smooth* if $\mathbf{r}'(t)$ is continuous and $\mathbf{r}'(t) \neq \mathbf{0}$ (except possibly at the endpoints of $I$). A curve $\gamma$ consisting of a finite number of smooth pieces is called *piecewise smooth*.

Example 5.4. For the helix $\mathbf{r}(t) = (\cos t)\mathbf{e}_1 + (\sin t)\mathbf{e}_2 + t\mathbf{e}_3$, $t \geq 0$, we have

$$\mathbf{r}'(t) = (-\sin t)\mathbf{e}_1 + (\cos t)\mathbf{e}_2 + \mathbf{e}_3 \neq \mathbf{0}.$$

Hence, the helix is a smooth curve. □

Example 5.5. Determine whether the curve $\gamma$ defined by the vector function

$$\mathbf{r}(t) = t^2\mathbf{e}_1 + t^3\mathbf{e}_2 + (\cos t)\mathbf{e}_3 \qquad (t \in \mathbb{R})$$

is smooth, piecewise smooth, or neither.

Solution. The derivative

$$\mathbf{r}'(t) = 2t\mathbf{e}_1 + 3t^2\mathbf{e}_2 - \sin t\mathbf{e}_3$$

vanishes when $t = 0$. Thus, $\gamma$ is not smooth. On the other hand, the curves

$$\gamma_1 : \quad \mathbf{r}(t) = t^2\mathbf{e}_1 + t^3\mathbf{e}_2 + (\cos t)\mathbf{e}_3 \qquad (t \geq 0)$$

and

$$\gamma_2 : \quad \mathbf{r}(t) = t^2\mathbf{e}_1 + t^3\mathbf{e}_2 + (\cos t)\mathbf{e}_3 \qquad (t \leq 0)$$

are smooth, since they fail the condition $\mathbf{r}'(t) \neq \mathbf{0}$ only at an endpoint of the intervals on which they are defined. Since $\gamma$ is the union of $\gamma_1$ and $\gamma_2$, it follows that $\gamma$ is piecewise smooth. □

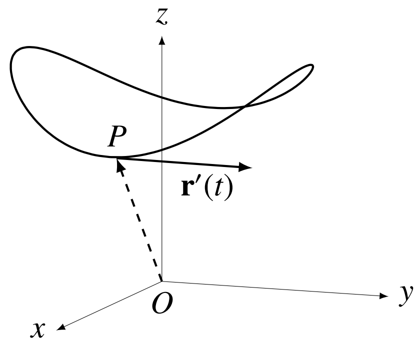



FIGURE 5.4. The tangent vector $\mathbf{r}'(t)$ at $P = \mathbf{r}(t)$

## 5.4. Tangent lines

The requirement that $\mathbf{r}'(t) \neq 0$ for a smooth curve is not arbitrary. To explain where it comes from, we now consider the following problem.

PROBLEM. *Let $\gamma$ be a smooth parametric curve with parametric equation*

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3 \qquad (t \in I).$$

*Find an equation of the tangent line to $\gamma$ at the point $\mathbf{r}(t_0) = (x_0, y_0, z_0)$, where $t_0$ is an internal point of the interval $I$.*

IDEA OF SOLUTION. The tangent line will have the parametric equations

$$x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct \qquad (t \in \mathbb{R}),$$

where $\mathbf{u} = \langle a, b, c\rangle$ is any vector parallel to the tangent line. In other words, we need to determine the direction of the tangent line.

Recall that the tangent line through $P(x_0, y_0, z_0)$ is the limit of the secant lines through $P$ and a point $Q(x, y, z)$ on the curve as $Q \to P$. If $Q = \mathbf{r}(t)$, then a vector parallel to the secant line through $P$ and $Q$ is

$$\overrightarrow{PQ} = \langle x - x_0, y - y_0, z - z_0\rangle = \mathbf{r}(t) - \mathbf{r}(t_0).$$

Thus,

$$\begin{aligned}
\text{direction of tangent line at } P &= \lim_{t\to t_0} \big[\text{direction of secant line through } P, Q\big] \\
&= \lim_{t\to t_0} \big[\text{direction of } \mathbf{r}(t) - \mathbf{r}(t_0)\big] \\
&= \lim_{t\to t_0} \big[\text{direction of normalized } \mathbf{r}(t) - \mathbf{r}(t_0)\big] \\
&= \lim_{t\to t_0} \frac{\mathbf{r}(t) - \mathbf{r}(t_0)}{\|\mathbf{r}(t) - \mathbf{r}(t_0)\|} = \lim_{t\to t_0} \frac{\big(\mathbf{r}(t) - \mathbf{r}(t_0)\big)/(t - t_0)}{\|\mathbf{r}(t) - \mathbf{r}(t_0)\|/(t - t_0)}.
\end{aligned}$$

By the definition of derivative, the limit of the numerator is $\mathbf{r}'(t_0)$. By a slightly more complicated argument (which we omit), the limit of the denominator is $\|\mathbf{r}'(t_0)\|$. Thus, a vector parallel to the tangent line is $\mathbf{r}'(t_0)/\|\mathbf{r}'(t_0)\|$. □

We now see the reason for the requirement of non-vanishing of the derivative in the definition of smoothness: with that requirement, a smooth curve has a tangent line at every point. Furthermore, the vector $\mathbf{r}'(t)$ is parallel to the tangent line to the curve at the endpoint $P$ of $\mathbf{r}(t)$. For this reason, $\mathbf{r}'(t)$ is called also the *tangent vector* to the curve $\mathbf{r}(t)$ (see Figure 5.4). The *unit tangent vector* to $\mathbf{r}(t)$ is merely the normalization of $\mathbf{r}'(t)$, and hence, is given by the formula

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}.$$

Example 5.6. For the helix, we have

$$\mathbf{r}'(t) = (-\sin t)\mathbf{e}_1 + (\cos t)\mathbf{e}_2 + \mathbf{e}_3,$$

so the unit tangent vector is

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|} = \frac{\langle -\sin t, \cos t, 1\rangle}{\sqrt{(-\sin t)^2 + (\cos t)^2 + 1^2}} = \left\langle \frac{-\sin t}{\sqrt{2}}, \frac{\cos t}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle.$$

□

Example 5.7. Find the equation of the tangent line to the helix at the point $P$ corresponding to the value of the parameter $t = \pi$.

Solution. We have $P(-1, 0, \pi)$. The tangent line passes through $P$ and is parallel to the tangent vector $\mathbf{r}'(\pi) = \langle 0, -1, 1\rangle$ (see Example 5.6). Hence, its parametric equation is

$$\mathbf{r}(t) = \langle -1, 0, \pi\rangle + t\,\langle 0, -1, 1\rangle = \langle -1 + t, -t, \pi + t\rangle \qquad (t \in \mathbb{R});$$

the symmetric equations of the tangent line are $x = -1$, $\dfrac{y}{-1} = \dfrac{z - \pi}{1}$.

□

Example 5.8. Find the angle between the helix and the curve $\gamma$ defined by the vector function

$$\mathbf{r}_2(t) = -t\mathbf{e}_1 + (t-1)^2\mathbf{e}_2 + \pi t\mathbf{e}_3 \qquad (t \geq 0)$$

at $P(-1, 0, \pi)$.

Solution. Let $\sigma$ denote the helix. If $P$ lies on both $\gamma$ and $\sigma$, the angle between $\gamma$ and $\sigma$ at $P$ is the angle between their tangent vectors. It is easy to see that both curves do pass through $P$: $\mathbf{r}_1(\pi) = \mathbf{r}_2(1) = \langle -1, 0, \pi\rangle$. We know from Example 5.7 that the tangent vector to the helix at $P$ is $\mathbf{v}_1 = \langle 0, -1, 1\rangle$. The tangent vector to $\gamma$ at $P$ is

$$\mathbf{v}_2 = \mathbf{r}_2'(1) = \langle -1, 2(t-1), \pi\rangle_{t=1} = \langle -1, 0, \pi\rangle\,.$$

Hence, if $\alpha$ is the angle between $\gamma$ and $\sigma$ at $P$, we have

$$\cos\alpha = \frac{\mathbf{v}_1 \cdot \mathbf{v}_2}{\|\mathbf{v}_1\|\,\|\mathbf{v}_2\|} = \frac{0\cdot(-1) + (-1)\cdot 0 + 1\cdot\pi}{\sqrt{0^2 + (-1)^2 + 1^2}\,\sqrt{(-1)^2 + 0^2 + \pi^2}} = \frac{\pi}{\sqrt{2 + 2\pi^2}}.$$

Thus,

$$\alpha = \arccos\left(\frac{\pi}{\sqrt{2 + 2\pi^2}}\right) \approx 0.8315.$$

□

## 5.5. Normal vectors*

We now want to find vectors that are perpendicular to the curve $\mathbf{r}(t)$, that is, we want to find vectors orthogonal to the tangent vector $\mathbf{r}'(t)$, or equivalently, to the unit tangent vector $\mathbf{T}(t)$. One such vector is the *unit normal vector*

$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{\|\mathbf{T}'(t)\|}.$$

In order to prove that $\mathbf{N}(t)$ and $\mathbf{T}(t)$ are orthogonal, we differentiate both sides of the identity

$$\mathbf{T}(t) \cdot \mathbf{T}(t) = \|\mathbf{T}(t)\|^2 = 1.$$

We obtain

$$\mathbf{T}'(t) \cdot \mathbf{T}(t) + \mathbf{T}(t) \cdot \mathbf{T}'(t) = 0 \qquad \Longrightarrow \qquad 2\mathbf{T}(t) \cdot \mathbf{T}'(t) = 0 \qquad \Longrightarrow \qquad \mathbf{T}(t) \cdot \mathbf{T}'(t) = 0.$$

That is, $\mathbf{T}'(t)$ is orthogonal to $\mathbf{T}(t)$.

Another vector orthogonal to $\mathbf{T}(t)$ is the *unit binormal vector*

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t).$$

By the properties of the cross product, $\mathbf{B}(t)$ is perpendicular to both $\mathbf{T}(t)$ and $\mathbf{N}(t)$ and has length equal to the area of the parallelogram with sides $\mathbf{T}(t)$ and $\mathbf{N}(t)$. Since this parallelogram is a square of side length 1, $\mathbf{B}(t)$ is a unit vector.

Remark. Note that it is important to use $\mathbf{T}(t)$ and not $\mathbf{r}'(t)$ when we compute the normal vector. In fact, $\mathbf{r}''(t)$ need not be orthogonal to $\mathbf{r}'(t)$, and in most cases it will not be. For example, if $\mathbf{r}(t) = \langle 1, t, t^2 \rangle$, we have $\mathbf{r}'(t) = \langle 0, 1, 2t \rangle$ and $\mathbf{r}''(t) = \langle 0, 0, 2 \rangle$, so

$$\mathbf{r}'(t) \cdot \mathbf{r}''(t) = 4t \neq 0 \qquad \text{unless } t = 0.$$

On the other hand,

$$\mathbf{T}(t) = \left\langle 0, (1+4t^2)^{-1/2}, 2t(1+4t^2)^{-1/2} \right\rangle, \qquad \mathbf{T}'(t) = \left\langle 0, -4t(1+4t^2)^{-3/2}, 2(1+4t^2)^{-3/2} \right\rangle,$$

and these *are* orthogonal.

## 5.6. Arc length

You should be familiar with the formula for arc length of a parametric plane curve: if $\gamma$ is a parametric curve in the $xy$-plane with parametric equations (5.1), where $I = [a, b]$, then its length $L(\gamma)$ is given by the formula

$$L(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2}\, dt. \tag{5.4}$$

In this section, we generalize this formula to space curves. We argue from scratch, so even if (5.4) appears new, this should not be an obstacle.

First, we must say what we mean by "length" of a curve. We shall focus on piecewise smooth curves given by parametric equations of the form (5.3). The intuitive idea of length is that we can take the curve, "straighten" it into a line segment, and then measure the length of that line segment. Now, imagine that we live in a "rigid" world, where we cannot straighten our curve (or bend a thread so it would lay on top of the curve). We could try to cut a number of short straight line segments and piece them together so that the resulting polygon lies close to our original curve; we can then measure the total length of the polygon and argue that the length of the curve is about the same. This is the basic idea behind our definition of arc length.

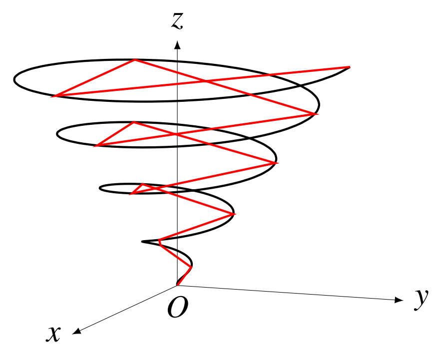



FIGURE 5.5. Polygonal approximation to a curve by a 13-gon

Let $\gamma$ be a parametric curve given by (5.3) with $I = [a, b]$. For each $n \geq 1$, set $\Delta_n = (b - a)/n$ and define the numbers that partition $[a, b]$ into $n$ subintervals of equal lengths:

$$t_0 = a, \quad t_1 = a + \Delta_n, \quad t_2 = a + 2\Delta_n, \quad \dots, \quad t_n = a + n\Delta_n = b.$$

We write $x_i = x(t_i)$, $y_i = y(t_i)$, $z_i = z(t_i)$ for the coordinates of the point $P_i(x_i, y_i, z_i)$ that corresponds to the value $t = t_i$ of the parameter. The points $P_0, P_1, P_2, \dots, P_n$ partition $\gamma$ into $n$ arcs (see Figure 5.5). Let $L_n$ be the length of the polygon with vertices at those points, that is,

$$L_n = \sum_{i=1}^{n} |P_{i-1}P_i| = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}.$$

If the sequence $\{L_n\}_{n=1}^{\infty}$ converges and $\lim_{n\to\infty} L_n = L$, we call $L$ the *arc length* of the given curve.

Using the above definition and the properties of smooth functions and of definite integrals, we can prove the following formula for the length of a piecewise smooth curve.

THEOREM 5.1. *Let $\gamma$ be a piecewise smooth parametric curve with equation*

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3 \qquad (t \in [a, b]).$$

*Then the arc length of $\gamma$ is*

$$L(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt = \int_a^b \|\mathbf{r}'(t)\|\, dt. \tag{5.5}$$

REMARKS. 1. Note that (5.4) is a special case of (5.5): namely, the case when $z(t) = 0$.

2. Note that (5.5) gives the length of the parametric curve and not the length of the graph. For example, the parametric equation

$$\mathbf{r}(t) = \cos t\mathbf{e}_1 + \sin t\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (0 \leq t \leq 4\pi)$$

represents the unit circle in the $xy$-plane. If we apply (5.5) to this curve, we will find that its length is $4\pi$ (and not $2\pi$). The explanation for this is that the above curve traces the unit circle twice.

EXAMPLE 5.9. If $f'$ is continuous on $[a, b]$, the length of the curve $y = f(x)$ is

$$L = \int_a^b \sqrt{1 + f'(x)^2}\, dx.$$

This formula follows from the theorem by representing the given graph as a parametric curve with parametrization $\mathbf{r}(t) = t\mathbf{e}_1 + f(t)\mathbf{e}_2 + 0\mathbf{e}_3$, $a \le t \le b$. □

Example 5.10. Find the length of the curve $\gamma$ given by $\mathbf{r}(t) = \langle 2t^{3/2}, \cos 2t, \sin 2t \rangle$, where $0 \le t \le 1$.

Solution. We have

$$\mathbf{r}'(t) = 3t^{1/2}\mathbf{e}_1 - 2\sin 2t\mathbf{e}_2 + 2\cos 2t\mathbf{e}_3,$$

so (5.5) yields

$$\begin{aligned} L(\gamma) &= \int_0^1 \|\mathbf{r}'(t)\|\, dt = \int_0^1 \sqrt{\left(3t^{1/2}\right)^2 + (-2\sin 2t)^2 + (2\cos 2t)^2}\, dt \\ &= \int_0^1 \sqrt{9t + 4\sin^2 2t + 4\cos^2 2t}\, dt = \int_0^1 \sqrt{9t+4}\, dt \\ &= \frac{1}{9}\int_4^{13} u^{1/2}\, du = \frac{2\left(13\sqrt{13} - 8\right)}{27} \approx 2.879. \end{aligned}$$

□

## 5.7. Space-filling curves*

In this section, we describe the general idea of the construction of Peano's space-filling curve. We start with a piecewise linear curve contained in the unit square. This curve is then shrunk to a fraction of its original size and several copies of it are placed throughout the unit square in such a way that they do not intersect and can be easily connected one to another with straight lines. This yields a more complicated polygon that also lies within the unit square. The whole process (shrinking, copying, connecting) is then repeated iteratively. The first two iterations of Peano's original constructions are displayed on Figure 5.6. The result of the iterative process is a sequence of curves. It can be proved that the sequence of continuous functions underlying this sequence of curves has a limit, which is also a continuous function. It should be intuitively clear (and this can be proved rigorously) that the limit curve fills the entire square.

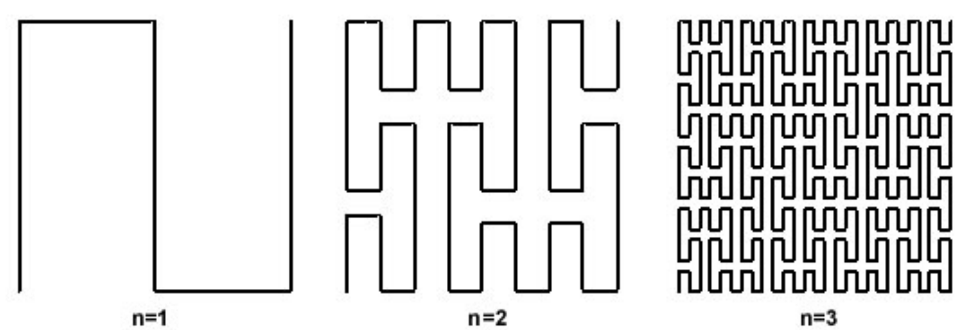



Figure 5.6. The first two iterations in Peano's construction

## 5.8. Space curves in *Mathematica*

*Mathematica* can be quite useful when trying to visualize space curves. *Mathematica*'s main built-in command for plotting a space curve is `ParametricPlot3D`. Its basic syntax is

```
ParametricPlot3D[f[t], {t,a,b}]
```

where `f` is a three-dimensional vector function and `a` and `b` are two numerical expressions describing the endpoints of the interval for the parameter `t`. The function `f` can be given by its name as above, or explicitly, by a vector whose components are functions of `t`. The command `ParametricPlot3D` has many additional options which allow the user to format its output. The best source to learn about all of those is *Mathematica*'s help system. Here, we include only a few examples illustrating the basic syntax and some of the most common options:

- To plot three loops of a helix, having already defined the function $\mathbf{r}(t) = \langle \cos t, \sin t, t/6 \rangle$:

```
r[t_] := {Cos[t], Sin[t], t/6}
ParametricPlot3D[r[t], {t,0,6Pi}]
```

- To plot two thick, black loops of a helix, using the explicit definition:

```
ParametricPlot3D[{Cos[t], Sin[t], t/6}, {t,0,4Pi},
 PlotStyle->{Black,Thick}]
```

- To plot the spiral on Figure 5.3, without a coordinate frame, thick and in red:

```
ParametricPlot3D[{t*Cos[t], t*Sin[t], t}, {t,0,8.5Pi},
 PlotStyle->{Red,Thick}, Axes->False, Boxed->False]
```

### Exercises

5.1. Does the point $P(0, \pi, 1)$ lie on the curve parametrized by the given vector function $\mathbf{r}(t)$? If so, for what value(s) of $t$?

(a) $\mathbf{r}(t) = \langle \sin t, 2\arccos t, t + 1 \rangle$ (b) $\mathbf{r}(t) = \langle \sin t, t, \cos 2t \rangle$ (c) $\mathbf{r}(t) = t\mathbf{e}_2 + \tan t\mathbf{e}_3$

5.2. Do the curves $\mathbf{r}_1(t) = \langle t, t^2, t^3 \rangle$ and $\mathbf{r}_2(t) = \langle t - 2, t^2 - 3t + 2, t^2 + t - 6 \rangle$ intersect?

5.3. Let two objects move in space along trajectories described by the vector functions $\mathbf{r}_1(t) = \langle t, t^2, t^3 \rangle$ and $\mathbf{r}_2(t) = \langle t - 2, t^2 - 3t + 2, t^2 + t - 6 \rangle$. Will the objects ever collide? That is, will they ever be at the same point in space at the same time?

5.4. Find the unit tangent vector $\mathbf{T}(t)$ to the curve $\mathbf{r}(t) = \langle \sin 3t, \cos 3t, \ln(t + 1) \rangle$ at the point with $t = 0$.

5.5. Find the equations of the tangent line to the curve $\mathbf{r}(t) = \langle \cos(\pi t), 3\sin(\pi t), t^3 \rangle$ at the point $(-1, 0, 1)$.

5.6. Find the equations of the tangent line to the curve $\mathbf{r}(t) = \langle t^3 - 2, t^2 + 1, 3t + 1 \rangle$ at the point $(-2, 1, 1)$.

5.7. Find the equations of the tangent line to the curve $\mathbf{r}(t) = \langle 5\cos t, 5\sin t, 2\cos 2t \rangle$ at the point $(-5, 0, 2)$.

5.8. Find the points on the curve $\mathbf{r}(t) = \langle 5\cos t, 5\sin t, 2\cos 2t \rangle$ where the tangent line is parallel to the $x$-axis.

Is the given curve smooth? If not, then is it piecewise smooth?

5.9. $\mathbf{r}(t) = \langle t^3 + t, \sin 2t, 3t^2 - 1 \rangle$

5.10. $\mathbf{r}(t) = \langle 1, t^2, t^3 \rangle$, $-1 \le t \le 1$

5.11. $\mathbf{r}(t) = \langle \cos t, t^2, \cos 2t \rangle$

5.12. $\mathbf{r}(t) = \langle \cos t, \cos 2t, \cos 3t \rangle$

5.13. Find the intersection point of the curves $\mathbf{r}_1(t) = \langle 2t, t^2 + t, t^3 \rangle$ and $\mathbf{r}_2(t) = \langle t, t^2 - 3t + 4, t^2 + t - 5 \rangle$. Find the angle between the curves at that point, that is, find the angle between the tangent vectors at the point.

5.14. The curves $\mathbf{r}_1(t) = \langle t^2, 2t, \ln(t^2 - 3) \rangle$ and $\mathbf{r}_2(t) = \langle t + 1, t^2 - 5, t^2 - 2t - 3 \rangle$ intersect at the point $(4, 4, 0)$. Find the cosine of the angle between the two curves at that point.

Find the arc length of the curve $\mathbf{r}(t)$.

5.15. $\mathbf{r}(t) = \langle 1, t^2, t^3 \rangle$, $-1 \le t \le 1$

5.16. $\mathbf{r}(t) = \langle t, \ln t, 2\sqrt{2t} \rangle$, $1 \le t \le 2$

5.17. Let $\mathbf{r}(t)$, $a \le t \le b$, be a parametrization of a smooth curve $\gamma$. Let $L = \int_a^b \|\mathbf{r}'(t)\|\, dt$ be the length of $\gamma$ and define the new function

$$s(t) = \int_a^t \|\mathbf{r}'(u)\|\, du \qquad (a \le t \le b).$$

(a) Evaluate $s(a)$ and $s(b)$.

(b) Use the Fundamental Theorem of Calculus to show that $s'(t) > 0$ for all $t$ with $a < t < b$.

(c) Let $t = t(s)$, $0 \le s \le L$, be the inverse function of the function $s(t)$, and define the vector function $\mathbf{f}(s) = \mathbf{r}(t(s))$, $0 \le s \le L$. Show that $\|\mathbf{f}'(s)\| = 1$ for all $s$ with $0 < s < L$.

(d) Convince yourself that the vector function $\mathbf{f}(s)$ from part (c) is another parametrization of the curve $\gamma$. This parametrization has a special property: $\mathbf{f}(s)$ is the position vector of the point on $\gamma$ that is $s$ units along the curve after the initial point. This parametrization is known as the *parametrization with respect to arc length* or the *natural parametrization* and plays an important role in the study of space curves.

(e) Parametrize with respect to arc length the curve $\mathbf{r}(t) = \langle e^{3t} \sin 4t, e^{3t} \cos 4t, 5 \rangle$, $0 \le t \le \frac{1}{3} \ln 4$.

LECTURE 6

# Multivariable Functions

## 6.1. Functions of two variables

In this lecture we introduce multivariable functions. These are functions that depend on multiple inputs (usually 2, 3, or 4), but produce only a single output. We start our discussion with the simplest example: functions of two variables. For example, the temperature $T$ at any point on the surface of the Earth at this very moment depends on two pieces of information: the longitude $x$ and the latitude $y$ of the point. Thus, we can think of $T$ as a function $T(x, y)$ of two variables $x$ and $y$ or as a function of the ordered pair $(x, y)$. We now give a formal definition.

Definition. A *function $f$ of two variables* is a rule that assigns to each ordered pair of real numbers $(x, y)$ in a set $D$ a unique real number denoted $f(x, y)$. The set $D$ is the *domain* of $f$ and the set of all values that $f$ takes on is its *range*, that is, the range of $f$ is $\{f(x, y) : (x, y) \in D\}$.

## 6.2. Functions of three and more variables

The example that we used above to motivate the introduction of functions of two variables is somewhat artificial. A more natural (and also more useful) quantity is the temperature $T$ at a point $(x, y, z)$ in space at this very moment. Then $T$ is a function of three variables—the coordinates $(x, y, z)$ of the point. Note that with this function in hand we don't have to assume that the temperature in a classroom and the temperature on the wing of an airplane flying 39,000 feet above that classroom are the same. An even more realistic example is that of the temperature $T$ at the point $(x, y, z)$ in space at time $t$. This quantity is a function $T(x, y, z, t)$ of four variables.

Definition. A *function $f$ of $n$ variables* $x_1, x_2, \dots, x_n$ is a rule that assigns to each ordered $n$-tuple of real numbers $(x_1, x_2, \dots, x_n)$ in a set $D$ a unique real number denoted $f(x_1, x_2, \dots, x_n)$. The set $D$ is the *domain* of $f$ and the set of all values that $f$ takes on is its *range*. In particular, a *function $f$ of three variables* is a rule that assigns to each ordered triple of real numbers $(x, y, z)$ in a set $D$ (in $\mathbb{R}^3$) a unique real number denoted $f(x, y, z)$.

## 6.3. Examples of domains and ranges of multivariable functions

Example 6.1. The function $f(x, y) = \sqrt{1 - xy}$ is defined whenever

$$1 - xy \geq 0 \quad \iff \quad xy \leq 1.$$

If $x > 0$, the last inequality is satisfied for $y \leq x^{-1}$; if $x < 0$, the inequality is satisfied for $y \geq x^{-1}$; and if $x = 0$, it is satisfied for all $y$. That is, the domain of $f$ contains the points in I quadrant on and below the graph of $y = x^{-1}$, the points in III quadrant on and above the graph of $y = x^{-1}$, and all the points in II and IV quadrants, including the $x$- and $y$-axes (see Figure 6.1).

The range of $f$ is $[0, \infty)$. Clearly, $f$ cannot attain negative values. On the other hand, if $k \geq 0$,

$$\sqrt{1 - xy} = k \quad \iff \quad 1 - xy = k^2 \quad \iff \quad xy = 1 - k^2.$$

Hence, if $k \geq 0$, $k \neq 1$, $f(x, y) = k$ for all points $(x, y)$ on the hyperbola $xy = 1 - k^2$; also, $f(x, y) = 1$ for all points $(x, y)$ on the coordinate axes. That is, every number $k \in [0, \infty)$ does belong to the range of $f$. □

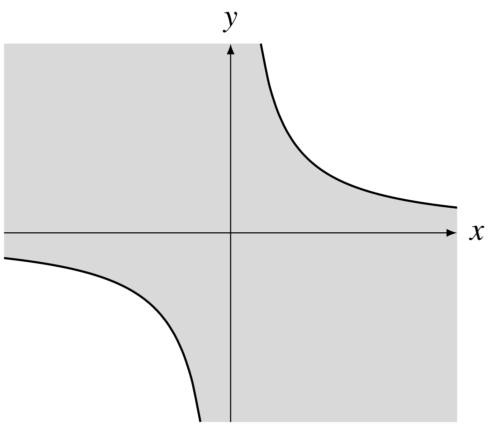



FIGURE 6.1. The domain of $z = \sqrt{1 - xy}$

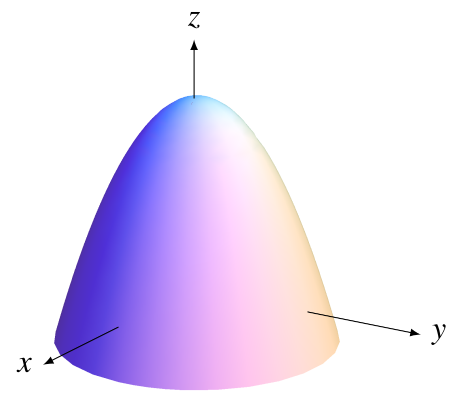



FIGURE 6.2. $z = 3 - x^2 - y^2$

EXAMPLE 6.2. The domain of the function $f(x, y, z) = \ln(16 - x^2 - y^2 - 9z^2)$ consists of the points $(x, y, z)$ in space whose coordinates satisfy

$$16 - x^2 - y^2 - 9z^2 > 0 \quad \iff \quad \frac{x^2}{16} + \frac{y^2}{16} + \frac{z^2}{16/9} < 1.$$

Since $\ln x$ is an increasing function, the range of $f$ consists of the logarithms of the positive numbers in the range of the function $w = 16 - x^2 - y^2 - 9z^2$. Because $x^2 + y^2 + 9z^2$ can attain any non-negative value and cannot attain any negative value, it follows that $w$ takes on the numbers in the interval $(-\infty, 16]$. Thus, the range of $f$ consists of the logarithms of the numbers in $(0, 16]$, that is, the range of $f$ is $(-\infty, \ln 16]$. □

## 6.4. Graphs

DEFINITION. If $f$ is a function of $n$ variables $x_1, x_2, \dots, x_n$ with domain $D$, then the *graph* of $f$ is the set of all points $(x_1, x_2, \dots, x_n, z)$ in $\mathbb{R}^{n+1}$ with $(x_1, x_2, \dots, x_n) \in D$ and $z = f(x_1, x_2, \dots, x_n)$, that is, the graph of $f$ is the set

$$\big\{(x_1, x_2, \dots, x_n, f(x_1, x_2, \dots, x_n)) : (x_1, x_2, \dots, x_n) \in D\big\}.$$

In particular, the graph of a function $f(x, y)$ defined on a set $D$ in $\mathbb{R}^2$ is the surface given by

$$\big\{(x, y, f(x, y)) : (x, y) \in D\big\}.$$

EXAMPLE 6.3. The graph of the function $f(x, y) = 3 - x^2 - y^2$ is the surface

$$z = 3 - x^2 - y^2.$$

This surface is displayed on Figure 6.2. □

## 6.5. Level curves and surfaces

While the graph of a function of two variables provides an excellent visualization of the function, it might be quite difficult to sketch without advanced technology. What can we do using only manual (and mental) labor? One answer is well-known to cartographers. When drawing topographic or termal maps, cartographers often produce two-dimensional drawings marked with curves of constant elevation or constant temperature. We can do the same for any function $f$ of two variables: we can graph $f$ in the $xy$-plane by marking its "level curves".

Definition. Let $f$ be a function of two veriables, denoted $x$ and $y$, and let $k$ be a real number. The *$k$-level curve* of $f$ is the set of all points $(x, y)$ in the domain $D$ of $f$ which satisfy the equation $f(x, y) = k$.

This idea becomes even more important in the case of functions of three variables. Note that the graph $w = f(x, y, z)$ of such a function is a four-dimensional object, which we can study analytically but cannot represent graphically in our three-dimensional universe. However, we can draw the "level surfaces" of $f$.

Definition. Let $f$ be a function of three veriables, denoted $x, y$ and $z$, and let $k$ be a real number. The *$k$-level surface* of $f$ is the set of all points $(x, y, z)$ in the domain $D$ of $f$ which satisfy the equation $f(x, y, z) = k$.

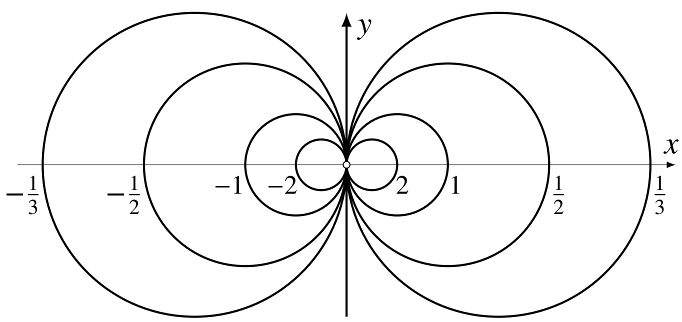



Figure 6.3. The level curves of $f(x, y) = x/(x^2 + y^2)$

Example 6.4. Describe the level curves of the function

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

Solution. We want to draw the curve

$$\frac{x}{x^2 + y^2} = k$$

for all $k$. If we exclude the origin from consideration (since the curve is undefined there), we can represent this curve by the equation

$$kx^2 + ky^2 - x = 0.$$

When $k = 0$, this equation represents the $y$-axis $x = 0$, so the level curve $f(x, y) = 0$ is the $y$-axis with the origin removed from it. When $k \neq 0$, we can transform the equation of the level curve further to

$$x^2 - \frac{x}{k} + y^2 = 0 \quad \iff \quad \left(x - \frac{1}{2k}\right)^2 + y^2 = \frac{1}{4k^2},$$

which is the equation of a circle centered at $(\frac{1}{2k}, 0)$ of radius $\frac{1}{2|k|}$ ($k$ can be negative). Since the origin does lie on this circle, the level curve $f(x, y) = k$ is the circle with the origin removed from it. A contour map of $f$ showing the level curves with $k = 0, \pm\frac{1}{3}, \pm\frac{1}{2}, \pm 1, \pm 2$ is shown on Figure 6.3. □

## 6.6. Limits of multivariable functions

Since all the important concepts in single-variable calculus were defined in terms of limits, if we are to generalize those concepts to functions of two and more variables, we must first define the notion of a limit of a multivariable function. The following definition generalizes the formal definition of limit to functions of two variables.

Definition. Let $f$ be a function of two variables whose domain $D$ includes points arbitrarily close to $(a, b)$. We say that the *limit of* $f(x, y)$ *as* $(x, y)$ *approaches* $(a, b)$ *is* $L$ and write

$$\lim_{(x,y)\to(a,b)} f(x, y) = L \qquad \text{or} \qquad f(x, y) \to L \text{ as } (x, y) \to (a, b),$$

if for any given number $\varepsilon > 0$, there is a corresponding number $\delta = \delta(\varepsilon) > 0$ such that

$$0 < \sqrt{(x - a)^2 + (y - b)^2} < \delta \qquad \Longrightarrow \qquad |f(x, y) - L| < \varepsilon.$$

This definition is *very* formal. It is also very useful in proofs (something we will not be concerned with) and very imposing at first sight. A more intuitive definition reads as follows.

Definition. Let $f$ be a function of two variables whose domain $D$ includes points arbitrarily close to $(a, b)$. We say that the *limit of* $f(x, y)$ *as* $(x, y)$ *approaches* $(a, b)$ *is* $L$ and write

$$\lim_{(x,y)\to(a,b)} f(x, y) = L \qquad \text{or} \qquad f(x, y) \to L \text{ as } (x, y) \to (a, b),$$

if $f(x, y)$ approaches $L$ as $(x, y)$ approaches $(a, b)$, independent of how $(x, y)$ approaches $(a, b)$.

We can define limits of functions of three or more variables in a similar fashion.

Definition. Let $f$ be a function of three variables whose domain $D$ includes points arbitrarily close to $(a, b, c)$. We say that the *limit of* $f(x, y, z)$ *as* $(x, y, z)$ *approaches* $(a, b, c)$ *is* $L$ and write

$$\lim_{(x,y,z)\to(a,b,c)} f(x, y, z) = L \qquad \text{or} \qquad f(x, y, z) \to L \text{ as } (x, y, z) \to (a, b, c),$$

if $f(x, y, z)$ approaches $L$ as $(x, y, z)$ approaches $(a, b, c)$, independent of how $(x, y, z)$ approaches $(a, b, c)$. More precisely, the limit of $f(x, y, z)$ as $(x, y, z)$ approaches $(a, b, c)$ is $L$, if for any given number $\varepsilon > 0$, there is a corresponding number $\delta = \delta(\varepsilon) > 0$ such that

$$0 < \sqrt{(x - a)^2 + (y - b)^2 + (z - c)^2} < \delta \qquad \Longrightarrow \qquad |f(x, y, z) - L| < \varepsilon.$$

Limits of multivariable functions can be much trickier than those of single-variable functions. The primary reason for that is the infinitude of ways in which the arguments $(x, y)$ (or $(x, y, z)$) can approach the point $(a, b)$ (or $(a, b, c)$). This point is best illustrated by examples of limits that do not exist.

Example 6.5. The limit

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

does not exist. Indeed, we know from Example 6.4 (see also Figure 6.3) that the level curves of this function all want to pass through $(0,0)$ (but since the point is not in the domain of $f$, they can't). Thus, if we approach $(0,0)$ along two distinct level curves (say, along $f(x,y) = 1$ and $f(x,y) = -1$), the function values will approach distinct values (1 and $-1$, respectively). However, if the limit existed, $f(x,y)$ would have to approach the same value independent of the way $(x,y)$ approaches $(0,0)$. Therefore, the limit does not exist. □

Example 6.6. Show that the limit $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy^4}{x^2+y^8}$ does not exist.

Solution. First, consider the behavior of the function as $(x,y) \to (0,0)$ along the straight line $y = x$. We have

$$\lim_{x\to 0} f(x,x) = \lim_{x\to 0} \frac{x^5}{x^2+x^8} = \lim_{x\to 0} \frac{x^3}{1+x^6} = 0.$$

On the other hand, if we let $(x,y) \to (0,0)$ along the curve $x = y^4$, we get

$$\lim_{y\to 0} f(y^4,y) = \lim_{y\to 0} \frac{y^8}{y^8+y^8} = \frac{1}{2}.$$

Since different approach paths lead to different limiting values, the limit does not exist. □

Example 6.7. For any numbers $a, b, c$, we have

$$\lim_{(x,y)\to(a,b)} x = a, \qquad \lim_{(x,y)\to(a,b)} y = b, \qquad \lim_{(x,y)\to(a,b)} c = c.$$

Using these facts and the two-variable versions of the limit laws:

$$\lim_{(x,y)\to(a,b)} [f(x,y) \pm g(x,y)] = \lim_{(x,y)\to(a,b)} f(x,y) \pm \lim_{(x,y)\to(a,b)} g(x,y), \qquad \text{etc.},$$

we can compute the limit of any polynomial in $x$ and $y$ at any point $(a,b)$ in the plane. For example,

$$\lim_{(x,y)\to(2,3)} xy = \Big(\lim_{(x,y)\to(2,3)} x\Big)\Big(\lim_{(x,y)\to(2,3)} y\Big) = 2\cdot 3 = 6.$$

□

Next, we look at some examples of evaluation of limits of functions of two and three variables.

Example 6.8. Evaluate $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2}$.

Solution. If we change the coordinates of the point $(x,y)$ from Cartesian to polar, the condition $(x,y) \to (0,0)$ is equivalent to $r \to 0$ and the limit takes the form

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2} = \lim_{r\to 0} \frac{\sin(r^2)}{r^2} = \lim_{r\to 0} \frac{2r\cos(r^2)}{2r} = 1,$$

where we used l'Hospital's rule to pass from the second limit to the third. □

Example 6.9. Evaluate $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{(x^2+y^2)z}{x^2+y^2+z^2}$.

Solution. This time, we change the coordinates of the point $(x,y,z)$ from Cartesian to spherical, and the condition $(x,y,z)\to(0,0,0)$ becomes $\rho\to 0$. Hence,

$$\begin{aligned}\lim_{(x,y,z)\to(0,0,0)} \frac{(x^2+y^2)z}{x^2+y^2+z^2} &= \lim_{\rho\to 0} \frac{(\rho\sin\phi)^2(\rho\cos\phi)}{\rho^2}\\ &= \lim_{\rho\to 0} \rho\sin^2\phi\cos\phi = 0.\end{aligned}$$

□

Example 6.10. Evaluate $\displaystyle\lim_{(x,y)\to(1,2)} \frac{x^2y-2x^2-2xy+6x+5y-12}{x^2+y^2-2x-4y+5}$.

Solution. Both the numerator and the denominator of the function vanish at the point $(1,2)$, so the limit is an indeterminate form of type $0/0$. We will try to argue similarly to the previous example, but before we can do so, we need to change the limit from one at the point $(1,2)$ to one at the origin. Let us define new variables $u$ and $v$ by

$$x = 1+u, \quad y = 2+v.$$

The condition $(x,y)\to(1,2)$ is equivalent to $(u,v)\to(0,0)$, so we can rewrite the original limit as

$$\begin{aligned}&\lim_{(x,y)\to(1,2)} \frac{x^2y-2x^2-2xy+6x+5y-12}{x^2+y^2-2x-4y+5}\\ &= \lim_{(u,v)\to(0,0)} \frac{(1+u)^2(2+v)-2(1+u)^2-2(1+u)(2+v)+6(1+u)+5(2+v)-12}{(1+u)^2+(2+v)^2-2(1+u)-4(2+v)+5}\\ &= \lim_{(u,v)\to(0,0)} \frac{u^2v+2u+4v}{u^2+v^2}.\end{aligned}$$

We can now change the coordinates of the point $(u,v)$ from Cartesian to polar. We find that

$$\begin{aligned}\lim_{(u,v)\to(0,0)} \frac{u^2v+2u+4v}{u^2+v^2} &= \lim_{r\to 0} \frac{r^3\cos\theta\sin\theta+2r\cos\theta+4r\sin\theta}{r^2}\\ &= \lim_{r\to 0} \left(r\cos\theta\sin\theta+(2\cos\theta+4\sin\theta)r^{-1}\right)\\ &= \lim_{r\to 0} (2\cos\theta+4\sin\theta)r^{-1}\\ &= \begin{cases} +\infty & \text{when } 2\cos\theta+4\sin\theta>0,\\ 0 & \text{when } 2\cos\theta+4\sin\theta=0,\\ -\infty & \text{when } 2\cos\theta+4\sin\theta<0.\end{cases}\end{aligned}$$

Hence, the given limit does not exist. □

## 6.7. Continuity of multivariable functions

We can use the notion of limit of a multivariable function to define the notion of continuity.

Definition. Let $f$ be a function of two variables whose domain $D$ includes $(a,b)$ and points arbitrarily close to $(a,b)$. We say that $f(x,y)$ is *continuous at* $(a,b)$, if

$$\lim_{(x,y)\to(a,b)} f(x,y) = f(a,b).$$

We say that $f$ is *continuous in $D$*, if it is continuous at every point $(a, b) \in D$.

Definition. Let $f$ be a function of three variables whose domain $D$ includes $(a, b, c)$ and points arbitrarily close to $(a, b, c)$. We say that $f$ is *continuous at* $(a, b, c)$, if

$$\lim_{(x,y,z)\to(a,b,c)} f(x, y, z) = f(a, b, c).$$

We say that $f$ is *continuous in $D$*, if it is continuous at every point $(a, b, c) \in D$.

We want also to extend the notion of piecewise continuity to multivariable functions, and in particular, to functions of two variables. Recall that a single-variable function $f$ is called piecewise continuous on $[a, b]$ if it is continuous everywhere in $[a, b]$ except for a finite number of points, where it can have only jump or removable discontinuities. The two-dimensional version of this definition reads as follows.

Definition. Let $f$ be a function of two variables with domain $D$. We say that $f$ is *piecewise continuous in $D$*, if it is continuous at every point $(a, b) \in D$, except possibly for the points on a finite number of smooth curves in $D$, say $\gamma_1, \gamma_2, \dots, \gamma_n$, where $f$ may have "jump discontinuities."

Remark. As in the case of single-variable functions, polynomials in several variables are continuous. Likewise, rational functions in several variables are continuous except for the points where the denominator is 0. Finally, if $f$ is a continuous function (of a single variable) and $g$ is a continuous function of the variables $x, y, \dots$, the composition $f(g(x, y, \dots))$ is continuous at all points $(x, y, \dots)$ where it is defined.

Example 6.11. The function

$$f(x, y) = \cos\left(\frac{x^2 - y^2}{x^2 + y^2}\right)$$

is continuous everywhere in its domain, which is all the points in $\mathbb{R}^2$ except $(0, 0)$. Indeed, $f$ is the composition of the continuous function $w = \cos t$ and the rational function

$$t = \frac{x^2 - y^2}{x^2 + y^2}.$$

Example 6.12. Show that the function

$$f(x, y) = \begin{cases} \dfrac{\sin(x^2 + y^2)}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 1 & \text{if } (x, y) = (0, 0), \end{cases}$$

is continuous everywhere.

Solution. Recall that

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

Hence, the function

$$g(x) = \begin{cases} \dfrac{\sin x}{x} & \text{if } x \neq 0, \\ 1 & \text{if } x = 0, \end{cases}$$

is continuous everywhere. Furthermore, the polynomial $x^2 + y^2$ is also continuous everywhere. Therefore, the function $f(x, y) = g(x^2 + y^2)$ is also continuous everywhere. □

## Exercises

Find the domain and the range of the given function.

6.1. $f(x, y) = y^2 e^{-x^2/2}$

6.2. $f(x, y) = \ln\left(x + y^2\right)$

6.3. $f(x, y, z) = \dfrac{z}{x^2 - y^2 - 1}$

Sketch the domain of the given function.

6.4. $f(x, y) = \ln\left(x + y^2\right)$

6.5. $f(x, y) = \dfrac{\sqrt{x + 2y}}{x^2 + y^2 - 1}$

For each function, draw the level curves for levels $-1, 0, 1, 2$, and 3; then sketch the graph $z = f(x, y)$.

6.6. $f(x, y) = 1 + y^2$

6.7. $f(x, y) = 2x + 3y - 5$

6.8. $f(x, y) = x^2 + y^2$

Find the given limit or show that it does not exist.

6.9. $\displaystyle\lim_{(x,y)\to(1,3)} \left(x^4 - 4xy^2 + 2x^3 y\right)$

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

6.11. $\displaystyle\lim_{(x,y,z)\to(1,0,2)} \frac{\sin(2xy)}{1 - zy^2}$

6.12. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3x^2 y}{x^4 + y^2}$

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

6.14. $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{1 + xy} - 1}$

Find the set of points in $\mathbb{R}^2$ or $\mathbb{R}^3$ where the given function is continuous.

6.15. $f(x, y) = \sin\left(x^4 + x^2 y^2\right)$

6.16. $f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0), \\ 1 & \text{if } (x, y) = (0, 0). \end{cases}$

6.17. $f(x, y, z) = \ln\left(4 - x^2 - y^2 - 4z^2\right)$

6.18. $f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{1 + xy} - 1} & \text{if } (x, y) \neq (0, 0), \\ 1 & \text{if } (x, y) = (0, 0). \end{cases}$

LECTURE 7

# Surfaces

## 7.1. Introduction

In single-variable calculus, the interplay between curves and functions is essential: we use geometrical questions about curves to motivate the study of calculus and then apply our knowledge of calculus to answer questions about plane curves. A similar symbiosis occurs in multivariable calculus between functions and surfaces. However, while we are familiar with a fair number of plane curves, our supply of known surfaces is limited essentially to planes, spheres, and graphs $z = f(x, y)$. We introduced also the level surfaces of functions of three variables: these are the sets

$$\Sigma = \left\{ (x, y, z) \in \mathbb{R}^3 : F(x, y, z) = k \right\}.$$

However, there are hardly any level surfaces that we recognize. The main goal of this lecture is to introduce some other surfaces that are commonly used in multivariable calculus.

## 7.2. Cylinders

Definition. A *cylinder* is a surface that consists of all lines—called *rullings*—that are parallel to a given line and pass through a given plane curve.

Example 7.1. We have seen already one example of a cylinder: in Example 1.8, we explained that the equation $x^2+y^2 = x$ describes the surface $\Sigma$ in space that consists of all the points on vertical lines that pass through a circle in the $xy$-plane. Note that by a similar argument, any equation of one of the forms

$$f(x, y) = 0, \qquad f(x, z) = 0, \qquad \text{or} \qquad f(y, z) = 0$$

represents a cylinder in space. For example, the equation $2y^2 + z^2 = 1$ represents the cylinder that passes through the ellipse $2y^2 + z^2 = 1$ in the $yz$-plane and has rullings parallel to the $x$-axis. Note that the three-dimensional equations of the ellipse are

$$2y^2 + z^2 = 1, \quad x = 0.$$

Similarly, in space, the equation $x = y^2$ represents the cylinder that passes through the parabola $x = y^2$ in the $xy$-plane and has rullings parallel to the $z$-axis. □

## 7.3. Quadric surfaces

Quadric surfaces are generalizations of the conic sections to three dimensions.

Definition. A *quadric surface* is the set of all points $(x, y, z) \in \mathbb{R}^3$ that satisfy a second-degree equation in the variables $x$, $y$, and $z$:

$$A_{11}x^2 + A_{12}xy + A_{13}xz + A_{22}y^2 + A_{23}yz + A_{33}z^2 + B_1x + B_2y + B_3z + B_0 = 0,$$

where $A_{11}, \dots, A_{33}, B_0, \dots, B_3$ are constants.

In other words, a quadric surface is a level surface of a quadratic polynomial in three variables. In the next couple of examples, we visualize some quadric surfaces by sketching their traces in planes parallel to the coordinate planes.

Example 7.2. Consider the quadric surface

$$\frac{x^2}{4} + y^2 + z^2 = 1.$$

The trace of this surface in the horizontal plane $z = k$ has equations

$$\frac{x^2}{4} + y^2 = 1 - k^2, \quad z = k,$$

which represent an ellipse, provided that

$$1 - k^2 > 0 \quad \Longleftrightarrow \quad -1 < k < 1.$$

Similarly, the traces of the surface in the planes $y = k$, $-1 < k < 1$, and $x = k$, $-2 < k < 2$, are also ellipses. Thus, the given surface is called an *ellipsoid*. □

Example 7.3. Consider the quadric surface

$$z = x^2 - 2y^2.$$

The trace of this surface in a plane $x = k$ is

$$z = k^2 - 2y^2, \quad x = k,$$

which is a parabola open downward. Likewise, the trace in a plane $y = k$ is a parabola open upward,

$$z = x^2 - 2k^2, \quad y = k.$$

Lastly, the trace in a horizontal plane $z = k$ is the hyperbola

$$x^2 - 2y^2 = k, \quad z = k.$$

This surface is called a *hyperbolic paraboloid*. □

Remark. At first, one might think that there is no neat classification of all quadric surfaces, but in fact there is. First of all, by a change of the coordinates, every quadric surface can be brought into the form

$$Ax^2 + By^2 + Cz^2 + Dz + E = 0,$$

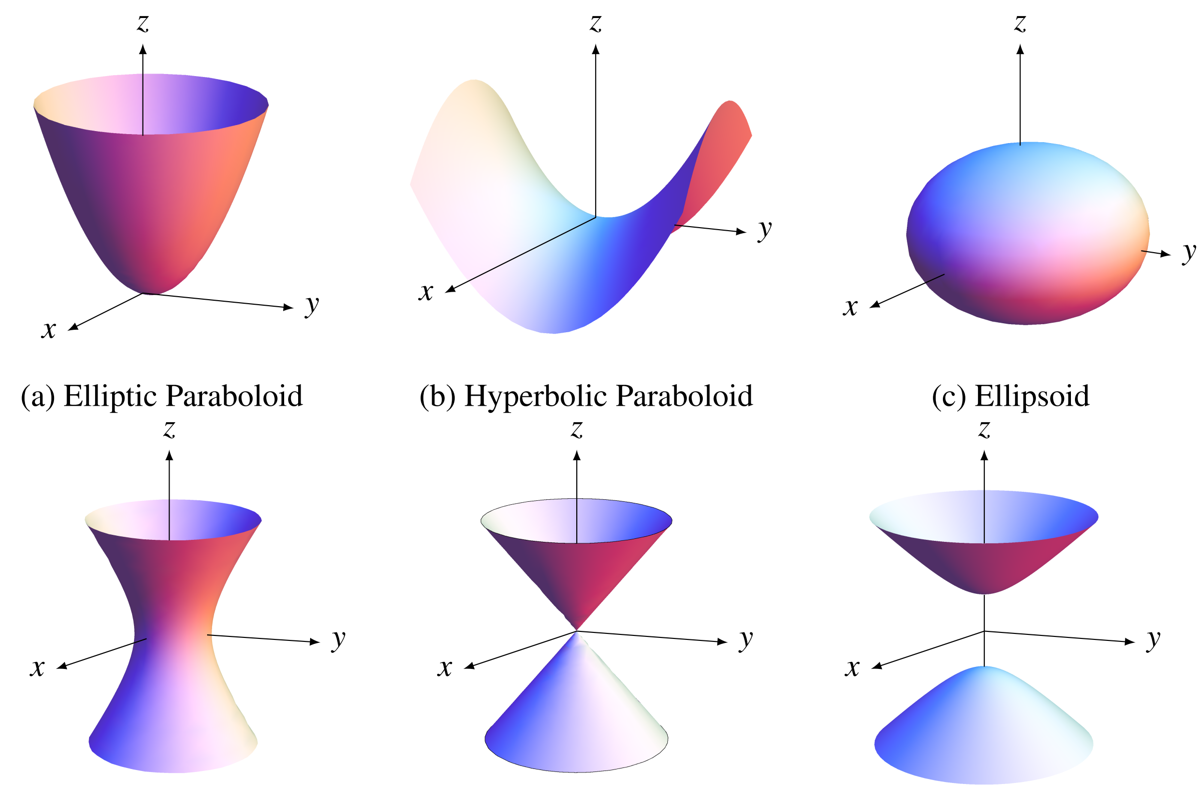
where either $D = 0$ or $C = E = 0$. If one of the variables is missing—say, if $A = 0$ or $B = 0$, then the quadric surface is a cylinder parallel to the respective coordinate axis. This leaves only six possibilities, which lead to the following six surfaces: *cone, ellipsoid, elliptic paraboloid, hyperbolic paraboloid, one-sheet hyperboloid*, and *two-sheet hyperboloid*. These surfaces in standard positions are displayed on Figure 7.1 and some basic facts about them are summarized in Table 7.1.

Example 7.4. Classify the quadric surface Σ with equation

$$4y^2 + z^2 - x + 16y - 4z + 19 = 0.$$

| Surface | Equation | Traces | | | Axis |
|---|---|---|---|---|---|
| | | $xy$ | $yz$ | $xz$ | |
| Elliptic Paraboloid | $\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ | Ellipses | Parabolas | Parabolas | $z$-axis |
| Hyperbolic Paraboloid | $\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ | Hyperbolas | Parabolas | Parabolas | $z$-axis |
| Ellipsoid | $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ | Ellipses | Ellipses | Ellipses | |
| One-Sheet Hyperboloid | $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ | Ellipses | Hyperbolas | Hyperbolas | $z$-axis |
| Cone | $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$ | Ellipses | Hyperbolas | Hyperbolas | $z$-axis |
| Two-Sheet Hyperboloid | $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$ | Ellipses | Hyperbolas | Hyperbolas | $z$-axis |

TABLE 7.1. The non-degenerate quadric surfaces




(a) Elliptic Paraboloid (b) Hyperbolic Paraboloid (c) Ellipsoid

(d) One-Sheet Hyperboloid (e) Cone (f) Two-Sheet Hyperboloid

FIGURE 7.1. The non-degenerate quadric surfaces

SOLUTION. We can rewrite this equation in the form

$$4(y+2)^2 + (z-2)^2 - (x+1) = 0.$$

If we now change the coordinates to

$$X = y + 2, \quad Y = z - 2, \quad Z = x + 1,$$

we find that a point in $P(x, y, z)$ in the $xyz$-space is on $\Sigma$ if and only if $P(X, Y, Z)$ is on the surface $Z = 4X^2 + Y^2$ in the $XYZ$-space. The latter surface is an elliptic paraboloid with $a = \frac{1}{4}$, $b = 1$, and $c = 1$. Note that in the $xyz$-space the vertex of the paraboloid is at $(-1, -2, 2)$ (these are the $xyz$-coordinates of the origin of the $XYZ$-coordinate system) and its axis is the line $y = -2$, $z = 2$ (these are equations of the $Z$-axis in $xyz$-coordinates). □

## 7.4. Parametric surfaces

By now, you have probably come to the realization that a "surface" is a three-dimensional geometric object that is essentially two-dimensional. In the case of graphs $z = f(x, y)$ this is quite obvious: the points $(x, y, z)$ on the graph are in one-to-one correspondence with the points $(x, y)$ in the domain of $f(x, y)$. For example, the points on the elliptic paraboloid $z = x^2 + 2y^2$ (see Figure 7.1(a)) are in a one-to-one correspondence with the points in the $xy$-plane. On the other hand, the two-dimensionality of a level surface, such as the ellipsoid $x^2 + 2y^2 + z^2 = 4$, is less obvious from its equation (though it is quite intuitive from Figure 7.1(c)). Level surfaces are also hard to plot, because there is no easy way to generate points on a level surface. On the other hand, level surfaces are the natural analytic representation for some basic examples: the cone $x^2 + y^2 - z^2 = 0$, the cylinder $x^2 + y^2 = 1$, and the sphere $x^2 + y^2 + z^2 = 1$ are all level surfaces that cannot be represented as graphs of functions of two variables. In this section, we introduce the notion of a "parametric surface".

DEFINITION. If $x = x(u, v)$, $y = y(u, v)$ and $z = z(u, v)$ are three continuous functions defined on a set $D$ in $\mathbb{R}^2$, then the set $\Sigma$ of the points $(x, y, z)$ satisfying the *parametric equations*

$$x = x(u, v), \quad y = y(u, v), \quad z = z(u, v) \qquad ((u, v) \in D), \tag{7.1}$$

is called a *parametric surface*; the variables $u, v$ are called the *parameters*. Equations (7.1) and the vector function $\mathbf{r} : \mathbb{R}^2 \to \mathbb{R}^3$, given by

$$\mathbf{r}(u, v) = x(u, v)\mathbf{e}_1 + y(u, v)\mathbf{e}_2 + z(u, v)\mathbf{e}_3 \qquad ((u, v) \in D), \tag{7.2}$$

are referred to as *parametrizations* of $\Sigma$.

REMARK. Like level surfaces, parametric surfaces generalize of the graphs $z = f(x, y)$ of functions of two variables. Indeed, suppose that $\Sigma$ is the surface $z = f(x, y)$, $(x, y) \in D$. Then we can parametrize $\Sigma$ by the equations

$$x = u, \quad y = v, \quad z = f(u, v) \qquad ((u, v) \in D).$$

On the other hand, a parametric surface is almost as easy to plot as the graph of a function: to plot $z = f(x, y)$, we let the point $(x, y)$ roam through the domain $D$ and plot the points $(x, y, f(x, y))$; to plot the surface $\Sigma$ parametrized by (7.2), we let $(u, v)$ roam through $D$ and plot the endpoints of the position vectors $\mathbf{r}(u, v)$.

Example 7.5. Identify the surface $\Sigma$, parametrized by the vector function

$$\mathbf{r}(u,v) = u\mathbf{e}_1 + u\cos v\mathbf{e}_2 + u\sin v\mathbf{e}_3,$$

where $-\infty < u < \infty$ and $0 \le v \le 2\pi$.

Solution. For a point $(x,y,z)$ on $\Sigma$, we have

$$y^2 + z^2 = (u\cos v)^2 + (u\sin v)^2 = u^2 = x^2.$$

This is the equation of a cone whose axis of symmetry is the $x$-axis. □

Example 7.6. Find a parametrization of the plane $\Pi$ through the points $A(1,2,3)$, $B(1,-1,-2)$, and $C(-2,0,1)$.

Solution. Let us first find an equation of the plane. Two vectors parallel to $\Pi$ are $\overrightarrow{AB} = \langle 0,-3,-5\rangle$ and $\overrightarrow{BC} = \langle -3,1,3\rangle$. Hence, the vector $\mathbf{n} = \overrightarrow{AB} \times \overrightarrow{BC}$ is normal to $\Pi$. We have

$$\begin{aligned}\mathbf{n} &= (-3\mathbf{e}_2 - 5\mathbf{e}_3) \times (-3\mathbf{e}_1 + \mathbf{e}_2 + 3\mathbf{e}_3)\\ &= 9(\mathbf{e}_2 \times \mathbf{e}_1) - 9(\mathbf{e}_2 \times \mathbf{e}_3) + 15(\mathbf{e}_3 \times \mathbf{e}_1) - 5(\mathbf{e}_3 \times \mathbf{e}_2) = -4\mathbf{e}_1 + 15\mathbf{e}_2 - 9\mathbf{e}_3,\end{aligned}$$

so an equation for $\Pi$ is

$$-4(x-1) + 15(y-2) - 9(z-3) = 0.$$

Let us solve this equation for one of the variables, say for $x$:

$$\begin{aligned}-4(x-1) + 15(y-2) - 9(z-3) = 0 \quad &\Longleftrightarrow \quad x - 1 = \tfrac{15}{4}(y-2) - \tfrac{9}{4}(z-3)\\ &\Longleftrightarrow \quad x = \tfrac{15}{4}y - \tfrac{9}{4}z + \tfrac{1}{4}.\end{aligned}$$

We see that $\Pi$ is the graph of the function

$$x = \tfrac{15}{4}y - \tfrac{9}{4}z + \tfrac{1}{4},$$

so it is parametrized by the vector function

$$\mathbf{r}(y,z) = \left(\tfrac{15}{4}y - \tfrac{9}{4}z + \tfrac{1}{4}\right)\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3,$$

where $-\infty < y, z < \infty$. □

Example 7.7. Let $\mathbf{a} = \langle a_1,a_2,a_3\rangle$ and $\mathbf{b} = \langle b_1,b_2,b_3\rangle$ be two vectors that are not parallel to each other. Find a parametrization of the plane $\Pi$ through the origin that is parallel to $\mathbf{a}$ and $\mathbf{b}$.

Solution. We can try to argue similarly to the previous problem. However, the equation of $\Pi$ is

$$\big(a_2b_3 - a_3b_2\big)x + \big(a_3b_1 - a_1b_3\big)y + \big(a_1b_2 - a_2b_1\big)z = 0,$$

and it is not clear which variable to solve for: any of the coefficients can be zero. Thus, we will use a different approach.

Let $P(x,y,z)$ be an arbitrary point in $\Pi$. We consider a parallelogram $OAPB$ in $\Pi$, whose sides $OA$ and $OB$ are parallel to $\mathbf{a}$ and $\mathbf{b}$, respectively (see Figure 7.2). Then $\overrightarrow{OA} = u\mathbf{a}$ and $\overrightarrow{OB} = v\mathbf{b}$ for some numbers $u, v$. If we denote the position vector of $P$ by $\mathbf{r}$, we conclude that $\mathbf{r} = u\mathbf{a} + v\mathbf{b}$. By varying $u, v$ we get the entire plane $\Pi$, so one possible parametrization is

$$\mathbf{r}(u,v) = (a_1u + b_1v)\mathbf{e}_1 + (a_2u + b_2v)\mathbf{e}_2 + (a_3u + b_3v)\mathbf{e}_3 \qquad (-\infty < u, v < \infty).$$ □

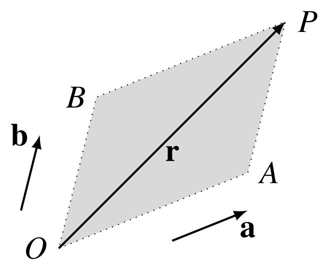



Figure 7.2. The parallelogram $OAPB$

Example 7.8. Find a parametrization of the elliptic cylinder

$$x^2 + 4z^2 = 1, \quad y \geq 5.$$

Solution. A parametrization of the ellipse $x^2+4z^2 = 1$ in the $xz$-plane is given by the functions

$$x = \cos u, \quad z = \tfrac{1}{2}\sin u \qquad (0 \leq u \leq 2\pi).$$

The points on the given cylinder must have $x$- and $z$-coordinates that satisfy these two equations and a $y$-coordinate with $y \geq 5$. Thus, we can parametrize the cylinder by

$$\mathbf{r}(u, v) = \cos u\mathbf{e}_1 + v\mathbf{e}_2 + \tfrac{1}{2}\sin u\mathbf{e}_3,$$

where $0 \leq u \leq 2\pi$ and $v \geq 5$. □

Example 7.9. Find a parametrization of the part of the plane $3x + 2y + z = 4$ that lies within the paraboloid $z = x^2 + 2y^2 + 1$.

Solution. By solving the equation of the plane for $z$, $z = 4 - 3x - 2y$, we can easily obtain a parametrization of the whole plane:

$$\mathbf{r}(u, v) = u\mathbf{e}_1 + v\mathbf{e}_2 + (4 - 3u - 2v)\mathbf{e}_3 \qquad (-\infty < u, v < \infty).$$

To parametrize just the given part of the plane, we restrict $u, v$ to its projection onto the $xy$-plane. The plane and the paraboloid intersect along the curve that contains the simultaneous solutions of the equations

$$3x + 2y + z = 4, \qquad z = x^2 + 2y^2 + 1.$$

Eliminating $z$, we find that the $x$- and $y$-coordinates of the points on this curve satisfy

$$4 - 3x - 2y = x^2 + 2y^2 + 1 \qquad \Longrightarrow \qquad \left(x + \tfrac{3}{2}\right)^2 + 2\left(y + \tfrac{1}{2}\right)^2 = \tfrac{23}{4}.$$

Thus, the intersection curve of the plane and the ellipsoid consists of all points on the plane whose $x$- and $y$-coordinates satisfy the last equation. Consequently, the intersection curve projects onto an ellipse in the $xy$-plane, and the part of the plane $3x+2y+z = 4$ that lies within the given paraboloid projects onto the interior of that ellipse. We obtain the parametrization

$$\mathbf{r}(u, v) = u\mathbf{e}_1 + v\mathbf{e}_2 + (4 - 3u - 2v)\mathbf{e}_3, \qquad \left(u + \tfrac{3}{2}\right)^2 + 2\left(v + \tfrac{1}{2}\right)^2 \leq \tfrac{23}{4}.$$ □

## 7.5. Plotting surfaces using *Mathematica*

*Mathematica* provides several commands for plotting surfaces, which we survey in this section.

**7.5.1. Graphs of functions of two variables: `Plot3D`.** To plot the surface $z = f(x, y)$, where $f$ is a function of two variables, we use the command `Plot3D`. The basic form of this command is

```
Plot3D[f[x,y], {x,a,b}, {y,c,d}]
```

where $a, b, c$ and $d$ are numbers or numerical expressions. This will generate the part of the graph of $f$ where $a \leq x \leq b$ and $c \leq y \leq d$. For example, the following command will plot the surface $z = x^2 - y^2$ over the rectangular region given by $-1 \leq x \leq 1$ and $-1 \leq y \leq 1$:

```
Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}]
```

Similarly to the `Plot` command, `Plot3D` has many options. You can get a list of these options and their default settings by executing

```
Options[Plot3D]
```

or by reading the help on `Plot3D`. Here we demonstrate a few of the more commonly used options.

- To label the axes, we use the `AxesLabel` option:

  ```
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, AxesLabel->{"x","y","z"}]
  ```

- To change the aspect ratio of the plot, we specify values to the parameter `BoxRatios`. The option `BoxRatios` specifies the ratio of the side lengths for the box that encloses the graph. The setting `BoxRatios->Automatic` gives identical scales on all three axes.

  ```
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, BoxRatios->{1,2,3}]
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, BoxRatios->Automatic]
  ```

- To suppress the box surrounding the surface we set the option `Boxed` to `False`. Similarly, the axes can be suppressed by setting `Axes` to `False`.

  ```
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, Boxed->False, Axes->False]
  ```

- The colors of the surface may be changed using the options `PlotStyle` and `ColorFunction` and the color-specifying commands `RGBColor` or `Hue` as shown in the examples below.

  ```
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, PlotStyle->Red]
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, PlotStyle->RGBColor[1,0,0]]
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1}, PlotStyle->Hue[0.5]]
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1},
   ColorFunction->Function[{x,y,z}, RGBColor[x^2,0,y^2]]]
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1},
   ColorFunction->Function[{x,y,z}, RGBColor[x^2,0,y^2]],
   ColorFunctionScaling->False]
  ```

- To eliminate the colors altogether, we set the option `ColorFunction` to `(White&)`. We can also suppress the mesh that *Mathematica* lays on the plot by setting the option `Mesh` to `False`.

  ```
  Plot3D[x^2-y^2, {x,-1,1}, {y,-1,1},
   ColorFunction->(White&), Mesh->False]
  ```

Rotating the surface around can, sometimes, reveal its shape and characteristics more clearly. To rotate a surface plot in *Mathematica*, simply left-click on it and start moving the mouse without releasing the mouse button. Sliding the mouse up and down while holding the `Ctrl` key allows you to zoom in and out the surface.

**7.5.2. Level surfaces: `ContourPlot3D`.** The basic syntax for plotting a level surface using *Mathematica* is

```
ContourPlot3D[f[x,y,z]==k, {x,x1,x2}, {y,y1,y2}, {z,z1,z2}]
```

This command plots the part of the surface $f(x, y, z) = k$ that lies within the box given by the inequalities

$$x_1 \le x \le x_2, \quad y_1 \le y \le y_2, \quad z_1 \le z \le z_2.$$

For convenience, the command `ContourPlot3D` also accepts surfaces given by the "more general" type of equation $f(x, y, z) = g(x, y, z)$:

```
ContourPlot3D[f[x,y,z]==g[x,y,z], {x,x1,x2}, {y,y1,y2}, {z,z1,z2}]
```

Finally, this command can plot simultaneously several level surfaces of the same function $f$. For example, to plot the $-1$-, 3- and 6-level surfaces of the function $x^2 - y^2 - 3z$ (in blue, red and green, respectively), we use

```
ContourPlot3D[x^2-y^2-3z, {x,-2,2}, {y,-2,2}, {z,-2,2},
 Contours->{{-1,Blue}, {3,Red}, {6,Green}}
```

**7.5.3. Parametric surfaces: `ParametricPlot3D`.** To plot the parametric surface with parametrization (7.1), we use a version of the command `ParametricPlot3D`, which we used in Lecture #5 to plot a parametric space curve. When used to plot a parametric surface, the basic syntax of this command is

```
ParametricPlot3D[{x[u,v],y[u,v],z[u,v]}, {u,u1,u2}, {v,v1,v2}]
```

Alternatively, if the surface is described in the vector form (7.2) and we have defined the vector function $\mathbf{r}(u, v)$ in *Mathematica* as `r[u,v]`, we can use the form

```
ParametricPlot3D[r[u,v], {u,u1,u2}, {v,v1,v2}]
```

Below we give examples of both forms of `ParametricPlot3D`. Both commands plot the surface with parametrization $\mathbf{r}(u, v) = u \cos v\mathbf{e}_1 + u \sin v\mathbf{e}_2 + (v/2)\mathbf{e}_3$, where $0 \le u \le 1$, $0 \le v \le 2\pi$.

```
ParametricPlot3D[{u*Cos[v], u*Sin[v], v/2}, {u,0,1}, {v,0,2Pi},
 PlotStyle->Red]

r[u_,v_] := {u*Cos[v], u*Sin[v], v/2}
ParametricPlot3D[r[u,v], {u,0,1}, {v,0,2Pi}, Mesh->None]
```

The above examples illustrate that `ContourPlot3D` and `ParametricPlot3D` accept some options that resemble those for `Plot` and `Plot3D`. This is true for most plotting options and the best way to learn about such matters is to take a look at *Mathematica*'s help and the examples there.

**7.5.4. Non-rectangular domains: `RegionFunction`.** Sometimes, one wants to plot a function or a parametric surface with a domain other than a rectangular box. Or, perhaps, one wants to plot a part of a level surface, but that part is restricted by a shape other than a box (for example, an ellipsoid). In such situation, one can use the option `RegionFunction` of the above plotting commands to restrict the variables to regions other than rectangular boxes. For example, the following command restricts the plot of the function $z = x^2 - y^2$ to the interior of the ellipse $x^2 + y^2/4 \le 1$:

```
Plot3D[x^2-y^2, {x,-3,3}, {y,-3,3},
 RegionFunction->Function[{x,y,z}, x^2+y^2/4<=1],
 PlotStyle->Red, Mesh->None]
```

Note that the value of RegionFunction is set to Function[{x,y,z}, ...] where in place of the dots we have included the restriction on the variables. That restriction can involve double inequalities or even several simultaneous inequalities (see the next example).

Here is another example of use of RegionFunction, this time inside a ContourPlot3D command:

```
ContourPlot3D[x^2+2y^2-z^2==1, {x,-3,3}, {y,-3,3}, {z,-3,3},
 RegionFunction->Function[{x,y,z}, 1<x^2+y^2<4 && x+z>-2], Mesh->None]
```

This command plots the part of the hyporboloid $x^2 + 2y^2 - z^2 = 1$ that lies between the cylinders $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ and above the plane $x + z = -2$.

The usage of RegionFunction to restrict the parameters of a parametric surface is similar, but somewhat more subtle: the region function generally takes the form

```
Function[{x,y,z,u,v}, ...]
```

with a restriction on the parameters $u$ and $v$ only. (Though it is possible to include also restrictions on $x$, $y$ and $z$, which would have the same effect as described above.) The dynamic command below plots the parametric surface

$$\mathbf{r}(u, v) = u \cos v \mathbf{e}_1 + u \sin v \mathbf{e}_2 + (v/2)\mathbf{e}_3,$$

where $0 \le u \le 1$, $0 \le v \le 2\pi$ and $u^2 + v^2/4 \le a$, with $a$ allowed to change between 0.1 and 10.

```
Manipulate[
 ParametricPlot3D[{u*Cos[v], u*Sin[v], v/2}, {u,0,1}, {v,0,2Pi},
  RegionFunction->Function[{x,y,z,u,v}, u^2+v^2/4<=a],
  PlotRange->{{-1,1}, {-1,1}, {0,Pi}}, Mesh->None],
 {a,1/10,10}]
```

It is instructive to execute it and experiment with various values of $a$ in order to see the effect on the plot from changing the restrictions on the parameters.

## Exercises

Find the traces of the given quadric surface in the planes $x = k$, $y = k$, and $z = k$. Use those traces and Table 7.1 to classify the surface.

7.1. $z = x^2 - y^2$ 7.2. $4y^2 - x^2 - z^2 = 1$ 7.3. $y^2 + 4z^2 = 5 - x^2$ 7.4. $y^2 + 4z^2 = 5 + x^2$

Rewrite the equation of the given quadric surface in a standard form and use the result and Table 7.1 to classify the surface.

7.5. $3x + 6y^2 - 3z^2 = 9$ 7.6. $x^2 + 4y^2 + z^2 - 4x + 2z = 4$

7.7. Parametrize the curve of intersection of the cylinder $x^2 + y^2 = 9$ and the plane $z = x + y + 2$.

Identify the given parametric surface.

7.8. $\mathbf{r}(u, v) = u \cos v \mathbf{e}_1 + u \sin v \mathbf{e}_2 + u^2\mathbf{e}_3$ 7.9. $\mathbf{r}(\theta, z) = 2 \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2 + z\mathbf{e}_3$

7.10. Parametrize the plane that contains the lines $\dfrac{x-1}{2} = \dfrac{y+1}{3} = \dfrac{z}{-2}$ and $\dfrac{x+2}{-3} = y = \dfrac{z-4}{4}$.

7.11. Parametrize the half of the hyperboloid $x^2 - 4y^2 - 4z^2 = 16$ that lies in the half-space $x \le 0$.

7.12. Parametrize the part of the hemisphere $x^2 + y^2 + z^2 = 5$, $z \ge 0$ that lies inside the cylinder $x^2 + y^2 = 1$.

7.13. Parametrize the part of the paraboloid $z = 4 - x^2 - y^2$ that lies above the $xy$-plane by using:
(a) $x$ and $y$ as parameters; (b) the polar coordinates in the plane as parameters.

7.14. Let $0 < a < b$. A *torus of radii a and b* is the surface of revolution obtained by rotating about $z$-axis the circle $(x - b)^2 + z^2 = a^2$ in the $xz$-plane. Parametrize the torus of radii $a$ and $b$.

LECTURE 8

# Partial Derivatives

In this lecture, we introduce the derivatives of multivariable functions—the so-called "partial derivatives."

## 8.1. First-order partial derivatives

DEFINITION. If $f$ is a function of two variables, its *partial derivatives* $f_x$ *and* $f_y$ *at* $(a, b)$ are defined by

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

and

$$f_y(a,b) = \frac{\partial f}{\partial y}(a,b) = \partial_y f(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h}.$$

If $f$ is a function of $n$ variables $x_1, x_2, \dots, x_n$, its *partial derivatives* $f_{x_1}, f_{x_2}, \dots, f_{x_n}$ *at* $(a_1, a_2, \dots, a_n)$ are defined by

$$\begin{aligned} f_{x_j}(a_1,a_2,\dots,a_n) &= \frac{\partial f}{\partial x_j}(a_1,a_2,\dots,a_n) = \partial_{x_j} f(a_1,a_2,\dots,a_n) \\ &= \lim_{h\to 0} \frac{f(a_1,\dots,a_j+h,\dots,a_n) - f(a_1,\dots,a_j,\dots,a_n)}{h} \qquad (j = 1,2,\dots,n). \end{aligned}$$

To illustrate the meaning of the partial derivatives, we take a closer look at the two-dimensional case. Given a function $f(x, y)$ and a point $(a, b)$ in its domain, we obtain two (single-variable) functions

$$F_1(x) = f(x,b) \qquad \text{and} \qquad F_2(y) = f(a,y),$$

by fixing, respectively, the variables $y$ and $x$ in $f(x, y)$. The points on the surface $z = f(x, y)$ with $y = b$ lie on the curve of intersection of that surface and the vertical plane $y = b$ (see the blue curve

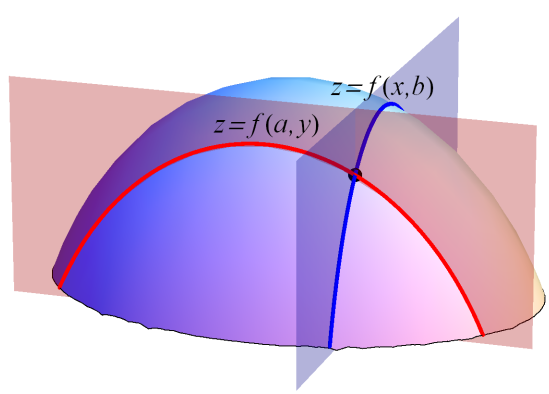



FIGURE 8.1. The partial derivatives as slopes

on Figure 8.1). That curve has exactly the same shape as the graph of the function $F_1(x)$ in the plane. Likewise, the graph of $F_2(y)$ can be visualized as the curve of intersection of the surface $z = f(x, y)$ and the vertical plane $x = a$ (see the red curve on Figure 8.1). Using the functions $F_1(x)$ and $F_2(y)$, we can express the partial derivatives of $f(x, y)$ at $(a, b)$ as

$$f_x(a,b) = \lim_{h\to 0} \frac{f(a+h,b) - f(a,b)}{h} = \lim_{h\to 0} \frac{F_1(a+h) - F_1(a)}{h} = F_1'(a) \tag{8.1}$$

and

$$f_y(a,b) = \lim_{h\to 0} \frac{f(a,b+h) - f(a,b)}{h} = \lim_{h\to 0} \frac{F_2(b+h) - F_2(b)}{h} = F_2'(b). \tag{8.2}$$

That is, the partial derivatives $f_x$ and $f_y$ of $f$ are the derivatives of the functions $F_1(x)$ and $F_2(y)$. Recall that $F_1'(a)$ and $F_2'(b)$ are the slopes of the graphs of those functions at $x = a$ and $y = b$, respectively. Looking at the copies of those graphs on the surface $z = f(x, y)$ (the blue and red curves on Figure 8.1), we see that:

*The partial derivatives $f_x(a, b)$ and $f_y(a, b)$ represent the slopes at $(a, b)$ of the sections of the graph $z = f(x, y)$ with the vertical planes $y = b$ and $x = a$, respectively. That is, $f_x(a, b)$ and $f_y(a, b)$ represent the rates of change of $f$ at $(a, b)$ in the directions of the $x$- and the $y$-axes, respectively.*

Furthermore, equations (8.1) and (8.2) can be extended to functions of more than two variables. In that more general form, they lead to the following practical observation.

*To compute the partial derivative $\partial_{x_j} f(x_1, \dots, x_n)$ with respect to one of the variables, one can treat all other variables as constants and differentiate the resulting single-variable function $F(x_j)$ using the familiar rules from single-variable calculus.*

The following example demonstrates how we use this simple observation to calculation of partial derivatives.

Example 8.1. Find the first-order partials of the functions

$$f(x,y) = \frac{x}{x^2+y^2}, \qquad g(x,y,z) = e^{xy}\sin xz.$$

Solution. The partial derivatives of $f$ are

$$f_x(x,y) = \frac{1\cdot(x^2+y^2) - x\cdot(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2}$$

and

$$f_y(x,y) = x\cdot\frac{-(2y)}{(x^2+y^2)^2} = \frac{-2xy}{(x^2+y^2)^2}.$$

The partials of $g$ are

$$g_x(x,y,z) = ye^{xy}\sin xz + e^{xy}(z\cos xz),$$
$$g_y(x,y,z) = xe^{xy}\sin xz, \qquad g_z(x,y,z) = e^{xy}(x\cos xz).$$

□

## 8.2. Higher-order partial derivatives

Just as we can differentiate the derivative $f'(x)$ of a single-variable function to obtain its second derivative, we can differentiate the partial derivatives of a multivariable function to obtain its higher-order partials. For example, a function $f$ of two variables has four *second-order partial derivatives* $f_{xx}$, $f_{xy}$, $f_{yx}$, *and* $f_{yy}$:

$$f_{xx} = \frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \qquad f_{xy} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right),$$
$$f_{yx} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad f_{yy} = \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

Example 8.2. Find the second-order partials of the function $f(x, y) = x^2 \cos y + y^2 \sin x$.

Solution. We have

$$f_x(x, y) = 2x \cos y + y^2 \cos x, \qquad f_y(x, y) = -x^2 \sin y + 2y \sin x.$$

The second-order partials of $f$ then are:

$$f_{xx} = 2 \cos y - y^2 \sin x, \qquad f_{xy} = -2x \sin y + 2y \cos x,$$
$$f_{yx} = -2x \sin y + 2y \cos x, \qquad f_{yy} = -x^2 \cos y + 2 \sin x. \qquad \square$$

Remark. Notice that the second-order partials of the function in the example satisfy $f_{xy} = f_{yx}$. In other words, for this function it is not important whether we differentiate first with respect to $x$ and then with respect to $y$ or vice versa. One might ask whether this is always true. The answer is: "only for 'nice' functions". The following theorem gives a necessary condition for this to hold.

Theorem 8.1. *Suppose that $f$ is a function of two variables defined near $(a, b)$ and having second-order partials $f_{xy}$ and $f_{yx}$ that are continuous at $(a, b)$. Then*

$$f_{xy}(a, b) = f_{yx}(a, b).$$

The following example illustrates the usefulness of the above theorem.

Example 8.3. Find the second-order partials of

$$f(x, y, z) = \sin(x^2 + 2y^2 - z).$$

Solution. The first-order partials of $f$ are:

$$f_x(x, y, z) = 2x \cos(x^2 + 2y^2 - z),$$
$$f_y(x, y, z) = 4y \cos(x^2 + 2y^2 - z),$$
$$f_z(x, y, z) = -\cos(x^2 + 2y^2 - z).$$

There are nine second-order partials, but we really have to compute only six of them. We have

$$f_{xx}(x, y, z) = 2 \cos(x^2 + 2y^2 - z) + 2x\big(-2x \sin(x^2 + 2y^2 - z)\big),$$
$$f_{yy}(x, y, z) = 4 \cos(x^2 + 2y^2 - z) + 4y\big(-4y \sin(x^2 + 2y^2 - z)\big),$$
$$f_{zz}(x, y, z) = -\sin(x^2 + 2y^2 - z).$$

For the mixed partials, we have

$$\begin{aligned} f_{xy}(x,y,z) &= 2x\big(-4y\sin(x^2+2y^2-z)\big) = -8xy\sin(x^2+2y^2-z),\\ f_{zy}(x,y,z) &= -\big(-4y\sin(x^2+2y^2-z)\big) = 4y\sin(x^2+2y^2-z),\\ f_{zx}(x,y,z) &= -\big(-2x\sin(x^2+2y^2-z)\big) = 2x\sin(x^2+2y^2-z). \end{aligned}$$

Note that the last three functions are continuous (they are products of polynomials in $x, y, z$ and compositions of sine functions and polynomials in $x, y, z$). Furthermore, it is clear that the partials $f_{yx}$, $f_{xz}$, and $f_{yz}$ will also be continuous. But then the theorem implies that, in fact, we have

$$f_{yx} = f_{xy}, \qquad f_{xz} = f_{zx}, \qquad f_{yz} = f_{zy}. \qquad \square$$

EXAMPLE 8.4. Find $f_{xy}(0,0)$ and $f_{yx}(0,0)$ for

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

SOLUTION. The first order partials of $f$ are:

$$f_x(x,y) = \begin{cases} \dfrac{y^5 - x^4y - 4x^2y^3}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0),\\ 0 & \text{if } (x,y) = (0,0), \end{cases}$$

and

$$f_y(x,y) = \begin{cases} \dfrac{xy^4 + 4x^3y^2 - x^5}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0),\\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

These formulas are straightforward everywhere except at $(x,y) = (0,0)$, where we use the definition of partial derivatives:

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

and

$$f_y(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0.$$

Then

$$f_{xy}(0,0) = \lim_{h\to 0} \frac{f_x(0,h) - f_x(0,0)}{h} = \lim_{h\to 0} \frac{h-0}{h} = 1$$

and

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h\to 0} \frac{-h-0}{h} = -1,$$

that is, $f$ fails to satisfy $f_{xy}(0,0) = f_{yx}(0,0)$. On the other hand, when $(x,y) \neq (0,0)$, we have

$$f_{xy}(x,y) = f_{yx}(x,y) = \frac{y^6 + 9x^2y^4 - 9x^4y^2 - x^6}{(x^2+y^2)^3}.$$

This is not surprising, considering that the function

$$\frac{y^6 + 9x^2y^4 - 9x^4y^2 - x^6}{(x^2+y^2)^3}$$

is continuous everywhere except at $(0,0)$ (since it is a rational function in $x$ and $y$, whose denominator vanishes only at $(0,0)$). □

### 8.3. Tangent planes

Recall that the equation of the tangent line to the curve $y = f(x)$ at a point $P(x_0, y_0)$ on the curve is given by

$$y - y_0 = f'(x_0)(x - x_0).$$

The following theorem is the analog of this formula for functions of two variables. It describes the equation of the tangent plane to the surface $z = f(x, y)$ at a point $P(x_0, y_0, z_0)$ on that surface in terms of the partials of $f$. Strictly speaking, we should first **define** what a tangent plane is, but we will leave that for Lecture #10. The same applies to the proof of the theorem.

THEOREM 8.2. *Suppose that $f$ has continuous partial derivatives at $(x_0, y_0)$. Then the equation of the tangent plane to the surface $z = f(x, y)$ at the point $P(x_0, y_0, z_0)$ is*

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

EXAMPLE 8.5. Find the tangent plane to the surface $z = e^{x^2-y^2}$ at $(-1, 0, e)$.

SOLUTION. First, observe that the point is indeed on the surface:

$$e = e^{(-1)^2-0^2}.$$

The partials of the given function are

$$\frac{\partial z}{\partial x} = 2xe^{x^2-y^2}, \qquad \frac{\partial z}{\partial y} = -2ye^{x^2-y^2},$$

whence

$$\frac{\partial z}{\partial x}(-1, 0) = -2e, \qquad \frac{\partial z}{\partial y}(-1, 0) = 0.$$

Applying the theorem, we find that the equation of the tangent plane is

$$z - e = -2e(x - (-1)) + 0(y - 0) \quad \Longleftrightarrow \quad 2ex + z + e = 0.$$ □

### 8.4. Linear approximations and differentiable functions*

Recall that for a single-variable function $f$ which has a derivative $f'(a)$ at $x = a$ (we called such functions "differentiable"), the tangent line at $(a, f(a))$ always exists and has equation

$$y = f(a) + f'(a)(x - a).$$

That is, the linear function $f(a) + f'(a)(x - a)$ provides a very good approximation to $f(x)$ when $x$ is near $a$:

$$f(x) \approx f(a) + f'(a)(x - a).$$

When we deal with functions of two variables, we may encounter functions $f(x, y)$ for which both partial derivatives $f_x(a, b)$ and $f_y(a, b)$ exist, but the plane

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

is **not** tangent to the surface $z = f(x, y)$ and does not provide a good approximation to $f(x, y)$ for $(x, y)$ near $(a, b)$. Thus, in the case of functions of two or more variables, we reserve the term "differentiable function" for functions for which this type of approximation (known as the *linear approximation to $f$ at $(a, b)$*) does hold.

Definition. A function $f$ of two variables is called *differentiable at* $(a, b)$, if near $(a, b)$ we have

$$f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + \varepsilon_1(x, y)(x - a) + \varepsilon_2(x, y)(y - b),$$

where

$$\lim_{(x,y)\to(a,b)} \varepsilon_1(x, y) = \lim_{(x,y)\to(a,b)} \varepsilon_2(x, y) = 0.$$

Similarly, a function $f$ of three variables is called *differentiable at* $(a, b, c)$, if near $(a, b, c)$ we have

$$\begin{aligned} f(x, y, z) = f(a, b, c) &+ f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) \\ &+ \varepsilon_1(x, y, z)(x - a) + \varepsilon_2(x, y, z)(y - b) + \varepsilon_3(x, y, z)(z - c), \end{aligned}$$

where

$$\lim_{(x,y,z)\to(a,b,c)} \varepsilon_1(x, y, z) = \lim_{(x,y,z)\to(a,b,c)} \varepsilon_2(x, y, z) = \lim_{(x,y,z)\to(a,b,c)} \varepsilon_3(x, y, z) = 0.$$

Example 8.6. Show that the function $f(x, y) = x^3 + xy - 2y^3$ is differentiable at $(1, -1)$.

Solution. We have

$$f(1, -1) = 2, \quad f_x(1, -1) = 2, \quad f_y(1, -1) = -5.$$

Hence, we are looking for functions $\epsilon_1(x, y)$ and $\epsilon_2(x, y)$ such that

$$f(x, y) = 2 + 2(x - 1) - 5(y + 1) + \epsilon_1(x, y)(x - 1) + \epsilon_2(x, y)(y + 1),$$
$$\lim_{(x,y)\to(1,-1)} \epsilon_1(x, y) = \lim_{(x,y)\to(1,-1)} \epsilon_2(x, y) = 0.$$

To find such functions $\epsilon_1(x, y)$ and $\epsilon_2(x, y)$, we will transform the expression

$$f(x, y) - 2 - 2(x - 1) + 5(y + 1).$$

Similarly to Example 6.10, it is convenient to rewrite this expression in terms of the variables $u = x - 1$ and $v = y + 1$. We have

$$\begin{aligned} f(u + 1, v - 1) - 2 - 2u + 5v &= u^3 - 2v^3 + 3u^2 + uv + 6v^2 \\ &= (u^2 + 3u + v)u + (-2v^2 + 6v)v. \end{aligned}$$

Returning back to the original variables $x$ and $y$, we now obtain

$$f(x, y) - 2 - 2(x - 1) + 5(y + 1) = \epsilon_1(x, y)(x - 1) + \epsilon_2(x, y)(y + 1),$$

where

$$\epsilon_1(x, y) = (x - 1)^2 + 3(x - 1) + (y + 1), \quad \epsilon_2(x, y) = -2(y + 1)^2 + 6(y + 1).$$

Thus, we have now found functions $\epsilon_1(x, y)$ and $\epsilon_2(x, y)$ satisfying (8.4). Moreover, it is easy to see that these functions satisfy also (8.4). Therefore, the function $f(x, y)$ is differentiable at $(1, -1)$. □

Next, we state an alternative definition of differentiability, which is sometimes more convenient to use. For example, this definition is easier to apply when one aims to show that a given function is not differentiable (see Exercises 8.30 and 8.31).

Definition. A function $f$ of two variables is called *differentiable at* $(a, b)$, if

$$\lim_{(x,y)\to(a,b)} \frac{f(x, y) - f(a, b) - f_x(a, b)(x - a) - f_y(a, b)(y - b)}{\sqrt{(x - a)^2 + (y - b)^2}} = 0.$$

Similarly, a function $f$ of three variables is called *differentiable at* $(a,b,c)$, if

$$\lim_{(x,y,z)\to(a,b,c)} \frac{f(x,y,z)-f(a,b,c)-f_x(a,b,c)(x-a)-f_y(a,b,c)(y-b)-f_z(a,b,c)(z-c)}{\sqrt{(x-a)^2+(y-b)^2+(z-c)^2}} = 0.$$

As we mentioned above (and as Exercises 8.30 and 8.31 demonstrate), the existence of partial derivatives does not guarantee the differentiability of $f$. It is therefore natural to ask what condition does? It turns out that the existence of **continuous** partials suffices. For the sake of simplicity, we state the relevant theorem only for functions of two variables.

Theorem 8.3. *If the partial derivatives $f_x$ and $f_y$ exist near $(a,b)$ and are continuous at $(a,b)$, then $f$ is differentiable at $(a,b)$.*

## 8.5. The chain rule

Recall that the chain rule in single-variable calculus gives a formula for the derivative of a composite function: if $y = f(x(t))$, then

$$\frac{dy}{dt} = \frac{df}{dx}\frac{dx}{dt}.$$

In this section, we discuss the generalizations of the chain rule to functions of two and more variables. Our first result concerns composite functions of the form $F(t) = f(x_1(t),\dots,x_n(t))$, where $f$ is a differentiable function of $n$ variables and $x_j = x_j(t)$, $1 \le j \le n$, are differentiable single-variable functions.

Theorem 8.4. *Let $f$ be a differentiable function of $n$ variables and let $x_1(t),\dots,x_n(t)$ be differentiable single-variable functions. Then $z = f(x_1(t),\dots,x_n(t))$ is a differentiable function of $t$, and its derivative is*

$$\frac{dz}{dt} = \frac{\partial f}{\partial x_1}\frac{dx_1}{dt} + \cdots + \frac{\partial f}{\partial x_n}\frac{dx_n}{dt}.$$

Remark. Take a note that we use "partial d's," $\frac{\partial}{\partial x_j}$, to denote derivatives of functions of two variables, but we use "straight d's," $\frac{d}{dt}$, to denote derivatives of functions of a single variable. It is important to follow this convention when dealing with a mixture of multivariable and single-variable functions.

Example 8.7. Let $z = f(x,y) = x\ln(x+2y)$, $x = \cos t$, and $y = \sin t$. Find $z'(0)$.

Solution. The partials of $f$ are

$$f_x(x,y) = \ln(x+2y) + \frac{x}{x+2y}, \qquad f_y(x,y) = \frac{2x}{x+2y},$$

and the derivatives of $x(t)$ and $y(t)$ are

$$x'(t) = -\sin t, \qquad y'(t) = \cos t.$$

Hence, by the above theorem, the derivative $z'(0)$ of the composite function $z = f(x(t),y(t))$ at $t=0$ is

$$\begin{aligned}\frac{dz}{dt}(0) &= \frac{\partial z}{\partial x}(1,0)\frac{dx}{dt}(0) + \frac{\partial z}{\partial y}(1,0)\frac{dy}{dt}(0)\\ &= \left(\ln(1+0) + \frac{1}{1+0}\right)(-0) + \left(\frac{2}{1+0}\right)(1) = 2.\end{aligned}$$

In general, we have

$$\frac{dz}{dt} = \left(\ln(\cos(t) + 2\sin t) + \frac{\cos t}{\cos t + 2\sin t}\right)(-\sin t) + \left(\frac{2\cos t}{\cos t + 2\sin t}\right)(\cos t).$$

□

We now state the general version of the chain rule for compositions of multivariable functions.

THEOREM 8.5. *Let $f$ be a differentiable function of $n$ variables and let each $x_j = x_j(t_1, \dots, t_m)$, $1 \le j \le n$, be a differentiable function of (the same) $m$ variables. Then*

$$u = f(x_1(t_1, \dots, t_m), \dots, x_n(t_1, \dots, t_m))$$

*is a differentiable function of $t_1, \dots, t_m$, with partial derivatives*

$$\frac{\partial u}{\partial t_i} = \frac{\partial f}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \cdots + \frac{\partial f}{\partial x_n}\frac{\partial x_n}{\partial t_i} \qquad (i = 1, \dots, m).$$

EXAMPLE 8.8. Let $z = x^2 - xy$, $x = s\sin t$, and $y = t\sin s$. Find $\partial z/\partial s$ and $\partial z/\partial t$.

SOLUTION. We have

$$z_x(x, y) = 2x - y, \qquad z_y(x, y) = -x,$$

and

$$x_s(s, t) = \sin t, \qquad x_t(s, t) = s\cos t, \qquad y_s(s, t) = t\cos s, \qquad y_t(s, t) = \sin s.$$

Hence, the partials of the composite function $z(x(s, t), y(s, t))$ are

$$\begin{aligned} z_s(s, t) &= z_x(x(s, t), y(s, t))x_s(s, t) + z_y(x(s, t), y(s, t))y_s(s, t) \\ &= (2s\sin t - t\sin s)\sin s + (-s\sin t)(t\cos s) \end{aligned}$$

and

$$\begin{aligned} z_t(s, t) &= z_x(x(s, t), y(s, t))x_t(s, t) + z_y(x(s, t), y(s, t))t_s(s, t) \\ &= (2s\sin t - t\sin s)(s\cos t) + (-s\sin t)\sin s. \end{aligned}$$

□

REMARK. Often, we do not substitute $x(s, t)$ and $y(s, t)$ in the formulas for the partials of $z(s, t)$. In such cases, the answer in the above example would take the form

$$z_s(s, t) = (2x - y)\sin s - xt\cos s, \qquad z_t(s, t) = (2x - y)s\cos t - x\sin s.$$

EXAMPLE 8.9. Find the partial derivatives of $w(x, y, t)$, where

$$w = \sqrt{u^2 + v^2}, \qquad u = y + x\cos t, \qquad v = x + y\sin t.$$

SOLUTION. The partial derivatives of these functions are:

$$\frac{\partial w}{\partial u} = \frac{u}{\sqrt{u^2 + v^2}}, \qquad \frac{\partial w}{\partial v} = \frac{v}{\sqrt{u^2 + v^2}},$$

$$\frac{\partial u}{\partial x} = \cos t, \qquad \frac{\partial u}{\partial y} = 1, \qquad \frac{\partial u}{\partial t} = -x\sin t,$$
$$\frac{\partial v}{\partial x} = 1, \qquad \frac{\partial v}{\partial y} = \sin t, \qquad \frac{\partial v}{\partial t} = x\cos t.$$

Hence, the partial derivatives of the composite function $w(x, y, t)$ are:

$$\begin{aligned}
\frac{\partial w}{\partial x} &= \frac{\partial w}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial x} = \frac{u\cos t}{\sqrt{u^2+v^2}} + \frac{v}{\sqrt{u^2+v^2}} = \frac{u\cos t + v}{\sqrt{u^2+v^2}},\\
\frac{\partial w}{\partial y} &= \frac{\partial w}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial y} = \frac{u}{\sqrt{u^2+v^2}} + \frac{v\sin t}{\sqrt{u^2+v^2}} = \frac{u + v\sin t}{\sqrt{u^2+v^2}},\\
\frac{\partial w}{\partial t} &= \frac{\partial w}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial w}{\partial v}\frac{\partial v}{\partial t} = \frac{-ux\sin t}{\sqrt{u^2+v^2}} + \frac{vx\cos t}{\sqrt{u^2+v^2}} = \frac{-ux\sin t + vx\cos t}{\sqrt{u^2+v^2}}.
\end{aligned}$$

□

EXAMPLE 8.10. Find the partial derivative $w_{yx}$ of $w(x, y, t)$, where

$$w = \sqrt{u^2+v^2}, \qquad u = y + x\cos t, \qquad v = x + y\sin t.$$

SOLUTION. To find a higher-order partial derivative of a composite function, we "mix" the chain rule with the other rules for partial differentiation. We have

$$\begin{aligned}
\frac{\partial^2 w}{\partial x\partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial w}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right)\\
&= \frac{\partial}{\partial u}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right)\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right)\frac{\partial v}{\partial x}.
\end{aligned}$$

We now compute the partials with respect to $u$ and $v$:

$$\begin{aligned}
\frac{\partial}{\partial u}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right) &= \frac{\partial}{\partial u}\left((u+v\sin t)(u^2+v^2)^{-1/2}\right)\\
&= (u^2+v^2)^{-1/2} + (u+v\sin t)(-u)(u^2+v^2)^{-3/2}\\
&= (u^2+v^2)^{-3/2}\left[(u^2+v^2) - u(u+v\sin t)\right] = \frac{v^2 - uv\sin t}{(u^2+v^2)^{3/2}},\\
\frac{\partial}{\partial v}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right) &= \frac{\partial}{\partial v}\left((u+v\sin t)(u^2+v^2)^{-1/2}\right)\\
&= (\sin t)(u^2+v^2)^{-1/2} + (u+v\sin t)(-v)(u^2+v^2)^{-3/2}\\
&= (u^2+v^2)^{-3/2}\left[(u^2+v^2)\sin t - v(u+v\sin t)\right] = \frac{u^2\sin t - uv}{(u^2+v^2)^{3/2}}.
\end{aligned}$$

Substituting these partials into the expression for $w_{yx}$, we get

$$\begin{aligned}
\frac{\partial^2 w}{\partial x\partial y} &= \frac{\partial}{\partial u}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right)\frac{\partial u}{\partial x} + \frac{\partial}{\partial v}\left(\frac{u+v\sin t}{\sqrt{u^2+v^2}}\right)\frac{\partial v}{\partial x}\\
&= \frac{v^2 - uv\sin t}{(u^2+v^2)^{3/2}}\cdot\cos t + \frac{u^2\sin t - uv}{(u^2+v^2)^{3/2}}\cdot 1\\
&= \frac{u^2\sin t - uv(1+\sin t\cos t) + v^2\cos t}{(u^2+v^2)^{3/2}}.
\end{aligned}$$

□

## 8.6. Partial derivatives in *Mathematica*

Recall that if $f(x)$ is a single-variable function, we can calculate its derivative in *Mathematica* using either of the following two commands:

```
f'[x]
D[f[x],x]
```

The `D` command that appears in the latter option is very versatile and is used also to calculate partial derivatives. Its basic form for multivariable functions is

`D[f[`$x_1, x_2, \dots, x_n$`], ` $x_i, x_j, \dots$`]`

For a function $f(x_1, \dots, x_n)$, this command will calculate the partial derivative $f_{x_i,x_j,\dots}(x_1, \dots, x_n)$. Here are some examples:

| | |
|---|---|
| `D[f[x,y,z],x]` | $f_x(x, y, z)$ |
| `D[f[x,y,z],x,z]` | $f_{xz}(x, y, z)$ |
| `D[f[x,y],x,x,y]` | $f_{xxy}(x, y)$ |
| `D[f[x,u,v],v,v,x,u]` | $f_{vvxu}(x, u, v)$ |

The `D` command has also two other forms that provides useful shortcuts for special kinds of partial derivatives. If $f$ is a function of several variables, one of which is $x$, then

```
D[f[x,y,...], {x,n}]
```

is a quick way to obtain the $n$-th order partial $\partial^n f/\partial x^n$; furthermore, the command

```
D[f[x,y,...,z], {{x,y,...,z}}]
```

will calculate the gradient $\nabla f(x, y, \dots, z)$ (see Lecture #9).

There is one use of the `D` command that requires extra care. Suppose that we have a function $f(x, y) = x^2 - y^2$ and we want to define a new function $g$ by $g(x, y) = f_x(x, y) + f_y(x, y)$. It may feel natural to use the following *Mathematica* commands:

```
f[x_,y_] := x^2 - y^2
g[x_,y_] := D[f[x,y],x] + D[f[x,y],y]
```

However, the function $g$ defined this way behaves rather bizarrely. For instance, asking *Mathematica* to execute simple commands like

```
g[0,0]
Plot3D[g[x,y], {x,-1,1}, {y,-1,1}]
```

results in cryptic error messages like "General::ivar: 0 is not a valid variable." The proper way to define $g$ in this case is to use "=" instead of "`:=`" in the above code

```
g[x_,y_] = D[f[x,y],x] + D[f[x,y],y]
```

Try the above commands with $g$ defined this way.

### Exercises

Find the first-order partial derivatives of the given function.

8.1. $f(x, y) = x^4 - 4x^2y^{3/2}$

8.2. $f(x, y, z) = e^{2xy} \sin xz$

8.3. $f(x, y, z) = \ln(3x + 7yz)$

8.4. $f(x, y) = \left(x^2 + xy + y^2\right)^x$

8.5. $f(x, z) = \sqrt{x^2 + z^2 + 1}$

8.6. $f(r, \theta) = r \tan^2 \theta$

Find the second-order partial derivatives of the given function.

8.7. $f(x,y) = x^4 - 4x^2y^{3/2}$ 8.8. $f(x,y) = \cos\left(x^2 + xy + y^2\right)$ 8.9. $f(x,z) = \sqrt{x^2 + z^2 + 1}$

Find the indicated partial derivative of the given function.

8.10. $f(x,y) = x^4 - 4x^2y^{3/2}$; $f_{xyx}$

8.11. $f(x,z) = \sqrt{x^2 + z^2 + 1}$; $\partial^3 f/\partial x\partial^2 z$

8.12. $f(x,y) = \cos\left(x^2 + xy + y^2\right)$; $f_{xxy}$

8.13. $f(x,y,z) = xz^2e^{2yz}$; $f_{xyz}$, $f_{zzy}$

Use the chain rule to calculate $dF/dt$ for the given function $F$.

8.14. $F(t) = f(x(t), y(t))$, where $f(x,y) = 2xy + y^3$, $x(t) = t^2 + 3t + 4$, $y(t) = t^2$

8.15. $F(t) = f(x(t), y(t))$, where $f(x,y) = xye^{x^2+4y^2}$, $x(t) = 2\cos t$, $y(t) = \sin t$

8.16. $F(t) = f(x(t), y(t), z(t))$, where $f(x,y,z) = x^2 + y^2 - 3xyz$, $x(t) = t\sin t$, $y(t) = t\cos t$, $z(t) = \frac{1}{2}t^2$

8.17. $F(t) = f(x(t), y(t), z(t))$, where $f(x,y,z) = \ln(xy + 2yz - 3xz)$, $x(t) = 2e^t$, $y(t) = 3e^{-t}$, $z(t) = e^{2t}$

Use the chain rule to calculate $\partial F/\partial s$ and $\partial F/\partial t$ for the given function $F$.

8.18. $F(s,t) = f(x(s,t), y(s,t))$, where $f(x,y) = 2x^2y - xy^2$, $x(s,t) = t^2 + 3s^2$, $y(s,t) = 3st$

8.19. $F(s,t) = f(x(s,t), y(s,t))$, where $f(x,y) = \arctan\left(x^2 - 4y^2\right)$, $x(s,t) = 2se^t + 2se^{-t}$, $y(s,t) = te^s - te^{-s}$

8.20. $F(s,t) = f(x(s,t), y(s,t), z(s,t))$, where $f(x,y,z) = \ln\left(x^2 + 2y^2 + 3z^2\right)$, $x(s,t) = s\cos t$, $y(s,t) = s\sin t$, $z(s,t) = t^2$

8.21. $F(s,t) = f(x(s,t), y(s,t), z(s,t))$, where $f(x,y,z) = \cos xy + y3^x + z^2$, $x(s,t) = s^2 + t$, $y(s,t) = s - t^2$, $z(s,t) = st + s + t$

Use the chain rule to calculate the given partial derivative at the given point.

8.22. $g_r(r,\theta)$ at $(r,\theta) = (2,\pi)$, where $g(x,y) = xy^2 - x^2y$, $x = r\cos\theta$, $y = r\sin\theta$

8.23. $h_u(u,v)$ at $(u,v) = (2,-3)$, where $h(x,y,z) = e^{x^2-y^2-z^2}$, $x = 2u + v$, $y = u - v$, $z = u + v + 1$

8.24. $\partial R/\partial x$ at $(x,y) = (1,1)$, where $R(u,v,w) = \ln(u^2 + v^2 + w^2)$, $u = x + 2y$, $v = 2x - y$, $w = 2xy$

8.25. Show that the function $u(x,y) = ae^y\cos x + be^y\sin x$ is a solution of the differential equation $u_{xx} + u_{yy} = 0$.

8.26. Show that any function $u(x,t)$ of the form

$$u(x,t) = f(x + at) + g(x - at)$$

is a solution of the differential equation $u_{tt} = a^2u_{xx}$.

Find the linear approximation to the given differentiable function near the given point.

8.27. $f(x,y) = \ln(2xy + x + y)$, $(2,1)$ 8.28. $f(x,y,z) = \arctan(xy + xz + 2yz)$, $(3,2,1)$

8.29. Show that the function $f(x,y) = xy$ is differentiable at $(1,2)$ by finding functions $\epsilon_1(x,y)$ and $\epsilon_2(x,y)$ that satisfy the definition of differentiability in §8.4.

8.30. (a) Calculate $f_x(0,0)$ and $f_y(0,0)$ for the function

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

(b) Show that $f(x,y)$ is not continuous at $(0,0)$.

(c) Show that $f(x,y)$ is not differentiable at $(0,0)$.

8.31. Let $a, b, c$, and $d$ be real numbers.

(a) Calculate $f_x(0,0)$ and $f_y(0,0)$ for the function

$$f(x,y) = \begin{cases} \dfrac{ax^3 + bx^2y + cxy^2 + dy^3}{x^2 + y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

(b) Show that $f(x,y)$ is continuous at $(0,0)$.

(c) Show that if $a \neq c$ or $b \neq d$, then $f(x,y)$ is not differentiable at $(0,0)$.

LECTURE 9

# Directional Derivatives and Gradients

## 9.1. Directional derivatives

Given a function of two variables $f(x, y)$, we know how to compute its rate of change in the $x$-direction and in the $y$-direction: the rate of change in the $x$-direction is given by the partial derivative with respect to $x$,

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

and the rate of change in the $y$-direction is given by the partial derivative with respect to $y$,

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

A natural question then is how to compute the rate of change of $f$ in other directions. This question leads to the notion of a "directional derivative."

Definition. If $f$ is a function of two variables and if $\mathbf{u} = \langle a, b \rangle$ is a unit vector, the *directional derivative of $f$ in the $\mathbf{u}$-direction* is

$$D_{\mathbf{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + ha, y + hb) - f(x, y)}{h}, \tag{9.1}$$

if the limit exists. If $f$ is a function of three variables and if $\mathbf{u} = \langle a, b, c \rangle$ is a unit vector, the *directional derivative of $f$ in the $\mathbf{u}$-direction* is

$$D_{\mathbf{u}} f(x, y, z) = \lim_{h \to 0} \frac{f(x + ha, y + hb, z + hc) - f(x, y, z)}{h}. \tag{9.2}$$

Remark. Note that if we introduce the function $g(t) = f(x + ta, y + tb)$, the limit on the right side of (9.1) is also

$$\lim_{h \to 0} \frac{g(h) - g(0)}{h} = g'(0);$$

that is, $D_{\mathbf{u}} f(x, y) = g'(0)$. On the other hand, by the chain rule,

$$g'(t) = f_x(x + ta, y + tb) \cdot a + f_y(x + ta, y + tb) \cdot b,$$
$$g'(0) = f_x(x, y)a + f_y(x, y)b.$$

This is (essentially) how one proves the following result.

Theorem 9.1. *If $f$ is a differentiable function of two variables and $\mathbf{u} = \langle a, b \rangle$ is a unit vector, then*

$$D_{\mathbf{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b.$$

*If $f$ is a differentiable function of three variables and $\mathbf{u} = \langle a, b, c \rangle$ is a unit vector, then*

$$D_{\mathbf{u}} f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c.$$

Example 9.1. Find the directional derivative of $f(x, y) = \dfrac{2x}{x - y}$ at the point $(1, 0)$ in the direction of $\mathbf{v} = \mathbf{e}_1 - \sqrt{3}\mathbf{e}_2$.

Solution. First, notice that $\mathbf{v}$ is not unit and that the unit vector in the direction of $\mathbf{v}$ is

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\langle 1, -\sqrt{3}\rangle}{\sqrt{1+3}} = \left\langle \frac{1}{2}, -\frac{\sqrt{3}}{2} \right\rangle.$$

The partials of $f$ are

$$f_x(x, y) = \frac{2(x - y) - 2x}{(x - y)^2} = \frac{-2y}{(x - y)^2}, \qquad f_y(x, y) = \frac{2x}{(x - y)^2},$$

so the theorem gives

$$D_{\mathbf{u}} f(x, y) = \frac{1}{2}\frac{-2y}{(x - y)^2} + \frac{-\sqrt{3}}{2}\frac{2x}{(x - y)^2} = \frac{-y - \sqrt{3}x}{(x - y)^2}, \qquad D_{\mathbf{u}} f(1, 0) = -\sqrt{3}.$$

□

Example 9.2. Find the directional derivative of $f(x, y, z) = \dfrac{x}{y + z}$ at the point $(2, 1, 1)$ in the direction of $\mathbf{v} = \langle 1, 2, 3\rangle$.

Solution. The unit vector in the direction of $\mathbf{v}$ is

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|} = \frac{\langle 1, 2, 3\rangle}{\sqrt{1 + 4 + 9}} = \left\langle \frac{1}{\sqrt{14}}, \frac{2}{\sqrt{14}}, \frac{3}{\sqrt{14}} \right\rangle,$$

and the partial derivatives of $f$ are

$$f_x(x, y, z) = \frac{1}{y + z}, \qquad f_y(x, y, z) = \frac{-x}{(y + z)^2}, \qquad f_z(x, y, z) = \frac{-x}{(y + z)^2}.$$

Hence,

$$D_{\mathbf{u}} f(2, 1, 1) = \frac{1}{2}\frac{1}{\sqrt{14}} - \frac{1}{2}\frac{2}{\sqrt{14}} - \frac{1}{2}\frac{3}{\sqrt{14}} = \frac{-2}{\sqrt{14}}.$$

□

## 9.2. The gradient vector

Definition. If $f$ is a function of two variables $x$ and $y$, its *gradient*, denoted $\nabla f$ or grad $f$, is the vector function

$$\nabla f = \langle f_x(x, y), f_y(x, y)\rangle = \frac{\partial f}{\partial x}\mathbf{e}_1 + \frac{\partial f}{\partial y}\mathbf{e}_2.$$

If $f$ is a function of three variables $x$, $y$, and $z$, its *gradient*, $\nabla f$ or grad $f$, is

$$\nabla f = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle = \frac{\partial f}{\partial x}\mathbf{e}_1 + \frac{\partial f}{\partial y}\mathbf{e}_2 + \frac{\partial f}{\partial z}\mathbf{e}_3.$$

Remark. Note that using the gradient vector, we can restate both formulas in Theorem 9.1 in the form

$$D_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u}, \tag{9.3}$$

where $\mathbf{x} = (x, y)$ or $\mathbf{x} = (x, y, z)$.

THEOREM 9.2. *Let $f$ be a differentiable function of two or three variables. The maximum value of the directional derivative $D_{\mathbf{u}} f(\mathbf{x})$ is $\|\nabla f(\mathbf{x})\|$; it occurs when $\mathbf{u}$ has the same direction as $\nabla f(\mathbf{x})$.*

PROOF. By (9.3) and the properties of the dot product,

$$D_{\mathbf{u}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \cdot \mathbf{u} = \|\nabla f(\mathbf{x})\| \, \|\mathbf{u}\| \cos\theta = \|\nabla f(\mathbf{x})\| \cos\theta,$$

where $\theta$ is the angle between $\nabla f(\mathbf{x})$ and $\mathbf{u}$. Since $\cos\theta \leq 1$, with equality when $\theta = 0$ (i.e., when $\nabla f(\mathbf{x})$ and $\mathbf{u}$ point in the same direction), we find that

$$D_{\mathbf{u}} f(\mathbf{x}) \leq \|\nabla f(\mathbf{x})\|,$$

with an equality when $\mathbf{u}$ points in the direction of the gradient. □

EXAMPLE 9.3. By Example 9.2, the gradient of the function

$$f(x, y, z) = \frac{x}{y+z}$$

is

$$\nabla f(x, y, z) = \left\langle \frac{1}{y+z}, \frac{-x}{(y+z)^2}, \frac{-x}{(y+z)^2} \right\rangle.$$

Hence, the largest directional derivative at $(2, 1, 1)$ occurs in the direction of

$$\nabla f(2, 1, 1) = \left\langle \frac{1}{2}, \frac{-1}{2}, \frac{-1}{2} \right\rangle.$$

Its value is $\|\nabla f(2, 1, 1)\| = \frac{\sqrt{3}}{2}$. □

## Exercises

Find the directional derivative of the given function in the given direction.

9.1. $f(x, y) = \left(x^2 + y^2\right) e^{2xy}$, in the direction of $\langle 1, 2 \rangle$

9.2. $f(x, y, z) = \ln\left(xy + z^2\right)$, in the direction of $\mathbf{e}_2 - 2\mathbf{e}_3$

9.3. Find the directional derivative of $F(x, y, z) = \sqrt{4x^2 - y^2 - z^2}$ at $P(2, -1, 3)$ in the direction of $\mathbf{v} = \langle -1, -2, 2 \rangle$.

9.4. Find the directional derivative of

$$G(x, y, z) = \arcsin\left(\frac{2x}{4x^2 + y^2 + z^2}\right)$$

at the point $P(1, 2, 0)$ in the direction of the point $Q(-1, 3, 1)$.

Find the maximum rate of change of $f$ at the given point and the direction in which it occurs.

9.5. $f(x, y) = \tan(xy + x + y)$, $P(2, -1)$

9.6. $f(x, y, z) = xy + yz^2$, $Q(1, -3, 1)$

LECTURE 10

# Tangent Planes and Normal Vectors

## 10.1. Tangent planes and normal vectors to level surfaces

We now fulfill a promise made in Lecture #8. We start by defining the tangent plane to a level surface.

Definition. Let $F(x, y, z)$ be a differentiable function, and let $P(x_0, y_0, z_0)$ be a point on the surface $\Sigma$ with equation $F(x, y, z) = k$. The *tangent plane to* $\Sigma$ *at* $P$ is the plane, if it exists, that contains $P$ and the tangent vectors to any smooth curve $\gamma$ on $\Sigma$ that passes through $P$.

The following theorem establishes a formula for the equation of the tangent plane to a level surface.

Theorem 10.1. *Let $F(x, y, z)$ be a differentiable function, and let $P(x_0, y_0, z_0)$ be a point on the surface $\Sigma$ with equation $F(x, y, z) = k$. If $\nabla F(x_0, y_0, z_0) \neq \mathbf{0}$, then the tangent plane to $\Sigma$ at the point $P$ exists and is perpendicular to $\nabla F(x_0, y_0, z_0)$. In particular, the tangent plane has equation*

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0. \tag{10.1}$$

Proof. Note that if $\gamma$ is a curve on $\Sigma$ given by the vector function $\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3$, then the component functions of $\mathbf{r}(t)$ must satisfy the equation of $\Sigma$:

$$F(x(t), y(t), z(t)) = k.$$

When we differentiate this equation using the chain rule, we get

$$\frac{\partial F}{\partial x}\frac{dx}{dt} + \frac{\partial F}{\partial y}\frac{dy}{dt} + \frac{\partial F}{\partial z}\frac{dz}{dt} = 0 \quad \Longleftrightarrow \quad \nabla F \cdot \mathbf{r}'(t) = 0.$$

In other words, the tangent vector to the curve, $\mathbf{r}'(t)$, is perpendicular to the gradient vector $\nabla F$. First of all, this proves the existence of the tangent plane—it is the plane through $P$ that is perpendicular to $\nabla F(x_0, y_0, z_0)$. Second, we get that (10.1) is the equation of the tangent plane. □

Example 10.1. Find the equation of the tangent plane at a point $P(x_0, y_0, z_0)$ of the one-sheet hyperboloid

$$F(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1.$$

Solution. We have

$$F_x(x_0, y_0, z_0) = \frac{2x_0}{a^2}, \qquad F_y(x_0, y_0, z_0) = \frac{2y_0}{b^2}, \qquad F_z(x_0, y_0, z_0) = -\frac{2z_0}{c^2},$$

so the equation of the tangent plane to the hyperboloid at a point $P(x_0, y_0, z_0)$ is

$$\frac{2x_0}{a^2}(x - x_0) + \frac{2y_0}{b^2}(y - y_0) - \frac{2z_0}{c^2}(z - z_0) = 0,$$

which is equivalent to

$$\frac{2xx_0}{a^2} + \frac{2yy_0}{b^2} - \frac{2zz_0}{c^2} = \frac{2x_0^2}{a^2} + \frac{2y_0^2}{b^2} - \frac{2z_0^2}{c^2}.$$

Since $P$ lies on the hyperboloid, the right side is equal to 2 and we can rewrite the last equation in the form

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} - \frac{zz_0}{c^2} = 1.$$

□

Example 10.2. Find the equation of the tangent plane at a point $P(x_0, y_0, z_0)$ of the surface $z = f(x, y)$, where $f(x, y)$ is a differentiable function.

Solution. We can rewrite the equation of the surface as

$$f(x, y) - z = 0,$$

which is of the form considered above with $F(x, y, z) = f(x, y) - z$ and $k = 0$. Since

$$F_x(x_0, y_0, z_0) = f_x(x_0, y_0), \qquad F_y(x_0, y_0, z_0) = f_y(x_0, y_0), \qquad F_z(x_0, y_0, z_0) = -1,$$

we deduce that the equation of the tangent plane to the surface at $(x_0, y_0, z_0)$ is

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

which is equivalent to the equation stated without proof in §8.3. Therefore, Theorem 8.2 is merely special case of Theorem 10.1. □

Finally, we define the normal line to a surface.

Definition. Let $F(x, y, z)$ be a differentiable function and let $P(x_0, y_0, z_0)$ be a point of the surface $\Sigma$ with equation $F(x, y, z) = k$. The *normal line to* $\Sigma$ *at* $P$ is the line that passes through $P$ and is orthogonal to the tangent plane at $P$.

The same reasoning that we used to establish (10.1) shows that the symmetric equations of the normal line to $\Sigma$ at $P$ is

$$\frac{x - x_0}{F_x(x_0, y_0, z_0)} = \frac{y - y_0}{F_y(x_0, y_0, z_0)} = \frac{z - z_0}{F_z(x_0, y_0, z_0)},$$

with the proper modifications, if some of the partials vanish at $(x_0, y_0, z_0)$.

## 10.2. Tangent planes and normal vectors to parametric surfaces

Suppose that $\Sigma$ is a surface with parametrization

$$\mathbf{r}(u, v) = x(u, v)\mathbf{e}_1 + y(u, v)\mathbf{e}_2 + z(u, v)\mathbf{e}_3 \qquad ((u, v) \in D)$$

and $P_0(x_0, y_0, z_0) = \mathbf{r}(u_0, v_0)$ is a point on $\Sigma$. We shall focus on curves $\gamma$ on $\Sigma$ that are the images of parametric curves

$$u = u(t), \quad v = v(t) \qquad (t \in I)$$

that lie in $D$ and pass through the point $(u_0, v_0)$.

Definition. The *tangent plane to* $\Sigma$ *at* $P$ is the plane, if it exists, that contains $P$ and the tangent vectors to any smooth curve $\gamma$ on $\Sigma$ of the above type.

We remark that this definition, at least at first glance, seems more restrictive than the definition we gave in §8.3 for the tangent plane to a level surface: we look only at some curves on $\Sigma$ instead of looking at all such curves. However, it can be shown that our definition and the definition that refers to all the curves on $\Sigma$ yield the same notion of tangent plane. The following theorem establishes the existence of the tangent plane at a "smooth" point of a parametric surface.

Theorem 10.2. *Let* $\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3$, $(u,v) \in D$, *be a differentiable vector function with partials*

$$\mathbf{r}_u = x_u\mathbf{e}_1 + y_u\mathbf{e}_2 + z_u\mathbf{e}_3, \quad \mathbf{r}_v = x_v\mathbf{e}_1 + y_v\mathbf{e}_2 + z_v\mathbf{e}_3,$$

*and let* $P(x_0, y_0, z_0) = \mathbf{r}(u_0, v_0)$ *be a point on the parametric surface* $\Sigma$ *given by* $\mathbf{r}(u,v)$. *If*

$$\mathbf{n} = \mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0) \neq \mathbf{0},$$

*then the tangent plane to* $\Sigma$ *at the point* $P$ *exists and is perpendicular to the vector* $\mathbf{n}$.

Proof. Let $\gamma$ be a curve on $\Sigma$ of the type considered in the above definition. It can be parametrized by means of the vector function

$$\mathbf{f}(t) = \mathbf{r}(u(t), v(t)) = x(u(t), v(t))\mathbf{e}_1 + y(u(t), v(t))\mathbf{e}_2 + z(u(t), v(t))\mathbf{e}_3 \qquad (t \in I).$$

Thus, we can use the chain rule to express its tangent vector $\mathbf{f}'(t)$ in terms of the derivatives of $\mathbf{r}(u,v)$, $u(t)$, and $v(t)$:

$$\begin{aligned} \frac{d\mathbf{f}}{dt} &= \frac{dx}{dt}\mathbf{e}_1 + \frac{dy}{dt}\mathbf{e}_2 + \frac{dz}{dt}\mathbf{e}_3 \\ &= \left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_1 + \left(\frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_2 + \left(\frac{\partial z}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt}\right)\mathbf{e}_3 \\ &= \left(\frac{\partial x}{\partial u}\frac{du}{dt}\mathbf{e}_1 + \frac{\partial y}{\partial u}\frac{du}{dt}\mathbf{e}_2 + \frac{\partial z}{\partial u}\frac{du}{dt}\mathbf{e}_3\right) + \left(\frac{\partial x}{\partial v}\frac{dv}{dt}\mathbf{e}_1 + \frac{\partial y}{\partial v}\frac{dv}{dt}\mathbf{e}_2 + \frac{\partial z}{\partial v}\frac{dv}{dt}\mathbf{e}_3\right) \\ &= \frac{du}{dt}\left(\frac{\partial x}{\partial u}\mathbf{e}_1 + \frac{\partial y}{\partial u}\mathbf{e}_2 + \frac{\partial z}{\partial u}\mathbf{e}_3\right) + \frac{dv}{dt}\left(\frac{\partial x}{\partial v}\mathbf{e}_1 + \frac{\partial y}{\partial v}\mathbf{e}_2 + \frac{\partial z}{\partial v}\mathbf{e}_3\right) = u'(t)\mathbf{r}_u + v'(t)\mathbf{r}_v. \end{aligned}$$

In particular, the tangent vector $\mathbf{f}'(t_0)$ at the point $P_0$ of $\gamma$ is

$$\mathbf{f}'(t_0) = u'(t_0)\mathbf{r}_u(u_0, v_0) + v'(t_0)\mathbf{r}_v(u_0, v_0).$$

In other words, the tangent vector is the sum of a vector parallel to $\mathbf{r}_u(u_0, v_0)$ and a vector parallel to $\mathbf{r}_v(u_0, v_0)$. It follows that the cross product $\mathbf{n}$, which is perpendicular to both $\mathbf{r}_u(u_0, v_0)$ and $\mathbf{r}_v(u_0, v_0)$, will be perpendicular to the tangent vector $\mathbf{f}'(t_0)$ to the curve $\gamma$. Note that the last vector is independent of the curve and is completely determined by the parametrization of the surface $\Sigma$. This establishes the existence of the tangent plane: it is the plane through the point $P_0$ with normal vector $\mathbf{n}$. □

Example 10.3. Find an equation of the tangent plane to the surface $\Sigma$ with parametrization

$$\mathbf{r}(u,v) = \cos u(4 + 2\cos v)\mathbf{e}_1 + \sin u(4 + 2\cos v)\mathbf{e}_2 + 2\sin v\mathbf{e}_3 \qquad (0 \le u, v \le 2\pi)$$

at the point $(4, 0, 2) = \mathbf{r}(0, \pi/2)$.

Solution. We have

$$\mathbf{r}_u(u,v) = -\sin u(4+2\cos v)\mathbf{e}_1 + \cos u(4+2\cos v)\mathbf{e}_2,$$
$$\mathbf{r}_v(u,v) = -2\cos u \sin v\mathbf{e}_1 - 2\sin u \sin v\mathbf{e}_2 + 2\cos v\mathbf{e}_3,$$

so the normal vector to $\Sigma$ at $(4,0,2)$ is

$$\mathbf{n} = (\mathbf{r}_u \times \mathbf{r}_v)(0,\pi/2) = \langle 0,4,0\rangle \times \langle -2,0,0\rangle = (4\mathbf{e}_2)\times(-2\mathbf{e}_1) = -8(\mathbf{e}_2\times\mathbf{e}_1) = 8\mathbf{e}_3.$$

Thus, the equation of the tangent plane is

$$8(z-2) = 0 \iff z = 2. \qquad \square$$

Example 10.4. Find the equation of the tangent plane at a point $P(x_0,y_0,z_0)$ of the surface $z = f(x,y)$, where $f(x,y)$ is a differentiable function.

Solution. We can parametrize the given surface as

$$\mathbf{r}(x,y) = x\mathbf{e}_1 + y\mathbf{e}_2 + f(x,y)\mathbf{e}_3 \qquad ((x,y)\in D),$$

so the normal vector $\mathbf{n}$ to it is

$$\mathbf{n} = \mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = \begin{vmatrix} 0 & f_x \\ 1 & f_y \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} 1 & f_x \\ 0 & f_y \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}\mathbf{e}_3 = -f_x\mathbf{e}_1 - f_y\mathbf{e}_2 + \mathbf{e}_3.$$

Thus, the equation of the tangent plane at $(x_0,y_0,z_0)$ is

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

which is equivalent to the equation we gave earlier in Theorem 8.2 and Example 10.2. $\square$

### 10.3. Smooth and piecewise smooth surfaces

Recall that we called the parametric curve $\gamma$ given by the vector function $\mathbf{r}(t)$, $t \in I$, smooth if the tangent vector to the curve, $\mathbf{r}'(t)$, does not vanish for $t \in I$. We later saw that this condition ensures that $\gamma$ has a tangent line at each point. In the next definition, we define a "smooth surface" so that it has a tangent plane at each point.

Definition. The level surface $\Sigma$ with equation $F(x,y,z) = k$ is called *smooth*, if the function $F(x,y,z)$ is differentiable at every point of $\Sigma$ and $\nabla F(x,y,z) \neq \mathbf{0}$, except possibly at the boundary of $\Sigma$. The parametric surface $\Sigma$, given by

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v)\in D), \tag{10.2}$$

is called *smooth*, if the partials $\mathbf{r}_u$ and $\mathbf{r}_v$ are defined for all $(u,v)\in D$ and $\mathbf{r}_u \times \mathbf{r}_v \neq \mathbf{0}$, except possibly at the boundary of $D$.

Note that it is an immediate consequence of this definition and Theorems 10.1 and 10.2 that a smooth surface (either level or parametric) has a tangent plane at each point.

Definition. The level surface $\Sigma$ with equation $F(x,y,z) = k$ is called *piecewise smooth*, if the function $F(x,y,z)$ is continuous and we can represent $\Sigma$ as the union of a finite number of smooth surfaces, say $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \cdots \cup \Sigma_n$. The parametric surface $\Sigma$, given by (10.2) is called *piecewise smooth*, if $\mathbf{r}(u,v)$ is continuous and we can represent $\Sigma$ as the union of a finite number of smooth surfaces, say $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \cdots \cup \Sigma_n$.

EXAMPLE 10.5. Determine whether the surface $\Sigma$ is smooth, piecewise smooth, or neither:

(a) $\Sigma$ is the elliptic paraboloid $z = x^2 + 4y^2$.
(b) $\Sigma$ is the cone $z^2 = x^2 + y^2$.

FIRST SOLUTION. (a) We view $\Sigma$ as a level surface of $F(x, y, z) = z - x^2 - 4y^2$. Since

$$\nabla F(x, y, z) = \langle -2x, -8y, 1\rangle \neq \mathbf{0},$$

the paraboloid is smooth.

(b) We view $\Sigma$ as a level surface of $F(x, y, z) = z^2 - x^2 - y^2$. Since

$$\nabla F(x, y, z) = \langle -2x, -2y, 2z\rangle = \mathbf{0} \quad \Longleftrightarrow \quad x = y = z = 0,$$

the cone is not smooth. However, we can break $\Sigma$ into two surfaces: the level surface $\Sigma_1$ of

$$F(x, y, z) = z^2 - x^2 - y^2 \qquad (z \geq 0),$$

and the level surface $\Sigma_2$ of

$$F(x, y, z) = z^2 - x^2 - y^2 \qquad (z \leq 0),$$

each of which is smooth, because the origin is a boundary point for them. Thus, the cone is piecewise smooth. □

SECOND SOLUTION. (b) We view $\Sigma$ as a parametric surface given by

$$\mathbf{r}(u, v) = v \cos u\mathbf{e}_1 + v \sin u\mathbf{e}_2 + v\mathbf{e}_3,$$

where $0 \leq u \leq 2\pi$, $-\infty < v < \infty$. Then

$$\mathbf{r}_u = -v \sin u\mathbf{e}_1 + v \cos u\mathbf{e}_2, \quad \mathbf{r}_v = \cos u\mathbf{e}_1 + \sin u\mathbf{e}_2 + \mathbf{e}_3,$$

and

$$\begin{aligned}\mathbf{r}_u \times \mathbf{r}_v &= (-v \sin u\mathbf{e}_1 + v \cos u\mathbf{e}_2) \times (\cos u\mathbf{e}_1 + \sin u\mathbf{e}_2 + \mathbf{e}_3)\\ &= -v \sin^2 u(\mathbf{e}_1 \times \mathbf{e}_2) - v \sin u(\mathbf{e}_1 \times \mathbf{e}_3) + v \cos^2 u(\mathbf{e}_2 \times \mathbf{e}_1) + v \cos u(\mathbf{e}_2 \times \mathbf{e}_3)\\ &= -v \sin^2 u\mathbf{e}_3 + v \sin u\mathbf{e}_2 - v \cos^2 u\mathbf{e}_3 + v \cos u\mathbf{e}_1 = v \cos u\mathbf{e}_1 + v \sin u\mathbf{e}_2 - v\mathbf{e}_3.\end{aligned}$$

Hence, $\mathbf{r}_u \times \mathbf{r}_v = \mathbf{0}$ when $v = 0$ and the cone is not smooth. However, we can break $\Sigma$ into two surfaces: the parametric surface $\Sigma_1$ given by

$$\mathbf{r}(u, v) = v \cos u\mathbf{e}_1 + v \sin u\mathbf{e}_2 + v\mathbf{e}_3 \qquad (0 \leq u \leq 2\pi,\ v \geq 0),$$

and the parametric surface $\Sigma_2$ given by

$$\mathbf{r}(u, v) = v \cos u\mathbf{e}_1 + v \sin u\mathbf{e}_2 + v\mathbf{e}_3 \qquad (0 \leq u \leq 2\pi,\ v \leq 0),$$

each of which is smooth, because the origin is a boundary point for them. Thus, the cone is piecewise smooth. □

## 10.4. Orientable surfaces

We conclude this lecture with the notion of orientation of a surface, which we shall need in our discussion of surface integrals later in these notes. Intuitively, we think of a surface as a two-dimensional object, which is placed in space and thus has an "upper side" and a "lower side." However, this is not always the case, even for nice, smooth surfaces. We use the concept of orientation of a surface to separate the intuitive, two-sided surfaces from the one-sided ones. However, before giving a formal definition, we present an example of a surprisingly simple one-sided surface.

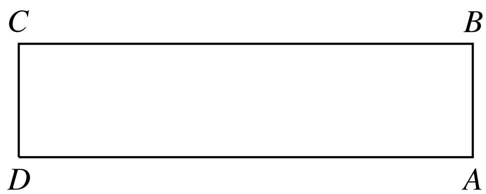



FIGURE 10.1. The Möbius strip before gluing

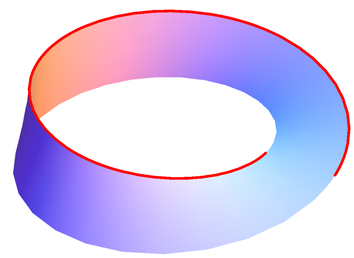


FIGURE 10.2. The Möbius strip after gluing

EXAMPLE 10.6. Take a long strip of paper in the form of a rectangle $ABCD$, the vertices listed counterclockwise so that $AB$ and $CD$ are the two short sides of the strip (see Figure 10.1). Then twist the paper half a turn and glue the short sides so that $A$ gets glued to $C$ and $B$ to $D$. The resulting surface (see Figure 10.2) has only one side! Indeed, if you take a paintbrush, started at the cut, and moved along the surface parallel to its edges, you will be able to paint the entire strip without lifting the paintbrush. This surface is known as the *Möbius strip*. You may think at first that the Möbius strip is an esoteric example and that "normal" surfaces, given by nice, simple formulas, don't behave this way. Think again! The Möbius strip *is* given by a nice, simple formula: its parametric equations are

$$x = (2 + v\cos(u/2))\cos u, \qquad y = (2 + v\cos(u/2))\sin u, \qquad z = v\sin(u/2), \tag{10.3}$$

where $0 \le u \le 2\pi$ and $-\frac{1}{2} \le v \le \frac{1}{2}$. □

DEFINITION. Let $\Sigma$ be a smooth surface in $\mathbb{R}^3$. For each nonboundary point $P(x, y, z)$ on the surface, let $\mathbf{n}_1(x, y, z)$ and $\mathbf{n}_2(x, y, z)$ be the two unit vectors at $P$ that are normal to $\Sigma$. We say that $\Sigma$ is *orientable*, if it is possible to select a unit normal vector $\mathbf{n}(x, y, z)$ at every nonboundary point of $\Sigma$ in such a way that $\mathbf{n}(x, y, z)$ varies continuously as $(x, y, z)$ varies over $\Sigma$. A choice of the unit normal vector $\mathbf{n}$ as described above is called an *orientation* of $\Sigma$.

REMARK. If a surface $\Sigma$ is orientable, then there are two distinct orientations of that surface. For example, let us choose a distinguished point $P_0(x_0, y_0, z_0)$ on $\Sigma$. Let us select and fix one of the two unit normal vectors at $P_0$; we denote it by $\mathbf{n}_0$. Then for each nonboundary point $P$ near $P_0$ only one of the unit normal vectors at $P$ will be close to $\mathbf{n}_0$, the other will be close to $-\mathbf{n}_0$. Thus, the continuity requirement on our choice of normal vectors settles the orientation of all normal vectors near $P_0$. Of course, once we have decided on the choice of normal vectors at points near $P_0$, we can repeat the same procedure to settle the orientation of the normal vectors at points that may not be near $P_0$ but are near points that are near $P_0$; after that, we can settle the choices at points that may not be near $P_0$ or near points that are near $P_0$, but are near points that are near points that are near $P_0$; etc. Ultimately, after we choose between the two unit normal vectors $\mathbf{n}_0$ and $-\mathbf{n}_0$ at $P_0$, we set a domino effect that yields a consistent choice between the unit normal vectors at all other nonboundary points of $\Sigma$.

```
r[u_,v_] := {2Cos[u]Sin[v], 2Sin[u]Sin[v], 2Cos[v]}
n[u_,v_] := 0.5*r[u,v]
p = ParametricPlot3D[r[u,v], {u,0,2Pi}, {v,0,Pi},
  Boxed -> False, Axes -> False, Mesh -> None, ViewPoint -> {6,1,4}];
Manipulate[
  Show[{p, vector[r[u,v], r[u,v] + n[u,v], Red]}, PlotRange -> All],
  {{u,0}, 0, 4Pi}, {{v,0}, 0, Pi}]

r[u_,v_] := {Cos[u](2 + v*Cos[u/2]), Sin[u](2 + v*Cos[u/2]), v*Sin[u/2]}
n[u_,v_] = 0.5*Cross[D[r[u,v], u], D[r[u,v], v]]
p = ParametricPlot3D[r[u,v], {u,0,2Pi}, {v,-1/2,1/2},
  Boxed -> False, Axes -> False, Mesh -> None, ViewPoint -> {6,1,4}];
Manipulate[
  Show[{p, vector[r[u,v], r[u,v] + n[u,v], Red]}, PlotRange -> All],
  {{u,0}, 0, 4Pi}, {{v,0}, -1/2, 1/2}]
```

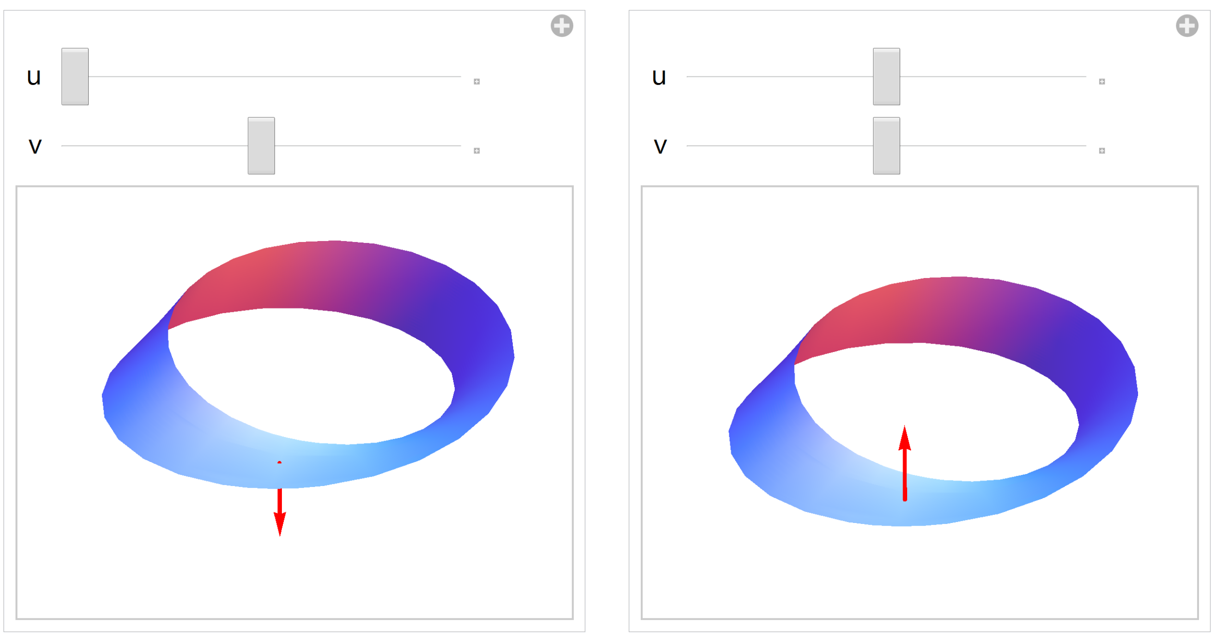



FIGURE 10.3. Normal vectors of orientable and non-orientable surfaces

EXAMPLE 10.7. A good way to illustrate the concept of orientability is to compare the normal vectors on an orientable surface to those on a non-orientable surface. Such a comparison is most successful, if it is undertaken using graphical software like *Mathematica*, so that is what we will do. In this example, we will compare the normal vectors to the sphere $x^2 + y^2 + z^2 = 4$ and the normal vectors to the Möbius strip with parametric equations (10.3).

The first block of *Mathematica* code in Figure 10.3 produces a dynamic plot of the sphere (parametrized as in Example 10.9 below) and of the outward unit normal vector at its north pole. One can then use the sliders to move the normal vector on the surface of the sphere. Execution that code[1] in *Mathematica* and experimentation with the sliders will demonstrate that as one moves the normal vector on the sphere, its direction changes but always points away from the center of the sphere. Furthermore, no matter how the normal vector meanders around, every time it returns to a specific point it points out. This is because the sphere is orientable (see Example 10.9 below); the normal vectors *Mathematica* displays are the vectors of its "outward orientation."

The second block of code in Figure 10.3 produces a similar dynamic plot of the Möbius strip and one of its normal vectors. However, experimentation with its output will demonstrate that as one moves the normal vector on the Möbius strip, the vector may be positioned at the same location with opposite directions: e.g., when $(u, v) = (0, 0)$ and when $(u, v) = (2\pi, 0)$ (see the screenshots in Figure 10.3). This is because the Möbius strip is non-orientable. □

Example 10.8. Let $\Sigma$ be the graph of a function $z = f(x, y)$, $(x, y) \in D$. In Example 10.4, we used the parametrization

$$\mathbf{r}(x, y) = x\mathbf{e}_1 + y\mathbf{e}_2 + f(x, y)\mathbf{e}_3 \qquad ((x, y) \in D)$$

to show that

$$\mathbf{r}_x \times \mathbf{r}_y = -f_x\mathbf{e}_1 - f_y\mathbf{e}_2 + \mathbf{e}_3$$

is a normal vector to the surface at the point $(x, y, z)$ of the surface. Note that this vector always has a positive third component. Thus, the normalized cross product

$$\mathbf{n} = \frac{\mathbf{r}_x \times \mathbf{r}_y}{\|\mathbf{r}_x \times \mathbf{r}_y\|} = \frac{-f_x\mathbf{e}_1 - f_y\mathbf{e}_2 + \mathbf{e}_3}{\left(1 + f_x^2 + f_y^2\right)^{1/2}},$$

will pick consistently one of the two unit normal vectors to the surface: namely, it will always pick the unit normal that points upward. Consequently, $\mathbf{n}$ represents an orientation of the surface; this orientation is often referred to as the *upward orientation* of the surface. □

Example 10.9. Let $\Sigma$ be the sphere $x^2 + y^2 + z^2 = 4$. Its spherical equation is $\rho = 2$, so we can describe $\Sigma$ as the set of points $(x, y, z)$ in space whose Cartesian coordinates are related to their spherical coordinates via the equations

$$x = 2\cos\theta\sin\phi, \quad y = 2\sin\theta\sin\phi, \quad z = 2\cos\phi,$$

where $0 \le \theta \le 2\pi$, $0 \le \phi \le \pi$. This provides a parametrization of $\Sigma$:

$$\mathbf{r}(\theta, \phi) = 2\cos\theta\sin\phi\,\mathbf{e}_1 + 2\sin\theta\sin\phi\,\mathbf{e}_2 + 2\cos\phi\,\mathbf{e}_3,$$

where $0 \le \theta \le 2\pi$, $0 \le \phi \le \pi$. Then

$$\mathbf{r}_\theta = -2\sin\theta\sin\phi\,\mathbf{e}_1 + 2\cos\theta\sin\phi\,\mathbf{e}_2, \quad \mathbf{r}_\phi = 2\cos\theta\cos\phi\,\mathbf{e}_1 + 2\sin\theta\cos\phi\,\mathbf{e}_2 - 2\sin\phi\,\mathbf{e}_3,$$

[1]The code uses the `vector` command described in §2.9, so that code needs to be included as well.

and

$$\begin{aligned}
\mathbf{r}_\theta \times \mathbf{r}_\phi &= (-2\sin\theta\sin\phi\,\mathbf{e}_1 + 2\cos\theta\sin\phi\,\mathbf{e}_2) \times (2\cos\theta\cos\phi\,\mathbf{e}_1 + 2\sin\theta\cos\phi\,\mathbf{e}_2 - 2\sin\phi\,\mathbf{e}_3) \\
&= -4\sin^2\theta\sin\phi\cos\phi(\mathbf{e}_1 \times \mathbf{e}_2) + 4\sin\theta\sin^2\phi(\mathbf{e}_1 \times \mathbf{e}_3) \\
&\quad + 4\cos^2\theta\sin\phi\cos\phi(\mathbf{e}_2 \times \mathbf{e}_1) - 4\cos\theta\sin^2\phi(\mathbf{e}_2 \times \mathbf{e}_3) \\
&= -4\cos\theta\sin^2\phi\,\mathbf{e}_1 - 4\sin\theta\sin^2\phi\,\mathbf{e}_2 - 4\big(\sin^2\theta + \cos^2\theta\big)\sin\phi\cos\phi\,\mathbf{e}_3 \\
&= -4\sin\phi\big(\cos\theta\sin\phi\,\mathbf{e}_1 + \sin\theta\sin\phi\,\mathbf{e}_2 + \cos\phi\,\mathbf{e}_3\big) = -2\sin\phi\,\mathbf{r}.
\end{aligned}$$

Hence, the normal vectors at a point of the sphere are parallel to the position vector **r** of the point. Since the length of **r** is the distance from the origin to a point on $\Sigma$, we have $\|\mathbf{r}\| = 2$. Thus, $\mathbf{n} = \frac{1}{2}\mathbf{r}$ is a unit normal vector at the point $P = \mathbf{r}(u, v)$ of $\Sigma$. Note that since **n** has the same direction as the $\overrightarrow{OP}$, it will consistently be the unit normal vector to the sphere that is pointed away from the origin. Therefore, **n** yields an orientation of $\Sigma$ that is often called the *outward orientation* of the sphere. □

## Exercises

Find the equations of the tangent plane and of the normal line to the given surface at the given point.

10.1. $z = xye^{x+2y}$, $P(2, -1, -2)$

10.2. $z = \log_2\left(xy^2 + x^2y\right)$, $P(1, -2, 1)$

10.3. $x^2 - 2y^2 + 2z^2 = 10$, $P(2, -1, -2)$

10.4. $\ln(xyz) = x + yz^{-1}$, $P(1, e, e)$

10.5. (a) Find the equation of the normal line to the sphere $x^2 + y^2 + z^2 = r^2$ at a point $(x_0, y_0, z_0)$ of the sphere.

(b) Show that the normal line from part (a) passes through the origin.

10.6. (a) Find the equation of the tangent plane to the surface $\sqrt{x} + \sqrt{y} + \sqrt{z} = 4$ at a point $P(x_0, y_0, z_0)$ on this surface.

(b) Show that the sum of the lengths of the line segments that the tangent plane from part (a) cuts from the coordinate axes is independent of the choice of $P$. (The line segment that a plane cuts from a coordinate axis is the segment whose endpoints are the origin and the intersection point of the plane and the axis.)

10.7. Let $P(x_0, y_0, z_0)$ be a point on the surface $xyz = 6$. Show that the volume of the tetrahedron that the tangent plane at $P$ cuts from the first octant is independent of the choice of $P$.

10.8. Find the points on the hyperboloid $x^2 - 2y^2 - z^2 = -2$ at which the tangent plane is parallel to the plane $2x - 3y + 2z + 7 = 0$.

10.9. Find the points on the hyperbolic paraboloid $x = y^2 - 4z^2$ at which the normal line is perpendicular to the plane $x + y + z = 1$.

10.10. Show that the hyperboloid $x^2 - 2y^2 - z^2 = -2$ and the ellipsoid $\left(x + \frac{4}{3}\right)^2 + \left(y + \frac{1}{3}\right)^2 + \frac{1}{4}\left(z + \frac{1}{3}\right)^2 = 1$ are tangent to each other at the point $(-1, -1, 1)$, that is, show that they have the same tangent plane at that point.

Find an equation of the tangent plane to the given parametric surface at the given point.

10.11. $\mathbf{r}(u, v) = \langle uv, u\sin v, u\cos v\rangle$, $(u, v) = (1, \pi/2)$

10.12. $\mathbf{r}(y, z) = \left\langle \sqrt{4 + y^2 + z^2}, y, z\right\rangle$, $(y, z) = (1, 1)$

10.13. $\mathbf{r}(u, v) = \left\langle uv, u^2 + v^2, u + v + 2\right\rangle$, $(u, v) = (-1, 0)$

10.14. $\mathbf{r}(u, v) = \left\langle u^2 - v^2, u + v, u^2 + 2v\right\rangle$, $(0, 2, 3)$

10.15. $\mathbf{r}(u, v) = \langle u\cos 2v, u\sin 2v, v\rangle$, $(1, \sqrt{3}, \pi/6)$

10.16. Let $\Sigma$ be the parametric surface given by

$$\mathbf{r}(u, v) = (u + v)\mathbf{e}_1 + (uv)\mathbf{e}_2 + (u^3 - v^3)\mathbf{e}_3 \qquad (-\infty < u, v < \infty).$$

(a) Find the point $P$ on $\Sigma$ that corresponds to the values $u = 1$ and $v = 1$ of the parameters.

(b) Find the equation of the tangent plane to $\Sigma$ at the point $P$ found in (a).

10.17. Let $\Sigma$ be the parametric surface given by

$$\mathbf{r}(u, v) = (u + v)\mathbf{e}_1 + (u^2 + v^2)\mathbf{e}_2 + (u^3 + v^3)\mathbf{e}_3 \qquad (-\infty < u, v < \infty).$$

Find the point(s) on $\Sigma$ where the tangent plane is perpendicular to the line $x = y = -z$.

Determine whether the given surface is smooth, piecewise smooth, or neither.

10.18. $x^2 + 3y^2 - 2xyz = 6$

10.19. $\mathbf{r}(u, v) = \left\langle 3u, u^2 - 2v, u^3 + v^2 \right\rangle, (u, v) \in \mathbb{R}^2$

10.20. $\mathbf{r}(u, v) = \langle u, uv, \sin uv \rangle, (u, v) \in \mathbb{R}^2$

10.21. $z^2 + \sin(xy) = 1$

10.22. Find the point(s) where the parametric surface given by

$$\mathbf{r}(u, v) = (u + v)\mathbf{e}_1 + (uv)\mathbf{e}_2 + (u^2 - v^2)\mathbf{e}_3 \qquad (-\infty < u, v < \infty)$$

is not smooth.

## LECTURE 11

# Extremal Values of Multivariable Functions

In this lecture, we shall discuss the local and absolute (or global) maximum and minimum values of a function of two variables $f(x, y)$.

### 11.1. Review of extremal values of single-variable functions

We start with a review of the extrema of single-variable functions. For example, let us classify the local and global extrema of the function

$$f(x) = x^4 - 8x^3 + 18x^2 - 10$$

in the interval $[-1, 2]$. First, we find the critical points of $f$: these are the zeros of

$$f'(x) = 4x^3 - 24x^2 + 36x = 4x(x^2 - 6x + 9) = 4x(x-3)^2.$$

We have

$$4x(x-3)^2 = 0 \quad \iff \quad x = 0 \text{ or } x = 3.$$

We determine whether the critical points yield local extrema by the second-derivative test:

$$f''(x) = 12x^2 - 48x + 36 \qquad \Longrightarrow \qquad f''(0) = 36 > 0, \quad f''(3) = 0,$$

that is, $f$ has a local minimum at $x = 0$, $(0, -10)$. (We can also show that $f$ has no extremum at $x = 3$, but that does not follow from the second derivative test.) Finally, to find the absolute extrema of $f$ in $[-1, 2]$, we compare the local extrema and the boundary values: $f(-1) = 17$, $f(0) = -10$, and $f(2) = 14$. We find that the absolute maximum of $f$ is 17 and its absolute minimum is $-10$.

### 11.2. Local extrema of a function of two variables

DEFINITION. Let $f$ be a function of two variables. If $(a, b)$ is a point in the domain of $f$, then $f$ has a *local maximum at* $(a, b)$ if $f(x, y) \leq f(a, b)$ for all points $(x, y)$ near $(a, b)$; $f$ has a *local minimum at* $(a, b)$ if $f(x, y) \geq f(a, b)$ for all points $(x, y)$ near $(a, b)$. A local maximum or minimum is called also a *local extremum*.

THEOREM 11.1. *If $f$ has a local extremum at $(a, b)$ and its first-order partials $f_x(a, b)$ and $f_y(a, b)$ exist, then*

$$f_x(a, b) = f_y(a, b) = 0.$$

PROOF. If $f(x, y)$ has a local maximum/minimum at $(a, b)$, then the single-variable function $g(x) = f(x, b)$ also has a local maximum/minimum at $x = a$. Hence,

$$g'(a) = 0 \qquad \Longrightarrow \qquad f_x(a, b) = 0.$$

Similarly, $f_y(a, b) = 0$ by considering $h(y) = f(a, y)$. □

Definition. Let $f$ be a function of two variables. A *critical point of* $f$ is a point $(a, b)$ such that either $f_x(a, b) = f_y(a, b) = 0$, or one of these partial derivatives does not exist. Those critical points $(a, b)$ of $f$ for which $f_x(a, b) = f_y(a, b) = 0$ are sometimes called *stationary points*.

Example 11.1. Find all critical points of the function

$$f(x, y) = x^3 + y^2 - 6xy + 6x + 3y - 2.$$

Solution. This function has first-order (and also second-, third-, and higher-order) partial derivatives at any point $(x, y)$ in the $\mathbb{R}^2$, so the only critical points will be the stationary points of $f$. Since

$$f_x(x, y) = 3x^2 - 6y + 6 \qquad \text{and} \qquad f_y(x, y) = 2y - 6x + 3,$$

the stationary points of $f$ are the solutions of the system

$$\begin{cases} 3x^2 - 6y + 6 = 0 \\ 2y - 6x + 3 = 0 \end{cases} \iff \begin{cases} x^2 - 2y + 2 = 0 \\ 2y = 6x - 3 \end{cases}.$$

Substituting $2y$ from the second equation into the first, we find that $x$ must satisfy the equation

$$x^2 - (6x - 3) + 2 = 0 \iff x^2 - 6x + 5 = 0 \iff x = 1 \text{ or } x = 5.$$

When we substitute $x = 1$ in the second equation of the system, we obtain $2y = 3$ so $y = 1.5$; when we substitute $x = 5$, we obtain $2y = 27$ so $y = 13.5$. Thus, the critical points of $f$ are $(1, 1.5)$ and $(5, 13.5)$. □

## 11.3. The second-derivative test

Note that the theorem in the previous section says that (just as in the case of functions of a single variable) a function of two variables can have a local extremum only at a critical point. However, we know that for a function of one variable a critical point is not necessarily an extremal point. Is this also the case for functions of two variables? And if so, then how can we determine whether a critical point is an extremal point or not? The following theorem is often sufficient to answer these questions. It is the two-variable version of the second-derivative test.

Theorem 11.2 (Second-derivative test). *Suppose that $f(x, y)$ has continuous second-order partial derivatives near the stationary point $(a, b)$ and define*

$$D = D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - f_{xy}(a, b)^2 = \begin{vmatrix} f_{xx}(a, b) & f_{xy}(a, b) \\ f_{yx}(a, b) & f_{yy}(a, b) \end{vmatrix}.$$

*Then:*

i) *if $D > 0$ and $f_{xx}(a, b) > 0$, $f$ has a local minimum at $(a, b)$;*
ii) *if $D > 0$ and $f_{xx}(a, b) < 0$, $f$ has a local maximum at $(a, b)$;*
iii) *if $D < 0$, $f$ does not have a local extremum at $(a, b)$.*

In case iii), the point $(a, b)$ is called a *saddle point*, since under those conditions the graph of $f$ near $(a, b)$ resembles a saddle.

Example 11.2. Find the local extrema of the function

$$f(x, y) = x^3 + y^2 - 6xy + 6x + 3y - 2.$$

Solution. We know from Example 9.1 that the critical points of $f$ are $(1, 1.5)$ and $(5, 13.5)$. The second-order partials of $f$ are:

$$f_{xx}(x, y) = 6x, \qquad f_{yy}(x, y) = 2, \qquad f_{xy}(x, y) = -6,$$

so $f_{xx}(5, 13.5) = 30$ and

$$D(1, 1.5) = \begin{vmatrix} 6 & -6 \\ -6 & 2 \end{vmatrix} = -24, \qquad D(5, 13.5) = \begin{vmatrix} 30 & -6 \\ -6 & 2 \end{vmatrix} = 24.$$

Hence, by the second-derivative test, $(1, 1.5)$ is a saddle point of $f$ and $f(5, 13.5) = -29$ is a local minimum. □

Example 11.3. Find the point on the plane $2x - y + 2z - 3 = 0$ that is closest to the origin. What is the distance between the plane and the origin?

Solution. We want to minimize the function

$$d(x, y, z) = x^2 + y^2 + z^2$$

for $(x, y, z)$ subject to the constraint

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

This is the same as minimizing the function

$$f(x, z) = x^2 + (2x + 2z - 3)^2 + z^2 = 5x^2 + 8xz + 5z^2 - 12x - 12z + 9$$

in the $xz$-plane. The first-order partials of $f$ are

$$f_x(x, z) = 10x + 8z - 12, \qquad f_z(x, z) = 8x + 10z - 12,$$

so the critical points are the solutions of the system

$$\begin{cases} 10x + \phantom{1}8z - 12 = 0 \\ \phantom{1}8x + 10z - 12 = 0 \end{cases} \iff \begin{cases} 5x + 4z = 6 \\ 4x + 5z = 6 \end{cases}.$$

Solving the last system, we find that the only critical point of $f$ is $(\frac{2}{3}, \frac{2}{3})$. The second-order partials of $f$ are

$$f_{xx}(x, z) = f_{zz}(x, z) = 10, \qquad f_{xz}(x, z) = f_{zx}(x, z) = 8,$$

so

$$D = D(\tfrac{2}{3}, \tfrac{2}{3}) = \begin{vmatrix} 10 & 8 \\ 8 & 10 \end{vmatrix} = 36.$$

Since $D > 0$ and $f_{xx} > 0$, we conclude that $f(\frac{2}{3}, \frac{2}{3}) = \frac{4}{3}$ is a local minimum of $f$. It is intuitively clear that this local minimum must actually be an absolute minimum of $f$, so the minimum distance occurs when the point $(x, y, z)$ on the plane is $(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})$. The distance from the origin to that point (which is the distance from the origin to the plane) is $\sqrt{(2/3)^2 + (-1/3)^2 + (2/3)^2} = 1$. □

### 11.4. Absolute extrema of a function of two variables

In real life, we are more likely to be interested in absolute (or global) extrema of functions than in local extrema. In the last example, we used practical considerations (i.e., intuition) to make the jump from the local to the global minimum. Is there another (hopefully, more rigorous) way to approach global extrema? The answer is "yes," if the function is defined on a "closed, bounded set."

First, we must say what does it mean for a set in the plane to be "closed and bounded." Instead of formal definitions we will be content with explanations by example. A set $R$ is *closed* if it contains all its "boundary points." For example, among the sets on Figure 11.1 (where solid boundary curve belongs to the set and a dotted curve does not), the sets in (a) and (d) are closed and the sets in (b), (c), and (e) are not. A set $R$ is *bounded* if it is contained inside some disk centered at the origin. For example, of the sets on Figure 11.1, only (d) is not bounded.

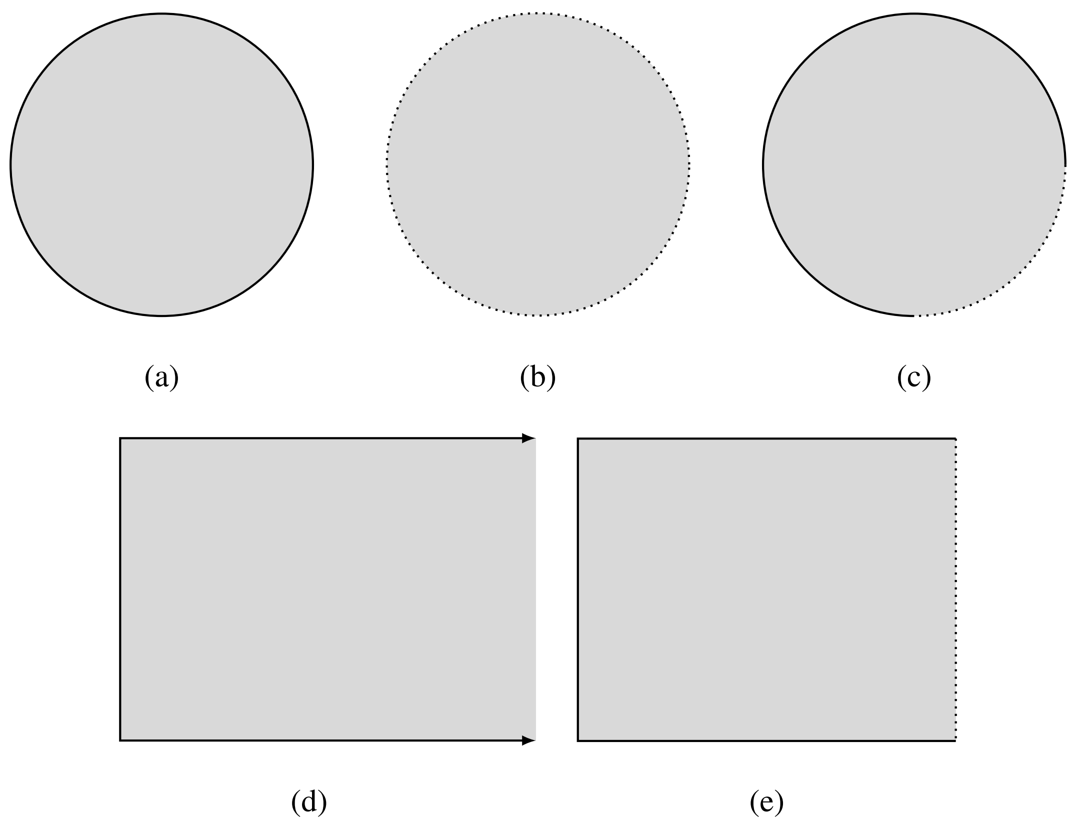



FIGURE 11.1. Examples of closed and bounded sets

We can now state the algorithm for finding the absolute extrema of a function of two variables on a closed, bounded set. It is a generalization of the closed interval method for finding absolute extrema of functions of a single variable on a closed interval. To find the absolute extrema of $f(x,y)$ on a closed, bounded set $R$:

1. Find the critical points of $f$ inside $R$ and the respective values of $f$.
2. Find the absolute extrema of $f$ on the boundary of $R$.

3. The absolute maximum of $f$ in $R$ is the largest among the numbers found in Steps 1 and 2; the absolute minimum of $f$ in $R$ is the smallest among the numbers found in Steps 1 and 2.

EXAMPLE 11.4. Find the absolute maximum and minimum values of the function

$$f(x,y) = x^3 - y^3 - 3x + 12y$$

in the closed and bounded region

$$R = \{(x,y) : 0 \le x \le 2,\ -3x \le y \le 0\}.$$

SOLUTION. 1. The first-order partials of $f$ are

$$f_x(x,y) = 3x^2 - 3, \qquad f_y(x,y) = -3y^2 + 12,$$

so the critical points of $f$ are the solutions of the system

$$\begin{cases} 3x^2 - \ \ 3 = 0 \\ -3y^2 + 12 = 0 \end{cases} \iff \begin{cases} x^2 = 1 \\ y^2 = 4 \end{cases} \iff \begin{cases} x = \pm 1 \\ y = \pm 2 \end{cases}.$$

That is, $f$ has four critical points: $(1,2)$, $(1,-2)$, $(-1,2)$, and $(-1,-2)$. However, of these only $P(1,-2)$ lies inside $R$. The respective value of $f$ is $f(1,-2) = -18$.

2. The boundary of $R$ consists of three line segments (see Figure 11.2):

$$\begin{aligned} \ell_1 : &\quad \{(x,y) : 0 \le x \le 2, y = 0\}, \\ \ell_2 : &\quad \{(x,y) : 0 \le x \le 2, y = -3x\}, \\ \ell_3 : &\quad \{(x,y) : x = 2, -6 \le y \le 0\}. \end{aligned}$$

2.1. When $(x,y)$ lies on $\ell_1$, $f(x,y)$ is really only a function of $x$:

$$g_1(x) = f(x,0) = x^3 - 3x \qquad (0 \le x \le 2).$$

The critical values of $g_1$ are the solutions of $3x^2 - 3 = 0$. These are $x = \pm 1$, but we are interested only in $x = 1$ (since $-1$ lies outside $[0,2]$). Thus, we need compare

$$g_1(0) = 0, \qquad g_1(1) = -2, \qquad \text{and} \qquad g_1(2) = 2.$$

We find that the absolute minimum and maximum of $f$ on $\ell_1$ are $f(1,0) = -2$ and $f(2,0) = 2$.

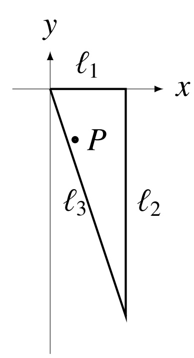



FIGURE 11.2. The boundary of $R$

2.2. When $(x, y)$ lies on $\ell_2$, $f(x, y)$ is again a function of $x$ only:

$$g_2(x) = f(x, -3x) = 28x^3 - 39x \qquad (0 \le x \le 2).$$

The critical values of $g_2$ are the solutions of $84x^2 - 39 = 0$: $x = \pm\sqrt{13/28}$. Again, we are interested only in one of these critical points: $x = \sqrt{13/28}$. We compare

$$g_2(0) = 0, \qquad g_2\big(\sqrt{13/28}\big) = -26\sqrt{13/28} = -17.71\ldots, \qquad \text{and} \qquad g_2(2) = 146.$$

We find that the absolute minimum and maximum of $f$ on $\ell_2$ are

$$f(\sqrt{13/28}, -3\sqrt{13/28}) = -17.71\ldots \qquad \text{and} \qquad f(2, -6) = 146.$$

2.3. When $(x, y)$ lies on $\ell_3$, $f(x, y)$ is only a function of $y$:

$$g_3(y) = f(2, y) = -y^3 + 12y + 2 \qquad (-6 \le y \le 0).$$

The critical values of $g_3$ are the solutions of $-3y^2 + 12 = 0$, $y = \pm 2$. After dropping $y = 2$, we compare

$$g_3(0) = 2, \qquad g_3(-2) = -14, \qquad \text{and} \qquad g_3(-6) = 146.$$

We find that the absolute minimum and maximum of $f$ on $\ell_3$ are, respectively, $f(2, -2) = -14$ and $f(2, -6) = 146$.

3. Comparing all the values found in Steps 1 and 2, we conclude that

$$\min_{(x,y)\in R} f(x, y) = \min\{-18, -2, 2, -17.71, 146, -14\} = -18 = f(1, -2),$$
$$\max_{(x,y)\in R} f(x, y) = \max\{-18, -2, 2, -17.71, 146, -14\} = 146 = f(2, -6). \qquad \square$$

Example 11.5. Find the absolute maximum and minimum values of the function

$$f(x, y) = x^2 - 2xy - 2x + 4y$$

in the closed region bounded by the ellipse $\gamma$ with parametric equations

$$x = 3\cos t, \quad y = 2\sin t \qquad (0 \le t \le 2\pi).$$

Solution. 1. The first-order partials of $f$ are

$$f_x(x, y) = 2x - 2y - 2, \qquad f_y(x, y) = -2x + 4,$$

so the critical points of $f$ are the solutions of the system

$$\begin{cases} 2x - 2y - 2 = 0 \\ -2x + 4 = 0 \end{cases} \iff \begin{cases} x = y + 1 \\ x - 2 = 0 \end{cases} \iff \begin{cases} y = x - 1 = 1 \\ x = 2 \end{cases}$$

Hence, $P(2, 1)$ is the only critical point of $f(x, y)$. It lies inside $\gamma$ (see Figure 11.3) and $f(2, 1) = 0$.

2. On $\gamma$, we have

$$f(x, y) = f(3\cos t, 2\sin t) = 9\cos^2 t - 12\cos t\sin t - 6\cos t + 8\sin t,$$

so we have to find the maximum and the minimum of

$$g(t) = 9\cos^2 t - 12\cos t\sin t - 6\cos t + 8\sin t \qquad (0 \le t \le 2\pi).$$

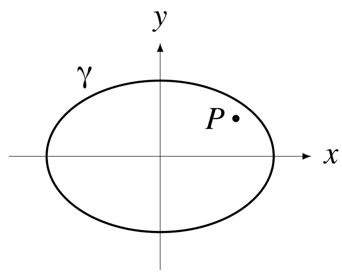



FIGURE 11.3. The ellipse γ and the critical point $P$

We have

$$\begin{aligned} g'(t) &= -18\cos t \sin t + 12\sin^2 t - 12\cos^2 t + 6\sin t + 8\cos t \\ &= 24\sin^2 t + 6\sin t - 12 + 2\cos t(4 - 9\sin t). \end{aligned}$$

Substituting $u = \sin t$ and $\cos t = \pm\sqrt{1-u^2}$, we can rewrite the equation $g'(t) = 0$ as

$$\begin{gathered} 24u^2 + 6u - 12 \pm 2\sqrt{1-u^2}(4 - 9u) = 0 \\ 12u^2 + 3u - 6 = \pm\sqrt{1-u^2}(9u - 4) \\ \left(12u^2 + 3u - 6\right)^2 = (1 - u^2)(9u - 4)^2 \\ 225u^4 - 200u^2 + 36u + 20 = 0. \end{gathered}$$

The last equation has four real roots, whose approximate values are:

$$\alpha_1 \approx -0.980, \quad \alpha_2 \approx -0.245, \quad \alpha_3 \approx 0.547, \quad \alpha_4 \approx 0.678.$$

These are the values of $\sin t$ at the critical numbers of $g(t)$. They yield a total of eight possible critical numbers: four pairs that have the same sines and opposite cosines. In fact, only one of the numbers in each pair is a critical number of $g$, but we will ignore this and will treat all eight numbers as critical. The values of all these critical (and pseudo-critical) numbers, their sines and cosines, and the respective values of $g(t)$ are listed in Table 11.1.

| $\alpha_k$ | $t$ | $\cos t$ | $\sin t$ | $g(t)$ |
|---|---|---|---|---|
| −0.980 | 4.510 | −0.201 | −0.980 | −8.630 |
| −0.980 | 4.915 | 0.201 | −0.980 | −6.316 |
| −0.245 | 3.389 | −0.970 | −0.245 | 9.470 |
| −0.245 | 6.036 | 0.970 | −0.245 | 3.533 |
| 0.547 | 0.579 | 0.837 | 0.547 | 0.166 |
| 0.547 | 2.563 | −0.837 | 0.547 | 21.201 |
| 0.678 | 0.745 | 0.735 | 0.678 | −0.104 |
| 0.678 | 2.397 | −0.735 | 0.678 | 20.681 |

TABLE 11.1. Values of $g(t)$ at its critical numbers

We also have $g(0) = g(2\pi) = 3$, so

$$\max_{0 \le t \le 2\pi} g(t) = g(\pi - \arcsin \alpha_3) \approx 21.201, \qquad \min_{0 \le t \le 2\pi} g(t) = g(\pi - \arcsin \alpha_1) \approx -8.630.$$

3. Comparing all the values found in Steps 1 and 2, we conclude that the absolute minimum of $f$ inside $\gamma$ is

$$\min\{0, -8.630, 21.201\} = -8.630,$$

and that the absolute maximum of $f$ is

$$\max\{0, -8.630, 21.201\} = 21.201. \qquad \square$$

## Exercises

Find and classify the critical points (as local maxima, local minima, or saddle points) of the given function.

11.1. $f(x, y) = 9 + x + 2y - x^2 + \frac{1}{2}y^2$
11.2. $f(x, y) = 2x - 4y - x^2 - 2y^2$
11.3. $f(x, y) = x^4 + y^4 - 4xy + 2$
11.4. $f(x, y) = xy(x + y + 1)$
11.5. $f(x, y) = e^{2y} \sin(xy)$
11.6. $f(x, y) = xye^{-x^2-y^2}$
11.7. $f(x, y) = x^2 - 6xy + 4y^3$
11.8. $f(x, y) = x^3 - 6xy + 8y^3$
11.9. $f(x, y) = x^3 + y^4 - 3xy + 4$

Find the absolute maxima and minima of the given function on the given set.
11.10. $f(x, y) = 1 + 2x + 3y$ on the triangle with vertices $(0, 3)$, $(2, 1)$, and $(5, 3)$
11.11. $f(x, y) = x^2 + 4y^2 - xy$ on the square $-1 \le x \le 1$, $-1 \le y \le 1$
11.12. $f(x, y) = x^3 + 2xy^2 - 2x$ on the disk $x^2 + y^2 \le 4$
11.13. $f(x, y) = x^3 + y^2 - 2xy + 3$ on the square $0 \le x \le 1$, $0 \le y \le 1$

11.14. Find the shortest distance from the point $(3, 2, -1)$ to the plane $x - y + 2z = 1$.

11.15. Find the point(s) on the cone $(z - 2)^2 = x^2 + y^2$ that is closest to the point $(1, 1, 2)$.

11.16. Among all the rectangular boxes of volume 125 ft$^3$, find the dimensions of that which has the least surface area.

11.17. Find the volume of the largest rectangular box with faces parallel to the coordinate planes that can be inscribed in the ellipsoid $x^2 + 4y^2 + 4z^2 = 16$.

LECTURE 12

# Lagrange Multipliers*

## 12.1. Introduction

Recall Example 11.3. In that example, we had to minimize the function

$$f(x, y, z) = x^2 + y^2 + z^2$$

subject to the constraint

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

We used the constraint to eliminate one of the variables and reduced the problem to that of minimizing a related function of two variables (without any constraints). On an abstract level, one expects that the same idea should work if one wants to minimize or maximize any "nice" function $f(x, y, z)$ subject to the constraint that $g(x, y, z) = k$, where $g$ is another "nice" function and $k$ is a constant. However, while such an approach is theoretically possible, for any practical purpose it is pretty much doomed unless $g$ is a very simple function (such as the linear function that appears in the above example). In this lecture, we describe the method of Lagrange multipliers, which allows us to solve the general problem.

## 12.2. Basic form of the method of Lagrange multipliers

Suppose that $P(x_0, y_0, z_0)$ is a point at which the function $f(x, y, z)$ is maximal subject to the constraint $g(x, y, z) = k$. This constraint describes a level surface, which we will denote by $\Sigma$. Thus, $P$ is a point on the surface $\Sigma$ such that $f(x_0, y_0, z_0)$ is maximal. Let $\gamma$ be a smooth parametric curve on $\Sigma$ parametrized by the vector function $\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$, $t \in I$, and such that $\mathbf{r}(t_0) = P$ (so that, $\gamma$ passes through $P$). Since $f(x_0, y_0, z_0)$ is maximal on $\Sigma$ and $\gamma$ lies on $\Sigma$, we have

$$f(x_0, y_0, z_0) \geq f(x, y, z) \qquad \text{for all } (x, y, z) \text{ on } \gamma.$$

Using the parametrization of $\gamma$, we get

$$f(x(t_0), y(t_0), z(t_0)) \geq f(x(t), y(t), z(t)) \qquad \text{for all } t \in I,$$

so $t = t_0$ is a local maximum of the single-variable function $h(t) = f(x(t), y(t), z(t))$, $t \in I$. Therefore, $h'(t_0) = 0$. On the other hand, by the chain rule,

$$\begin{aligned} h'(t) &= f_x(x(t), y(t), z(t))x'(t) + f_y(x(t), y(t), z(t))y'(t) + f_z(x(t), y(t), z(t))z'(t) \\ &= \nabla f(x(t), y(t), z(t)) \cdot \mathbf{r}'(t), \end{aligned}$$

so the equation $h'(t_0) = 0$ can be rewritten as

$$\nabla f(x_0, y_0, z_0) \cdot \mathbf{r}'(t_0) = 0.$$

That is, the gradient of $f$ at the point $P$ is perpendicular to the tangent vector $\mathbf{r}'(t_0)$ to $\gamma$ at $P$. Since $\gamma$ was an arbitrary curve on $\Sigma$, this shows that the gradient $\nabla f(P)$ is perpendicular to all tangent vectors to $\Sigma$. Hence, $\nabla f(P)$ is parallel to the normal vector to $\Sigma$ at $P$—which is $\nabla g(P)$. We conclude that if $P(x_0, y_0, z_0)$ is on $\Sigma$ and $f(x_0, y_0, z_0)$ is maximal, then

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0) \qquad \text{for some number } \lambda.$$

We note that if we had started with a point $P(x_0, y_0, z_0)$ on $\Sigma$ such that $f(x_0, y_0, z_0)$ is minimal, we could have used a similar argument to reach the same conclusion. These observation motivate the following approach towards constrained optimization.

To find the extrema of a function $f(x, y, z)$ subject to the constraint $g(x, y, z) = k$ (when $\nabla g \neq \mathbf{0}$ on the surface $g(x, y, z) = k$):

1. Find all the values $x, y, z$, and $\lambda$ such that

$$\begin{cases} \nabla f(x, y, z) = \lambda \nabla g(x, y, z), \\ g(x, y, z) = k. \end{cases}$$

   That is, solve this system of four equations in the four unknowns $x, y, z, \lambda$.
2. Evaluate $f$ at all the points $(x, y, z)$ found in Step 1. The maximum of $f$ is the largest among these numbers; the minimum of $f$ is the smallest among them.

Example 12.1. Find the maximum and minimum values of the function $f(x, y, z) = 8x - 4z$ on the ellipsoid $x^2 + 10y^2 + z^2 = 5$.

Solution. Let $g(x, y, z) = x^2 + 10y^2 + z^2$. We have

$$f_x(x, y, z) = 8, \quad f_y(x, y, z) = 0, \quad f_z(x, y, z) = -4,$$

and

$$g_x(x, y, z) = 2x, \quad g_y(x, y, z) = 20y, \quad g_z(x, y, z) = 2z,$$

so the system in Step 1 of the method of Lagrange multipliers is

$$\begin{cases} 8 = 2\lambda x \\ 0 = 20\lambda y \\ -4 = 2\lambda z \\ x^2 + 10y^2 + z^2 = 5 \end{cases} \iff \begin{cases} x = 4/\lambda \\ y = 0 \\ z = -2/\lambda \\ x^2 + 10y^2 + z^2 = 5 \end{cases}$$

Here we have used that $\lambda \neq 0$, because of the first or third equations. Substituting the expressions for $x$, $y$, and $z$ in the last equation of the system, we find that $\lambda$ must be a solution of

$$16/\lambda^2 + 0 + 4/\lambda^2 = 5 \iff \lambda^2 = 4 \iff \lambda = \pm 2.$$

The respective points $(x, y, z)$ are $(2, 0, -1)$ and $(-2, 0, 1)$, so the conditional maximum of $f$ is

$$\max\{f(2, 0, -1), f(-2, 0, 1)\} = \max\{20, -20\} = 20 = f(2, 0, -1),$$

and the conditional minimum of $f$ is

$$\min\{f(2, 0, -1), f(-2, 0, 1)\} = \min\{20, -20\} = -20 = f(-2, 0, 1). \qquad \square$$

Example 12.2. Find the maximum and minimum values of the function $f(x, y, z) = x^3 + y^3 + z^3$ on the sphere $x^2 + y^2 + z^2 = 1$.

SOLUTION. Let $g(x, y, z) = x^2 + y^2 + z^2$. We have

$$f_x(x, y, z) = 3x^2, \quad f_y(x, y, z) = 3y^2, \quad f_z(x, y, z) = 3z^2,$$

and

$$g_x(x, y, z) = 2x, \quad g_y(x, y, z) = 2y, \quad g_z(x, y, z) = 2z,$$

so the system in Step 1 of the method of Lagrange multipliers is

$$\begin{cases} 3x^2 = 2\lambda x \\ 3y^2 = 2\lambda y \\ 3z^2 = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases} \iff \begin{cases} x(x - 2\lambda/3) = 0 \\ y(y - 2\lambda/3) = 0 \\ z(z - 2\lambda/3) = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases} \iff \begin{cases} x = 0, 2\lambda/3 \\ y = 0, 2\lambda/3 \\ z = 0, 2\lambda/3 \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

For each of the eight choices for $x, y, z$ from the first three equations, we now solve the fourth equation for $\lambda$:

1. $x = y = z = 0$. Then the fourth equation fails and we get no solution of the system.
2. $x = y = 0$, $z = 2\lambda/3$. Then the fourth equation becomes
$$4\lambda^2/9 = 1 \iff \lambda = \pm 3/2,$$
and we obtain the points $(0, 0, 1)$ and $(0, 0, -1)$.
3. $x = z = 0$, $y = 2\lambda/3$. Then the fourth equation becomes
$$4\lambda^2/9 = 1 \iff \lambda = \pm 3/2,$$
and we obtain the points $(0, 1, 0)$ and $(0, -1, 0)$.
4. $y = z = 0$, $x = 2\lambda/3$. Then the fourth equation becomes
$$4\lambda^2/9 = 1 \iff \lambda = \pm 3/2,$$
and we obtain the points $(1, 0, 0)$ and $(-1, 0, 0)$.
5. $x = 0$, $y = z = 2\lambda/3$. Then the fourth equation becomes
$$8\lambda^2/9 = 1 \iff \lambda = \pm 3\sqrt{2}/4,$$
and we obtain the points $(0, \sqrt{2}/2, \sqrt{2}/2)$ and $(0, -\sqrt{2}/2, -\sqrt{2}/2)$.
6. $y = 0$, $x = z = 2\lambda/3$. Then the fourth equation becomes
$$8\lambda^2/9 = 1 \iff \lambda = \pm 3\sqrt{2}/4,$$
and we obtain the points $(\sqrt{2}/2, 0, \sqrt{2}/2)$ and $(-\sqrt{2}/2, 0, -\sqrt{2}/2)$.
7. $z = 0$, $x = y = 2\lambda/3$. Then the fourth equation becomes
$$8\lambda^2/9 = 1 \iff \lambda = \pm 3\sqrt{2}/4,$$
and we obtain the points $(\sqrt{2}/2, \sqrt{2}/2, 0)$ and $(-\sqrt{2}/2, -\sqrt{2}/2, 0)$.
8. $x = y = z = 2\lambda/3$. Then the fourth equation becomes
$$12\lambda^2/9 = 1 \iff \lambda = \pm\sqrt{3}/2,$$
and we obtain the points $(\sqrt{3}/3, \sqrt{3}/3, \sqrt{3}/3)$ and $(-\sqrt{3}/3, -\sqrt{3}/3, -\sqrt{3}/3)$.

A list of the values of $f$ at the above points is

$$1,\ -1,\ 1/\sqrt{2},\ -1/\sqrt{2},\ 1/\sqrt{3},\ -1/\sqrt{3},$$

so the minimum and the maximum of $f$ on the sphere are $-1$ and 1, respectively. The minimum is attained at the points of intersection of the sphere with the negative directions of the coordinate axes and the maximum is attained at the points of intersection of the sphere with the positive directions of the coordinate axes. □

### 12.3. Lagrange multipliers in the case of two constraints**

Sometimes, one needs to find the extrema of a function $f(x, y, z)$ subject to multiple constraints: say, $g(x, y, z) = k$, $h(x, y, z) = m$. The variant of the method of Lagrange multipliers that addresses this question is as follows:

1. Find all the values $x, y, z$, $\lambda$, and $\mu$ such that

$$\begin{cases} \nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \\ g(x, y, z) = k \\ h(x, y, z) = m \end{cases}.$$

   That is, solve this system of five equations in the five unknowns $x, y, z, \lambda, \mu$.
2. Evaluate $f$ at all the points $(x, y, z)$ found in Step 1. The maximum of $f$ is the largest among these numbers; the minimum of $f$ is the smallest among them.

Example 12.3. Find the maximum and minimum values of

$$f(x, y, z) = 3x - y - 3z$$

on the intersection curve of the plane $x + y - z = 0$ and the elliptic cylinder $x^2 + 2z^2 = 1$.

Solution. Let

$$g(x, y, z) = x + y - z, \quad h(x, y, z) = x^2 + 2z^2.$$

The partials of $f, g, h$ are

$$\begin{aligned} f_x &= 3, & f_y &= -1, & f_z &= -3, \\ g_x &= 1, & g_y &= 1, & g_z &= -1, \\ h_x &= 2x, & h_y &= 0, & h_z &= 4z, \end{aligned}$$

so we have to solve the system

$$\begin{cases} 3 = \lambda + 2\mu x \\ -1 = \lambda \\ -3 = -\lambda + 4\mu z \\ x + y - z = 0 \\ x^2 + 2z^2 = 1 \end{cases} \iff \begin{cases} \lambda = -1 \\ 2\mu x = 4 \\ 4\mu z = -4 \\ y = z - x \\ x^2 + 2z^2 = 1 \end{cases} \iff \begin{cases} \lambda = -1 \\ x = 2\mu^{-1} \\ z = -\mu^{-1} \\ y = -3\mu^{-1} \\ x^2 + 2z^2 = 1 \end{cases}$$

Substituting the expressions for $x$ and $z$ into the fifth equation, we find that $\mu$ must satisfy the equation

$$4\mu^{-2} + 2\mu^{-2} = 1 \quad \iff \quad \mu^2 = 6 \quad \iff \quad \mu = \pm\sqrt{6}.$$

These give the points $(2/\sqrt{6}, -3/\sqrt{6}, -1/\sqrt{6})$ and $(-2/\sqrt{6}, 3/\sqrt{6}, 1/\sqrt{6})$. Since

$$f\left(\frac{2}{\sqrt{6}}, \frac{-3}{\sqrt{6}}, \frac{-1}{\sqrt{6}}\right) = \frac{12}{\sqrt{6}} = 2\sqrt{6} \qquad \text{and} \qquad f\left(\frac{-2}{\sqrt{6}}, \frac{3}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right) = -2\sqrt{6},$$

we conclude that the maximum of $f$ on the given curve is $2\sqrt{6}$ and that its minimum is $-2\sqrt{6}$. □

## Exercises

Use the method of Lagrange multipliers to find the maximum and minimum values of the function $f$ subject to the given constraint.

12.1. $f(x, y) = x^2 y$; $x^2 + y^4 = 1$

12.2. $f(x, y, z) = 3x - 2y + z$; $3x^2 + 2y^2 + z^2 = 24$

12.3. $f(x, y, z) = xyz$; $x^2 + 2y^2 + z^2 = 3$

12.4. Use the method of Lagrange multipliers to find the minimum of the function $f(x, y, z) = xy + yz + zx$ subject to the constraints $xy = 1$, $z = x^2 + y^2 + 1$, and $x > 0$. What can you say about the maximum of $f$ subject to the same constraints?

12.5. Use the method of Lagrange multipliers to find the point(s) on the cone $(z - 2)^2 = x^2 + y^2$ that is closest to the point $(1, 1, 2)$.

12.6. Use the method of Lagrange multipliers to find three positive numbers whose sum is $S$, where $S > 0$, and whose product is maximal.

12.7. Maximize $\sum_{i=1}^{n} x_i y_i$ subject to the constraints $\sum_{i=1}^{n} x_i^2 = 1$ and $\sum_{i=1}^{n} y_i^2 = 1$.

LECTURE 13

# Double Integrals over Rectangles

In this lecture, we generalize the idea of the definite integral (see Appendix B) to functions of two variables.

## 13.1. Definition

Let $f(x, y)$ be a bounded function on the rectangle $R$,

$$R = [a, b] \times [c, d] = \big\{(x, y) : a \le x \le b,\ c \le y \le d\big\}.$$

For each $n \ge 1$, let

$$x_0 = a, \quad x_1 = a + \Delta x, \quad x_2 = a + 2\Delta x, \quad \dots, \quad x_n = a + n\Delta x = b, \qquad \Delta x = (b - a)/n,$$

be the points that divide the interval $[a, b]$ into $n$ subintervals of equal lengths; and let

$$y_0 = c, \quad y_1 = c + \Delta y, \quad y_2 = c + 2\Delta y, \quad \dots, \quad y_n = c + n\Delta y = d, \qquad \Delta y = (d - c)/n,$$

be the points that divide the interval $[c, d]$ into $n$ subintervals of equal lengths. Also, pick $n^2$ sample points $(x^*_{ij}, y^*_{ij})$ so that

$$x_{i-1} \le x^*_{ij} \le x_i, \quad y_{j-1} \le y^*_{ij} \le y_j \qquad (1 \le i, j \le n).$$

Then the *double integral of $f$ over $R$* is defined by

$$\iint_R f(x, y)\, dA = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x^*_{ij}, y^*_{ij})\Delta A, \qquad \Delta A = \Delta x \Delta y, \tag{13.1}$$

provided that the limit exists. The sums on the right side of (13.1) are called *Riemann sums of $f$ on $R$*.

Remarks. 1. Note that the vertical lines $x = x_i$ and $y = y_j$ split the rectangle $R$ into $n^2$ congruent subrectangles, which we may label according to their upper right corner points: if $1 \le i, j \le n$, then $R_{ij}$ is the subrectangle whose upper right corner is the point $(x_i, y_j)$. With this labelling of the subrectangles, we can report that each sample point $(x^*_{ij}, y^*_{ij})$ lies in the subrectangle $R_{ij}$ with the same indices and there is exactly one sample point per subrectangle (see the left side of Figure 13.1). Furthermore, the quantity $\Delta A$ in (13.1) is equal to the common area of the rectangles $R_{ij}$: one $n^2$th part of the area of $R$. Thus, each term in the Riemann sum of $f$ is the product of a sample value of $f$ in one of the rectangles $R_{ij}$ and the area of that rectangle. As the latter product is the volume of a rectangular box with base $R_{ij}$ and height $f(x^*_{ij}, y^*_{ij})$, the entire Riemann sum equals the volume of a solid built by sticking together $n^2$ rectangular boxes (see the right side of Figure 13.1).

2. The limit in (13.1) is supposed to exist and be independent of the choice of the sample points. This is not necessarily true for any old function, but it does hold for functions that are piecewise continuous in $R$ (see §6.7 for the definition of piecewise continuity).

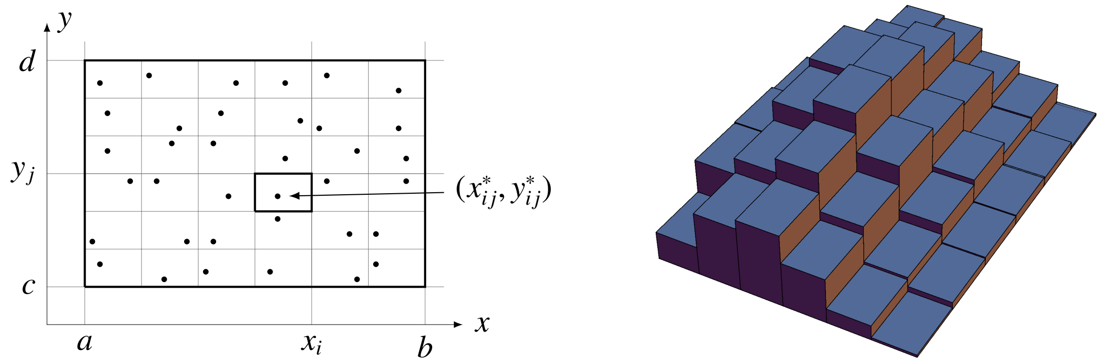



FIGURE 13.1. The sample points and solid volume of a Riemann sum

3. If $f(x, y) \geq 0$ on $R$, then $\iint_R f(x, y)\, dA$ represents the volume of the solid whose upper boundary is the surface $z = f(x, y)$ and whose lower boundary is $R$. That is,

$$\begin{aligned}\iint_R f(x, y)\, dA &= \text{vol.}\big\{(x, y, z) : (x, y) \in R,\ 0 \leq z \leq f(x, y)\big\}\\ &= \text{vol.}\big\{(x, y, z) : a \leq x \leq b,\ c \leq y \leq d,\ 0 \leq z \leq f(x, y)\big\}.\end{aligned}$$

4. From now on, we shall use the notation $|D|$ to denote the area of a plane region $D$. Let us divide both sides of (13.1) by $|R| = (b - a)(c - d)$. Since $\Delta A = |R|/n^2$, we have $\Delta A/|R| = 1/n^2$, so we get

$$\frac{1}{|R|} \iint_R f(x, y)\, dA = \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x^*_{ij}, y^*_{ij}).$$

Since $\frac{1}{n^2} \sum_i \sum_j f(x^*_{ij}, y^*_{ij})$ is the average of the sample values $f(x^*_{ij}, y^*_{ij})$, we interpret the last equation in terms of the average value of $f$:

$$\underset{(x,y)\in R}{\text{average}}\, f(x, y) = \frac{1}{|R|} \iint_R f(x, y)\, dA. \tag{13.2}$$

EXAMPLE 13.1. Evaluate the double integral of the function

$$f(x, y) = \begin{cases} 5 - 2x - 3y & \text{if } 5 - 2x - 3y \geq 0,\\ 0 & \text{if } 5 - 2x - 3y \leq 0,\end{cases}$$

over the rectangle $R = [0, 3] \times [0, 2]$.

SOLUTION. We shall use the interpretation of the double integral as volume. By Remark 3 above, the value of the integral

$$\iint_R f(x, y)\, dA$$

is the volume of the solid $E$ that lies between $R$ and the part of the surface $z = f(x, y)$ with $(x, y) \in R$. The graph of $f(x, y)$ (before the restriction to $R$) consists of the part of the plane $z = 5 - 2x - 3y$ with $5-2x-3y \geq 0$ and the part of the $xy$-plane $z = 0$ with $5-2x-3y \leq 0$. The line $\ell$ with equation $5 - 2x - 3y = 0$ splits the rectangle $R$ in two parts: a triangle with vertices $(0, 0)$, $(\frac{5}{2}, 0)$, and $(0, \frac{5}{3})$,

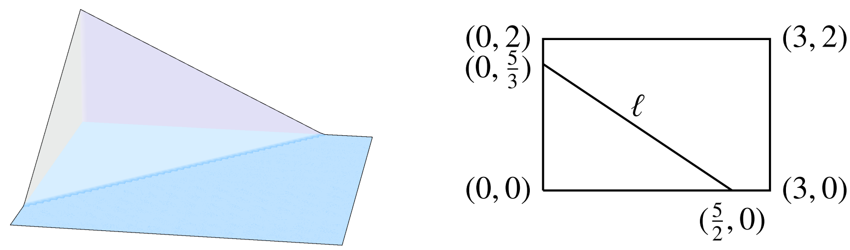



FIGURE 13.2. The pyramid $E$ and its base

where $5 - 2x - 3y \geq 0$; and a pentagon with vertices $(\frac{5}{2}, 0)$, $(3, 0)$, $(3, 2)$, $(0, 2)$, and $(0, \frac{5}{3})$, where $5 - 2x - 3y \leq 0$ (see Figure 13.2). For points $(x, y)$ lying in the pentagon, the graph $z = f(x, y)$ and the $xy$-plane coincide, and for $(x, y)$ in the triangle, the graph $z = f(x, y)$ is a triangle in the plane $z = 5 - 2x - 3y$. Thus, the solid $E$ is the pyramid with vertices $O(0, 0, 0)$, $A(\frac{5}{2}, 0, 0)$, $B(0, \frac{5}{3}, 0)$, and $C(0, 0, 5)$. Its volume is

$$\frac{1}{3}|OC| \cdot \text{area}(\triangle OAB) = \frac{1}{3}|OC| \cdot \frac{1}{2}|OA|\,|OB| = \frac{1}{3}5 \cdot \frac{1}{2}\frac{5}{2}\frac{5}{3} = \frac{125}{36}.$$

Hence,

$$\iint_R f(x, y)\, dA = \frac{125}{36}.$$

□

## 13.2. Iterated integrals

In practice, we want to reverse the order of things in the above example, so that we are able to compute volumes, averages, etc. by interpreting them as double integrals and then evaluating the latter using some "standard" machinery independent of our interpretation. One such method is the method of iterated integrals. An *iterated integral* is an expression of one of the forms

$$\int_c^d \left[ \int_a^b f(x, y)\, dx \right] dy \qquad \text{or} \qquad \int_a^b \left[ \int_c^d f(x, y)\, dy \right] dx.$$

Sometimes, we write these integrals as

$$\int_c^d \int_a^b f(x, y)\, dxdy \qquad \text{and} \qquad \int_a^b \int_c^d f(x, y)\, dydx,$$

respectively. Note that these are definite integrals of functions defined by means of definite integrals and *not* double integrals. Thus, we can evaluate such integrals using the familiar methods from single-variable calculus.

EXAMPLE 13.2. Evaluate

$$\int_0^1 \int_0^3 x^3 y\, dydx \qquad \text{and} \qquad \int_0^3 \int_0^1 x^3 y\, dxdy.$$

SOLUTION. We have

$$\int_0^1 \int_0^3 x^3 y\, dydx = \int_0^1 \left[ \frac{x^3 y^2}{2} \right]_0^3 dx = \frac{9}{2} \int_0^1 x^3 dx = \left[ \frac{9x^4}{8} \right]_0^1 = \frac{9}{8},$$

and

$$\int_0^3 \int_0^1 x^3 y \, dxdy = \int_0^3 \left[\frac{x^4 y}{4}\right]_0^1 dy = \frac{1}{4}\int_0^3 y \, dy = \left[\frac{y^2}{8}\right]_0^3 = \frac{9}{8}.$$

□

## 13.3. Fubini's theorem

Notice that in the last example, we evaluated the two possible iterated integrals of $x^3y$ over $[0,1]\times[0,3]$ and obtained the same value for both despite following different routes. Is this always the case? The answer to this question is in the affirmative. Moreover, the common value of the two iterated integrals is the value of the double integral of $f$ over the rectangle. This fact is known as *Fubini's theorem*.

Theorem 13.1 (Fubini). *Let $f(x,y)$ be piecewise continuous in the rectangle $R = [a,b]\times[c,d]$. Then*

$$\iint_R f(x,y)\, dA = \int_a^b \int_c^d f(x,y)\, dydx = \int_c^d \int_a^b f(x,y)\, dxdy.$$

Example 13.3. Find the volume of the solid bounded above by the graph of $f(x,y) = x^3y$ and below by the rectangle $R = [0,1]\times[0,3]$.

Solution. We can express this volume by the double integral

$$\iint_R x^3 y \, dA.$$

By Fubini's theorem and by Example 13.2,

$$\iint_R x^3 y \, dA = \int_0^1 \int_0^3 x^3 y \, dydx = \int_0^3 \int_0^1 x^3 y \, dxdy = \frac{9}{8},$$

so the volume of the solid is $9/8$. □

Example 13.4. Evaluate

$$\iint_R e^y \sqrt{x + e^y}\, dA, \qquad \text{where } R = [0,4]\times[0,1].$$

Solution. By Fubini's theorem,

$$\iint_R e^y \sqrt{x+e^y}\, dA = \int_0^4 \int_0^1 e^y \sqrt{x+e^y}\, dydx.$$

By the substitution $u = x + e^y$, $du = e^y\, dy$, the inner integral is

$$\int_0^1 e^y \sqrt{x+e^y}\, dy = \int_{1+x}^{e+x} u^{1/2} du = \left[\frac{u^{3/2}}{3/2}\right]_{1+x}^{e+x} = \frac{2}{3}\big((x+e)^{3/2} - (x+1)^{3/2}\big).$$

Hence,

$$\begin{aligned}\iint_R e^y \sqrt{x+e^y}\, dA &= \frac{2}{3}\int_0^4 \big((x+e)^{3/2} - (x+1)^{3/2}\big)\, dx \\ &= \frac{2}{3}\frac{2}{5}\left[(x+e)^{5/2} - (x+1)^{5/2}\right]_0^4 \\ &= \frac{4}{15}\big((e+4)^{5/2} - e^{5/2} - 5^{5/2} + 1\big) \approx 13.308.\end{aligned}$$

□

Second solution. Let us try to see how things change, if we use the other iterated integral. We have

$$\int_0^4 \sqrt{x+e^y}\,dx = \int_{e^y}^{4+e^y} \sqrt{u}\,du = \left[\frac{u^{3/2}}{3/2}\right]_{e^y}^{4+e^y} = \frac{2}{3}\left((4+e^y)^{3/2} - e^{3y/2}\right).$$

Hence, Fubini's theorem yields

$$\begin{aligned}\iint_R e^y \sqrt{x+e^y}\,dA &= \int_0^1 \int_0^4 e^y \sqrt{x+e^y}\,dxdy = \int_0^1 e^y \left[\int_0^4 \sqrt{x+e^y}\,dx\right] dy \\ &= \frac{2}{3}\int_0^1 e^y\left((4+e^y)^{3/2} - e^{3y/2}\right) dy = \frac{2}{3}\int_1^e \left((4+u)^{3/2} - u^{3/2}\right) du \\ &= \frac{2}{3}\frac{2}{5}\left[(4+u)^{5/2} - u^{5/2}\right]_1^e = \frac{4}{15}\left((4+e)^{5/2} - e^{5/2} - 5^{5/2} + 1\right).\end{aligned}$$

□

## Exercises

13.1. Use a Riemann sum with $n = 3$ to estimate the value of $\iint_R (xy + y^2)\,dA$, where $R = [1,2]\times[-1,1]$. Take the sample points to be:

(a) the lower right corners of the subrectangles; (b) the centers of the subrectangles.

Evaluate the given double integral by identifying it as the volume of a solid and calculating that volume using geometric means.

13.2. $\displaystyle\iint_R 2\,dA,\ R = [-1,2]\times[0,3]$

13.3. $\displaystyle\iint_R (3-x)\,dA,\ R = [0,1]\times[0,1]$

Evaluate the given iterated integral.

13.4. $\displaystyle\int_1^2\int_1^3 \left(1+2xy^2\right) dxdy$

13.5. $\displaystyle\int_1^4\int_1^3 \left(\frac{x}{y} + y\ln x\right) dydx$

13.6. $\displaystyle\int_0^1\int_0^1 3^{x+2y}\,dydx$

Evaluate the given double integral.

13.7. $\displaystyle\iint_R \left(xy+y^2\right) dA,\ R = [1,2]\times[-1,1]$

13.8. $\displaystyle\iint_R \frac{xy}{x^2+y^2+1}\,dA,\ R = [0,1]\times[0,1]$

13.9. $\displaystyle\iint_R \sin(3x+2y)\,dA,\ R = [-\pi,0]\times[-\tfrac{1}{2}\pi, \tfrac{1}{4}\pi]$

13.10. $\displaystyle\iint_R \frac{2+y}{1+x^2}\,dA,\ R = [0,1]\times[-2,3]$

13.11. $\displaystyle\iint_R (x+y)e^{2xy}\,dA,\ R = [1,2]\times[0,1]$

13.12. $\displaystyle\iint_R xy\cos\left(x^2+y^2\right) dA,\ R = \left[0,\sqrt{\pi}\right]\times\left[0,\sqrt{\pi}\right]$

LECTURE 14

# Double Integrals over General Regions

To define the double integral of a bounded function $f(x, y)$ on a bounded region $D$ in the plane, we use that the definition given in §13.1 applies to any piecewise continuous function $F(x, y)$ defined on a rectangle (recall Remark 2 in §13.1).

## 14.1. Definition

Let $D$ be the closed region in $\mathbb{R}^2$ bounded by a piecewise smooth curve $\gamma$ and let $f(x, y)$ be piecewise continuous inside $D$. Define the function

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

Then the *double integral of $f$ over $D$* is defined by

$$\iint_D f(x, y)\, dA = \iint_R F(x, y)\, dA,$$

where $R$ is any rectangle containing $D$.

Remark. Note that with the assumption that $f(x, y)$ is piecewise continuous in $D$, $F(x, y)$ is continuous everywhere except possibly at the points of $\gamma$ and the (possible) points inside $D$ where $f(x, y)$ is discontinuous. Thus, if $f(x, y)$ is actually continuous in $D$, the only discontinuities of $F(x, y)$ lie on $\gamma$. If $f(x, y)$ is a genuine piecewise continuous function, with discontinuities on some smooth curves $\gamma_1, \dots, \gamma_n$ inside $D$, then the discontinuities of $F(x, y)$ lie on the curves $\gamma, \gamma_1, \dots, \gamma_n$. In either case, $F(x, y)$ is piecewise continuous and $\iint_R F(x, y)\, dA$ is defined by (13.1).

## 14.2. Properties of the double integral

We now state several properties of the double integral which are two-dimensional versions of the properties of the definite integral listed in Appendix B.

Theorem 14.1. *Let $f(x, y), g(x, y)$ be piecewise continuous functions and $a, b$ denote constants. Then*

i) $\displaystyle\iint_D [af(x, y) + bg(x, y)]\, dA = a \iint_D f(x, y)\, dA + b \iint_D g(x, y)\, dA.$

ii) *If $f(x, y) \le g(x, y)$, then*

$$\iint_D f(x, y)\, dA \le \iint_D g(x, y)\, dA.$$

iii) *If $D = D_1 \cup D_2,\ D_1 \cap D_2 = \varnothing$, then*

$$\iint_D f(x, y)\, dA = \iint_{D_1} f(x, y)\, dA + \iint_{D_2} f(x, y)\, dA.$$

iv) $\iint_D 1\, dA = |D|$, *$|D|$ being the area of $D$.*

Furthermore, similar to the double integral over a rectangle, the double integral over a general region $D$ can be interpreted as the volume of a solid or the average value of the function. When $f(x, y) \geq 0$,

$$\iint_D f(x, y)\, dA = \text{vol.}\left\{(x, y, z) : (x, y) \in D,\ 0 \leq z \leq f(x, y)\right\}.$$

Also, for any function $f(x, y)$ that is piecewise continuous on $D$, we have

$$\underset{(x,y)\in D}{\text{average}} f(x, y) = \frac{1}{|D|} \iint_D f(x, y)\, dA. \tag{14.1}$$

## 14.3. Evaluation of double integrals

We now state two formulas which follow from Fubini's theorem and which we shall use to evaluate double integrals. First, consider a region $D$ of the form

$$D = \left\{(x, y) : a \leq x \leq b,\ g_1(x) \leq y \leq g_2(x)\right\},$$

where $g_1$ and $g_2$ are continuous functions of $x$. We call such regions *type I regions*. If $f$ is a continuous function on $D$, we have

$$\iint_D f(x, y)\, dA = \int_a^b \left[\int_{g_1(x)}^{g_2(x)} f(x, y)\, dy\right] dx. \tag{14.2}$$

We also define *type II regions*, which a regions of the form

$$D = \left\{(x, y) : c \leq y \leq d,\ g_1(y) \leq x \leq g_2(y)\right\},$$

where $g_1$ and $g_2$ are continuous functions of $y$. If $f$ is a continuous function on such a region, we have

$$\iint_D f(x, y)\, dA = \int_c^d \left[\int_{g_1(y)}^{g_2(y)} f(x, y)\, dx\right] dy. \tag{14.3}$$

Example 14.1. Evaluate

$$\iint_D ye^x\, dA,$$

where $D = \left\{(x, y) : 0 \leq x \leq 1,\ \sqrt{x} \leq y \leq 1\right\}$.

Solution. Since $D$ is a type I region, (14.2) yields

$$\begin{aligned}\iint_D ye^x\, dA &= \int_0^1 \int_{\sqrt{x}}^1 ye^x\, dy dx = \int_0^1 \left[\frac{y^2 e^x}{2}\right]_{\sqrt{x}}^1 dx = \frac{1}{2}\int_0^1 (1 - x)e^x\, dx \\ &= \frac{1}{2}\left(\left[(1 - x)e^x\right]_0^1 + \int_0^1 e^x\, dx\right) = \frac{1}{2}\left(-1 + \left[e^x\right]_0^1\right) = \frac{e - 2}{2}.\end{aligned}$$

□

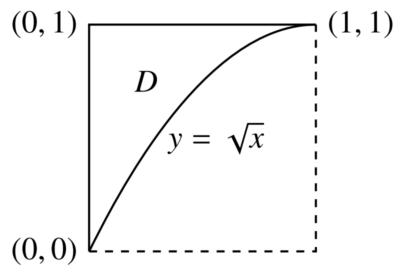



FIGURE 14.1. The region $D$

SECOND SOLUTION. Alternatively, we can try to describe $D$ as a type II region and then use (14.3). We start by sketching $D$: see Figure 14.1. In our type I description of $D$, we first let $x$ vary in the range $0 \le x \le 1$, and then for each fixed $x$, we restrict $y$ to the range $\sqrt{x} \le y \le 1$. To obtain a type II description, we interchange the roles of $x$ and $y$: first, we let $y$ vary in $0 \le y \le 1$, and then for each fixed $y$, we restrict $x$ to the range $0 \le x \le y^2$. Here, the limit $y^2$ comes from the description of the curve $y = \sqrt{x}$ in the form $x = x(y)$:

$$y = \sqrt{x} \quad \Longleftrightarrow \quad x = y^2.$$

Thus,

$$D = \left\{(x, y) : 0 \le y \le 1,\ 0 \le x \le y^2\right\},$$

and (14.3) yields

$$\begin{aligned} \iint_D f(x, y)\, dA &= \int_0^1 \int_0^{y^2} ye^x \, dxdy = \int_0^1 y\left[e^x\right]_0^{y^2} dy = \int_0^1 y\left(e^{y^2} - 1\right) dy \\ &= \int_0^1 \left(e^u - 1\right)\left(\tfrac{1}{2}du\right) = \left[\frac{e^u - u}{2}\right]_0^1 = \frac{e-2}{2}. \end{aligned}$$

□

EXAMPLE 14.2. Evaluate the double integral

$$\iint_D e^{y^2}\, dA,$$

where $D = \left\{(x, y) : 0 \le x \le 1,\ x \le y \le 1\right\}$.

SOLUTION. Our initial impulse is to write

$$\iint_D e^{y^2}\, dA = \int_0^1 \int_x^1 e^{y^2} dydx,$$

but this approach is doomed, because the indefinite integral $\int e^{y^2} dy$ is not solvable in elementary functions. On the other hand, if we observe that

$$D = \left\{(x, y) : 0 \le y \le 1,\ 0 \le x \le y\right\},$$

we have

$$\iint_D e^{y^2}\,dA = \int_0^1 \int_0^y e^{y^2}\,dxdy = \int_0^1 e^{y^2}\left[x\right]_0^y dy = \int_0^1 e^{y^2} y\,dy$$
$$= \int_0^1 e^u \left(\tfrac{1}{2}du\right) = \left[\frac{e^u}{2}\right]_0^1 = \frac{e-1}{2}.$$

□

Example 14.3. Evaluate the double integral

$$\iint_D 2xy\,dA,$$

where $D$ is the triangle with vertices $(0,0)$, $(0,3)$, and $(1,2)$.

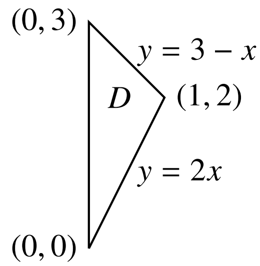



Figure 14.2. The triangle $D$

Solution. The sides of the triangle $D$ are the lines in the plane determined by the pairs of points $(0,0)$ and $(0,3)$, $(0,0)$ and $(1,2)$, and $(0,3)$ and $(1,2)$. These are the lines with equations

$$x = 0, \quad y = 2x, \quad \text{and} \quad y = 3 - x,$$

respectively (see Figure 14.2). Hence, we can express $D$ as a type I region:

$$D = \left\{(x,y) : 0 \le x \le 1,\ 2x \le y \le 3 - x\right\}.$$

Thus, the given double integral is

$$\iint_D 2xy\,dA = \int_0^1 \int_{2x}^{3-x} 2xy\,dydx = \int_0^1 x\left[(3-x)^2 - 4x^2\right]\,dx$$
$$= \int_0^1 \left(9x - 6x^2 - 3x^3\right)\,dx = \left[\frac{9x^2}{2} - 2x^3 - \frac{3x^4}{4}\right]_0^1 = \frac{7}{4}.$$

□

Example 14.4. Find the volume of the solid $E$ in the first octant that is bounded by the cylinder $y^2 + z^2 = 4$ and by the plane $x = 2y$.

Solution. We start by sketching the solid $E$: the left side of Figure 14.3 shows a joint display of the parts of the cylinder $y^2 + z^2 = 4$ and the plane $x = 2y$ that lie in the first octant and how they intersect. From that display, we see that $E$ is bounded from below by a triangle $D$ in the $xy$-plane and from above by the part of the cylinder that lies above $D$. The triangle $D$ (see the right side of

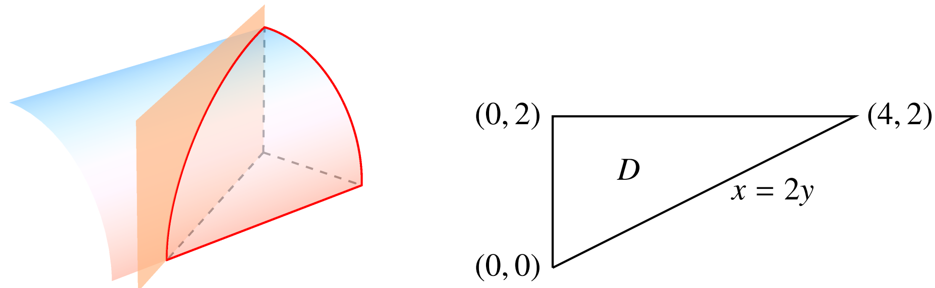



FIGURE 14.3. The solid $E$ and its base

Figure 14.3) is the region in the $xy$-plane bounded by the $y$-axis, by the line $x = 2y$, and by the line $y = 2$, along which the cylinder intersects tha $xy$-plane. Thus, we see that

$$D = \left\{(x, y) : 0 \le x \le 4,\ \tfrac{1}{2}x \le y \le 2\right\} = \left\{(x, y) : 0 \le y \le 2,\ 0 \le x \le 2y\right\}.$$

Since the surface of the cylinder is the graph of $z = \sqrt{4 - y^2}$, it follows that

$$E = \left\{(x, y, z) : (x, y) \in D,\ 0 \le z \le \sqrt{4 - y^2}\right\}.$$

Hence, by Remark 3 and Fubini's theorem,

$$\text{vol.}(E) = \iint_D \sqrt{4 - y^2}\, dA = \int_0^4 \int_{x/2}^2 \sqrt{4 - y^2}\, dydx = \int_0^2 \int_0^{2y} \sqrt{4 - y^2}\, dxdy.$$

We shall use the latter iterated integral, because its evaluation is less technical:

$$\text{vol.}(E) = \int_0^2 \int_0^{2y} \sqrt{4 - y^2}\, dxdy = \int_0^2 2y\sqrt{4 - y^2}\, dy = \int_4^0 \sqrt{u}\,(-du) = \left[\frac{u^{3/2}}{3/2}\right]_0^4 = \frac{16}{3}.$$

□

EXAMPLE 14.5. Reverse the order of integration in the iterated integral

$$\int_0^1 \int_0^{\arctan x} f(x, y)\, dydx.$$

FIRST SOLUTION. We have

$$\int_0^1 \int_0^{\arctan x} f(x, y)\, dydx = \iint_D f(x, y)\, dA,$$

where $D = \left\{(x, y) : 0 \le x \le 1,\ 0 \le y \le \arctan x\right\}$ is the region displayed on Figure 14.4.

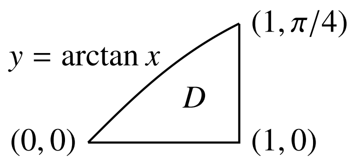



FIGURE 14.4. The region $D$

Since

$$y = \arctan x \quad \Longleftrightarrow \quad x = \tan y,$$

a reversal of the roles of $x$ and $y$ yields a representation of $D$ as a type II region:

$$D = \{(x, y) : 0 \le y \le \pi/4,\ \tan y \le x \le 1\}.$$

Thus,

$$\iint_D f(x, y)\, dA = \int_0^{\pi/4} \int_{\tan y}^{1} f(x, y)\, dxdy.$$

□

Second solution. We now demonstrate an alternative way for changing $D$ into a type II set. Initially, $D$ is described by the inequalities

$$0 \le x \le 1, \quad 0 \le y \le \arctan x.$$

Since $\arctan x \le \pi/4$ for all $x$, $0 \le x \le 1$, we can add to the above inequalities the redundant inequality $y \le \pi/4$. Thus, we now have $D$ described by

$$0 \le x \le 1, \quad 0 \le y \le \arctan x, \quad y \le \pi/4,$$

or equivalently, by

$$0 \le x \le 1, \quad 0 \le y \le \pi/4, \quad y \le \arctan x.$$

We observe that

$$y \le \arctan x \quad \Longleftrightarrow \quad \tan y \le x \quad \Longleftrightarrow \quad x \ge \tan y,$$

so we can now replace the definition of $D$ by

$$0 \le x \le 1, \quad 0 \le y \le \pi/4, \quad x \ge \tan y.$$

Since $\tan y \ge 0$ for all $y$ in the range $0 \le y < \pi/2$, we find that the condition $x \ge 0$ is superfluous and can be dropped. This yields

$$D = \{(x, y) : 0 \le y \le \pi/4,\ \tan y \le x \le 1\},$$

whence

$$\int_0^1 \int_0^{\arctan x} f(x, y)\, dydx = \int_0^{\pi/4} \int_{\tan y}^{1} f(x, y)\, dxdy.$$

□

Example 14.6. Reverse the order of integration in the iterated integral

$$\int_0^2 \int_{x/2}^{3-x} f(x, y)\, dydx.$$

First solution. We have

$$\int_0^2 \int_{x/2}^{3-x} f(x, y)\, dydx = \iint_D f(x, y)\, dA,$$

where $D = \{(x, y) : 0 \le x \le 2,\ x/2 \le y \le 3 - x\}$ is the region displayed on Figure 14.5. We can easily convert the equations of the lines to the form $x = g(y)$ (see the figure). However, this time we have to be careful, because the function $x = g(y)$ that bounds $D$ from the right is neither $x = 3 - y$ nor $x = 2y$. The precise representation of $D$ as a type II region is

$$D = \{(x, y) : 0 \le y \le 3,\ 0 \le x \le g(y)\},$$

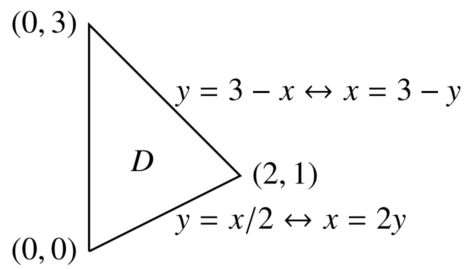



Figure 14.5. The region $D$

where

$$g(y) = \begin{cases} 2y & \text{if } 0 \le y \le 1, \\ 3 - y & \text{if } 1 \le y \le 3. \end{cases}$$

However, this function is not particularly convenient to work with. To obtain a "friendlier" answer, we split $D$ in two using the horizontal line $y = 1$. Then each of the two subregions we get is bounded from the right by a single line, say:

$$D_1 = \{(x, y) : 0 \le y \le 1,\ 0 \le x \le 2y\}, \quad D_2 = \{(x, y) : 1 \le y \le 3,\ 0 \le x \le 3 - y\}.$$

Thus,

$$\int_0^2 \int_{x/2}^{3-x} f(x, y)\, dydx = \int_0^1 \int_0^{2y} f(x, y)\, dxdy + \int_1^3 \int_0^{3-y} f(x, y)\, dxdy.$$

□

Second solution. Initially, $D$ is described by the inequalities

$$0 \le x \le 2, \quad x/2 \le y \le 3 - x.$$

On noting that $3 - x \le 3$ and $x/2 \ge 0$, we can add to the above inequalities the redundant condition $0 \le y \le 3$, that is, we have $D$ described by

$$0 \le x \le 2, \quad 0 \le y \le 3, \quad x/2 \le y \le 3 - x.$$

Next, we observe that

$$x/2 \le y \iff x \le 2y, \qquad y \le 3 - x \iff x \le 3 - y,$$

so we can further transform the definition of $D$ to

$$0 \le y \le 3, \quad 0 \le x \le 2, \quad x \le 2y, \quad x \le 3 - y,$$

or equivalently to

$$0 \le y \le 3, \quad 0 \le x \le \min(2, 2y, 3 - y).$$

This is the equivalent of the description of $D$ using $g(y)$ in our first solution. By comparing 2, $2y$, and $3 - y$, we find that the minimum equals $2y$ or $3 - y$ according as $y \le 1$ or $y \ge 1$. Therefore, $D$ is the union of the region $D_1$ described by

$$0 \le y \le 1, \quad 0 \le x \le 2y$$

and the region $D_2$ described by

$$1 \le y \le 3, \quad 0 \le x \le 3 - y.$$

Thus,

$$\int_0^2 \int_{x/2}^{3-x} f(x,y)\,dydx = \int_0^1 \int_0^{2y} f(x,y)\,dxdy + \int_1^3 \int_0^{3-y} f(x,y)\,dxdy.$$

□

## Exercises

Evaluate the given iterated integral.

14.1. $\int_0^2 \int_y^{4-y} \left(y^2 - 3x\right) dxdy$ 14.2. $\int_1^2 \int_0^{2x} \sqrt[3]{4 - x^2}\,dydx$ 14.3. $\int_0^1 \int_0^y y\cos(x-y)\,dxdy$

Evaluate the given double integral.

14.4. $\iint_D yx^2\,dA,\ D = \left\{(x,y) : 1 \le y \le 3, y \le x \le y^2\right\}$

14.5. $\iint_D 3x\sin 2y\,dA,\ D = \left\{(x,y) : 0 \le x \le \pi, x/2 \le y \le x\right\}$

14.6. $\iint_D \frac{2y}{x^2+1}\,dA,\ D = \left\{(x,y) : 0 \le x \le 1, 0 \le y \le \sqrt{x}\right\}$

14.7. $\iint_D e^{x/y}\,dA,\ D = \left\{(x,y) : 0 \le y \le 1, y^2 \le x \le y\right\}$

14.8. $\iint_D x\sin y\,dA,\ D = \left\{(x,y) : 0 \le x \le 1, 0 \le y \le x^2\right\}$

14.9. $\iint_D \left(x^2 + y^3\right) dA$, $D$ is the region bounded by $y = 2x$ and $y = x^2$

14.10. $\iint_D (3x - 4y)\,dA$, $D$ is the triangle with vertices $(1,-1)$, $(3,1)$, and $(1,2)$

14.11. $\iint_D xy^2\,dA$, $D$ is the right half of the disk $x^2 + y^2 \le 4$

14.12. $\iint_D x^2\,dA$, $D$ is the triangle with vertices $(0,-1)$, $(1,1)$, and $(0,2)$

14.13. $\iint_D \frac{6x^2}{y^4+2}\,dA,\ D = \left\{(x,y) : 1 \le y \le 2, 0 \le x \le 2y\right\}$

Sketch the region of integration and reverse the order of integration.

14.14. $\int_0^1 \int_{\sqrt{x}}^1 f(x,y)\,dydx$ 14.16. $\int_0^2 \int_{x^2/2}^{4-x} f(x,y)\,dydx$ 14.18. $\int_1^2 \int_0^{\ln y} f(x,y)\,dxdy$

14.15. $\int_0^2 \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} f(x,y)\,dydx$ 14.17. $\int_0^2 \int_2^{4-y} f(x,y)\,dxdy$ 14.19. $\int_0^{\pi/2} \int_{\sin x}^1 f(x,y)\,dydx$

Evaluate the given integral by reversing the order of integration.

14.20. $\int_0^1 \int_{y^2}^1 ye^{-x^2}\,dxdy$ 14.21. $\int_0^8 \int_{\sqrt[3]{y}}^2 \frac{dxdy}{x^4+1}$ 14.22. $\int_0^4 \int_{\sqrt{x}}^2 e^{y^3}\,dydx$

14.23. Find the average value of the function $f(x,y) = y^3$ over the triangle with vertices $(0,2)$, $(1,1)$, and $(3,2)$.

14.24. Find the volume of the solid lying below the hyperbolic paraboloid $z = 4 + x^2 - y^2$ and above the rectangle $R = [0,1] \times [0,2]$.

14.25. Find the volume of the solid in the first octant bounded by the cylinder $x^2 + z^2 = 1$, the plane $y = 3$, and the coordinate planes.

14.26. Find the volume of the solid lying below the plane $4x - y + 2z = 6$ and above the region bounded by the parabola $x = y^2 - 1$ and the right half of the circle $x^2 + y^2 = 1$.

14.27. Find the volume of the solid lying below the paraboloid $z = 3xy + 2y + 1$ and above the region bounded by the parabola $y = x^2$ and the line $y = 2x + 3$.

14.28. Find the volume of the solid lying below the parabolic cylinder $z = y^2$ and above the region bounded by the parabola $y = x^2$ and the line $y = x + 2$.

14.29. Find the volume of the solid bounded by the cylinders $x^2 + y^2 = 1$ and $x^2 + z^2 = 1$.

14.30. Evaluate $\iint_D \sqrt{1 - x^2 - y^2}\, dA$, where $D$ is the unit disk $x^2 + y^2 \leq 1$, by identifying it as the volume of a solid and calculating that volume using geometric means.

14.31. Evaluate $\iint_D \left(x^8 \sin x + y^9 + 1\right) dA$, where $D$ is the disk $x^2 + y^2 \leq 4$.

LECTURE 15

# Double Integrals in Polar Coordinates

## 15.1. The formula

THEOREM 15.1. *Let $f(x, y)$ be piecewise continuous in a region $D$ in the (Cartesian) xy-plane and let $R$ be the same region in polar coordinates, that is,*

$$R = \{(r, \theta) : (r\cos\theta, r\sin\theta) \in D\}.$$

*Then*

$$\iint_D f(x, y)\, dA = \iint_R f(r\cos\theta, r\sin\theta) r\, dA. \tag{15.1}$$

The formal proof of (15.1) lies beyond the scope of these notes, but we should give at least some motivation for this formula. Let $a, b, \alpha, \beta$ be non-negative numbers such that $a < b$ and $0 \le \alpha < \beta \le 2\pi$. Consider the case when $D$ is the region bounded by the circles $x^2 + y^2 = a^2$ and $x^2 + y^2 = b^2$ and by the half-lines through the origin forming angles $\alpha$ and $\beta$ with the positive $x$-axis (see Figure 15.1). Note that $D$ contains the points $(x, y)$ in the plane whose polar coordinates $(r, \theta)$ satisfy the inequalities $a \le r \le b$ and $\alpha \le \theta \le \beta$. We recall that the area of $D$ is given by the formula

$$\text{area}(D) = \tfrac{1}{2}(\beta - \alpha)\big(b^2 - a^2\big) = (\beta - \alpha)(b - a)\left(\frac{b + a}{2}\right). \tag{15.2}$$

For each integer $n \ge 1$, let

$$r_0 = a, \quad r_1 = a + h, \quad r_2 = a + 2h, \quad \dots, \quad r_n = a + nh = b, \qquad h = (b - a)/n,$$

be the points that divide the interval $[a, b]$ into $n$ subintervals of equal lengths; and let

$$\theta_0 = \alpha, \quad \theta_1 = \alpha + \delta, \quad \theta_2 = \alpha + 2\delta, \quad \dots, \quad \theta_n = \alpha + n\delta = \beta, \qquad \delta = (\beta - \alpha)/n,$$

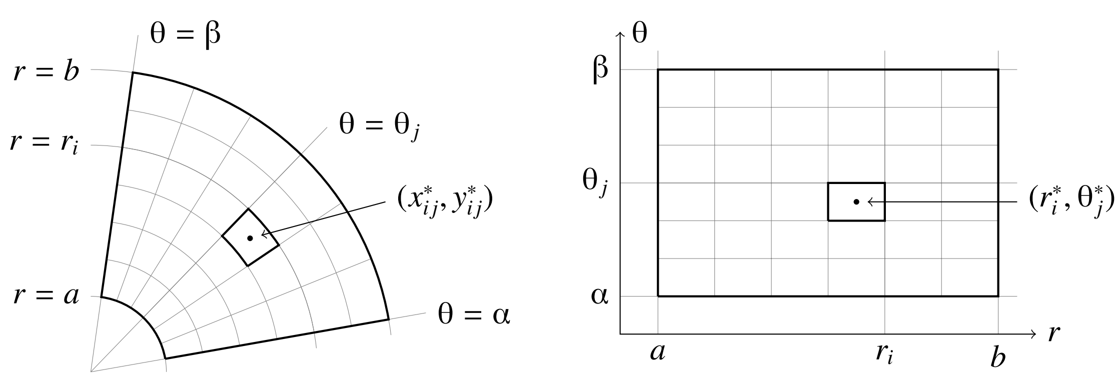



FIGURE 15.1. The regions $D$ and $R$, with partitions and sample points

be the points that divide the interval $[\alpha, \beta]$ into $n$ subintervals of equal lengths. We now subdivide the region $D$ into $n^2$ subregions of similar shapes bounded by the circles of radii $a, r_1, r_2, \ldots, r_{n-1}, b$ centered at the origin and the half-lines through the origin forming angles $\alpha, \theta_1, \theta_2, \ldots, \theta_{n-1}, \beta$ with the positive $x$-axis (see Figure 15.1). Let $D_{ij}$ denote the subregion bounded by the circles of radii $r_{i-1}$ and $r_i$ and by the half-lines at angles $\theta_{j-1}$ and $\theta_j$ with the positive $x$-axis: in polar coordinates, $D_{ij}$ is given by the inequalities

$$r_{i-1} \leq r \leq r_i, \quad \theta_{j-1} \leq \theta \leq \theta_j.$$

We now pick a sample point $(x_{ij}^*, y_{ij}^*)$ in each subregion $D_{ij}$: we choose $(x_{ij}^*, y_{ij}^*)$ to be the point in the plane whose polar coordinates are $r_i^* = (r_{i-1} + r_i)/2$ and $\theta_j^* = (\theta_{j-1} + \theta_j)/2$. Using the subregions $D_{ij}$ and the sample points $(x_{ij}^*, y_{ij}^*)$, we form the sum

$$\sum_{i=1}^{n} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)|D_{ij}|, \qquad |D_{ij}| = \text{area}(D_{ij}).$$

Note that this sum bears a strong resemblance to a Riemann sum for $f(x, y)$. Indeed, it can be proved that if $f(x, y)$ is continuous, then

$$\iint_D f(x, y)\, dA = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)|D_{ij}|. \tag{15.3}$$

We now express the double sum in (15.3) in terms of the $r_i$'s and the $\theta_j$'s. By (15.2) with $a = r_{i-1}$, $b = r_i$, $\alpha = \theta_{j-1}$ and $\beta = \theta_j$, we have

$$|D_{ij}| = (\theta_j - \theta_{j-1})(r_i - r_{i-1}) \left( \frac{r_i + r_{i-1}}{2} \right) = \delta h r_i^*.$$

Furthermore, since the polar coordinates of $(x_{ij}^*, y_{ij}^*)$ are $(r_i^*, \theta_j^*)$, we have

$$x_{ij}^* = r_i^* \cos \theta_j^*, \quad y_{ij}^* = r_i^* \sin \theta_j^*,$$

whence $f(x_{ij}^*, y_{ij}^*) = f(r_i^* \cos \theta_j^*, r_i^* \sin \theta_j^*)$. Therefore,

$$\sum_{i=1}^{n} \sum_{j=1}^{n} f(x_{ij}^*, y_{ij}^*)|D_{ij}| = \sum_{i=1}^{n} \sum_{j=1}^{n} f(r_i^* \cos \theta_j^*, r_i^* \sin \theta_j^*) r_i^* \delta h.$$

Substituting this expression into the right side of (15.3), we obtain

$$\iint_D f(x, y)\, dA = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(r_i^* \cos \theta_j^*, r_i^* \sin \theta_j^*) r_i^* \delta h. \tag{15.4}$$

The double sum on the right side of (15.4) is a Riemann sum for the function $f(r \cos \theta, r \sin \theta) r$ on the rectangle $R = [a, b] \times [\alpha, \beta]$ when the sample points $(r_i^*, \theta_j^*)$ are always chosen at the centers of the subrectangles (see Figure 15.1). Hence,

$$\lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(r_i^* \cos \theta_j^*, r_i^* \sin \theta_j^*) r_i^* \delta h = \iint_R f(r \cos \theta, r \sin \theta) r\, dA.$$

The latter equality and (15.4) justify the (15.1) in the special case when $D$ is a sector of the above kind.

## 15.2. Examples

EXAMPLE 15.1. Evaluate the integral

$$\iint_D (x+y)\,dA,$$

where $D$ is the region lying between the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$ and to the left of the $y$-axis.

SOLUTION. The polar form of $D$ is

$$R = \{(r,\theta) : 1 \le r \le 2,\ \pi/2 \le \theta \le 3\pi/2\}.$$

Hence,

$$\begin{aligned}\iint_D (x+y)\,dA &= \iint_R (r\cos\theta + r\sin\theta) r\,dA = \int_1^2 \left[\int_{\pi/2}^{3\pi/2} \left(r^2\cos\theta + r^2\sin\theta\right) d\theta\right] dr \\ &= \int_1^2 r^2 \big[\sin\theta - \cos\theta\big]_{\pi/2}^{3\pi/2} dr = -2\int_1^2 r^2 dr = -\frac{14}{3}.\end{aligned}$$

□

EXAMPLE 15.2. Find the area enclosed by the curve $r = 4 + 3\cos\theta$.

SOLUTION. Since $4 + 3\cos\theta \ge 1$ for all $\theta$, the given curve "loops" around the origin, enclosing the region (see Figure 15.2)

$$R = \{(r,\theta) : 0 \le \theta \le 2\pi,\ 0 \le r \le 4 + 3\cos\theta\}.$$

If $D$ is the region $R$ in Cartesian coordinates, then we know from property (iv) in Theorem 14.1 that its area is

$$|D| = \iint_D 1\,dA.$$

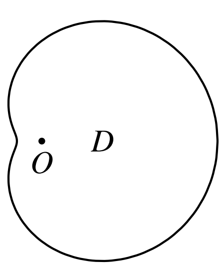



FIGURE 15.2. The region $r \le 4 + 3\cos\theta$

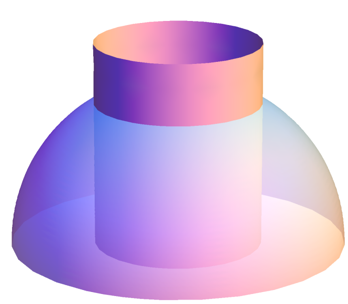


FIGURE 15.3. The solid in Example 15.3

Hence, by (15.1),

$$
\begin{aligned}
|D| &= \iint_D 1\,dA = \iint_R r\,dA = \int_0^{2\pi}\left[\int_0^{4+3\cos\theta} r\,dr\right]d\theta \\
&= \frac{1}{2}\int_0^{2\pi}(4+3\cos\theta)^2\,d\theta = \frac{1}{2}\int_0^{2\pi}\left(16+24\cos\theta+9\cos^2\theta\right)d\theta \\
&= \int_0^{2\pi}\left[8+12\cos\theta+\frac{9}{4}(1+\cos 2\theta)\right]d\theta = \left[\frac{41\theta}{4}+12\sin\theta+\frac{9}{8}\sin 2\theta\right]_0^{2\pi} = \frac{41\pi}{2}.
\end{aligned}
$$

□

Example 15.3. Find the volume of the solid that lies inside the sphere $x^2 + y^2 + z^2 = 16$ and outside the cylinder $x^2 + y^2 = 4$.

Solution. The given solid is symmetric with respect to the $xy$-plane, so it is twice as large as the solid $E$ that lies inside the sphere, outside the cylinder, and above the $xy$-plane (see Figure 15.3). This solid is the part of the space between the upper hemisphere and the annular region $D$ in the $xy$-plane bound by the sphere and the cylinder:

$$D = \left\{(x,y) : 4 \le x^2 + y^2 \le 16\right\}.$$

Since the upper hemisphere of $x^2 + y^2 + z^2 = 16$ can be described as the graph of the function $z = \sqrt{16 - x^2 - y^2}$, we can express $E$ as

$$E = \left\{(x,y,z) : (x,y) \in D,\ 0 \le z \le \sqrt{16 - x^2 - y^2}\right\}.$$

Hence,

$$\text{vol.}(E) = \iint_D \sqrt{16 - x^2 - y^2}\,dA.$$

The polar form of $D$ is

$$R = \left\{(r,\theta) : 2 \le r \le 4,\ 0 \le \theta \le 2\pi\right\},$$

and

$$\sqrt{16 - (r\cos\theta)^2 - (r\sin\theta)^2} = \sqrt{16 - r^2(\cos^2\theta + \sin^2\theta)} = \sqrt{16 - r^2}.$$

Hence, (15.1) yields

$$
\begin{aligned}
\text{vol.}(E) &= \iint_D \sqrt{16 - x^2 - y^2}\,dA = \iint_R r\sqrt{16 - r^2}\,dA \\
&= \int_2^4\left[\int_0^{2\pi} r\sqrt{16 - r^2}\,d\theta\right]dr = 2\pi\int_2^4 r\sqrt{16 - r^2}\,dr.
\end{aligned}
$$

We evaluate the last integral using the substitution $u = 16 - r^2$. Then

$$du = -2r\,dr, \quad u(2) = 12, \quad u(4) = 0,$$

so

$$\text{vol.}(E) = 2\pi\int_{12}^0 u^{1/2}\left(-\tfrac{1}{2}du\right) = \pi\int_0^{12} u^{1/2}\,du = \left[\frac{2\pi}{3}u^{3/2}\right]_0^{12} = 16\sqrt{3}\,\pi.$$

Thus, the volume of the original solid is $32\sqrt{3}\,\pi$. □

## Exercises

Evaluate the given integral by changing to polar coordinates.

15.1. $\iint_D (x+3y)\,dA$, $D$ is the disk $x^2+y^2 \le 4$

15.2. $\iint_D e^{x^2+y^2}\,dA$, $D$ is the right half of the disk $x^2+y^2 \le 1$

15.3. $\iint_D \sqrt{16-x^2-y^2}\,dA$, $D$ is the part of the annulus $1 \le x^2+y^2 \le 9$ lying in the first quadrant

15.4. $\iint_D \dfrac{xy}{4+x^2+y^2}\,dA$, $D$ is the part of the disk $x^2+y^2 \le 9$ lying below the line $y=x$ and above the $x$-axis

15.5. $\iint_D xye^{-x^2-y^2}\,dA$, $D$ is the larger region bounded by the $y$-axis, the line $x+y=0$, and the semicircle $x=\sqrt{4-y^2}$

15.6. Find the volume of the solid enclosed by the paraboloid $z=4-x^2-y^2$ and the $xy$-plane.

15.7. Find the volume of the solid enclosed by the cone $z^2=x^2+y^2$ and the plane $z=3$.

15.8. Find the volume of the solid lying inside both the cylinder $y^2+z^2=4$ and the sphere $x^2+y^2+z^2=9$.

15.9. Find the volume lying above the paraboloid $z=\frac{1}{2}\left(x^2+y^2\right)$ and inside the ellipsoid $x^2+y^2+4z^2=2$.

15.10. Find the volume lying below the hemisphere $z=\sqrt{4-x^2-y^2}$ and above the region $D$ in the $xy$-plane where $x^2+y^2 \le 2x$.

15.11. Find the volume of the solid lying below the plane $z=x+y$ and above the disk $x^2+y^2 \le x+y$.

15.12. Find the volume of the solid lying inside all three cylinders $x^2+y^2=1$, $x^2+z^2=1$, and $y^2+z^2=1$.

LECTURE 16

# Triple Integrals

## 16.1. Definition

Let $f(x, y, z)$ be a bounded function on the rectangular box $B$,

$$[a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] = \{(x, y, z) : a_1 \le x \le b_1,\ a_2 \le y \le b_2,\ a_3 \le z \le b_3\}.$$

For each $n \ge 1$, let

$$x_0 = a_1, \quad x_1 = a_1 + \Delta x, \quad x_2 = a_1 + 2\Delta x, \quad \dots, \quad x_n = a_1 + n\Delta x = b_1, \qquad \Delta x = (b_1 - a_1)/n,$$

be the points that divide the interval $[a_1, b_1]$ into $n$ subintervals of equal lengths; let

$$y_0 = a_2, \quad y_1 = a_2 + \Delta y, \quad y_2 = a_2 + 2\Delta y, \quad \dots, \quad y_n = a_2 + n\Delta y = b_2, \qquad \Delta y = (b_2 - a_2)/n,$$

be the points that divide the interval $[a_2, b_2]$ into $n$ subintervals of equal lengths; and let

$$z_0 = a_3, \quad z_1 = a_3 + \Delta z, \quad z_2 = a_3 + 2\Delta z, \quad \dots, \quad z_n = a_3 + n\Delta z = b_3, \qquad \Delta z = (b_3 - a_3)/n,$$

be the points that divide the interval $[a_3, b_3]$ into $n$ subintervals of equal lengths. Also, pick $n^3$ sample points $(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})$ so that

$$x_{i-1} \le x^*_{ijk} \le x_i, \quad y_{j-1} \le y^*_{ijk} \le y_j, \quad z_{k-1} \le z^*_{ijk} \le z_k \qquad (1 \le i, j, k \le n).$$

Then the *triple integral of $f$ over $B$* is defined by

$$\iiint_B f(x, y, z)\, dV = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} f(x^*_{ijk}, y^*_{ijk}, z^*_{ijk}) \Delta V, \qquad \Delta V = \Delta x \Delta y \Delta z, \tag{16.1}$$

provided that the limit exists. The sums on the right side of (16.1) are called *Riemann sums of $f$ on $B$*.

More generally, if $S$ is a solid region in $\mathbb{R}^3$ bounded by a piecewise smooth surface $\Sigma$ and $f$ is a piecewise continuous function on $S$, we define the *triple integral of $f$ over $S$* as

$$\iiint_S f(x, y, z)\, dV = \iiint_B F(x, y, z)\, dV,$$

where $B$ is any rectangular box containing $S$ and $F$ is the function

$$F(x, y, z) = \begin{cases} f(x, y, z) & \text{if } (x, y, z) \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Remark. As in the cases of definite and double integrals, the above definitions define the triple integral for all piecewise continuous functions over solids bounded by piecewise smooth surfaces.

Furthermore, similarly to the definite and double integrals, the triple integral of a function is related to its average in the region of integration:

$$\underset{(x,y,z)\in S}{\text{average}}\, f(x,y,z) = \frac{1}{\text{vol.}(S)} \iiint_S f(x,y,z)\, dV. \tag{16.2}$$

## 16.2. Properties of the triple integral

The following theorem lists the three-dimensional versions of the properties of double integrals from Theorem 14.1.

THEOREM 16.1. *Let $f(x,y,z), g(x,y,z)$ be piecewise continuous functions and $a, b$ denote constants. Then*

i) $\displaystyle\iiint_S [af(x,y,z) + bg(x,y,z)]\, dV = a \iiint_S f(x,y,z)\, dV + b \iiint_S g(x,y,z)\, dV.$

ii) *If $f(x,y,z) \le g(x,y,z)$, then*

$$\iiint_S f(x,y,z)\, dV \le \iiint_S g(x,y,z)\, dV.$$

iii) *If $S = S_1 \cup S_2,\ S_1 \cap S_2 = \varnothing$, then*

$$\iiint_S f(x,y,z)\, dV = \iiint_{S_1} f(x,y,z)\, dV + \iiint_{S_2} f(x,y,z)\, dV.$$

iv) $\displaystyle\iiint_S 1\, dV = \text{vol.}(S).$

There is also a three-dimensional version of Fubini's theorem, which we shall use to evaluate triple integrals. Before stating it, we need to introduce some terminology. We say that a solid region $S$ in $\mathbb{R}^3$ is of *type I*, if it can be written in the form

$$S = \left\{(x,y,z) : (x,y) \in D,\ g_1(x,y) \le z \le g_2(x,y)\right\}$$

for some region $D$ in the $xy$-plane and some continuous functions $g_1$ and $g_2$ on $D$. We say that $S$ is a *type II region*, if it can be written in the form

$$S = \left\{(x,y,z) : (y,z) \in D,\ g_1(y,z) \le x \le g_2(y,z)\right\}$$

for some region $D$ in the $yz$-plane. We say that $S$ is a *type III region*, if it can be written in the form

$$S = \left\{(x,y,z) : (x,z) \in D,\ g_1(x,z) \le y \le g_2(x,z)\right\}$$

for some region $D$ in the $xz$-plane.

THEOREM 16.2. *Let $f(x,y,z)$ be a piecewise continuous function on a solid region $S$ in $\mathbb{R}^3$.*

i) *If $S$ is of type I, then*

$$\iiint_S f(x,y,z)\, dV = \iint_D \left[ \int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\, dz \right] dA; \tag{16.3}$$

ii) *If $S$ is of type II, then*

$$\iiint_S f(x,y,z)\, dV = \iint_D \left[ \int_{g_1(y,z)}^{g_2(y,z)} f(x,y,z)\, dx \right] dA; \tag{16.4}$$

iii) *If $S$ is of type III, then*

$$\iiint_S f(x,y,z)\,dV = \iint_D \left[ \int_{g_1(x,z)}^{g_2(x,z)} f(x,y,z)\,dy \right] dA. \tag{16.5}$$

## 16.3. Evaluation of triple integrals

EXAMPLE 16.1. Evaluate the triple integral

$$\iiint_S yz\cos(x^5)\,dV,$$

where $S = \{(x,y,z) : 0 \le x \le 1,\ 0 \le y \le x,\ x \le z \le 2x\}$.

SOLUTION. Let $D = \{(x,y) : 0 \le x \le 1,\ 0 \le y \le x\}$. By (16.3),

$$\begin{aligned}
\iiint_S yz\cos(x^5)\,dV &= \iint_D \left[ \int_x^{2x} yz\cos(x^5)\,dz \right] dA \\
&= \frac{1}{2}\iint_D y\cos(x^5)\left[(2x)^2 - x^2\right] dA = \frac{3}{2}\iint_D yx^2\cos(x^5)\,dA \\
&= \frac{3}{2}\int_0^1 \left[ \int_0^x yx^2\cos(x^5)\,dy \right] dx = \frac{3}{4}\int_0^1 x^4\cos(x^5)\,dx \\
&= \frac{3}{20}\int_0^1 \cos u\,du = \frac{3}{20}\big[\sin u\big]_0^1 = \frac{3}{20}\sin 1 \approx 0.126.
\end{aligned}$$

□

EXAMPLE 16.2. Evaluate the triple integral

$$\iiint_S \sqrt{x^2+z^2}\,dV,$$

where $S$ is the solid bounded by the paraboloid $y = x^2 + z^2$ and the plane $y = 4$.

SOLUTION. We can write $S$ as a type III region:

$$S = \{(x,y,z) : x^2+z^2 \le 4,\ x^2+z^2 \le y \le 4\}.$$

Hence, by (16.5),

$$\iiint_S \sqrt{x^2+z^2}\,dV = \iint_{x^2+z^2\le 4} \left[ \int_{x^2+z^2}^4 \sqrt{x^2+z^2}\,dy \right] dA = \iint_{x^2+z^2\le 4} \left(4 - x^2 - z^2\right)\sqrt{x^2+z^2}\,dA.$$

We can easily compute the last integral by changing to polar coordinates in the $xz$-plane:

$$\begin{aligned}
\iint_{x^2+z^2\le 4} \left(4 - x^2 - z^2\right)\sqrt{x^2+z^2}\,dA &= \iint_{0\le r\le 2} \left(4 - r^2\right) r\,r\,dA = \int_0^2 \int_0^{2\pi} r^2\left(4 - r^2\right) d\theta dr \\
&= 2\pi\int_0^2 \left(4r^2 - r^4\right) dr = 2\pi\left[\frac{4r^3}{3} - \frac{r^5}{5}\right]_0^2 = \frac{128\pi}{15}.
\end{aligned}$$

□

Example 16.3. Compute the volume of the solid $T$ described by the inequalities

$$x \geq 0, \quad z \geq 0, \quad y \geq 2x, \quad x + 2y + 2z \leq 2.$$

Solution. In fact, $T$ is the tetrahedron with vertices $(0,0,0)$, $(0,1,0)$, $(0,0,1)$, and $(\frac{2}{5}, \frac{4}{5}, 0)$, but we shall ignore this fact and make a point of using triple integrals and inequalities. By solving the inequality $x + 2y + 2z \leq 2$ for $z$, we can rewrite the definition of $T$ as

$$x \geq 0, \quad y \geq 2x, \quad 0 \leq z \leq 1 - x/2 - y. \tag{16.6}$$

Note that implicit in the inequality for $z$ is the condition

$$1 - x/2 - y \geq 0 \quad \Longleftrightarrow \quad y \leq 1 - x/2.$$

Adding the latter inequality to those in (16.6), we get

$$x \geq 0, \quad 2x \leq y \leq 1 - x/2, \quad 0 \leq z \leq 1 - x/2 - y$$

We now note that the condition $x \leq \frac{2}{5}$ is implicit in the inequalities for $y$. Making this condition explicit, we finally have $T$ described as a type I set:

$$T = \left\{(x,y,z) : 0 \leq x \leq \tfrac{2}{5},\ 2x \leq y \leq 1 - x/2,\ 0 \leq z \leq 1 - x/2 - y\right\}.$$

We can now compute the volume of $T$:

$$\begin{aligned}
\text{vol.}(T) &= \iiint_T 1\,dV = \int_0^{2/5} \int_{2x}^{1-x/2} \int_0^{1-x/2-y} 1\,dzdydx \\
&= \int_0^{2/5} \int_{2x}^{1-x/2} (1 - x/2 - y)\,dydx = \int_0^{2/5} \left[(1 - x/2)y - \tfrac{1}{2}y^2\right]_{2x}^{1-x/2} dx \\
&= \int_0^{2/5} \left((1 - x/2)^2 - \tfrac{1}{2}(1 - x/2)^2 - 2x(1 - x/2) + 2x^2\right) dx \\
&= \int_0^{2/5} \left(\tfrac{1}{2} - \tfrac{5}{2}x + \tfrac{25}{8}x^2\right) dx = \int_0^1 \left(\tfrac{1}{2} - u + \tfrac{1}{2}u^2\right)\left(\tfrac{2}{5}du\right) = \frac{1}{15}.
\end{aligned}$$

□

## 16.4. Double and triple integrals in *Mathematica*

*Mathematica*'s built-in commands `Integrate` and `NIntegrate` can evaluate not only definite integrals in one variable, but also double, triple, and even higher-dimensional integrals. There are two main ways to use these commands to evaluate double and triple integrals.

Suppose first that the integral can be easily converted to an iterated one: say,

$$\int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dydx \qquad \text{or} \qquad \int_a^b \int_{g_1(z)}^{g_2(z)} \int_{h_1(y,z)}^{h_2(y,z)} f(x,y,z)\,dxdydz.$$

If the integrations are within *Mathematica*'s reach, we can use `Integrate` as follows

```
Integrate[f[x,y], {x,a,b}, {y,g1[x],g2[x]}]
Integrate[f[x,y,z], {z,a,b}, {y,g1[z],g2[z]}, {x,h1[y,z],h2[y,z]}]
```

If the functions are such that *Mathematica* cannot evaluate the integrals exactly, we can use `NIntegrate` in a similar fashion to obtain a numeric approximation. Here are some examples:

```
 In[1]:= Integrate[x*y^2, {x,0,1}, {y,0,1}]
Out[1]:= 1/6

 In[2]:= Integrate[x*y^2, {x,0,1}, {y,0,x}]
Out[2]:= 1/15

 In[3]:= Integrate[y*z*Cos[x^5], {x,0,1}, {y,0,x}, {z,x,2x}]
Out[3]:= 3Sin[1]/20

 In[4]:= Integrate[Sin[x+y]/(x+y), {x,1,2}, {y,1,2}]
Out[4]:= Cos[2] - 2Cos[3] + Cos[4] + 2SinIntegral[2]
         - 6SinIntegral[3] + 4SinIntegral[4]

 In[5]:= NIntegrate[Sin[x+y]/(x+y), {x,1,2}, {y,1,2}]
Out[5]:= 0.0619179
```

Next, suppose that we are interested in an integral over two- or three-dimensional domain $D$ that is described by one or more inequalities. In such situations, we use a familiar trick: we replace the original function $f$ by another which is zero outside $D$. *Mathematica* provides two commands to execute this trick: `If[]` and `Boole[]`. For example, suppose that we want to integrate the function $x^2 - xy + y^2$ over the region $D$ above the $x$-axis and the line $y = -x$ and inside the ellipse $2x^2 + y^2 = 3$ (see Figure 16.1). To define the function

$$f(x, y) = \begin{cases} x^2 - xy + y^2 & \text{if } (x, y) \in D, \\ 0 & \text{if } (x, y) \notin D, \end{cases}$$

we can use either of the following two *Mathematica* syntaxes:

```
f[x_,y_] := If[2x^2+y^2<=3 && x+y>=0 && y>=0, x^2-x*y+y^2, 0]
f[x_,y_] := (x^2-x*y+y^2)Boole[2x^2+y^2<=3 && x+y>=0 && y>=0]
```

Thus, to evaluate $\iint_D (x^2 - xy + y^2)\, dA$, we can use (the condition $y \geq 0$ is incorporated in the range of $y$)

```
Integrate[(x^2-x*y+y^2)Boole[2x^2+y^2<=3 && x+y>=0], {x,-1,2}, {y,0,2}]
```

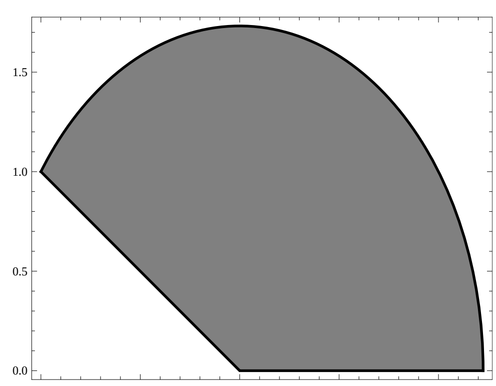



FIGURE 16.1. The region above $y = 0$ and $y = -x$ and insider $2x^2 + y^2 = 3$

Here are some examples:

```
 In[6]:= Integrate[(x*y+y^2)Boole[x^2+y^2<1], {x,-1,1}, {y,-1,1}]
Out[6]:= Pi/4

 In[7]:= Integrate[x*y*z*Boole[x^2+2y^2+z^2<4],
  {x,0,2}, {y,0,2}, {z,0,2}]
Out[7]:= 2/3

 In[8]:= Integrate[x*y*z*Boole[x+2y^2+3z^3<2],
  {x,0,2}, {y,0,2}, {z,0,2}]
Out[8]:= "Some gibberish"

 In[9]:= NIntegrate[x*y*z*Boole[x+2y^2+3z^3<2],
  {x,0,2}, {y,0,2}, {z,0,2}]
Out[9]:= 0.0468292
```

## Exercises

Evaluate the given iterated integral.

16.1. $\displaystyle\int_0^1 \int_{\sqrt{x}}^1 \int_{x+y}^{xy+1} 6xyz\, dzdydx$

16.2. $\displaystyle\int_0^2 \int_0^z \int_0^{yz} 4e^{-z^4}\, dxdydz$

Evaluate the given triple integral.

16.3. $\displaystyle\iiint_S \left(2x^2+3y+2z\right) dV,\ S = \left\{(x,y,z) : 0 \le x \le 1, 0 \le y \le x, y \le z \le 2y\right\}$

16.4. $\displaystyle\iiint_S e^{-y}\, dV,\ S = \left\{(x,y,z) : 1 \le z \le 2, 0 \le x \le 1/z, 0 \le y \le xz\right\}$

16.5. $\displaystyle\iiint_S xy\, dV$, $S$ lies below the plane $z = x + y + 2$ and above the right half of the disk $x^2 + y^2 \le 4$

16.6. $\displaystyle\iiint_S 4xz\, dV$, $S$ is the tetrahedron bounded by the plane $3x + 2y + z = 6$ and the coordinate planes

16.7. $\displaystyle\iiint_S x\cos z\, dV$, $S$ is the solid bounded by the cylinder $y = x^2$ and the planes $z = 0$, $y = 0$, $z = y$, $x = 1$

16.8. $\displaystyle\iiint_S y\left(x^2+z^2\right) dV$, $S$ is the solid bounded by the cylinder $x^2 + z^2 = 2$ and the planes $y = 1$ and $y = 3$

16.9. Use a triple integral to find the volume of the solid bounded by the cylinder $2y = x^2$ and the planes $z = -1$, $z = \frac{1}{2}x$, and $y = 2$.

LECTURE 17

# Triple Integrals in Cylindrical and Spherical Coordinates

## 17.1. Triple integrals in cylindrical coordinates

THEOREM 17.1. *Let $f(x, y, z)$ be piecewise continuous in a solid region $S$ in (Cartesian) space and let $R$ be the same region in cylindrical coordinates, that is,*

$$R = \{(r, \theta, z) : (r\cos\theta, r\sin\theta, z) \in S\}.$$

*Then*

$$\iiint_S f(x, y, z)\, dV = \iiint_R f(r\cos\theta, r\sin\theta, z) r\, dV. \tag{17.1}$$

The target applications of this theorem are triple integrals over type I regions

$$S = \{(x, y, z) : (x, y) \in D,\ u_1(x, y) \le z \le u_2(x, y)\},$$

where $D$ is a two-dimensional region with a simple polar form. In such situations, we can appeal to (16.3) to get

$$\iiint_S f(x, y, z)\, dV = \iint_D \left[ \int_{u_1(x,y)}^{u_2(x,y)} f(x, y, z)\, dz \right] dA.$$

Then, after evaluating the integral over $z$, we can try to evaluate the resulting double integral using polar coordinates in the $xy$-plane. What Theorem 17.1 does is change coordinates in the $xy$-plane from Cartesian to polar before the integration over $z$ has been performed.

EXAMPLE 17.1. Find the volume of the solid $S$ that lies within the cylinder $x^2 + y^2 = 4$ and the sphere $x^2 + y^2 + z^2 = 9$.

SOLUTION. The given solid lies within the cylinder, below the upper half of the sphere, and above the lower half of the sphere (see Figure 17.1). Since the upper and lower hemispheres are the respective graphs of the functions

$$z = \sqrt{9 - x^2 - y^2} \qquad \text{and} \qquad z = -\sqrt{9 - x^2 - y^2},$$

we can write $S$ as a type I region:

$$S = \left\{(x, y, z) : x^2 + y^2 \le 4,\ -\sqrt{9 - x^2 - y^2} \le z \le \sqrt{9 - x^2 - y^2}\right\}.$$

The cylindrical coordinates description of the solid then is

$$R = \left\{(r, \theta, z) : 0 \le r \le 2,\ -\sqrt{9 - r^2} \le z \le \sqrt{9 - r^2}\right\}.$$

Therefore, by (17.1),

$$\text{vol.}(S) = \iiint_S 1\, dV = \iiint_R r\, dV = \iint_D \left[ \int_{-\sqrt{9-r^2}}^{\sqrt{9-r^2}} r\, dz \right] dA = \iint_D 2r\sqrt{9 - r^2}\, dA,$$

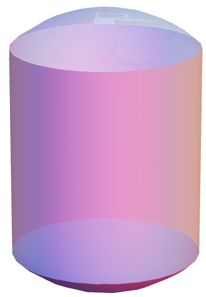


FIGURE 17.1. The solid $S$ in Example 17.1

where $D$ is the polar rectangle (that is, a disk) $D = [0, 2] \times [0, 2\pi]$. Applying Fubini's theorem, we get

$$\iint_D 2r\sqrt{9 - r^2}\, dA = \int_0^2 \int_0^{2\pi} 2r\sqrt{9 - r^2}\, d\theta dr = \int_0^2 4\pi r\sqrt{9 - r^2}\, dr.$$

We now make a substitution $u = 9 - r^2$. We have

$$du = -2r\, dr, \qquad u(0) = 9, \qquad u(2) = 5,$$

so the volume of $S$ is

$$\int_0^2 4\pi r\sqrt{9 - r^2}\, dr = \int_9^5 2\pi u^{1/2}(-du) = \left[\frac{2\pi u^{3/2}}{3/2}\right]_5^9 = \frac{4\pi(27 - 5\sqrt{5})}{3} \approx 66.265.$$

□

EXAMPLE 17.2. Evaluate

$$\iiint_B \sqrt{x^2 + y^2}\, dV,$$

where $B$ is the ball $x^2 + y^2 + z^2 \le 4$.

SOLUTION. In cylindrical coordinates, the ball is

$$\begin{aligned} R &= \left\{(r, \theta, z) : r^2 + z^2 \le 4\right\} \\ &= \left\{(r, \theta, z) : 0 \le r \le 2,\ -\sqrt{4 - r^2} \le z \le \sqrt{4 - r^2}\right\}. \end{aligned}$$

Thus, by (17.1) and (16.3),

$$\begin{aligned} \iiint_B \sqrt{x^2 + y^2}\, dV = \iiint_R r\, r\, dV = \iint_D \left[\int_{-\sqrt{4-r^2}}^{\sqrt{4-r^2}} r^2\, dz\right] dA = \iint_D 2r^2\sqrt{4 - r^2}\, dA \\ = \int_0^2 \int_0^{2\pi} 2r^2\sqrt{4 - r^2}\, d\theta dr = 2\pi \int_0^2 2r^2\sqrt{4 - r^2}\, dr. \end{aligned}$$

We now substitute $r = 2\sin t$. Since

$$\sqrt{4 - r^2} = \sqrt{4 - 4\sin^2 t} = 2\cos t, \qquad dr = (2\cos t)dt, \qquad t(0) = 0, \qquad t(2) = \pi/2,$$

we get

$$
\begin{aligned}
2\pi \int_0^2 2r^2 \sqrt{4-r^2}\, dr &= 4\pi \int_0^{\pi/2} (2\sin t)^2 (2\cos t)\,(2\cos t) dt \\
&= 16\pi \int_0^{\pi/2} (2\sin t\cos t)^2\, dt = 16\pi \int_0^{\pi/2} (\sin 2t)^2\, dt \\
&= 8\pi \int_0^{\pi/2} (1-\cos 4t)\, dt = 8\pi \big[t - \tfrac{1}{4}\sin 4t\big]_0^{\pi/2} = 4\pi^2.
\end{aligned}
$$

□

## 17.2. Triple integrals in spherical coordinates

THEOREM 17.2. *Let $f(x, y, z)$ be piecewise continuous in a solid region $S$ in (Cartesian) space and let $R$ be the same region in spherical coordinates, that is,*

$$R = \big\{(\rho, \theta, \phi) : (\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi) \in S\big\}.$$

*Then*

$$\iiint_S f(x, y, z)\, dV = \iiint_R f(\rho\cos\theta\sin\phi, \rho\sin\theta\sin\phi, \rho\cos\phi)\rho^2 \sin\phi\, dV. \tag{17.2}$$

EXAMPLE 17.3. Evaluate

$$\iiint_B \sqrt{x^2+y^2+z^2}\, e^{-x^2-y^2-z^2}\, dV,$$

where $B$ is the ball $x^2 + y^2 + z^2 \le a^2$.

SOLUTION. In spherical coordinates, $x^2 + y^2 + z^2 = \rho^2$ and

$$B = \big\{(\rho, \theta, \phi) : \rho \le a\big\} = [0, a] \times [0, 2\pi] \times [0, \pi],$$

so (17.2) yields

$$
\begin{aligned}
\iiint_B \sqrt{x^2+y^2+z^2}\, e^{-x^2-y^2-z^2}\, dV &= \int_0^a \int_0^{2\pi} \int_0^{\pi} \rho e^{-\rho^2} \rho^2 \sin\phi\, d\phi d\theta d\rho \\
&= \int_0^a \int_0^{2\pi} \rho^3 e^{-\rho^2} \big[-\cos\phi\big]_0^{\pi} d\theta d\rho \\
&= \int_0^a \int_0^{2\pi} 2\rho^3 e^{-\rho^2}\, d\theta d\rho = 4\pi \int_0^a \rho^3 e^{-\rho^2}\, d\rho.
\end{aligned}
$$

We substitute $u = \rho^2$ in the integral over $\rho$ and get

$$
\begin{aligned}
\iiint_B \sqrt{x^2+y^2+z^2}\, e^{-x^2-y^2-z^2}\, dV &= 2\pi \int_0^{a^2} u e^{-u}\, du = 2\pi \int_0^{a^2} u\, d\big(-e^{-u}\big) \\
&= \big[-2\pi u e^{-u}\big]_0^{a^2} + 2\pi \int_0^{a^2} e^{-u}\, du = 2\pi\big(1 - e^{-a^2} - a^2 e^{-a^2}\big).
\end{aligned}
$$

□

Example 17.4. Evaluate

$$\iiint_S y^2 z\, dV,$$

where $S$ is the solid that lies between the spheres $\rho = 1$ and $\rho = 2$ and below the cone $\phi = \pi/4$ (see Figure 17.2).

Solution. In spherical coordinates, the conditions describing $S$ can be expressed as

$$1 \le \rho \le 2, \qquad \pi/4 \le \phi \le \pi,$$

so in spherical coordinates $S$ is the rectangular box $[1,2] \times [0,2\pi] \times [\pi/4, \pi]$. Then, by (17.2),

$$\begin{aligned}
\iiint_S y^2 z\, dV &= \int_1^2 \int_0^{2\pi} \int_{\pi/4}^{\pi} (\rho \sin\phi \sin\theta)^2 (\rho\cos\phi)\rho^2 \sin\phi \, d\phi d\theta d\rho \\
&= \int_0^{2\pi} \int_{\pi/4}^{\pi} \left[ \sin^3\phi \cos\phi \sin^2\theta \int_1^2 \rho^5 \, d\rho \right] d\phi d\theta \\
&= \int_0^{2\pi} \int_{\pi/4}^{\pi} \sin^3\phi \cos\phi \sin^2\theta \left[ \tfrac{1}{6}\rho^6 \right]_1^2 d\phi d\theta \\
&= \frac{21}{2} \int_0^{2\pi} \left[ \sin^2\theta \int_{\pi/4}^{\pi} \sin^3\phi \cos\phi \, d\phi \right] d\theta.
\end{aligned}$$

We substitute $u = \sin\phi$ in the integral over $\phi$ and get

$$\int_{\pi/4}^{\pi} \sin^3\phi \cos\phi \, d\phi = \int_{1/\sqrt{2}}^{0} u^3 \, du = -\left[ \frac{u^4}{4} \right]_0^{1/\sqrt{2}} = -\frac{1}{16}.$$

Hence,

$$\iiint_S y^2 z\, dV = -\frac{21}{32} \int_0^{2\pi} \sin^2\theta \, d\theta = -\frac{21}{64} \int_0^{2\pi} (1 - \cos 2\theta)\, d\theta = -\frac{21\pi}{32}.$$

□

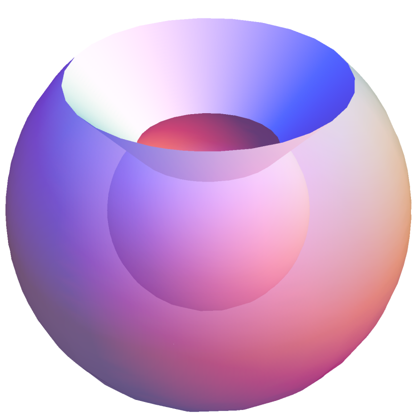


Figure 17.2. The solid in Example 17.4

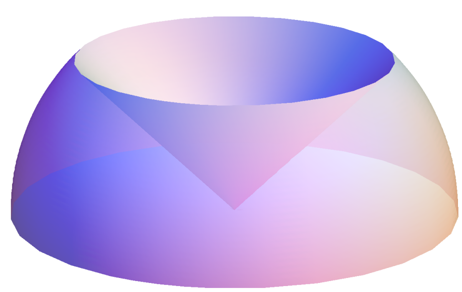


Figure 17.3. The solid in Example 17.5

EXAMPLE 17.5. Find the volume of the solid $S$ that lies inside the sphere $x^2 + y^2 + z^2 = 4$, above the $xy$-plane, and below the cone $z = \sqrt{x^2 + y^2}$.

SOLUTION. We first derive the spherical description of $S$. In Cartesian coordinates, the solid is defined by the inequalities

$$x^2 + y^2 + z^2 \leq 4, \quad 0 \leq z \leq \sqrt{x^2 + y^2}.$$

In spherical coordinates, these become

$$\rho^2 \leq 4, \quad 0 \leq \rho\cos\phi \leq \rho\sin\phi,$$

or equivalently,

$$0 \leq \rho \leq 2, \quad 0 \leq \cos\phi \leq \sin\phi.$$

The conditions on $\phi$ restrict it to the range $\pi/4 \leq \phi \leq \pi/2$, so the spherical description of $S$ is $[0, 2] \times [0, 2\pi] \times [\pi/4, \pi/2]$. The solid is displayed at Figure 17.3. We now use Theorem 16.1 to express the volume of $S$ as a triple integral and then appeal to (17.2) to convert the triple integral to spherical coordinates. We obtain

$$\begin{aligned} \text{vol.}(S) &= \iiint_S 1\,dV = \int_{\pi/4}^{\pi/2}\int_0^{2\pi}\int_0^2 \rho^2 \sin\phi\, d\rho d\theta d\phi \\ &= \frac{8}{3}\int_{\pi/4}^{\pi/2}\int_0^{2\pi} \sin\phi\, d\theta d\phi = \frac{16\pi}{3}\int_{\pi/4}^{\pi/2} \sin\phi\, d\phi \\ &= \frac{16\pi}{3}\big[-\cos\phi\big]_{\pi/4}^{\pi/2} = \frac{8\sqrt{2}\pi}{3}. \end{aligned}$$

□

## Exercises

Evaluate the given integral by changing to cylindrical coordinates.

17.1. $\iiint_S \sqrt{x^2 + y^2}\, dV$, $S$ is bounded by the $xy$-plane, the cylinder $x^2 + y^2 = 1$, and by the paraboloid $z = 4 - x^2 - y^2$

17.2. $\iiint_S y\, dV$, $S$ lies to the right of the $xz$-plane and is bounded by the cylinders $x^2 + y^2 = 1$ and $x^2 + y^2 = 9$ and by the planes $z = 1$ and $z = 4 + x$

17.3. $\iiint_S 1\, dV$, $S$ is the smaller solid bounded by the plane $z = 1$ and the sphere $x^2 + y^2 + z^2 = 2$

Evaluate the given integral by changing to spherical coordinates.

17.4. $\iiint_S \sqrt{x^2 + y^2}\, dV$, $S$ lies above the $xy$-plane and within the sphere $x^2 + y^2 + z^2 = 4$

17.5. $\iiint_S xz\, dV$, $S$ lies between the spheres $x^2 + y^2 + z^2 = a^2$ and $x^2 + y^2 + z^2 = 2a^2$ in the first octant

17.6. $\iiint_S 1\, dV$, $S$ lies above the cone $z^2 = x^2 + y^2$ and below the sphere $x^2 + y^2 + (z - 2)^2 = 4$

17.7. $\iiint_S (xy + z)\, dV$, $S$ lies above the cone $z = \sqrt{x^2 + y^2}$ and below the sphere $x^2 + y^2 + z^2 = 4$

Evaluate the given iterated integral by changing to cylindrical or spherical coordinates.

17.8. $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{x^2+y^2}^{2-x^2-y^2} (x^2 + y^2)^{3/2}\, dzdydx$

17.9. $\displaystyle\int_0^2 \int_0^{\sqrt{4-y^2}} \int_0^{\sqrt{4-x^2-y^2}} z\, dzdxdy$

17.10. $\displaystyle\int_0^2 \int_0^{\sqrt{4-x^2}} \int_{\sqrt{4-x^2-y^2}}^{\sqrt{9-x^2-y^2}} xyz\, dzdydx$

Use a triple integral to find the volume of the given solid.

17.11. The solid bounded by the paraboloid $y = -x^2 - z^2$ and the plane $y = -9$.

17.12. The smaller solid bounded by the plane $z = 1$ and the sphere $x^2 + y^2 + z^2 = 2$.

17.13. The solid above the cone $\phi = \pi/6$ and below the sphere $\rho = 4\cos\phi$.

LECTURE 18

# Applications of Double and Triple Integrals

We have already seen one application of the double and triple integrals: both can be used to compute volumes of solids. In this lecture, we describe some other applications.

## 18.1. Surface area

Let $\Sigma$ be a smooth parametric surface, given by

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v) \in D).$$

We can use a double integral to calculate the surface area of $\Sigma$. However, we first should *define* what "surface area" is. For that, we draw inspiration from the definition of arc length in §5.6.

For simplicity, we shall assume that $D$ is a rectangle in the $uv$-plane, say: $[a,b] \times [c,d]$. We can treat the general case similarly, but there are more technical details to pay attention to and little extra insight to gain. For each $n \geq 1$, let $a = u_0 < u_1 < \cdots < u_n = b$ be the points that divide the interval $[a,b]$ into $n$ subintervals of equal lengths; and let $c = v_0 < v_1 < \cdots < v_n = d$ be the points that divide the interval $[c,d]$ into $n$ subintervals of equal lengths. The points

$$\begin{array}{cccccc}
(u_n,v_0) & (u_n,v_1) & \cdots & (u_n,v_j) & \cdots & (u_n,v_n), \\
\vdots & \vdots & & \vdots & & \vdots \\
(u_i,v_0) & (u_i,v_1) & \cdots & (u_i,v_j) & \cdots & (u_i,v_n), \\
\vdots & \vdots & & \vdots & & \vdots \\
(u_1,v_0) & (u_1,v_1) & \cdots & (u_1,v_j) & \cdots & (u_1,v_n), \\
(u_0,v_0) & (u_0,v_1) & \cdots & (u_0,v_j) & \cdots & (u_0,v_n),
\end{array}$$

define a partition of the rectangle $D$ into $n^2$ congruent subrectangles. The vector function $\mathbf{r}$ maps each of those small rectangles to a small "patch" on $\Sigma$ and each grid point $(u_i, v_j)$ to a point $P_{ij}$ on the surface. The patches form a partition of $\Sigma$ into $n^2$ pieces, which we label as $\Sigma_{ij}$, $1 \leq i, j \leq n$, so that $\Sigma_{ij}$ is the image of the rectangle that has $(u_i, v_j)$ as its upper right corner point. We can approximate the area of $\Sigma_{ij}$ by the area of the parallelogram $\Pi_{ij}$ three of whose vertices are the points $P_{i-1,j}$, $P_{i,j-1}$, and $P_{ij}$. Recall that the area of the parallelogram $\Pi_{ij}$ is given by

$$|\Pi_{ij}| = \big\|\overrightarrow{P_{ij}P}_{i-1,j} \times \overrightarrow{P_{ij}P}_{i,j-1}\big\|.$$

It follows that the total area $A_n$ of all the parallelograms $\Pi_{ij}$ is

$$A_n = \sum_{i=1}^{n}\sum_{j=1}^{n} |\Pi_{ij}| = \sum_{i=1}^{n}\sum_{j=1}^{n} \big\|\overrightarrow{P_{ij}P}_{i-1,j} \times \overrightarrow{P_{ij}P}_{i,j-1}\big\|.$$

We define the *surface area* of $\Sigma$ as the limit of the sequence $\{A_n\}_{n=1}^{\infty}$, if it exists. Using the above definition and the properties of differentiable functions and double integrals, we can prove the following formula for the surface area of a piecewise smooth surface.

THEOREM 18.1. *Let $\Sigma$ be a smooth parametric surface, given by*

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v) \in D).$$

*The surface area of $\Sigma$ is given by*

$$\text{area}(\Sigma) = \iint_D \|\mathbf{r}_u \times \mathbf{r}_v\| \, dA. \tag{18.1}$$

If $\Sigma$ is the graph of $z = f(x,y)$, $(x,y) \in D$, then using the parametrization

$$\mathbf{r}(x,y) = x\mathbf{e}_1 + y\mathbf{e}_2 + f(x,y)\mathbf{e}_3 \qquad ((x,y) \in D),$$

we can express the area of $\Sigma$ as

$$\text{area}(\Sigma) = \iint_D \sqrt{1 + f_x^2 + f_y^2} \, dA. \tag{18.2}$$

EXAMPLE 18.1. Find the area of the part of the plane $2x+5y+z = 10$ that lies inside the cylinder $x^2 + y^2 = 9$.

SOLUTION. We apply (18.2) with $f(x,y) = 10 - 2x - 5y$ and $D$ the disk $x^2 + y^2 \leq 9$. The desired area is then

$$A = \iint_D \sqrt{1 + (-2)^2 + (-5)^2} \, dA = \iint_D \sqrt{30} \, dA = \sqrt{30} \cdot 9\pi,$$

where $9\pi$ is the area of the disk $x^2 + y^2 \leq 9$. □

EXAMPLE 18.2. Find the area of the surface $\Sigma$ with parametrization

$$\mathbf{r}(u,v) = \cos u(4 + 2\cos v)\mathbf{e}_1 + \sin u(4 + 2\cos v)\mathbf{e}_2 + 2\sin v\mathbf{e}_3 \qquad (0 \leq u, v \leq 2\pi).$$

This surface is called a *torus* and resembles a donut; it is displayed on Figure 18.1

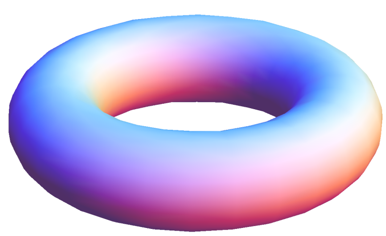


FIGURE 18.1. A torus

SOLUTION. We have

$$\mathbf{r}_u(u,v) = -\sin u(4 + 2\cos v)\mathbf{e}_1 + \cos u(4 + 2\cos v)\mathbf{e}_2,$$
$$\mathbf{r}_v(u,v) = -2\cos u \sin v\mathbf{e}_1 - 2\sin u \sin v\mathbf{e}_2 + 2\cos v\mathbf{e}_3,$$

so

$$\begin{aligned}
\mathbf{r}_u \times \mathbf{r}_v &= 2\sin^2 u \sin v(4+2\cos v)(\mathbf{e}_1 \times \mathbf{e}_2) - 2\sin u \cos v(4+2\cos v)(\mathbf{e}_1 \times \mathbf{e}_3) \\
&\quad - 2\cos^2 u \sin v(4+2\cos v)(\mathbf{e}_2 \times \mathbf{e}_1) + 2\cos u \cos v(4+2\cos v)(\mathbf{e}_2 \times \mathbf{e}_3) \\
&= 2\cos u \cos v(4+2\cos v)\mathbf{e}_1 + 2\sin u \cos v(4+2\cos v)\mathbf{e}_2 + 2\sin v(4+2\cos v)\mathbf{e}_3 \\
&= (8+4\cos v)\left(\cos u \cos v\mathbf{e}_1 + \sin u \cos v\mathbf{e}_2 + \sin v\mathbf{e}_3\right), \\
\|\mathbf{r}_u \times \mathbf{r}_v\| &= |8+4\cos v|\sqrt{\cos^2 u \cos^2 v + \sin^2 u \cos^2 v + \sin^2 v} = 8+4\cos v.
\end{aligned}$$

Thus, the surface area of $\Sigma$ is

$$\int_0^{2\pi}\int_0^{2\pi}(8+4\cos v)\,du dv = 2\pi\int_0^{2\pi}(8+4\cos v)\,dv = 32\pi^2.$$

□

Example 18.3. Express the area of the ellipsoid $\Sigma$ with equation $x^2+y^2+4z^2=1$ in terms of the number

$$\mu = \int_0^1 \sqrt{1+3w^2}\,dw = 1 + \frac{\ln(2+\sqrt{3})}{2\sqrt{3}}.$$

Solution. First, we have to parametrize $\Sigma$. The unit sphere $x^2+y^2+z^2=1$ is the set of all points $(x,y,z)$ whose spherical coordinate $\rho$ is 1:

$$x = \cos u \sin v, \qquad y = \sin u \sin v, \qquad z = \cos v,$$

where $0 \le u \le 2\pi$ and $0 \le v \le \pi$ ($u$ and $v$ play the roles of the spherical coordinates $\theta$ and $\phi$). Since $(x,y,z)$ lies on $\Sigma$ exactly when $(x,y,2z)$ lies on the unit sphere, we obtain the following parametrization for the ellipsoid:

$$\mathbf{r}(u,v) = \cos u \sin v\mathbf{e}_1 + \sin u \sin v\mathbf{e}_2 + \tfrac{1}{2}\cos v\mathbf{e}_3,$$

where $0 \le u \le 2\pi$ and $0 \le v \le \pi$. Then

$$\begin{aligned}
\mathbf{r}_u(u,v) &= -\sin u \sin v\mathbf{e}_1 + \cos u \sin v\mathbf{e}_2 + 0\mathbf{e}_3, \\
\mathbf{r}_v(u,v) &= \cos u \cos v\mathbf{e}_1 + \sin u \cos v\mathbf{e}_2 - \tfrac{1}{2}\sin v\mathbf{e}_3; \\
\mathbf{r}_u \times \mathbf{r}_v &= -\sin^2 u \sin v \cos v(\mathbf{e}_1 \times \mathbf{e}_2) + \tfrac{1}{2}\sin u \sin^2 v(\mathbf{e}_1 \times \mathbf{e}_3) \\
&\quad + \cos^2 u \sin v \cos v(\mathbf{e}_2 \times \mathbf{e}_1) - \tfrac{1}{2}\cos u \sin^2 v(\mathbf{e}_2 \times \mathbf{e}_3) \\
&= -\tfrac{1}{2}\cos u \sin^2 v\mathbf{e}_1 - \tfrac{1}{2}\sin u \sin^2 v\mathbf{e}_2 - \sin v \cos v\mathbf{e}_3 \\
&= -\tfrac{1}{2}\sin v\left(\cos u \sin v\mathbf{e}_1 + \sin u \sin v\mathbf{e}_2 + 2\cos v\mathbf{e}_3\right); \\
\|\mathbf{r}_u \times \mathbf{r}_v\| &= \tfrac{1}{2}|\sin v|\sqrt{\cos^2 u \sin^2 v + \sin^2 u \sin^2 v + 4\cos^2 v} \\
&= \tfrac{1}{2}\sin v\sqrt{\sin^2 v + 4\cos^2 v} = \tfrac{1}{2}\sin v\sqrt{1+3\cos^2 v}.
\end{aligned}$$

Thus, the surface area of $\Sigma$ is

$$\begin{aligned}
A &= \int_0^{\pi}\int_0^{2\pi}\tfrac{1}{2}\sin v\sqrt{1+3\cos^2 v}\,du dv = \pi\int_0^{\pi}\sin v\sqrt{1+3\cos^2 v}\,dv \\
&= \pi\int_1^{-1}\sqrt{1+3w^2}\,(-dw) = 2\pi\int_0^1\sqrt{1+3w^2}\,dw = 2\pi\mu.
\end{aligned}$$

□

EXAMPLE 18.4. Find the volume and the surface area of the solid bounded by the paraboloids $z = 3x^2 + 3y^2$ and $z = 4 - x^2 - y^2$.

SOLUTION. The given solid has the paraboloid $z = 4 - x^2 - y^2$ as its upper boundary and the paraboloid $z = 3x^2 + 3y^2$ as its lower boundary, so it can be described as

$$S = \{(x, y, z) : (x, y) \in D,\ 3x^2 + 3y^2 \le z \le 4 - x^2 - y^2\},$$

where $D$ is the projection of the solid on the $xy$-plane (see Figure 18.2). Since the intersection curve of the two paraboloids is

$$\begin{aligned} \gamma &= \{(x, y, z) : z = 3x^2 + 3y^2 = 4 - x^2 - y^2\} \\ &= \{(x, y, z) : x^2 + y^2 = 1,\ z = 3\}, \end{aligned}$$

we find that $D$ is the unit disk $x^2 + y^2 \le 1$. Hence, the volume of $E$ is

$$\text{vol.}(S) = \iiint_S 1\, dV = \iint_D \left[ \int_{3x^2+3y^2}^{4-x^2-y^2} 1\, dz \right] dA = \iint_D \left(4 - 4x^2 - 4y^2\right) dA.$$

The surface area $A$ of the solid is the sum of the surface areas of the two paraboloids, each of which can be computed using (18.2):

$$A = \iint_D \sqrt{1 + (-2x)^2 + (-2y)^2}\, dA + \iint_D \sqrt{1 + (6x)^2 + (6y)^2}\, dA.$$

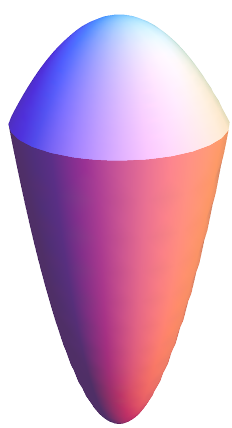


FIGURE 18.2. The solid in Example 18.4

Let $R = [0,1] \times [0,2\pi]$ be the polar form of the disk $D$. Converting all three integrals to polar coordinates, we obtain

$$\begin{aligned}
\text{vol.}(S) &= \iint_D \big(4 - 4(x^2+y^2)\big)\,dA = \iint_R r\big(4-4r^2\big)\,dA \\
&= \int_0^1 \left[\int_0^{2\pi} \big(4r - 4r^3\big)\,d\theta\right] dr = 2\pi \int_0^1 \big(4r-4r^3\big)\,dr = 2\pi\big[2r^2 - r^4\big]_0^1 = 2\pi, \\
A &= \iint_D \big(\sqrt{1+4(x^2+y^2)} + \sqrt{1+36(x^2+y^2)}\big)\,dA \\
&= \iint_R r\big(\sqrt{1+4r^2} + \sqrt{1+36r^2}\big)\,dA = \int_0^1 \left[\int_0^{2\pi} r\big(\sqrt{1+4r^2}+\sqrt{1+36r^2}\big)d\theta\right] dr \\
&= 2\pi \int_0^1 r\big(\sqrt{1+4r^2}+\sqrt{1+36r^2}\big)\,dr = 2\pi\left[\int_1^5 u^{1/2}\left(\tfrac{1}{8}du\right) + \int_1^{37} u^{1/2}\left(\tfrac{1}{72}du\right)\right] \\
&= 2\pi\left(\left[\frac{u^{3/2}}{12}\right]_1^5 + \left[\frac{u^{3/2}}{108}\right]_1^{37}\right) = \frac{\pi}{54}\big(37\sqrt{37} + 45\sqrt{5} - 10\big) \approx 18.366.
\end{aligned}$$

□

### 18.2. Density and mass

Double and triple integrals also represent various physical quantities. In this section, we discuss applications related to the computation of mass and the coordinates of the center of mass of laminas and solids. Suppose that a lamina occupies a region $D$ in the $xy$-plane and has *density* $\delta(x,y)$ at a point $(x,y) \in D$. If $\delta(x,y)$ is the constant function $\delta(x,y) = c$, the total mass of the lamina is $c|D|$, where $|D|$ denotes the area of $D$. Note that we can think of this number also as

$$\iint_D c\,dA = \iint_D \delta(x,y)\,dA.$$

Using the definition of the double integral as the limit of Riemann sums, we can prove that the latter formula applies to any continuous function $\delta(x,y)$. That is, if the density function $\delta$ is continuous in $D$, then the mass $m$ of the lamina is given by the formula

$$m = \iint_D \delta(x,y)\,dA.$$

We can also give formulas for the coordinates of the *center of mass* of the lamina: if $(\bar{x},\bar{y})$ is the center of mass, then

$$\bar{x} = \frac{1}{m}\iint_D x\delta(x,y)\,dA, \qquad \bar{y} = \frac{1}{m}\iint_D y\delta(x,y)\,dA.$$

Similarly, if a solid occupies a region $S$ in space and has *density* $\delta(x,y,z)$ at a point $(x,y,z) \in S$, then the mass $m$ of the solid is given by the integral

$$m = \iiint_S \delta(x,y,z)\,dV.$$

We can also give formulas for the coordinates of the *center of mass* of the solid: if $(\bar{x}, \bar{y}, \bar{z})$ is the center of mass of $S$, then

$$\bar{x} = \frac{1}{m} \iiint_S x\delta(x, y, z)\, dV, \quad \bar{y} = \frac{1}{m} \iiint_S y\delta(x, y, z)\, dV, \quad \bar{z} = \frac{1}{m} \iiint_S z\delta(x, y, z)\, dV.$$

Example 18.5. A flat lamina occupies the disk $x^2+y^2 \le 4$ and has density $\delta(x, y) = x+y+x^2+y^2$. Find its mass and the coordinates of its center of mass.

Solution. Let $R$ denote the polar form of the disk $D$, that is, $R = [0, 2] \times [0, 2\pi]$. By (15.1), the mass $m$ of the lamina is

$$\begin{aligned} m &= \iint_D \left(x + y + x^2 + y^2\right) dA = \iint_R r\left(r\cos\theta + r\sin\theta + r^2\right) dA \\ &= \int_0^{2\pi} \left[ \int_0^2 \left(r^2(\cos\theta + \sin\theta) + r^3\right) dr \right] d\theta = \int_0^{2\pi} \left(\tfrac{8}{3}(\cos\theta + \sin\theta) + 4\right) d\theta = 8\pi. \end{aligned}$$

Similarly, the coordinates of the center of mass are:

$$\begin{aligned} \bar{x} &= \frac{1}{m} \iint_D x(x + y + x^2 + y^2)\, dA \\ &= \frac{1}{8\pi} \iint_R r^2 \cos\theta\left(r\cos\theta + r\sin\theta + r^2\right) dA \\ &= \frac{1}{8\pi} \int_0^{2\pi} \left[ \int_0^2 \cos\theta\left(r^3(\cos\theta + \sin\theta) + r^4\right) dr \right] d\theta \\ &= \frac{1}{8\pi} \int_0^{2\pi} \cos\theta\left(4(\cos\theta + \sin\theta) + \tfrac{32}{5}\right) d\theta \\ &= \frac{1}{8\pi} \int_0^{2\pi} \left(2(1 + \cos 2\theta) + 2\sin 2\theta + \tfrac{32}{5}\cos\theta\right) d\theta = \frac{1}{2}, \\ \bar{y} &= \frac{1}{m} \iint_D y(x + y + x^2 + y^2)\, dA \\ &= \frac{1}{8\pi} \iint_R r^2 \sin\theta\left(r\cos\theta + r\sin\theta + r^2\right) dA \\ &= \frac{1}{8\pi} \int_0^{2\pi} \left[ \int_0^2 \sin\theta\left(r^3(\cos\theta + \sin\theta) + r^4\right) dr \right] d\theta \\ &= \frac{1}{8\pi} \int_0^{2\pi} \sin\theta\left(4(\cos\theta + \sin\theta) + \tfrac{32}{5}\right) d\theta \\ &= \frac{1}{8\pi} \int_0^{2\pi} \left(2\sin 2\theta + 2(1 - \cos 2\theta) + \tfrac{32}{5}\sin\theta\right) d\theta = \frac{1}{2}. \end{aligned}$$

□

Example 18.6. Find the center of mass of the solid $H$ that occupies the upper half of the ball $x^2 + y^2 + z^2 \le 4$ and has density $\delta(x, y, z) = z\sqrt{x^2 + y^2 + z^2}$.

Solution. The solid $H$ is described by inequalities

$$x^2 + y^2 + z^2 \le 4, \qquad z \ge 0.$$

In spherical coordinates these become $\rho \le 2$ and $0 \le \phi \le \pi/2$, respectively. Thus, by (17.2), the mass of the solid is

$$\begin{aligned} m &= \iiint_H z\sqrt{x^2+y^2+z^2}\,dV = \int_0^2\int_0^{2\pi}\int_0^{\pi/2}(\rho^2\cos\phi)(\rho^2\sin\phi)\,d\phi d\theta d\rho \\ &= \int_0^{\pi/2}\int_0^{2\pi}\sin\phi\cos\phi\left[\int_0^2\rho^4\,d\rho\right]d\theta d\phi = \frac{32}{5}\int_0^{\pi/2}\int_0^{2\pi}\sin\phi\cos\phi\,d\theta d\phi \\ &= \frac{64\pi}{5}\int_0^{\pi/2}\sin\phi\cos\phi\,d\phi = \frac{64\pi}{5}\int_0^1 u\,du = \frac{32\pi}{5}. \end{aligned}$$

The coordinates of the center of mass then are

$$\begin{aligned} \bar{x} &= \frac{1}{m}\iiint_H xz\sqrt{x^2+y^2+z^2}\,dV \\ &= \frac{5}{32\pi}\int_0^2\int_0^{2\pi}\int_0^{\pi/2}(\rho^3\cos\theta\sin\phi\cos\phi)(\rho^2\sin\phi)\,d\phi d\theta d\rho \\ &= \frac{5}{32\pi}\int_0^2\int_0^{\pi/2}\rho^5\sin^2\phi\cos\phi\left[\int_0^{2\pi}\cos\theta\,d\theta\right]d\phi d\rho = 0, \\ \bar{y} &= \frac{1}{m}\iiint_H yz\sqrt{x^2+y^2+z^2}\,dV \\ &= \frac{5}{32\pi}\int_0^2\int_0^{2\pi}\int_0^{\pi/2}(\rho^3\sin\theta\sin\phi\cos\phi)(\rho^2\sin\phi)\,d\phi d\theta d\rho \\ &= \frac{5}{32\pi}\int_0^2\int_0^{\pi/2}\rho^5\sin^2\phi\cos\phi\left[\int_0^{2\pi}\sin\theta\,d\theta\right]d\phi d\rho = 0, \\ \bar{z} &= \frac{1}{m}\iiint_H z^2\sqrt{x^2+y^2+z^2}\,dV \\ &= \frac{5}{32\pi}\int_0^2\int_0^{2\pi}\int_0^{\pi/2}(\rho^3\cos^2\phi)(\rho^2\sin\phi)\,d\phi d\theta d\rho \\ &= \frac{5}{32\pi}\int_0^{\pi/2}\int_0^{2\pi}\sin\phi\cos^2\phi\left[\int_0^2\rho^5\,d\rho\right]d\theta d\phi \\ &= \frac{5}{3\pi}\int_0^{\pi/2}\int_0^{2\pi}\sin\phi\cos^2\phi\,d\theta d\phi \\ &= \frac{10}{3}\int_0^{\pi/2}\sin\phi\cos^2\phi\,d\phi = \frac{10}{3}\int_1^0 u^2\,(-du) = \frac{10}{9}. \end{aligned}$$

□

## 18.3. An improper integral*

In this section, we use double integrals and polar coordinates to show that

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}.$$

Proof. Let $D_b$ denote the disk $x^2 + y^2 \leq b^2$ and let $S_b$ denote the square with vertices $(\pm b, \pm b)$ (see Figure 18.3). As $b \to \infty$, both $D_b$ and $S_b$ expand to fill the entire $xy$-plane, so we figure that

$$\lim_{b\to\infty} \iint_{D_b} e^{-x^2-y^2}\, dA = \lim_{b\to\infty} \iint_{S_b} e^{-x^2-y^2}\, dA = \iint_{\mathbb{R}^2} e^{-x^2-y^2}\, dA. \tag{18.3}$$

(This can be proved rigorously, but it is also intuitively clear, so we won't dwell on it.) We now consider each limit separately.

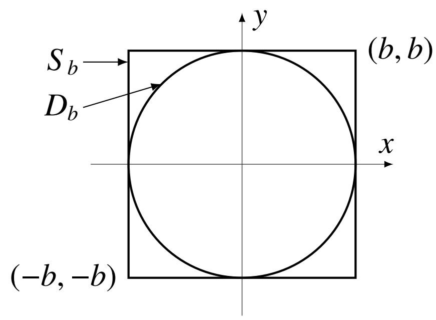



Figure 18.3. The disk $D_b$ and the square $S_b$

First, we consider the integral over the disk $D_b$. In polar coordinates, $D_b$ is $[0, b] \times [0, 2\pi]$, so

$$\iint_{D_b} e^{-x^2-y^2}\, dA = \int_0^{2\pi} \left[ \int_0^b r e^{-r^2}\, dr \right] d\theta = \int_0^{2\pi} \left[ \frac{1}{2} \int_0^{b^2} e^{-u}\, du \right] d\theta = \pi\left(1 - e^{-b^2}\right).$$

Taking limits as $b \to \infty$, we deduce that

$$\lim_{b\to\infty} \iint_{D_b} e^{-x^2-y^2}\, dA = \lim_{b\to\infty} \pi\left(1 - e^{-b^2}\right) = \pi. \tag{18.4}$$

On the other hand, if we define

$$I(b) = \int_{-b}^{b} e^{-x^2}\, dx,$$

we have

$$\iint_{S_b} e^{-x^2-y^2}\, dA = \int_{-b}^{b} \left[ \int_{-b}^{b} e^{-x^2-y^2}\, dx \right] dy = \int_{-b}^{b} I(b) e^{-y^2}\, dy = I(b)I(b) = I(b)^2.$$

Furthermore, since the function $e^{-x^2}$ is even, we have

$$I(b) = \int_{-b}^{b} e^{-x^2}\, dx = 2 \int_0^b e^{-x^2}\, dx.$$

Hence,

$$\lim_{b\to\infty} \iint_{S_b} e^{-x^2-y^2}\, dA = \lim_{b\to\infty} \left( 2 \int_0^b e^{-x^2}\, dx \right)^2 = 4 \left( \int_0^\infty e^{-x^2}\, dx \right)^2. \tag{18.5}$$

Combining (18.3)–(18.5), we obtain that

$$4 \left( \int_0^\infty e^{-x^2}\, dx \right)^2 = \pi,$$

which is equivalent to desired conclusion. □

REMARK. Note that in the above proof we used that the limit

$$\lim_{b\to\infty}\int_0^b e^{-x^2}dx$$

exists, or equivalently, that the improper integral $\int_0^\infty e^{-x^2}\,dx$ converges. This can be proved using the comparison test for improper integrals.

## Exercises

Find the area of the given surface $\Sigma$.

18.1. $\Sigma$ is the part of the plane $x + 2y + z = 3$ that lies in the first octant
18.2. $\Sigma$ is the part of the paraboloid $z = x^2 - y^2$ that lies between the cylinders $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$
18.3. $\Sigma$ is the part of the sphere $x^2 + y^2 + z^2 = 4$ that lies inside the cylinder $x^2 + y^2 = 2$
18.4. $\Sigma$ is the part of the surface $z = xy$ that lies inside the cylinder $x^2 + y^2 = 1$
18.5. $\Sigma$ is the parametric surface given by $\mathbf{r}(u, v) = (uv)\mathbf{e}_1 + (u + v)\mathbf{e}_2 + (u - v)\mathbf{e}_3$, $u^2 + v^2 \le 1$
18.6. $\Sigma$ is the parametric surface given by $\mathbf{r}(u, v) = (u\cos v)\mathbf{e}_1 + (u\sin v)\mathbf{e}_2 + v\mathbf{e}_3$, $0 \le u \le 1, 0 \le v \le \pi$

Find the mass and the center of mass of the lamina that occupies the region $D$ and has density given by $\delta(x, y)$.

18.7. $D$ is bounded by the curve $y = e^x$, the line $x = 1$, and the coordinate axes; $\delta(x, y) = y$
18.8. $D$ is the disk $x^2 + y^2 \le 2y$; $\delta(x, y)$ is twice the distance from $(x, y)$ to the origin

Find the mass and the center of mass of the given solid $S$ if its density is given by $\delta(x, y, z)$.

18.9. $S$ is the cube $0 \le x, y, z \le 2$; $\delta(x, y, z) = x^2 + y^2 + z^2$
18.10. $S$ is the hemisphere $x^2 + y^2 + z^2 \le 4$, $z \ge 0$; $\delta(x, y, z)$ is half the distance from $(x, y, z)$ to the $z$-axis
18.11. $S$ is the hemisphere $x^2 + y^2 + z^2 \le 1$, $z \ge 0$; $\delta(x, y, z) = 4z$

18.12. Use the value of the integral computed in §18.3 to evaluate the integrals

$$I_n = \int_0^\infty x^{2n}e^{-x^2}\,dx \qquad (n = 1, 2, 3, \ldots).$$

18.13. Use polar coordinates and the value of the integral computed in §18.3 to evaluate

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |ax + by|e^{-(x^2+y^2)/2}\,dA.$$

LECTURE 19

# Change of Variables in Multiple Integrals*

In this lecture, we describe the general theory behind the transformations of a double integral from Cartesian to polar coordinates and of a triple integral from Cartesian to cylindrical or spherical coordinates. We also show how the "$u$-substitution" from single-variable calculus fits within the same general theory.

## 19.1. Transformations of $\mathbb{R}^2$ and $\mathbb{R}^3$

Definition. Let $D$ and $R$ be regions in $\mathbb{R}^2$. A function $T : D \to R$ is called a *transformation of D onto R* if $R = T(D)$, the range of $T$. That is, a transformation of $D$ is given by a pair of functions

$$x = f(u, v), \quad y = g(u, v) \qquad ((u, v) \in D) \tag{19.1}$$

such that each point $(x, y) \in R$ satisfies $(x, y) = T(u, v)$ for some $(u, v) \in D$. The transformation $T$ is called *one-to-one* if for any choice of $(x, y) \in R$ the equations (19.1) have exactly one solution $(u, v)$.

Let $D$ and $R$ be solid regions in $\mathbb{R}^3$. A function $T : D \to R$ is called a *transformation of D onto R* if $R = T(D)$, the range of $T$. That is, a transformation of $D$ is given by a triple of functions

$$x = f(u, v, w), \quad y = g(u, v, w), \quad z = h(u, v, w) \qquad ((u, v, w) \in D) \tag{19.2}$$

such that each point $(x, y, z) \in R$ satisfies $(x, y, z) = T(u, v, w)$ for some $(u, v, w) \in D$. The transformation $T$ is called *one-to-one* if for any choice of $(x, y, z) \in R$ the equations (19.2) have exactly one solution $(u, v, w)$.

Example 19.1. The functions

$$x = u + v, \quad y = 2u - 3v \qquad ((u, v) \in \mathbb{R}^2)$$

define a transformation of $\mathbb{R}^2$ onto itself. This transformation is one-to-one. Indeed, if we solve the above equations for $u$ and $v$, we find that

$$u = 0.6x + 0.2y, \quad v = 0.4x - 0.2y.$$

Hence, in this case, equations (19.1) have exactly one solution for every choice of $(x, y)$. □

Example 19.2. The functions

$$x = u \cos v, \quad y = u \sin v \qquad ((u, v) \in \mathbb{R}^2) \tag{19.3}$$

define another transformation of $\mathbb{R}^2$ onto itself. However, this transformation is not one-to-one, because for any given point $(x, y)$, we can find multiple solutions $(u, v)$ of the above equations. Indeed, let $r = \sqrt{x^2 + y^2}$ and $\theta \in [0, 2\pi)$ be the polar coordinates of the point $(x, y)$. Then $(u, v) = (r, \theta)$, $(u, v) = (r, \theta + 4\pi)$, and $(u, v) = (-r, \theta + \pi)$ are three different solutions of (19.3).

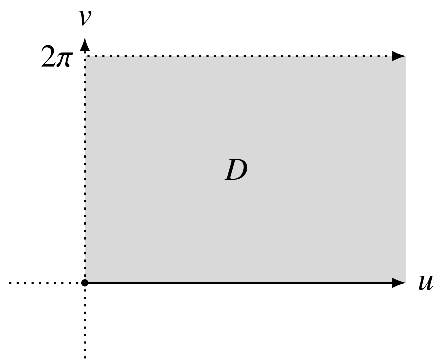



FIGURE 19.1. Polar coordinates as the domain of a one-to-one transformation

On the other hand, if we restrict $(u, v)$ to the set (see Figure 19.1: on that figure, the origin is part of the shaded area but the rest of the $v$-axis is not)

$$D = \left\{(u, v) \in \mathbb{R}^2 : u > 0, 0 \leq v < 2\pi\right\} \cup \{(0, 0)\},$$

then the functions in (19.3) do define a one-to-one transformation of $D$ onto $\mathbb{R}^2$. Indeed, for any $(x, y) \neq (0, 0)$, the only solution of (19.3) that lies in $D$ are the (unique) polar coordinates $(r, \theta)$ of the point $(x, y)$. If $(x, y) = (0, 0)$, then any solution of (19.3) must have $u = 0$, and the only such point in $D$ is the origin. □

EXAMPLE 19.3. The functions

$$x = u \cos v \sin w, \quad y = u \sin v \sin w, \quad z = u \cos w \tag{19.4}$$

define a transformation $T$ of the solid region (the set of spherical coordinates $(u, v, w)$ of points in space)

$$D = \left\{(u, v, w) \in \mathbb{R}^3 : u \geq 0, 0 \leq v \leq 2\pi, 0 \leq w \leq \pi\right\}$$

onto $\mathbb{R}^3$. Note that this transformation is not one-to-one, because any point $(0, v, w)$ is a solution of (19.4) when $x = y = z = 0$ and any point $(z, v, 0)$ is a solution of (19.4) when $x = y = 0$ and $z > 0$. The trasformation is one-to-one on the region

$$D' = \left\{(u, v, w) \in \mathbb{R}^3 : u > 0, 0 < v < 2\pi, 0 < w < \pi\right\},$$

but it is then not a transformation onto $\mathbb{R}^3$. Indeed, the points on the half-plane $y = 0$, $x \geq 0$ are then not part of the range of $T$; any other point in space has unique spherical coordinates that belong to $D'$. Hence, $T$ is a one-to-one transformation of $D'$ onto $\mathbb{R}^3$ with the above half-plane removed. □

## 19.2. The Jacobian

For the remainder of this lecture, we shall focus on transformations whose component functions have continuous partial derivatives in the *inside* their domain (that is, we exclude any boundary points of $D$ from this requirement).

DEFINITION. Let $T : D \to R$ be a (two-dimensional) transformation given by the functions (19.1). If the functions $f$ and $g$ have continuous first-order partials inside $D$, then the *Jacobian* of

$T$ is the determinant

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \partial_u f(u,v) & \partial_v f(u,v) \\ \partial_u g(u,v) & \partial_v g(u,v) \end{vmatrix}.$$

Let $T : D \to R$ be a (three-dimensional) transformation given by the functions (19.2). If the functions $f, g$ and $h$ have continuous first-order partials inside $D$, then the *Jacobian* of $T$ is the determinant

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \partial_u f(u,v,w) & \partial_v f(u,v,w) & \partial_w f(u,v,w) \\ \partial_u g(u,v,w) & \partial_v g(u,v,w) & \partial_w g(u,v,w) \\ \partial_u h(u,v,w) & \partial_v h(u,v,w) & \partial_w h(u,v,w) \end{vmatrix}.$$

Example 19.4. Compute the Jacobians of the transformations in Examples 19.2 and 19.3.

Solution. For the transformation (19.3), we have

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \partial_u(u\cos v) & \partial_v(u\cos v) \\ \partial_u(u\sin v) & \partial_v(u\sin v) \end{vmatrix} = \begin{vmatrix} \cos v & -u\sin v \\ \sin v & u\cos v \end{vmatrix} = u(\cos v)^2 + u(\sin v)^2 = u.$$

For the transformation (19.4), we have

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \begin{vmatrix} \cos v\sin w & -u\sin v\sin w & u\cos v\cos w \\ \sin v\sin w & u\cos v\sin w & u\sin v\cos w \\ \cos w & 0 & -u\sin w \end{vmatrix} = -u^2\sin w.$$

□

### 19.3. Change of variables in double and triple integrals

In this section, we reach the focal point the lecture: the substitution rules for double and triple integrals. However, before we state the main results, we need to adjust slightly our notation. So far in the lecture, we wrote our transformations as $T : D \to R$, which was a rather reasonable choice of names for a transformation and its domain and range. However, in most applications to double and triple integrals, we will be interested in transformations $T$ with a predetermined range, which will already be named $D$! Hence, from now on, we will reverse the roles of the letters $D$ and $R$ and write our transformations as $T : R \to D$. Our first theorem states the general form of the substitution rule for double integrals.

Theorem 19.1. *Let $D$ and $R$ be regions in $\mathbb{R}^2$, let $F : D \to \mathbb{R}$ be a continuous function, and let $T : R \to D$ be a transformation of $R$ onto $D$ given by functions* (19.1) *with continuous first-order partials inside $R$. Suppose also that $T$ is one-to-one inside $R$.*[1] *Then*

$$\iint_D F(x,y)\,dA = \iint_R F(f(u,v), g(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| dA. \tag{19.5}$$

Remark. The expression $|\partial(x,y)/\partial(u,v)|$ in (19.5) represents the *absolute value* of the Jacobian of the transformation $T$.

Example 19.5. Let us apply formula (19.5) to the transformation (19.3) from Example 19.2. By Example 19.4, we have

$$\left|\frac{\partial(x,y)}{\partial(u,v)}\right| = |u|.$$

[1]This restriction is allowed to fail at boundary points of $R$.

Starting with a set $D$ in the $xy$-plane, let $R$ be the set of polar coordinates of points in $D$ (in Lecture #15, we called $R$ the polar description of $D$). Then $R$ is essentially a subset of the region $D$ in Figure 19.1 where the transformation (19.3) is one-to-one ($R$ may have boundary points that are excluded on Figure 19.1). Hence, $T$ is a transformation of $R$ onto $D$ that is one-to-one inside $R$. Since $u \geq 0$ for $(u, v) \in R$, formula (19.5) becomes

$$\iint_D F(x, y)\, dA = \iint_R F(u \cos v, u \sin v) u\, dA.$$

Thus, the formula (15.1) for changing double integrals to polar coordinates is a special case of Theorem 19.1. □

Example 19.6. Evaluate the integral

$$\iint_D e^{(x+y)/(x-y)}\, dA,$$

Where $D$ is the quadrilateral with vertices $(1, 0)$, $(2, 0)$, $(0, -2)$, and $(0, -1)$.

Solution. The substitution

$$u = x + y, \quad v = x - y$$

will simplify the integrand. With this in mind, we solve the above equations for $x$ and $y$ to obtain the transformation $T$ given by

$$x = (u + v)/2, \quad y = (u - v)/2.$$

The Jacobian of $T$ is

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = \left( -\tfrac{1}{4} - \tfrac{1}{4} \right) = -\tfrac{1}{2}.$$

Hence, by (19.5),

$$\iint_D e^{(x+y)/(x-y)}\, dA = \iint_R e^{u/v} \left(\tfrac{1}{2}\right) dA,$$

where $R$ is the set in the $uv$-plane whose image under $T$ is the original quadrilateral $D$.

Next, we figure the set $R$. The quadrilateral $D$ in the $xy$-plane is shown on the right in Figure 19.2: it is bounded by the $x$- and $y$-axes and by the lines $y = x - 2$ and $y = x - 1$. The equations of the two oblique lines are easily described in $uv$-coordinates. For example, since

$$y = x - 2 \quad \Longleftrightarrow \quad x - y = 2 \quad \Longleftrightarrow \quad v = 2,$$

the line $y = x - 2$ is described by the equation $v = 2$; $y = x - 1$ is described by $v = 1$. Further, the $x$-axis has equation

$$y = 0 \quad \Longleftrightarrow \quad (u - v)/2 = 0 \quad \Longleftrightarrow \quad v = u,$$

and the $y$-axis has equation

$$x = 0 \quad \Longleftrightarrow \quad (u + v)/2 = 0 \quad \Longleftrightarrow \quad v = -u.$$

Hence, $R$ is the quadrilateral in the $uv$-plane bounded by the lines $v = 1$, $v = 2$, $v = u$, and $v = -u$. This quadrilateral is shown on the left in Figure 19.2.

For the purpose of evaluating the double integral over $R$, we now express $R$ as a type II region in the $uv$-plane:

$$R = \left\{ (u, v) : 1 \leq v \leq 2, \ -v \leq u \leq v \right\}.$$

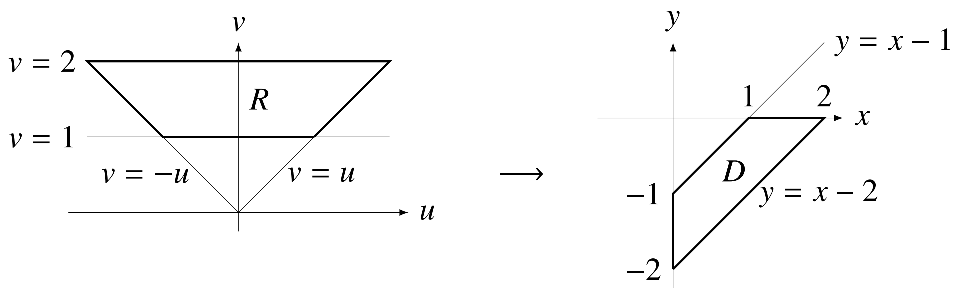



FIGURE 19.2. The quadrilateral $D$ and its $uv$-description $R$

Hence,

$$\begin{aligned}\iint_D e^{(x+y)/(x-y)}\,dA &= \frac{1}{2}\iint_R e^{u/v}\,dA = \frac{1}{2}\int_1^2 \int_{-v}^{v} e^{u/v}\,dudv\\ &= \frac{1}{2}\int_1^2 \left[ve^{u/v}\right]_{u=-v}^{u=v} dv = \frac{1}{2}\int_1^2 \left(ve - ve^{-1}\right)dv\\ &= \frac{e-e^{-1}}{2}\int_1^2 v\,dv = \frac{e-e^{-1}}{2}\left[\tfrac{1}{2}v^2\right]_1^2 = \frac{3(e-e^{-1})}{4}.\end{aligned}$$

□

Next, we state the version of Theorem 19.1 for triple integrals.

THEOREM 19.2. *Let $D$ and $R$ be solid regions in $\mathbb{R}^3$, let $F : D \to \mathbb{R}$ be a continuous function, and let $T : R \to D$ be a transformation of $R$ onto $D$ given by functions (19.2) with continuous first order partials inside $R$. Suppose also that $T$ is one-to-one inside $R$.*[2] *Then*

$$\iiint_D F(x,y,z)\,dV = \iiint_R F(f(u,v,w), g(u,v,w), h(u,v,w)) \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| dV. \tag{19.6}$$

EXAMPLE 19.7. Let us apply formula (19.6) to the transformation (19.4) from Example 19.3. By Example 19.4, we have

$$\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| = \left| -u^2 \sin w\right| = u^2|\sin w|.$$

Starting with a set $D$ in the $xyz$-space, let $R$ be the set of spherical coordinates of points in $D$ (the spherical description of $D$). With the possible exception of some boundary points, $R$ is a subset of the region $D'$ in Example 19.3, where the transformation is one-to-one. Furthermore, since $0 \le w \le \pi$, we have $|\sin w| = \sin w$. Hence, formula (19.6) becomes

$$\iiint_D F(x,y,z)\,dV = \iiint_R F(u\cos v \sin w, u \sin v \sin w, u\cos w)u^2 \sin w\,dV.$$

Save for the labeling of the solids and of the variables on the right side of this formula, this is equation (17.2). We see that the formula for changing triple integrals to spherical coordinates is a special case of (19.6). □

[2]As in Theorem 19.1, this restriction is allowed to fail at boundary points of $R$.

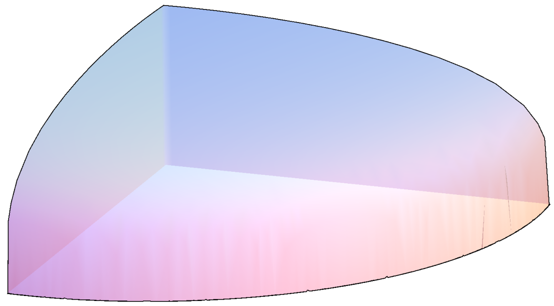


FIGURE 19.3. The solid $4x^2 + 9y^2 + z^4 \leq 1$, $x \geq 0$, $y \geq 0$, $z \geq 0$

EXAMPLE 19.8. Use a triple integral to find the volume of the solid $E$ (see Figure 19.3) in the first octant bounded by the coordinate planes and the surface $4x^2 + 9y^2 + z^4 = 1$.

SOLUTION. The solid $E$ is described by the inequalities

$$x \geq 0, \quad y \geq 0, \quad z \geq 0, \quad 4x^2 + 9y^2 + z^4 \leq 1. \tag{19.7}$$

The substitution

$$u = 2x, \quad v = 3y, \quad w = z^2$$

transforms the above inequalities to

$$u \geq 0, \quad v \geq 0, \quad w \geq 0, \quad u^2 + v^2 + w^2 \leq 1. \tag{19.8}$$

The latter inequalities can be written quite simply by using spherical coordinates in the $uvw$ space: if $(\rho, \theta, \phi)$ are the spherical coordinates of the Cartesian point $(u, v, w)$, then we can rewrite (19.8) as

$$0 \leq \theta, \phi \leq \pi/2, \quad 0 \leq \rho \leq 1. \tag{19.9}$$

Hence, solving (19.7) for $x, y, z$ and replacing $(u, v, w)$ by their spherical coordinates, we find that the transformation $T$ given by

$$x = \tfrac{1}{2}\rho \sin\phi \cos\theta, \quad y = \tfrac{1}{3}\rho \sin\phi \sin\theta, \quad z = (\rho \cos\phi)^{1/2}$$

maps the rectangular box $R$ described by (19.9) onto $E$. Furthermore, as before, it follows from the properties of the spherical coordinates that this transformation is one-to-one inside $R$.

Next we calculate the Jacobian of the above transformation. We have

$$\begin{aligned} \partial_\rho x &= x\rho^{-1}, & \partial_\theta x &= -x\tan\theta, & \partial_\phi x &= x\cot\phi, \\ \partial_\rho y &= y\rho^{-1}, & \partial_\theta y &= y\cot\theta, & \partial_\phi y &= \tfrac{2}{3}y\cot\phi, \\ \partial_\rho z &= \tfrac{1}{2}z\rho^{-1}, & \partial_\theta y &= 0, & \partial_\phi z &= -\tfrac{1}{2}z\tan\phi. \end{aligned}$$

Thus, using some properties of determinants from linear algebra, we find

$$\begin{aligned}
\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)} &= \begin{vmatrix} x\rho^{-1} & -x\tan\theta & x\cot\phi \\ y\rho^{-1} & y\cot\theta & y\cot\phi \\ \frac{1}{2}z\rho^{-1} & 0 & -\frac{1}{2}z\tan\phi \end{vmatrix} \\
&= \tfrac{1}{2}xyz\rho^{-1} \begin{vmatrix} 1 & -\tan\theta & \cot\phi \\ 1 & \cot\theta & \cot\phi \\ 1 & 0 & -\tan\phi \end{vmatrix} = \tfrac{1}{2}xyz\rho^{-1} \begin{vmatrix} 1 & -\tan\theta & \cot\phi \\ 0 & \cot\theta+\tan\theta & 0 \\ 1 & 0 & -\tan\phi \end{vmatrix} \\
&= \tfrac{1}{2}xyz\rho^{-1}(\cot\theta+\tan\theta) \begin{vmatrix} 1 & \cot\phi \\ 1 & -\tan\phi \end{vmatrix} = \tfrac{1}{2}xyz\rho^{-1}(\cot\theta+\tan\theta)(-\tan\phi-\cot\phi) \\
&= -\tfrac{1}{2}xyz\rho^{-1}(\sin\theta\cos\theta)^{-1}(\sin\phi\cos\phi)^{-1} = -\tfrac{1}{18}\rho^{3/2}\sin\phi(\cos\phi)^{-1/2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\text{vol.}(E) &= \iiint_E 1\,dV = \frac{1}{18}\iiint_R \rho^{3/2}\sin\phi(\cos\phi)^{-1/2}\,dV \\
&= \frac{1}{18}\int_0^{\pi/2}\int_0^{\pi/2}\int_0^1 \rho^{3/2}\sin\phi(\cos\phi)^{-1/2}\,d\rho d\theta d\phi \\
&= \frac{1}{18}\int_0^{\pi/2}\int_0^{\pi/2} \sin\phi(\cos\phi)^{-1/2}\left[\tfrac{2}{5}\rho^{5/2}\right]_0^1 d\theta d\phi = \frac{\pi}{90}\int_0^{\pi/2}\sin\phi(\cos\phi)^{-1/2}d\phi.
\end{aligned}$$

Note that the last integral is improper near $\phi = \pi/2$. The substitution $u = \cos\phi$ changes it to

$$\begin{aligned}
\text{vol.}(E) &= \frac{\pi}{90}\int_1^0 u^{-1/2}\,(-du) = \frac{\pi}{90}\lim_{b\to0^+}\int_b^1 u^{-1/2}\,du \\
&= \frac{\pi}{90}\lim_{b\to0^+}\left[2u^{1/2}\right]_b^1 = \frac{\pi}{90}\lim_{b\to0^+}\left(2-2b^{1/2}\right) = \frac{\pi}{45}.
\end{aligned}$$

□

## Exercises

Find the region $R$ such that the given funcion $T : D \to \mathbb{R}^2$ is a transformation of $D$ onto $R$.

19.1. $T(u,v) = (2u, v)$, $D : u^2 + v^2 \le 1$

19.2. $T(u,v) = (v, u/2)$, $D : u^2 - 2v^2 \le 2$

19.3. $T(u,v) = (u+v, v)$, $D : u^2 + v^2 \le 2$

19.4. $T(u,v) = (u+2v, 2u-v)$, $D : 2u^2 - v^2 \ge 1$

19.5. $T(u,v) = (u+v, u-3v)$, $D : 0 \le u, v \le 1$

19.6. Let $D$ be the triangle in the $xy$-plane enclosed by the coordinate axes and the line $x + y = 1$, and let $S$ be the square in the $uv$-plane where $0 \le u, v \le 1$.

(a) Show that the transformation $T$ given by $x = u - uv$, $y = uv$ transforms $S$ onto $D$. [Hint: $x + y = u$.]
(b) Show that $T$ is one-to-one inside $S$ (that is, when $0 < u, v < 1$).
(c) Find the Jacobian of $T$.

Use Theorem 19.1 or 19.2 to evaluate the given integral using the suggested change of variables. You may have to solve the given equations for $x$ and $y$ to obtain the transformation explicitly.

19.7. $\iint_D x^2y\,dA$, where $D$ is the quadrilateral with vertices $(0,0)$, $(1,1)$, $(1,-3)$, and $(2,-2)$; $x = u + v$, $y = u - 3v$

19.8. $\iint_D \cos\sqrt{x+2y}\,dA$, where $D$ is in the first quadrant and below $x + 2y = 2$; $x + 2y = u^2$, $y = v$

19.9. $\iiint_D (1+x+y+z)^{-3}\,dV$, where $D$ is in the first octant and below $x + y + z = 1$; $x + y + z = u$, $y = v$, $z = w$

19.10. Let $D$ be the quadrilateral in the $xy$-plane with vertices $(0,0)$, $(1,1)$, $(5,-1)$, and $(2,-1)$.

(a) Find a transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ such that $T$ maps the axis $u = 0$ to the line $y = x$ and the axis $v = 0$ to the line in the $xy$-plane that passes through the origin and the point $(2,-1)$.

(b) Let $T$ be the transformation $T$ found in part (a). Find the region $R$ in the $uv$-plane that $T$ maps onto the quadrilateral $D$.

(c) Let $T$ and $R$ be as in parts (a) and (b), and let $f : D \to \mathbb{R}$ be a continuous function. Use Theorem 19.1 and the transformation $T$ to transform the integral $\iint_D f(x,y)\,dA$ into an integral of the form $\iint_R g(u,v)\,dA$.

(d) Use part (c) to evaluate $\iint_D \ln(x + 2y + 1)\,dA$.

19.11. Let $R$ be the region in the first quadrant of the $xy$-plane bounded by the curves $xy = 1$, $xy = 4$, $2x^2 - y^2 = 1$, and $2x^2 - y^2 = 6$.

(a) Find a region $D$ in the $uv$-plane such that the functions $u = xy$ and $v = 2x^2 - y^2$ define a tranformation $T$ of $R$ onto $D$.

(b) Let $u > 0$ and $v > 0$. Show that if $(x, y)$ is a solution of the system of equations

$$xy = u, \qquad 2x^2 - y^2 = v,$$

then $x$ must be a solution of the equation $2x^4 - vx^2 - u^2 = 0$. Show that the latter equation has a unique positive real solution, and deduce that $T$ is a one-to-one transformation of $R$ onto $D$.

(c) Find the Jacobian of $T$.

(d) Use the transformation $T$ to evaluate $\iint_R (2x^2 + y^2)\,dA$. [Hint: Switch the usual roles of $x, y$ and $u, v$.]

19.12. In this exercise, we will find the volume of the solid $E$ enclosed by the surface $\Sigma : x^{2/3} + y^{2/3} + z^{2/3} = 4$.

(a) Let $x = u^3$, $y = v^3$, and $z = w^3$. Observe that the point $(x, y, z)$ is on/inside $\Sigma$ if and only if the point $(u, v, w)$ is on/inside the sphere $u^2 + v^2 + w^2 = 4$. Use this observation and spherical coordinates in the $uvw$-space to show that the formulas

$$x = \rho^3 \cos^3\theta \sin^3\phi, \quad y = \rho^3 \sin^3\theta \sin^3\phi, \quad z = \rho^3 \cos^3\phi$$

define a transformation $T$ of $B$ onto $E$, where $B$ is the rectangular box $0 \le \rho \le 2$, $0 \le \theta \le 2\pi$, $0 \le \phi \le \pi$. Show that $T$ is one-to-one inside $B$.

(b) Find the Jacobian of the transformation $T$ from part (a).

(c) Express the volume of $E$ as a triple integral and use the transformation $T$ above to evaluate that triple integral.

19.13. Let $p, q, r, s$ be non-negative integers, and let $S$ be the triangular pyramid (called also *simplex*) defined by the inequalities $x \ge 0$, $y \ge 0$, $z \ge 0$, and $x + y + z \le 1$.

(a) Let $U$ denote the cube of side length 1 in the $uvw$-space defined by the inequalities $0 \le u, v, w \le 1$. Show that the equations

$$x + y + z = u, \qquad y + z = uv, \qquad z = uvw$$

define a transformation $T$ of $U$ onto $S$ and that this transformation is one-to-one inside $U$. Is $T$ one-to-one on $U$; if not, then at which boundary points does $T$ fail to be one-to-one?

(b) Find the Jacobian of the transformation $T$ from part (a).

(c) Use Theorem 19.2 and the transformation $T$ above to evaluate

$$\iiint_S x^p y^q z^r (1 - x - y - z)^s \, dV.$$

[Hint: You may want first to evaluate $\int_0^1 t^a (1-t)^b \, dt$, where $a, b$ are non-negative integers.]

LECTURE 20

# Vector Fields

## 20.1. Definition

If $D$ is a region in $\mathbb{R}^2$, a *vector field* on $D$ is a function $\mathbf{F}$ that assigns to each point $(x, y)$ in $D$ a two-dimensional vector $\mathbf{F}(x, y)$. If $D$ is a solid region in $\mathbb{R}^3$, a *vector field* on $D$ is a function $\mathbf{F}$ that assigns to each point $(x, y, z) \in D$ a three-dimensional vector $\mathbf{F}(x, y, z)$.

Note that we can express each two-dimensional vector field $\mathbf{F}$ in terms of a pair of functions of two variables:

$$\mathbf{F}(x, y) = \langle P(x, y), Q(x, y)\rangle = P(x, y)\mathbf{e}_1 + Q(x, y)\mathbf{e}_2.$$

Similarly, we can express each three-dimensional vector field $\mathbf{F}$ in terms of a triple of functions of three variables:

$$\mathbf{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z)\rangle = P(x, y, z)\mathbf{e}_1 + Q(x, y, z)\mathbf{e}_2 + R(x, y, z)\mathbf{e}_3.$$

Example 20.1. The function $\mathbf{F}(x, y, z) = x^2\mathbf{e}_1 + y^2\mathbf{e}_2 + z^2\mathbf{e}_3$ is a vector field on $\mathbb{R}^3$. The function $\mathbf{F}(x, y) = \cos(xy)\mathbf{e}_1 + \sin(xy)\mathbf{e}_2$ is a vector field on $\mathbb{R}^2$. □

The standard way to visualize a vector field is to draw the arrows representing $\mathbf{F}(x, y)$ (or $\mathbf{F}(x, y, z)$) applied at the points $(x, y)$ (or $(x, y, z)$) for a number of choices for $(x, y)$ (or $(x, y, z)$) from $D$.

Example 20.2. Sketch twenty vectors of the vector field $\mathbf{F}(x, y) = \frac{1}{2}\mathbf{e}_1 - x\mathbf{e}_2$.

Answer. See Figure 20.1. □

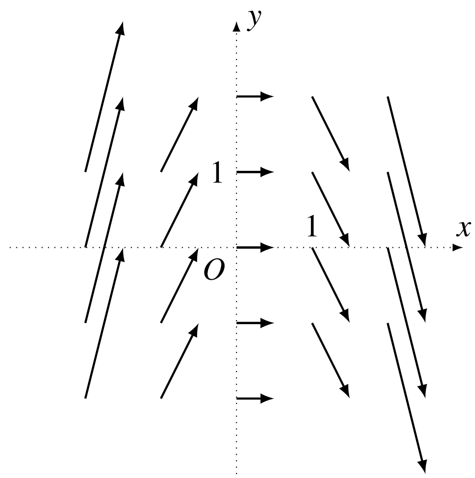



Figure 20.1. The vector field $\mathbf{F}(x, y) = \frac{1}{2}\mathbf{e}_1 - x\mathbf{e}_2$

## 20.2. Gradient fields and potential functions

We are already familiar with an important class of vector fields. Suppose that $f(x,y)$ is a differentiable function on $D$. In Lecture #9, we defined the gradient $\nabla f$ of $f$:

$$\nabla f(x,y) = f_x(x,y)\mathbf{e}_1 + f_y(x,y)\mathbf{e}_2.$$

We now see that $\nabla f$ is a two-dimensional vector field on $D$. Similarly, if $f(x,y,z)$ is a differentiable function on a solid $D$ in $\mathbb{R}^3$, then $\nabla f(x,y,z)$ is a three-dimensional vector field on $D$. From now on, we will refer to the gradient of a function $f$ as the *gradient vector field* of $f$.

Example 20.3. Find the gradient vector field of the function $f(x,y) = x^2 + y^2$.

Solution. We have

$$\nabla f(x,y) = f_x(x,y)\mathbf{e}_1 + f_y(x,y)\mathbf{e}_2 = 2x\mathbf{e}_1 + 2y\mathbf{e}_2. \qquad \square$$

In the subsequent lectures, we will see that vector fields $\mathbf{F}$ which are the gradient fields of some function play an important role in vector calculus. Such vector fields are known as *conservative*, that is, $\mathbf{F}$ is conservative if there is a function $f$ such that $\mathbf{F} = \nabla f$. If $\mathbf{F} = \nabla f$, then $f$ is called a *potential function* of $\mathbf{F}$, or simply a *potential*.

Example 20.4. By Example 20.3, the vector field $\mathbf{F}(x,y) = 2x\mathbf{e}_1 + 2y\mathbf{e}_2$ is conservative: $\mathbf{F} = \nabla f$, where $f(x,y) = x^2 + y^2$. Similarly, $\mathbf{F}(x,y,z) = 2y\mathbf{e}_1 + 2x\mathbf{e}_2 + \mathbf{e}_3$ is conservative, because $\mathbf{F} = \nabla g$, where $g(x,y,z) = 2xy + z$. □

Not every vector field is conservative, and it is not very difficult to give an example of a vector field that is not conservative.

Example 20.5. Show that the vector field $\mathbf{F}(x,y) = (x^2 + y)\mathbf{e}_1 + y^3\mathbf{e}_2$ is not conservative.

Solution. Suppose that $\mathbf{F}$ is conservative, that is, there is some function $f(x,y)$ such that

$$f_x(x,y) = x^2 + y, \quad f_y(x,y) = y^3.$$

Then

$$f_{xy}(x,y) = \partial_y(x^2 + y) = 1, \quad f_{yx}(x,y) = \partial_x(y^3) = 0.$$

Since both mixed partials are constants, they are continuous everywhere. Thus, by Theorem 8.1, we must have $f_{xy} = f_{yx}$. However, they are not equal, a contradiction. Therefore, our assumption that $\mathbf{F}$ is conservative must be false. □

We shall use this example as a starting point of our discussion of conservative vector fields. However, before delving into that, we need to fix some terminology.

## 20.3. Terminology

Recall that in §11.4 we defined a closed region in the plane to be a region containing all of its boundary. We call a region $D$ in the plane or in space *open*, if it contains none of the points on its boundary. Some examples of open and non-open sets in the plane are shown on Figure 20.2. On the figure, solid and dashed curves indicate boundary that does and does not belong to the set, respectively. The first set is open, because it does not contain any of its boundary; the second set is not open, because it contains all of its boundary (it is closed); and the last set is not open, because it contains a part of its boundary.

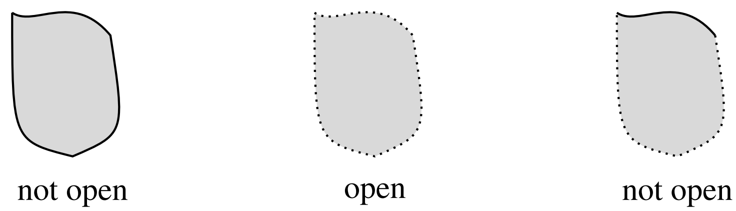



FIGURE 20.2. Open and non-open sets

We say that a set $D$, in the plane or in space, is *connected*, if we can connect every two points of $D$ with a continuous path that lies entirely in $D$. Geometrically, this means that $D$ "has only one piece". Several examples of connected and disconnected plane sets are displayed on Figure 20.3.

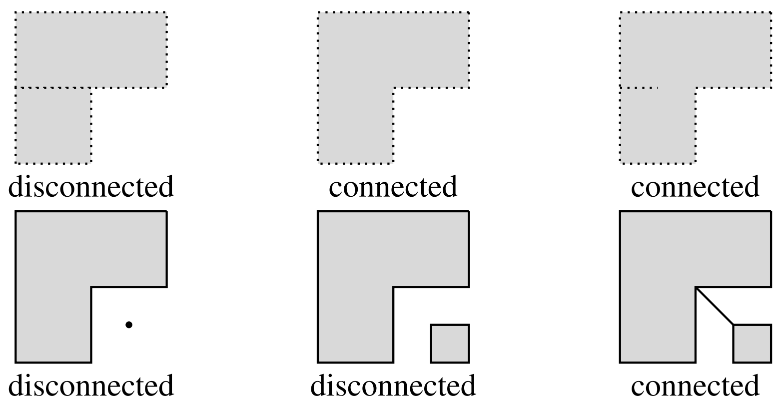



FIGURE 20.3. Connected and disconnected sets

We say that a set $D$ in the plane is *simply-connected*, if it is connected and every curve $\gamma$ that lies entirely in $D$ can be shrunk to a point in $D$ without leaving $D$. We say that a set $D$ in space is *simply-connected*, if it is connected and every surface $\Sigma$ that lies entirely in $D$ can be shrunk to a point in $D$ without leaving $D$. In both cases, it is "closed" curves and surfaces that we really need to test. Both in two and in three dimensions, simple-connectedness means that $D$ "has only one piece and no holes". Also, in both cases, we can check whether a connected set is simply-connected by looking at the compliment of $D$ to the whole plane or the whole space. If we denote that set by $D'$, then we must be able to connect each point of $D'$ to any other point of $D'$ and to "infinity" by a path that lies entirely in $D'$. Some examples of simply-connected and not simply-connected plane sets are displayed on Figure 20.4.

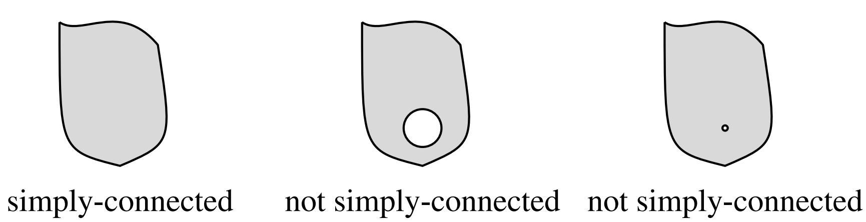



FIGURE 20.4. Simply-connected and not simply-connected sets

## 20.4. Conservative vector fields in $\mathbb{R}^2$

We can easily generalize the solution of Example 20.5. Let $\mathbf{F}(x,y) = P(x,y)\mathbf{e}_1 + Q(x,y)\mathbf{e}_2$ be a conservative vector field and let $f(x,y)$ be its potential, that is,

$$P(x,y) = f_x(x,y), \qquad Q(x,y) = f_y(x,y).$$

Then

$$P_y = f_{xy} \quad \text{and} \quad Q_x = f_{yx} \qquad \Longrightarrow \qquad P_y = Q_x.$$

This is the basic idea underlying the following theorem.

Theorem 20.1. *If $\mathbf{F}(x,y) = P(x,y)\mathbf{e}_1 + Q(x,y)\mathbf{e}_2$ is a conservative vector field and $P(x,y)$ and $Q(x,y)$ have continuous first order partials in a domain $D$, then throughout $D$ we have*

$$\partial_y P = \partial_x Q. \tag{20.1}$$

Note that we can use this theorem to show that a given vector field is **not** conservative. However, we would rather have a tool for establishing that a vector field **is** conservative. The next theorem says that equation (20.1) implies that $\mathbf{F}$ is conservative, provided that the region $D$ has a special property.

Theorem 20.2. *If $\mathbf{F}(x,y) = P(x,y)\mathbf{e}_1 + Q(x,y)\mathbf{e}_2$ is a vector field that satisfies (20.1) on an open, simply-connected region $D$, then $\mathbf{F}$ is conservative.*

Example 20.6. Determine whether or not the vector field is conservative:

(a) $\mathbf{F}(x,y) = (2x + y^2)\mathbf{e}_1 + (2y + x^2)\mathbf{e}_2$;
(b) $\mathbf{F}(x,y) = (x^3 + 3xy^2)\mathbf{e}_1 + (y^3 + 3x^2y)\mathbf{e}_2$;
(c) $\mathbf{F}(x,y) = \dfrac{-y}{x^2+y^2}\mathbf{e}_1 + \dfrac{x}{x^2+y^2}\mathbf{e}_2$.

Solution. (a) Let $P(x,y) = 2x + y^2$ and $Q(x,y) = 2y + x^2$. Then

$$P_y = 2y \neq 2x = Q_x,$$

so the vector field is not conservative.

(b) Let $P(x,y) = x^3 + 3xy^2$ and $Q(x,y) = y^3 + 3x^2y$. Then

$$P_y = 6xy = Q_x \qquad \text{for all } (x,y) \in \mathbb{R}^2.$$

Since $\mathbb{R}^2$ is open and simply-connected, it follows that the vector field is conservative.

(c) We have

$$\frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{(-1)(x^2+y^2) - (-y)(2y)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$
$$\frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{(1)(x^2+y^2) - (x)(2x)}{(x^2+y^2)^2} = \frac{y^2-x^2}{(x^2+y^2)^2},$$

for all $(x,y) \neq (0,0)$. However, the set of all $(x,y) \neq (0,0)$ is not simply-connected. Thus, Theorem 20.2 says nothing about $\mathbf{F}$. On the other hand, if we consider $\mathbf{F}$ on the set $D$ obtained by cutting the $xy$-plane along a half-line containing the origin (e.g., we can cut the plane along the positive $x$-axis, or the negative $y$-axis, or the part of the line $y = 2x$ with $x \leq 1$), then the theorem does say that $\mathbf{F}$ is conservative on $D$ (which is open and simply-connected). Later we shall show that this vector field is not conservative on the set of all $(x,y) \neq 0$. □

## 20.5. Curl and conservative vector fields in $\mathbb{R}^3$

DEFINITION. If $\mathbf{F}(x,y,z) = P(x,y,z)\mathbf{e}_1 + Q(x,y,z)\mathbf{e}_2 + R(x,y,z)\mathbf{e}_3$ is a vector field on $\mathbb{R}^3$ such that $P, Q, R$ are differentiable, then the *curl* of $\mathbf{F}$ is the vector field on $\mathbb{R}^3$ defined by

$$\operatorname{curl}\mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{e}_1 + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{e}_2 + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{e}_3.$$

REMARK. A formula for $\operatorname{curl}\mathbf{F}$ that is easier to remember is that

$$\operatorname{curl}\mathbf{F} = \nabla\times\mathbf{F} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ P & Q & R \end{vmatrix}.$$

In reality, this is a terrible abuse of notation, but it helps remember the formula, just as (2.6) helps us remember the definition of the cross product.

EXAMPLE 20.7. Find $\operatorname{curl}\mathbf{F}$ for $\mathbf{F}(x,y,z) = xe^z\mathbf{e}_1 + ye^z\mathbf{e}_2 + (x^2z + y^2z)\mathbf{e}_3$.

SOLUTION. We have

$$\begin{aligned}
\operatorname{curl}\mathbf{F} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ xe^z & ye^z & x^2z + y^2z \end{vmatrix} \\
&= \begin{vmatrix} \partial_y & \partial_z \\ ye^z & x^2z + y^2z \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} \partial_x & \partial_z \\ xe^z & x^2z + y^2z \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} \partial_x & \partial_y \\ xe^z & ye^z \end{vmatrix}\mathbf{e}_3 \\
&= \big(\partial_y(x^2z + y^2z) - \partial_z(ye^z)\big)\mathbf{e}_1 - \big(\partial_x(x^2z + y^2z) - \partial_z(xe^z)\big)\mathbf{e}_2 + \big(\partial_x(ye^z) - \partial_y(xe^z)\big)\mathbf{e}_3 \\
&= (2yz - ye^z)\mathbf{e}_1 - (2xz - xe^z)\mathbf{e}_2 + (0-0)\mathbf{e}_3 = y(2z - e^z)\mathbf{e}_1 + x(e^z - 2z)\mathbf{e}_2.
\end{aligned}$$

□

We can use the curl of a vector field to state three-dimensional versions of the theorems in the previous section.

THEOREM 20.3. *If $f(x,y,z)$ has continuous second-order partials, then*

$$\operatorname{curl}(\nabla f) = \mathbf{0}.$$

*In particular, if $\mathbf{F}$ is a conservative three-dimensional vector field, then* $\operatorname{curl}\mathbf{F} = \mathbf{0}$.

PROOF. The proof is straightforward:

$$\begin{aligned}
\operatorname{curl}(\nabla f) &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ \partial_x f & \partial_y f & \partial_z f \end{vmatrix} = \begin{vmatrix} \partial_y & \partial_z \\ \partial_y f & \partial_z f \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} \partial_x & \partial_z \\ \partial_x f & \partial_z f \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} \partial_x & \partial_y \\ \partial_x f & \partial_y f \end{vmatrix}\mathbf{e}_3 \\
&= \left(\frac{\partial^2 f}{\partial y\partial z} - \frac{\partial^2 f}{\partial z\partial y}\right)\mathbf{e}_1 - \left(\frac{\partial^2 f}{\partial x\partial z} - \frac{\partial^2 f}{\partial z\partial x}\right)\mathbf{e}_2 + \left(\frac{\partial^2 f}{\partial x\partial y} - \frac{\partial^2 f}{\partial y\partial x}\right)\mathbf{e}_3 = \mathbf{0}.
\end{aligned}$$

□

The converse theorem is not universally true, but as in the case of two-dimensional fields, it can be proved for special sets.

THEOREM 20.4. *If $\mathbf{F}$ is a three-dimensional vector field that has continuous partials and satisfies*

$$\operatorname{curl}\mathbf{F} = \mathbf{0}$$

*on an open, simply-connected set in $\mathbb{R}^3$, then $\mathbf{F}$ is conservative.*

EXAMPLE 20.8. Determine whether the vector field $\mathbf{F} = ye^{2xy+z^2}\mathbf{e}_1 + xe^{2xy+z^2}\mathbf{e}_2 + ze^{2xy+z^2}\mathbf{e}_3$ is conservative.

SOLUTION. We have

$$\begin{aligned}\operatorname{curl}\mathbf{F} &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ ye^{2xy+z^2} & xe^{2xy+z^2} & ze^{2xy+z^2} \end{vmatrix} \\ &= \begin{vmatrix} \partial_y & \partial_z \\ xe^{2xy+z^2} & ze^{2xy+z^2} \end{vmatrix}\mathbf{e}_1 - \begin{vmatrix} \partial_x & \partial_z \\ ye^{2xy+z^2} & ze^{2xy+z^2} \end{vmatrix}\mathbf{e}_2 + \begin{vmatrix} \partial_x & \partial_y \\ ye^{2xy+z^2} & xe^{2xy+z^2} \end{vmatrix}\mathbf{e}_3 \\ &= (z(2x) - x(2z))e^{2xy+z^2}\mathbf{e}_1 - (z(2y) - y(2z))e^{2xy+z^2}\mathbf{e}_2 + (x(2y) - y(2x))e^{2xy+z^2}\mathbf{e}_3 = \mathbf{0}.\end{aligned}$$

Since the partials of $\mathbf{F}$ are continuous everywhere in $\mathbb{R}^3$ (a simply-connected set), it follows that $\mathbf{F}$ is conservative. □

REMARK. Note that when $\mathbf{F}(x, y, z) = P(x, y)\mathbf{e}_1 + Q(x, y)\mathbf{e}_2 + 0\mathbf{e}_3$ (that is, when $\mathbf{F}(x, y, z)$ is a two-dimensional vector field that "pretends" to be three-dimensional), then

$$\operatorname{curl}\mathbf{F} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{e}_3,$$

and the condition $\operatorname{curl}\mathbf{F} = \mathbf{0}$ is equivalent to the condition discussed in §20.4.

## 20.6. Finding potentials

EXAMPLE 20.9. Find a potential for the vector field $\mathbf{F}(x, y) = (x^3 + 3xy^2)\mathbf{e}_1 + (y^3 + 3x^2y)\mathbf{e}_2$, if it exists.

SOLUTION. We know that this field is conservative from Example 20.6, so a potential exists. If we denote it by $f(x, y)$, then $f$ must satisfy the conditions

$$f_x(x, y) = x^3 + 3xy^2, \quad f_y(x, y) = y^3 + 3x^2y.$$

In other words, for a fixed $y$, $f(x, y)$ is an antiderivative of $x^3 + 3xy^2$, and for a fixed $x$, $f(x, y)$ is an antiderivative of $y^3 + 3x^2y$. Hence,

$$f(x, y) = \int x^3 + 3xy^2\,dx = \frac{x^4}{4} + \frac{3x^2y^2}{2} + g(y).$$

Note that our answer infolves an arbtrary function of $y$ instead of an arbitrary constant. In order to find out more about $g(y)$, we differentiate the above expression for $f(x, y)$ with respect to $y$:

$$f_y(x, y) = \partial_y\left(\tfrac{1}{4}x^4 + \tfrac{3}{2}x^2y^2 + g(y)\right) = 3x^2y + g'(y).$$

Comparing this with the equality $f_y = y^3 + 3x^2y$, we find that

$$3x^2y + g'(y) = y^3 + 3x^2y \iff g'(y) = y^3 \iff g(y) = \frac{y^4}{4} + C.$$

We conclude that $f(x, y) = \frac{1}{4}x^4 + \frac{3}{2}x^2y^2 + \frac{1}{4}y^4 + C$. □

Note that in the above example we can easily check our answer: direct differentiation gives $\nabla f(x, y) = \mathbf{F}(x, y)$, so $f(x, y)$ is potential for $\mathbf{F}(x, y)$ by the definition of potential.

EXAMPLE 20.10. Find a potential for the vector field $\mathbf{F}(x, y) = \dfrac{x}{x^2 + y^2}\mathbf{e}_1 + \dfrac{y}{x^2 + y^2}\mathbf{e}_2$, if it exists.

SOLUTION. The potential $f(x, y)$, if it exists, must satisfy the conditions

$$f_x(x, y) = \frac{x}{x^2 + y^2}, \quad f_y(x, y) = \frac{y}{x^2 + y^2}.$$

Thus,

$$f(x, y) = \int \frac{x}{x^2 + y^2}\, dx = \frac{1}{2} \int \frac{d(x^2 + y^2)}{x^2 + y^2} = \frac{\ln(x^2 + y^2)}{2} + g(y).$$

To find $g(y)$, we use $f_y(x, y)$:

$$f_y(x, y) = \frac{1}{2} \frac{2y}{x^2 + y^2} + g'(y) = \frac{y}{x^2 + y^2} \quad \Longleftrightarrow \quad g'(y) = 0 \quad \Longleftrightarrow \quad g(y) = C.$$

We conclude that $f(x, y) = \frac{1}{2} \ln(x^2 + y^2) + C$. □

REMARK. Note that the function $f(x, y) = \frac{1}{2} \ln(x^2 + y^2) + C$ in the previous example is a potential for the vector field $\mathbf{F}(x, y)$ for all $(x, y) \neq (0, 0)$. Therefore, $\mathbf{F}$ is conservative in the set $\mathbb{R}^2 - \{(0, 0)\}$. On the other hand, if we tried to apply Theorem 20.2 to it, we would run into the same trouble as in the solution of Example 20.6, part (c): this set is not simply-connected.

EXAMPLE 20.11. Find a potential for the vector field $\mathbf{F}(x, y, z) = (2xz + y^2)\mathbf{e}_1 + 2xy\mathbf{e}_2 + (x^2 + 3z^2)\mathbf{e}_3$, if it exists.

SOLUTION. The potential $f(x, y, z)$, if it exists, must satisfy the conditions

$$f_x(x, y, z) = 2xz + y^2, \quad f_y(x, y, z) = 2xy, \quad f_z(x, y, z) = x^2 + 3z^2.$$

Thus, arguing similarly to the last two examples, we find that

$$f(x, y, z) = \int (2xz + y^2)\, dx = x^2 z + y^2 x + g(y, z),$$

$$f_y(x, y, z) = 2yx + g_y(y, z) = 2xy \quad \Longleftrightarrow \quad g_y(y, z) = 0 \quad \Longleftrightarrow \quad g(y, z) = h(z),$$

$$f_z(x, y, z) = x^2 + h'(z) = x^2 + 3z^2 \quad \Longleftrightarrow \quad h'(z) = 3z^2 \quad \Longleftrightarrow \quad h(z) = z^3 + C.$$

We conclude that $f(x, y, z) = x^2 z + y^2 x + z^3 + C$. □

EXAMPLE 20.12. Find a potential for the vector field $\mathbf{F}(x, y, z) = e^{2xy}\mathbf{e}_1 + e^{2xy}\mathbf{e}_2 + z^2\mathbf{e}_3$, if it exists.

SOLUTION. Any potential $f(x, y, z)$ must satisfy

$$f_x(x, y, z) = e^{2xy}, \quad f_y(x, y, z) = e^{2xy}, \quad f_z(x, y, z) = z^2.$$

From the first of these equations, we find

$$f(x, y, z) = \int e^{2xy}\, dx = \frac{e^{2yx}}{2y} + g(y, z).$$

Next, we differentiate the last function with respect to $y$ and get

$$f_y(x, y, z) = \frac{2xe^{2xy}}{2y} - \frac{e^{2xy}}{2y^2} + g_y(y, z) = e^{2xy} \quad \Longleftrightarrow \quad g_y(y, z) = e^{2xy}\left(1 - \frac{x}{y} + \frac{1}{2y^2}\right).$$

We seem to be running in circles, and we indeed are, because the field is non-conservative. For example,

$$\partial_x\left(e^{2xy}\right) = 2ye^{2xy} \neq 2xe^{2xy} = \partial_y\left(e^{2xy}\right),$$

so the third component of curl $\mathbf{F}$ is non-zero. □

## 20.7. The divergence of a vector field

DEFINITION. If $\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{e}_1 + Q(x, y, z)\mathbf{e}_2 + R(x, y, z)\mathbf{e}_3$ is a vector field on $\mathbb{R}^3$ such that $P, Q, R$ are differentiable, then the *divergence* of $\mathbf{F}$ is the function defined by

$$\operatorname{div}\mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

EXAMPLE 20.13. Find $\operatorname{div}\mathbf{F}$ for $\mathbf{F} = e^z\sin(xy)\mathbf{e}_1 + e^z\cos(xy)\mathbf{e}_2 + xyz\mathbf{e}_3$.

SOLUTION. We have

$$\operatorname{div}\mathbf{F} = \partial_x(e^z\sin(xy)) + \partial_y(e^z\cos(xy)) + \partial_z(xyz) = ye^z\cos(xy) - xe^z\sin(xy) + xy.$$ □

THEOREM 20.5. *If* $\mathbf{F} = P\mathbf{e}_1 + Q\mathbf{e}_2 + R\mathbf{e}_3$ *has continuous second-order partials, then*

$$\operatorname{div}\operatorname{curl} F = 0.$$

PROOF. The proof is straightforward:

$$\begin{aligned}\operatorname{div}\operatorname{curl} F &= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\\ &= \frac{\partial^2 R}{\partial x\partial y} - \frac{\partial^2 Q}{\partial x\partial z} + \frac{\partial^2 P}{\partial y\partial z} - \frac{\partial^2 R}{\partial y\partial x} + \frac{\partial^2 Q}{\partial z\partial x} - \frac{\partial^2 P}{\partial z\partial y} = 0.\end{aligned}$$

□

### Exercises

20.1. Sketch the vector field $\mathbf{F}(x, y) = (x - y)\mathbf{e}_1 + x\mathbf{e}_2$ by drawing arrows in a similar fashion to Example 20.2.

Find the gradient vector field of the given function.

20.2. $f(x, y) = e^{2y}\sin(xy)$

20.3. $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$

20.4. $f(x, y, z) = \ln\left(x^2 + yz\right)$

Determine whether or not the given set is open, connected, and simply-connected.

20.5. $\{(x, y) : x > 0, y > 0\}$

20.6. $\{(x, y) : x > 0, y \geq 0\}$

20.7. $\{(x, y) : 1 < x^2 + y^2 < 2\}$

20.8. $\{(x, y) : x^2 + y^2 \leq 1\}$

20.9. $\{(x, y) : x^2 + y^2 < 1 \text{ or } x^2 + y^2 > 2\}$

20.10. $\{(x, y) : xy \leq 0\}$

Determine whether the vector field $\mathbf{F}(x, y)$ is conservative. If it is, find a potential function for it.

20.11. $\mathbf{F}(x, y) = (x - y)\mathbf{e}_1 + (x + y)\mathbf{e}_2$

20.12. $\mathbf{F}(x, y) = y\mathbf{e}_1 + (x + 3y)\mathbf{e}_2$

20.13. $\mathbf{F}(x, y) = e^y\mathbf{e}_1 + xe^y\mathbf{e}_2$

20.14. $\mathbf{F}(x, y) = \dfrac{y^2}{1 + x^2}\mathbf{e}_1 + (2y\arctan x)\mathbf{e}_2$

20.15. $\mathbf{F}(x, y) = (4x + \sin y)\mathbf{e}_1 + (\sin y + x\cos y)\mathbf{e}_2$

Find the curl and the divergence of the given vector field.

20.16. $\mathbf{F}(x, y, z) = x^2yz\mathbf{e}_1 - 4xyz^2\mathbf{e}_3$

20.17. $\mathbf{F}(x, y, z) = 2\arcsin y\mathbf{e}_1 + \dfrac{2x}{\sqrt{1 - y^2}}\mathbf{e}_2 - 3z^2\mathbf{e}_3$

20.18. $\mathbf{F}(x, y, z) = \cos xz\mathbf{e}_1 + 2\sin xy\mathbf{e}_2 + ze^{2xy}\mathbf{e}_3$

Determine whether the vector field $\mathbf{F}(x, y, z)$ is conservative. If it is, find a potential function for it.

20.19. $\mathbf{F}(x, y, z) = (yz + x^2)\mathbf{e}_1 + (xz + 4y)\mathbf{e}_2 + (xy - 3z^2)\mathbf{e}_3$

20.20. $\mathbf{F}(x, y, z) = y^2\cos z\mathbf{e}_1 + 2xy\cos z\mathbf{e}_2 - xy^2\sin z\mathbf{e}_3$

20.21. $\mathbf{F}(x, y, z) = xe^y\mathbf{e}_1 + ye^x\mathbf{e}_2 + xyz\mathbf{e}_3$

20.22. $\mathbf{F}(x, y, z) = 2\arcsin y\mathbf{e}_1 + \dfrac{2x}{\sqrt{1 - y^2}}\mathbf{e}_2 - 3z^2\mathbf{e}_3$

20.23. Is there a vector field $\mathbf{F}$ on $\mathbb{R}^3$ such that $\operatorname{curl}\mathbf{F} = xy^2\mathbf{e}_1 + yz^2\mathbf{e}_2 + zx^2\mathbf{e}_3$? Explain.

LECTURE 21

# Line Integrals

## 21.1. Line integrals of scalar functions

We start with a smooth parametric curve $\gamma$, given by a vector function

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3 \qquad (a \le t \le b).$$

Recall that the smoothness of $\gamma$ means that $\mathbf{r}'(t) = x'(t)\mathbf{e}_1 + y'(t)\mathbf{e}_2 + z'(t)\mathbf{e}_3$ is continuous and $\mathbf{r}'(t) \neq \mathbf{0}$ when $a < t < b$ (see §5.3). Further, suppose that $f(x, y, z)$ is a continuous function on a set $S$ in $\mathbb{R}^3$ that contains the curve $\gamma$. In particular, $f(x, y, z)$ is defined at each point $(x, y, z)$ on $\gamma$.

For each $n \ge 1$, set $\Delta_n = (b-a)/n$ and define the numbers that partition $[a, b]$ into $n$ subintervals of equal lengths:

$$t_0 = a, \quad t_1 = a + \Delta_n, \quad t_2 = a + 2\Delta_n, \quad \ldots, \quad t_n = a + n\Delta_n = b.$$

We write $x_i = x(t_i)$, $y_i = y(t_i)$, $z_i = z(t_i)$ for the coordinates of the point $P_i(x_i, y_i, z_i)$ that corresponds to the value $t = t_i$ of the parameter. The points $P_0, P_1, P_2, \ldots, P_n$ partition $\gamma$ into $n$ arcs: $\gamma_1, \gamma_2, \ldots, \gamma_n$, where $\gamma_i$ has $P_{i-1}$ and $P_i$ as its endpoints. Further, we pick $n$ values of the parameter $t_1^*, t_2^*, \ldots, t_n^*$ so that

$$t_0 \le t_1^* \le t_1, \quad t_1 \le t_2^* \le t_2, \quad \ldots, \quad t_{n-1} \le t_n^* \le t_n.$$

These yield $n$ *sample points* $P_i^*(x_i^*, y_i^*, z_i^*)$ on the curve with coordinates $x_i^* = x(t_i^*)$, $y_i^* = y(t_i^*)$, $z_i^* = z(t_i^*)$. Then the *line integral of $f$ along* $\gamma$ is defined by

$$\int_\gamma f(x, y, z)\, ds = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*, z_i^*)\Delta s_i, \qquad \Delta s_i = \text{length}(\gamma_i), \tag{21.1}$$

provided that the limit exists.

Recall that, by Theorem 5.1,

$$\Delta s_i = \int_{t_{i-1}}^{t_i} \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt.$$

Using this formula, the above definition, the properties of continuous functions, and the definition of the definite integral, it can be shown that for continuous functions $f$ the limit (21.1) always exists and the line integral can be expressed as a definite integral.

THEOREM 21.1. *Let $\gamma$ be a smooth parametric curve, given by the vector function*

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + z(t)\mathbf{e}_3 \qquad (a \le t \le b).$$

*If $f(x, y, z)$ is continuous on $\gamma$, then*

$$\int_\gamma f(x, y, z)\, ds = \int_a^b f(x(t), y(t), z(t)) \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt. \tag{21.2}$$

Remarks. 1. Using more extensively the language of vector functions, we can write (21.2) in the form

$$\int_\gamma f(x, y, z)\, ds = \int_a^b f(\mathbf{r}(t))\|\mathbf{r}'(t)\|\, dt. \tag{21.3}$$

2. As in the cases of definite and multiple integrals, the line integral of a continuous function along a curve is related to its average value on the points of the curve:

$$\operatorname*{average}_{(x,y,z)\in\gamma} f(x, y, z) = \frac{1}{\text{length}(\gamma)} \int_\gamma f(x, y, z)\, ds.$$

Example 21.1. Evaluate

$$\int_\gamma \left(x^2 + y^2 + z^2\right)^{-1} ds$$

where $\gamma$ is the helix

$$x = \cos t, \quad y = \sin t, \quad z = t \qquad (0 \le t \le T).$$

Solution. We have

$$\begin{aligned}\int_\gamma \left(x^2 + y^2 + z^2\right)^{-1} ds &= \int_0^T \left((\cos t)^2 + (\sin t)^2 + t^2\right)^{-1} \sqrt{(-\sin t)^2 + (\cos t)^2 + 1^2}\, dt \\ &= \int_0^T (1 + t^2)^{-1} \sqrt{1 + 1}\, dt = \int_0^T \frac{\sqrt{2}}{1 + t^2}\, dt = \sqrt{2} \arctan T.\end{aligned}$$

□

We have already seen on several instances that properties of plane curves or two-dimensional vectors often appear as special cases of analogous properties of space curves or three-dimensional vectors. This general principle also applies to line integrals. We just defined the line integral of a function of three variables along a space curve. Is it possible to define the line integral of a function of two variables along a plane curve? The answer to this question is in the affirmative. One way to develop the theory of such line integrals is by repeating the above definitions and reproving the above theorems with one less variable. Alternatively, we can define two-dimensional line integrals as a special case of three-dimensional line integrals. We take the latter route.

Observe that a two-variable function $f(x, y)$ is also a three-variable function $f(x, y, z)$ (which does not really depend on $z$) and that a plane curve, given by

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 \qquad (a \le t \le b), \tag{21.4}$$

is also a space curve, given by

$$\mathbf{r}_0(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (a \le t \le b). \tag{21.5}$$

If $\gamma$ is the plane curve (21.4) and $f(x, y)$ is continuous on $\gamma$, we define the *line integral of $f$ along* $\gamma$ by

$$\int_\gamma f(x, y)\, ds = \int_{\gamma_0} g(x, y, z)\, ds,$$

where $\gamma_0$ is the space curve (21.5) and $g(x, y, z) = f(x, y)$ for all $(x, y)$ in the domain of $f$ and all $z$. In particular, by (21.2), we have

$$\begin{aligned}\int_\gamma f(x, y)\, ds &= \int_a^b g(x(t), y(t), 0)\, \sqrt{x'(t)^2 + y'(t)^2 + 0^2}\, dt \\ &= \int_a^b f(x(t), y(t))\, \sqrt{x'(t)^2 + y'(t)^2}\, dt.\end{aligned}$$

That is, we have the formula

$$\int_\gamma f(x, y)\, ds = \int_a^b f(\mathbf{r}(t))\|\mathbf{r}'(t)\|\, dt = \int_a^b f(x(t), y(t))\, \sqrt{x'(t)^2 + y'(t)^2}\, dt. \tag{21.6}$$

Note that the vector form of (21.6) and the vector form (21.3) of (21.2) look identical; the only difference is that the vector function $\mathbf{r}(t)$ is two-dimensional in (21.6) and three-dimensional in (21.3).

Example 21.2. Evaluate

$$\int_\gamma xe^y\, ds,$$

where $\gamma$ is the arc of the curve $x = e^y$ from $(1, 0)$ to $(e, 1)$.

Solution. The curve can be parametrized by the equations

$$x = e^t, \quad y = t \qquad (0 \le t \le 1).$$

Hence, by (21.6),

$$\int_\gamma xe^y\, ds = \int_0^1 e^t e^t\, \sqrt{(e^t)^2 + 1^2}\, dt = \int_0^1 e^{2t}\, \sqrt{e^{2t} + 1}\, dt.$$

Using the substitution $u = e^{2t} + 1$, we find that

$$du = 2e^{2t}\, dt, \quad u(0) = 2, \quad u(1) = e^2 + 1,$$

and

$$\int_0^1 e^{2t}\, \sqrt{e^{2t} + 1}\, dt = \int_2^{e^2+1} u^{1/2}\left(\tfrac{1}{2}du\right) = \frac{(e^2 + 1)^{3/2} - 2^{3/2}}{3} \approx 7.157.$$

□

We can easily extend the defintion of line integrals to piecewise smooth curves. If $\gamma$ is a piecewise smooth parametric curve that is the union of the smooth curves $\gamma_1, \gamma_2, \dots, \gamma_n$, $\gamma = \gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_n$, we define

$$\int_\gamma f(x, y, z)\, ds = \int_{\gamma_1} f(x, y, z)\, ds + \int_{\gamma_2} f(x, y, z)\, ds + \cdots + \int_{\gamma_n} f(x, y, z)\, ds.$$

Note that all the line integrals on the right side of the formula are along smooth curves and can be evaluated via the limit in (21.1) or via the integral formulas (21.2), (21.3), or (21.6).

Example 21.3. Evaluate the line integral

$$\int_\gamma xy^2\, ds$$

where $\gamma$ is the triangle with vertices $(0, 1, 2)$, $(1, 0, 3)$, and $(0, -1, 0)$.

Solution. We split $\gamma$ into three line segments:

- $\gamma_1$ is the segment connecting $(0,1,2)$ and $(1,0,3)$:
$$\mathbf{r}(t) = t\langle 0,1,2\rangle + (1-t)\langle 1,0,3\rangle = \langle 1-t, t, 3-t\rangle \qquad (0 \le t \le 1);$$
- $\gamma_2$ is the segment connecting $(0,1,2)$ and $(0,-1,0)$:
$$\mathbf{r}(t) = t\langle 0,1,2\rangle + (1-t)\langle 0,-1,0\rangle = \langle 0, 2t-1, 2t\rangle \qquad (0 \le t \le 1);$$
- $\gamma_3$ is the segment connecting $(1,0,3)$ and $(0,-1,0)$:
$$\mathbf{r}(t) = t\langle 1,0,3\rangle + (1-t)\langle 0,-1,0\rangle = \langle t, t-1, 3t\rangle \qquad (0 \le t \le 1).$$

By (21.2),
$$\int_{\gamma_1} xy^2\,ds = \int_0^1 (1-t)t^2\sqrt{(-1)^2+1^2+(-1)^2}\,dt = \sqrt{3}\int_0^1 \left(t^2 - t^3\right)dt = \frac{\sqrt{3}}{12};$$
$$\int_{\gamma_2} xy^2\,ds = \int_0^1 0(2t-1)^2\sqrt{0^2+2^2+2^2}\,dt = 0;$$
$$\int_{\gamma_3} xy^2\,ds = \int_0^1 t(t-1)^2\sqrt{1^2+1^2+3^2}\,dt = \sqrt{11}\int_0^1 \left(t^3 - 2t^2 + t\right)dt = \frac{\sqrt{11}}{12}.$$

Hence,
$$\int_\gamma xy^2\,ds = \int_{\gamma_1} xy^2\,ds + \int_{\gamma_2} xy^2\,ds + \int_{\gamma_3} xy^2\,ds = \frac{\sqrt{3}}{12} + \frac{\sqrt{11}}{12} \approx 0.421.$$

□

## 21.2. Line integrals of vector fields

Let $\mathbf{F}$ be a continuous vector field defined on a piecewise smooth curve $\gamma$ given by a vector function $\mathbf{r}(t)$, $a \le t \le b$. We define the *line integral of* $\mathbf{F}$ *along* $\gamma$ by
$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_\gamma (\mathbf{F}\cdot\mathbf{T})\,ds,$$
where $\mathbf{T}$ is the unit tangent vector to the curve $\gamma$. Recall from §5.4 that the unit tangent vector $\mathbf{T}(t)$ at the point of $\gamma$ corresponding to the value $t$ of the parameter can be expressed in terms of $\mathbf{r}(t)$ by the formula
$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}.$$
Therefore, by (21.3) or (21.6) (depending on the dimension of $\mathbf{F}$ and $\gamma$),
$$\int_\gamma (\mathbf{F}\cdot\mathbf{T})\,ds = \int_a^b \left(\mathbf{F}(\mathbf{r}(t))\cdot\frac{\mathbf{r}'(t)}{\|\mathbf{r}'(t)\|}\right)\|\mathbf{r}'(t)\|\,dt = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt.$$
That is, we have
$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_\gamma (\mathbf{F}\cdot\mathbf{T})\,ds = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt. \tag{21.7}$$

What is the meaning of this kind of line integral? Based on the interpretation that we gave to the line integral of a function, we can say that

$$\frac{1}{\text{length}(\gamma)}\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \frac{1}{\text{length}(\gamma)}\int_\gamma (\mathbf{F}\cdot\mathbf{T})\,ds = \underset{\gamma}{\text{average}}\,(\mathbf{F}\cdot\mathbf{T}).$$

Recall from §2.7 that the dot product $\mathbf{F}\cdot\mathbf{T}$ is the size of the projection of $\mathbf{F}$ onto the direction of $\mathbf{T}$, taken with positive or negative sign according as the projection and $\mathbf{T}$ point in the same direction or in opposite directions. Since $\mathbf{T}(t)$ is the direction of the curve at the point corresponding to the value $t$ of the parameter, we conclude that

$$\frac{1}{\text{length}(\gamma)}\int_\gamma \mathbf{F}\cdot d\mathbf{r}$$

is the "average projection of $\mathbf{F}$ in the direction of the curve".

Example 21.4. Evaluate

$$\int_\gamma \big(y\mathbf{e}_1 + z\mathbf{e}_2 + x\mathbf{e}_3\big)\cdot d\mathbf{r},$$

where $\gamma$ is the helix $\mathbf{r}(t) = \langle 2\cos t, 2\sin t, 3t\rangle$, $0 \le t \le 2\pi$.

Solution. By (21.7),

$$\begin{aligned}
\int_\gamma \big(y\mathbf{e}_1 + z\mathbf{e}_2 + x\mathbf{e}_3\big)\cdot d\mathbf{r} &= \int_0^{2\pi} \langle 2\sin t, 3t, 2\cos t\rangle \cdot \langle (2\cos t)', (2\sin t)', (3t)'\rangle\; dt \\
&= \int_0^{2\pi} \langle 2\sin t, 3t, 2\cos t\rangle \cdot \langle -2\sin t, 2\cos t, 3\rangle\; dt \\
&= \int_0^{2\pi} \big(-4\sin^2 t + 6t\cos t + 6\cos t\big)\,dt \\
&= -2\int_0^{2\pi}(1-\cos 2t)\,dt + 6\int_0^{2\pi}(t+1)\,d(\sin t) \\
&= \big[-2t + \sin 2t\big]_0^{2\pi} + \big[6(t+1)\sin t\big]_0^{2\pi} - 6\int_0^{2\pi}\sin t\,dt = -4\pi.
\end{aligned}$$

□

In the above example, the parametrization of $\gamma$ was given explicitly. However, we often have only a geometric description of the curve and have to obtain the parametrization ourselves. In such situations, it is important to remember that (21.7) is sensitive to the *orientation* of the curve, that is, the direction in which $\mathbf{r}(t)$ traces the curve as $t$ increases. Indeed, if $\gamma_1$ and $\gamma_2$ are two parametric curves that represent the same geometric curve $\gamma$ but with opposite orientations, then

$$\int_{\gamma_1} \mathbf{F}\cdot d\mathbf{r} = -\int_{\gamma_2} \mathbf{F}\cdot d\mathbf{r}. \tag{21.8}$$

That is, if we change the orientation of the curve, the line integral changes sign. This fact goes back to the property of definite integrals that

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Because of (21.8), a geometric description of the curve must always specify its orientation.

Example 21.5. Evaluate

$$\int_\gamma \big(x^2\mathbf{e}_1 + y^2\mathbf{e}_2\big)\cdot d\mathbf{r},$$

where $\gamma$ is the right half of the ellipse $x^2 + 4y^2 = 4$ oriented in the counterclockwise direction.

Solution. The point $(a, b)$ is on the given ellipse exactly when the point $(a, 2b)$ is on the circle $x^2 + y^2 = 4$. Thus, we can parametrize $\gamma$ by

$$x = 2\cos t, \quad y = \sin t \qquad (-\pi/2 \le t \le \pi/2).$$

Note that as $t$ increases from $-\pi/2$ to $\pi/2$, the point $(x(t), y(t))$ moves along $\gamma$ in the counterclockwise direction. Applying (21.7) to the above parametrization, we get

$$\begin{aligned}\int_\gamma \big(x^2\mathbf{e}_1 + y^2\mathbf{e}_2\big)\cdot d\mathbf{r} &= \int_{-\pi/2}^{\pi/2} \big((2\cos t)^2\mathbf{e}_1 + (\sin t)^2\mathbf{e}_2\big)\cdot\big((2\cos t)'\mathbf{e}_1 + (\sin t)'\mathbf{e}_2\big)\,dt\\ &= \int_{-\pi/2}^{\pi/2} \big(4\cos^2 t\mathbf{e}_1 + \sin^2 t\mathbf{e}_2\big)\cdot\big(-2\sin t\mathbf{e}_1 + \cos t\mathbf{e}_2\big)\,dt\\ &= \int_{-\pi/2}^{\pi/2} \big(-8\cos^2 t\sin t + \sin^2 t\cos t\big)\,dt\\ &= -8\int_{-\pi/2}^{\pi/2}\cos^2 t\sin t\,dt + \int_{-\pi/2}^{\pi/2}\sin^2 t\cos t\,dt\\ &= 0 + 2\int_0^{\pi/2}\sin^2 t\cos t\,dt = 2\int_0^1 u^2\,du = \frac{2}{3} \qquad (u = \sin t).\end{aligned}$$

□

Suppose that $\mathbf{F}(x, y, z) = f(x, y, z)\mathbf{e}_1$. Then (21.7) yields

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_a^b \big(f(\mathbf{r}(t))\mathbf{e}_1\big)\cdot\big(x'(t)\mathbf{e}_1 + y'(t)\mathbf{e}_2 + z'(t)\mathbf{e}_3\big)\,dt = \int_a^b f(\mathbf{r}(t))x'(t)\,dt.$$

Another notation for this integral is $\int_\gamma f(x, y, z)\,dx$, that is,

$$\int_\gamma f(x, y, z)\,dx = \int_\gamma (f\mathbf{e}_1)\cdot d\mathbf{r} = \int_a^b f(\mathbf{r}(t))x'(t)\,dt. \tag{21.9}$$

Similarly, applying (21.7) to the vector fields $\mathbf{F}(x, y, z) = f(x, y, z)\mathbf{e}_2$ and $\mathbf{F}(x, y, z) = f(x, y, z)\mathbf{e}_3$, we obtain

$$\int_\gamma f(x, y, z)\,dy = \int_\gamma (f\mathbf{e}_2)\cdot d\mathbf{r} = \int_a^b f(\mathbf{r}(t))y'(t)\,dt \tag{21.10}$$

and

$$\int_\gamma f(x, y, z)\,dz = \int_\gamma (f\mathbf{e}_3)\cdot d\mathbf{r} = \int_a^b f(\mathbf{r}(t))z'(t)\,dt. \tag{21.11}$$

These three integrals are called the *line integrals of f with respect to x, y, and z*, respectively. Since they are derived from (21.7), these integrals are also sensitive to the orientation of the curve $\gamma$

and change signs when we reverse the orientation of the curve. For example, if $\gamma_1$ and $\gamma_2$ are two parametric curves that represent the same geometric curve with opposite orientations, then

$$\int_{\gamma_1} f(x, y, z)\,dx = -\int_{\gamma_2} f(x, y, z)\,dx.$$

Note that if $\mathbf{F}(x, y, z) = P(x, y, z)\mathbf{e}_1 + Q(x, y, z)\mathbf{e}_2 + R(x, y, z)\mathbf{e}_3$, then we can use (21.9)–(21.11) to write

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_\gamma P(x, y, z)\,dx + \int_\gamma Q(x, y, z)\,dy + \int_\gamma R(x, y, z)\,dz.$$

In other words, the line integral of a vector field decomposes into a sum of the line integrals of its components with respect to the corresponding variables. The right side of the last identity is often written as $\int_\gamma P\,dx + Q\,dy + R\,dz$. With this notation, we can rewrite that identity in the form

$$\int_\gamma (P\mathbf{e}_1 + Q\mathbf{e}_2 + R\mathbf{e}_3)\cdot d\mathbf{r} = \int_\gamma P\,dx + Q\,dy + R\,dz. \tag{21.12}$$

Example 21.6. Evaluate

$$\int_\gamma y\,dx + x\,dy + xyz\,dz,$$

where $\gamma$ is the space curve

$$x = e^t + e^{-t}, \quad y = e^t - e^{-t}, \quad z = t^2 \qquad (0 \le t \le 1).$$

Solution. By (21.9)–(21.11),

$$\begin{aligned}
\int_\gamma y\,dx + x\,dy + xyz\,dz &= \int_0^1 \left[\left(e^t - e^{-t}\right)\left(e^t + e^{-t}\right)' + \left(e^t + e^{-t}\right)\left(e^t - e^{-t}\right)' + \left(e^{2t} - e^{-2t}\right)t^2(2t)\right] dt \\
&= \int_0^1 \left[\left(e^t - e^{-t}\right)^2 + \left(e^t + e^{-t}\right)^2 + 2t^3\left(e^{2t} - e^{-2t}\right)\right] dt \\
&= \int_0^1 \left(e^{2t} - 2 + e^{-2t} + e^{2t} + 2 + e^{-2t}\right) dt + \int_0^1 t^3\,d\left(e^{2t} + e^{-2t}\right) \\
&= 2\int_0^1 \left(e^{2t} + e^{-2t}\right) dt + \left[t^3\left(e^{2t} + e^{-2t}\right)\right]_0^1 - \int_0^1 \left(e^{2t} + e^{-2t}\right)(3t^2)\,dt \\
&= \left[e^{2t} - e^{-2t}\right]_0^1 + e^2 + e^{-2} - \frac{3}{2}\int_0^1 t^2\,d\left(e^{2t} - e^{-2t}\right) \\
&= 2e^2 - \frac{3}{2}\left[t^2\left(e^{2t} - e^{-2t}\right)\right]_0^1 + \frac{3}{2}\int_0^1 \left(e^{2t} - e^{-2t}\right)(2t)\,dt \\
&= \frac{e^2 + 3e^{-2}}{2} + \frac{3}{2}\int_0^1 t\,d\left(e^{2t} + e^{-2t}\right) \\
&= \frac{e^2 + 3e^{-2}}{2} + \frac{3}{2}\left[t\left(e^{2t} + e^{-2t}\right)\right]_0^1 - \frac{3}{2}\int_0^1 \left(e^{2t} + e^{-2t}\right) dt \\
&= 2e^2 + 3e^{-2} - \frac{3}{4}\left[e^{2t} - e^{-2t}\right]_0^1 = \frac{5e^2 + 15e^{-2}}{4}.
\end{aligned}$$

□

Example 21.7. Evaluate
$$\int_\gamma \left(y^2 - xy\right) dx,$$
where $\gamma$ is the curve $y = \ln x$ from $(1,0)$ towards $(e,1)$.

Solution. We can parametrize $\gamma$ as
$$x = t, \quad y = \ln t \qquad (1 \le t \le e).$$
Hence, (21.9) gives
$$\int_\gamma \left(y^2 - xy\right) dx = \int_1^e \left(\ln^2 t - t\ln t\right) dt.$$
Using integration by parts, we find that
$$\int_1^e \ln^2 t\,dt = \left[t\ln^2 t\right]_1^e - \int_1^e t\,d\left(\ln^2 t\right) = e - \int_1^e t\left(2t^{-1}\ln t\right) dt = e - 2\int_1^e \ln t\,dt$$
$$= e - 2\left[t\ln t\right]_1^e + 2\int_1^e t\left(t^{-1}\right) dt = -e + 2(e-1) = e - 2;$$
$$\int_1^e t\ln t\,dt = \int_1^e \ln t\,d\left(\tfrac{1}{2}t^2\right) = \left[\tfrac{1}{2}t^2\ln t\right]_1^e - \int_1^e \tfrac{1}{2}t^2\left(t^{-1}\right) dt = \tfrac{1}{2}e^2 - \tfrac{1}{2}\int_1^e t\,dt = \tfrac{1}{4}\left(e^2+1\right).$$
Thus, the given line integral equals $e - 2 - \frac{1}{4}\left(e^2+1\right) \approx -1.379$. □

## Exercises

Evaluate the given line integral.

21.1. $\int_\gamma x^3 y\,ds$; $\gamma$ is given by $\mathbf{r}(t) = t\mathbf{e}_1 + t^2\mathbf{e}_2$, $0 \le t \le 1$

21.2. $\int_\gamma x^2 y\,ds$; $\gamma$ is the upper half of the circle $x^2 + y^2 = 4$

21.3. $\int_\gamma (xy + \ln x)\,dy$; $\gamma$ is the arc of the curve $y = x^{3/2}$ from $(1,1)$ to $(0,0)$

21.4. $\int_\gamma xy\,dx - x^2\,dy$; $\gamma$ is the arc of the circle $x^2 + y^2 = 4$ from $(2,0)$ to $(0,2)$

21.5. $\int_\gamma \left(xy + z^2\right) ds$; $\gamma$ is given by $\mathbf{r}(t) = \cos 2t\mathbf{e}_1 + \sin 2t\mathbf{e}_2 + 3t\mathbf{e}_3$, $0 \le t \le \pi$.

21.6. $\int_\gamma z^2\,dx + y\,dz$; $\gamma$ is given by $\mathbf{r}(t) = (t - \sin t)\mathbf{e}_1 + (1 - \cos t)\mathbf{e}_2 + \sin(t/2)\mathbf{e}_3$, $0 \le t \le 2\pi$

21.7. $\int_\gamma yz\,dx + 2xz\,dy - xy\,dz$; $\gamma$ consists of the line segments from $(1,0,0)$ to $(1,0,2)$ to $(2,2,2)$

21.8. $\int_\gamma y^2\,dx + x^2\,dy + xyz\,dz$; $\gamma$ consists of the upper half of the circle $x^2 + y^2 = 1$ in the $xy$-plane taken clockwise and the line segment from $(1,0,0)$ to $(0,1,1)$

Evaluate the line integral $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$ for the given vector field $\mathbf{F}$ and the given curve $\gamma$.

21.9. $\mathbf{F}(x,y,z) = x^2 z\,\mathbf{e}_2 - (x+y)\mathbf{e}_3$; $\gamma$ is the ellipse $\mathbf{r}(t) = 2\mathbf{e}_1 + 3\sin t\mathbf{e}_2 + \cos t\mathbf{e}_3$, $0 \le t \le 2\pi$
21.10. $\mathbf{F}(x,y,z) = \sin x\,\mathbf{e}_1 + \cos y\,\mathbf{e}_2 - 2xyz\mathbf{e}_3$; $\gamma$ is given by $\mathbf{r}(t) = t^3\mathbf{e}_1 + 2t^2\mathbf{e}_2 + \sqrt{t}\mathbf{e}_3$, $0 \le t \le 1$
21.11. $\mathbf{F}(x,y) = x\sin y\mathbf{e}_1 + y\mathbf{e}_2$; $\gamma$ is the arc of the parabola $x = y^2$ from $(1,-1)$ to $(4,2)$

LECTURE 22

# The Fundamental Theorem for Line Integrals

## 22.1. The main theorem

So far, we have been reducing line integrals to definite integrals. In the case of conservative vector fields, there is another, faster way to evaluate their line integrals. That alternative approach is based on the following theorem, known as the *fundamental theorem for line integrals*.

Theorem 22.1. *Let $\gamma$ be a piecewise smooth curve given by a vector function $\mathbf{r}(t)$, $a \le t \le b$, and let $f(x, y, z)$ be a differentiable function, whose gradient vector field $\nabla f$ is continuous on $\gamma$. Then*

$$\int_\gamma \nabla f \cdot d\mathbf{r} = f(\mathbf{r}(b)) - f(\mathbf{r}(a)). \tag{22.1}$$

Proof. We have

$$\begin{aligned}
\int_\gamma \nabla f \cdot d\mathbf{r} &= \int_\gamma f_x\, dx + f_y\, dy + f_z\, dz && \text{by (21.12)} \\
&= \int_a^b \big( f_x(\mathbf{r}(t))x'(t) + f_y(\mathbf{r}(t))y'(t) + f_z(\mathbf{r}(t))z'(t) \big)\, dt && \text{by (21.9)–(21.11)} \\
&= \int_a^b \frac{d}{dt}\big[ f(\mathbf{r}(t)) \big]\, dt && \text{by the chain rule} \\
&= f(\mathbf{r}(b)) - f(\mathbf{r}(a)) && \text{by the FTC.}
\end{aligned}$$

□

Suppose that $\mathbf{F}(x, y, z)$ is a conservative vector field with potential function $f(x, y, z)$, that is, $\mathbf{F} = \nabla f$. Theorem 22.1 then gives

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_\gamma \nabla f \cdot d\mathbf{r} = f(\mathbf{r}(b)) - f(\mathbf{r}(a)).$$

This demonstrates the importance of conservative vector fields. If $\mathbf{F}$ is conservative and its potential $f$ is known, the evaluation of the line integral of $\mathbf{F}$ is rather easy: all we need to do is calculate the net change of $f$ along the curve. The next example illustrates this.

Example 22.1. Evaluate

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r},$$

where $\mathbf{F}(x, y, z) = 2y\mathbf{e}_1 + 2x\mathbf{e}_2 + \mathbf{e}_3$ and $\gamma$ is the curve given by $\mathbf{r}(t) = \langle \sin 2t, \cos 2t, \ln(t + 1) \rangle$, $0 \le t \le \pi$.

Solution. We observe that $\mathbf{F} = \nabla g$, where $g(x, y, z) = 2xy + z$. Hence, by (22.1),

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_\gamma \nabla g \cdot d\mathbf{r} = g(\mathbf{r}(\pi)) - g(\mathbf{r}(0)) = g(0, 1, \ln(\pi + 1)) - g(0, 1, 0) = \ln(\pi + 1).$$

Observe that the answer will stay the same, if we replace $\gamma$ by **any** other piecewise smooth curve connecting the points $(0,1,0)$ and $(0,1,\ln(\pi+1))$! For example, the line integral of $\mathbf{F}$ along the curve displayed on Figure 22.1 (I have no idea what its parametrization may be!) will also equal $\ln(\pi+1)$. □

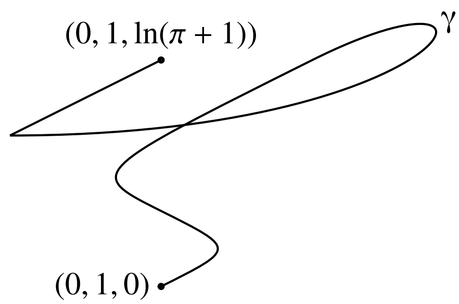



FIGURE 22.1. A curve with $\int_\gamma (2y\mathbf{e}_1 + 2x\mathbf{e}_2 + \mathbf{e}_3)\cdot d\mathbf{r} = \ln(\pi+1)$

## 22.2. Path independence

We just saw that if $\mathbf{F}(x,y,z)$ is a conservative vector field, Theorem 22.1 implies that the line integral $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$ depends only on the endpoints of $\gamma$. Thus, if $\gamma_1$ and $\gamma_2$ are two curves in the domain of $\mathbf{F}$ that share the same endpoints, then

$$\int_{\gamma_1} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_2} \mathbf{F}\cdot d\mathbf{r},$$

that is, $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$ is *path independent*. The next theorem says that the converse of this observation is also true: if the line integrals of $\mathbf{F}$ are path independent, then $\mathbf{F}$ is conservative.

THEOREM 22.2. *Let* $\mathbf{F}$ *be a vector field that is continuous in a open, connected region D. If the line integral* $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$ *is path independent in D, then* $\mathbf{F}$ *is conservative in D.*

Note that unlike Theorems 20.2 and 20.4, this theorem requires less of the set $D$: only that it be open and connected (as opposed to simply connected). On the other hand, the test that it provides is much more elusive: we must determine that **all** (infinitely many) line integrals of $\mathbf{F}$ are path independent. This hardly seems to be a practical test. And yet, if you look up the proofs of Theorems 20.2 and 20.4 in an advanced calculus text, you will discover that those results are deduced from Theorem 22.2. Thus, every time we show that some vector field is conservative by means of Theorems 20.2 or 20.4, we implicitly make use of Theorem 22.2 too.

We now return to the vector field

$$\mathbf{F}(x,y) = \frac{-y}{x^2+y^2}\mathbf{e}_1 + \frac{x}{x^2+y^2}\mathbf{e}_2,$$

which we studied in Example 20.6(c). Recall that this field is defined and differentiable everywhere but at the origin. Furthermore, we have

$$\frac{\partial}{\partial y}\left(\frac{-y}{x^2+y^2}\right) = \frac{\partial}{\partial x}\left(\frac{x}{x^2+y^2}\right) = \frac{y^2-x^2}{(x^2+y^2)^2},$$

also for all $(x, y) \neq (0, 0)$. Back in Lecture #20, we showed that this vector field is conservative on the plane cut along a half-line passing through the origin, but were unable to decide whether $\mathbf{F}$ is conservative in its entire domain. The next example shows that, in fact, this vector field is not conservative in its full domain.

Example 22.2. The vector field $\mathbf{F}(x, y) = \dfrac{-y}{x^2 + y^2}\mathbf{e}_1 + \dfrac{x}{x^2 + y^2}\mathbf{e}_2$ is not conservative.

Solution. Let $\gamma$ be the unit circle

$$x = \cos t, \qquad y = \sin t \quad (0 \le t \le 2\pi).$$

By (21.9) and (21.10),

$$\begin{aligned}
\int_\gamma \mathbf{F} \cdot d\mathbf{r} &= \int_\gamma \frac{-y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dy \\
&= \int_0^{2\pi} \left( \frac{(-\sin t)(\cos t)'}{\cos^2 t + \sin^2 t} + \frac{(\cos t)(\sin t)'}{\cos^2 t + \sin^2 t} \right) dt \\
&= \int_0^{2\pi} \frac{(-\sin t)^2 + (\cos t)^2}{\cos^2 t + \sin^2 t}\, dt = 2\pi.
\end{aligned}$$

On the other hand, if $\mathbf{F}$ was conservative, with potential function, say $f(x, y)$, Theorem 22.1 would give

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = f(\cos 2\pi, \sin 2\pi) - f(\cos 0, \sin 0) = f(1, 0) - f(1, 0) = 0.$$

Since this value differs from the value of the integral we got via direct calculation, $\mathbf{F}$ must be non-conservative. □

Example 22.3. Evaluate

$$\int_\gamma \frac{-y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dy,$$

where $\gamma$ is the circle $(x - 2)^2 + (y + 1)^2 = 25$, oriented counterclockwise.

First solution (not recommended). We just showed that the vector field

$$\mathbf{F}(x, y) = \frac{-y}{x^2 + y^2}\mathbf{e}_1 + \frac{x}{x^2 + y^2}\mathbf{e}_2$$

is conservative on the plane with a half-line removed and non-conservative on the plane with just the origin removed. Thus, $\mathbf{F}$ is conservative on curves that don't loop around the origin and non-conservative on curves that do. In particular, $\mathbf{F}$ is not conservative on $\gamma$, since the origin lies inside it. This means that we can't use Theorem 22.1 to evaluate the given integral, so we go back to the methods from the last lecture.

A point $(a, b)$ is on $\gamma$ exactly when $(a - 2, b + 1)$ is on the circle $x^2 + y^2 = 25$. Thus, we can derive the parametrization of $\gamma$ from that of the latter circle: the parametric equations of $\gamma$ are

$$x = 2 + 5\cos t, \quad y = -1 + 5\sin t \qquad (0 \le t \le 2\pi).$$

Denote the given line integral by $I$. By (21.9) and (21.10),

$$\begin{aligned} I &= \int_0^{2\pi} \left( \frac{(1-5\sin t)(2+5\cos t)'}{(2+5\cos t)^2 + (-1+5\sin t)^2} + \frac{(2+5\cos t)(-1+5\sin t)'}{(2+5\cos t)^2 + (-1+5\sin t)^2} \right) dt \\ &= \int_0^{2\pi} \frac{(1-5\sin t)(-5\sin t) + (2+5\cos t)(5\cos t)}{4 + 20\cos t + 25\cos^2 t + 1 - 10\sin t + 25\sin^2 t}\, dt \\ &= \int_0^{2\pi} \frac{25 - 5\sin t + 10\cos t}{30 - 10\sin t + 20\cos t}\, dt = \int_0^{2\pi} \frac{5 - \sin t + 2\cos t}{6 - 2\sin t + 4\cos t}\, dt. \end{aligned}$$

Most likely, you have not seen integrals like the last in single-variable calculus. It can be reduced to an integral of a rational function by the substitution $u = \tan(t/2)$, but we must be extremely careful, because this substitution is not "well-behaved". We must even be careful, if we use mathematical software or an integral table. For example, *Mathematica* reports that

$$\int \frac{5 - \sin t + 2\cos t}{6 - 2\sin t + 4\cos t}\, dt = \frac{t}{2} + \arctan\left( \frac{2}{1 - \tan(t/2)} \right) + C,$$

which would suggest that

$$\int_0^{2\pi} \frac{5 - \sin t + 2\cos t}{6 - 2\sin t + 4\cos t}\, dt = \left[ \frac{t}{2} + \arctan\left( \frac{2}{1 - \tan(t/2)} \right) \right]_0^{2\pi} = \pi,$$

... if it wasn't for the fact that *Mathematica* also proudly proclaims that

$$\int_0^{2\pi} \frac{5 - \sin t + 2\cos t}{6 - 2\sin t + 4\cos t}\, dt = 2\pi.$$

In fact, both answers that *Mathematica* returns are correct and our calculation between them is flawed, but this comes to show that we may want to avoid this approach... □

Second solution (recommended). Recall that in the solution of the last example, we showed that

$$\int_{\gamma_0} \frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2}\, dy = 2\pi,$$

where $\gamma_0$ is the unit circle taken in the counterclockwise direction. It turns out that, although we cannot appeal to Theorem 22.1 to evaluate the given line integral directly, we can use it to show that

$$\int_{\gamma} \frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2}\, dy = \int_{\gamma_0} \frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2}\, dy.$$

Let $\gamma^+$ be the part of $\gamma$ that lies above the $x$-axis and let $\tau^+$ be the curve that has the same endpoints as $\gamma^+$: $(2 \pm \sqrt{24}, 0)$, but consists of the line segment from $(2+\sqrt{24}, 0)$ to $(1,0)$, the upper half $\gamma_0^+$ of the unit circle $\gamma_0$, and the line segment from $(-1,0)$ to $(2-\sqrt{24}, 0)$ (see Figure 22.2). Also, let $D^+$ be the plane cut along the negative $y$-axis (dashed line). Since $\gamma^+$ and $\tau^+$ lie in $D^+$ and $\mathbf{F}$ is conservative in $D^+$, Theorem 22.2 gives

$$\int_{\gamma^+} \frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2}\, dy = \int_{\tau^+} \frac{-y}{x^2+y^2}\, dx + \frac{x}{x^2+y^2}\, dy.$$

Next, let $\gamma^-$ be the part of $\gamma$ that lies below the $x$-axis and let $\tau^-$ be the curve that consists of the line segment from $(2-\sqrt{24}, 0)$ to $(-1,0)$, the lower half $\gamma_0^-$ of the unit circle $\gamma_0$, and the line

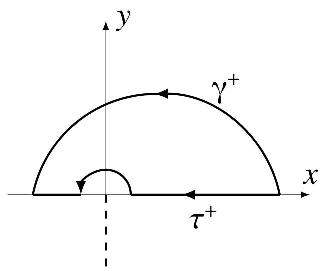



FIGURE 22.2. The curves $\gamma^+$ and $\tau^+$ in $D^+$

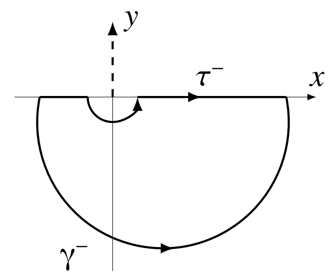



FIGURE 22.3. The curves $\gamma^-$ and $\tau^-$ in $D^-$

segment from $(1,0)$ to $(2+\sqrt{24},0)$ (see Figure 22.3). Also, let $D^-$ be the plane cut along the positive $y$-axis (dashed line). Since $\gamma^-$ and $\tau^-$ lie in $D^-$ and $\mathbf{F}$ is conservative in $D^-$, Theorem 22.2 gives

$$\int_{\gamma^-} \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = \int_{\tau^-} \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy.$$

Combining this identity with the one before it, we get

$$\int_{\gamma} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma^+} \mathbf{F}\cdot d\mathbf{r} + \int_{\gamma^-} \mathbf{F}\cdot d\mathbf{r} = \int_{\tau^+} \mathbf{F}\cdot d\mathbf{r} + \int_{\tau^-} \mathbf{F}\cdot d\mathbf{r}.$$

We break the integrals over $\tau^+$ and $\tau^-$ into six integrals over $\gamma_0^+$, $\gamma_0^-$, and the two horizontal edges. Each horizontal edge appears in the sum twice: once oriented from left to right and once oriented from right to left. Thus, those contributions cancel out (recall (21.8)) and we get

$$\int_{\tau^+} \mathbf{F}\cdot d\mathbf{r} + \int_{\tau^-} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_0^+} \mathbf{F}\cdot d\mathbf{r} + \int_{\gamma_0^-} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_0} \mathbf{F}\cdot d\mathbf{r}.$$

It follows that, as promised,

$$\int_{\gamma} \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = \int_{\gamma_0} \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 2\pi.$$

□

## Exercises

Find a potential for the given vector field and then use the fundamental theorem to evaluate $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$.

22.1. $\mathbf{F}(x,y) = e^y\mathbf{e}_1 + xe^y\mathbf{e}_2$; $\gamma$ is a smooth curve connecting $(1,0)$ and $(2,-1)$

22.2. $\mathbf{F}(x,y) = (4x+\sin y)\mathbf{e}_1 + (\sin y + x\cos y)\mathbf{e}_2$; $\gamma$ is given by $\mathbf{r}(t) = (t-\sin t)\mathbf{e}_1 + (1-\cos t)\mathbf{e}_2$, $0\le t\le 2\pi$

22.3. $\mathbf{F}(x,y) = y\mathbf{e}_1 + (x+3y)\mathbf{e}_2$; $\gamma$ follows the circle $x^2+y^2=5$ from $(1,2)$ to $(-\sqrt{5},0)$ and then the line segment from $(-\sqrt{5},0)$ to $(1,-2)$

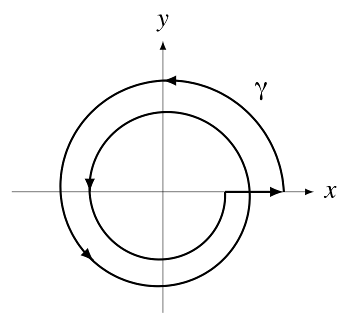
22.4. $\mathbf{F}(x,y,z) = (yz+x^2)\mathbf{e}_1 + (xz+4y)\mathbf{e}_2 + (xy-3z^2)\mathbf{e}_3$; $\gamma$ consists of the line segments from $(1,2,3)$ to $(0,2,1)$ to $(-1,3,0)$ to $(3,2,1)$

22.5. $\mathbf{F}(x,y,z) = y^2\cos z\,\mathbf{e}_1 + 2xy\cos z\,\mathbf{e}_2 - xy^2\sin z\,\mathbf{e}_3$; $\gamma$ is given by $\mathbf{r}(t) = t^2\mathbf{e}_1 + t^3\mathbf{e}_2 - 2t^2\mathbf{e}_3$, $0\le t\le 2$

22.6. Evaluate

$$\int_\gamma \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy,$$

where $\gamma$ is the curve from the figure to the right.



LECTURE 23

# Green's Theorem

In this lecture we discuss the first of the three major theorems of vector calculus. It is known as *Green's theorem* and provides an alternative technique for evaluation of certain line integrals.

## 23.1. Some definitions

First, we need to fix some terminology. A curve $\gamma$, given by $\mathbf{r}(t)$, $a \le t \le b$, is called *closed*, if $\mathbf{r}(a) = \mathbf{r}(b)$. In other words, the curve is closed, if its initital point is also its terminal point (see Figure 23.1). A curve $\gamma$ is called *simple*, if $\mathbf{r}(t_1) \ne \mathbf{r}(t_2)$ whenever $t_1 \ne t_2$, except that we

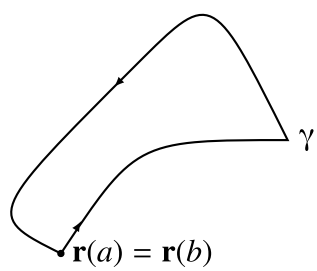



FIGURE 23.1. A closed, positively oriented curve

allow $\mathbf{r}(a) = \mathbf{r}(b)$. In other words, a simple curve does not self-intersect and does not touch itself, but can be closed. For example, among the curves displayed on Figure 23.2, those in (a) and (e) are simple and those in (b)–(d) are not: the curve in (b) self-intersects at point $A$; the curve in (c) passes through its initial point before reaching its terminal point; and the curve in (d) touches itself at point $B$.

It is intuitively clear (though quite difficult to prove) that a piecewise smooth, simple, closed plane curve $\gamma$ divides the plane into two sets, which we can call intuitively the *interior* and the *exterior* of $\gamma$, the interior being the bounded set and the exterior the unbounded. This fact is known as *Jordan's theorem*, named after the French mathematician Camille Jordan for his work on plane curves, including this result. However, the first full proof of Jordan's theorem was given by an American: Oswald Veblen.

Let $\gamma$ be a piecewise smooth, simple, closed curve. We say that $\gamma$ is *positively oriented*, if its interior lies always to the left of the point $\mathbf{r}(t)$ as it traces the curve in the direction of increasing of $t$. For "round" curves, such as a circle or the curve on Figure 23.2(e), the positive orientation of the curve is the counterclockwise orientation. However, this is not the case for all curves: the positive orientation goes clockwise for the part of the curve on Figure 23.1 that follows immediately the initial point $\mathbf{r}(a)$.

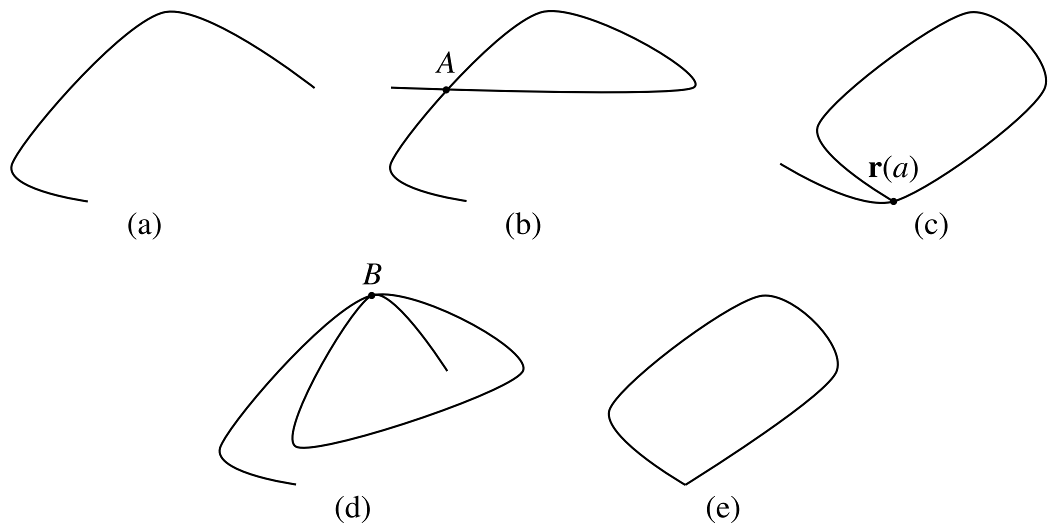



FIGURE 23.2. Simple and non-simple curves

## 23.2. Green's theorem

THEOREM 23.1 (Green). *Let $P(x, y)$ and $Q(x, y)$ have continuous first-order partials on an open set $U$. Let $\gamma$ be a positively oriented, piecewise smooth, simple, closed plane curve such that $\gamma$ and* ***all*** *of its interior $D$ are contained inside $U$. Then*

$$\int_{\gamma} P\,dx + Q\,dy = \iint_{D} \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA. \tag{23.1}$$

In other words, Green's theorem relates a line integral along a simple, closed plane curves to a double integral over the interior of the curve. Since we know how to evaluate double integrals this yields another possible integration technique for line integrals along closed curves. However, there is a catch: we absolutely **must** check the conditions of Green's theorem before we use it. For example, we showed in Example 20.6(c) that the functions

$$P(x, y) = \frac{-y}{x^2 + y^2}, \quad Q(x, y) = \frac{x}{x^2 + y^2}$$

are differentiable and satisfy the condition $\partial P/\partial y = \partial Q/\partial x$ everywhere except at the origin. Thus, if $\gamma$ is a simple, closed curve whose interior $D$ does not contain the origin, Green's theorem gives

$$\int_{\gamma} P\,dx + Q\,dy = \iint_{D} \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA = 0.$$

Note that for such a curve the application of Green's theorem is legitimate. On the other hand, if the origin is in the interior of the curve $\gamma$, the conditions of Green's theorem are violated (at one point) and the theorem is not applicable. Were we to apply it anyways, the result would again be that the line integral along $\gamma$ is 0. For example, such an illegitimate application of Green's theorem yields

$$\int_{\gamma_0} \frac{-y}{x^2 + y^2}\,dx + \frac{x}{x^2 + y^2}\,dy = 0$$

when $\gamma_0$ is the unit circle $x^2 + y^2 = 1$. However, we know from Example 22.2 that the last integral equals $2\pi$, not 0. We see that if the functions $P(x, y)$ and $Q(x, y)$ or some of their derivatives are not defined at even a single point in the interior of the curve, Green's theorem is not applicable and should not be used. Now, let us take a look at some of its legitimate uses.

Example 23.1. Evaluate

$$\int_\gamma \left(e^{4x} \sin x + 2y\right) dx + \left(x^2 + \arctan y\right) dy,$$

where $\gamma$ is the rectangle with vertices $(1, 2)$, $(5, 2)$, $(5, 4)$, $(1, 4)$.

Solution. We have $P(x, y) = e^{4x} \sin x + 2y$ and $Q(x, y) = x^2 + \arctan y$. The partials of these functions are:

$$P_x(x, y) = (4 \sin x + \cos x)e^{4x}, \quad P_y(x, y) = 2, \quad Q_x(x, y) = 2x, \quad Q_y(x, y) = \frac{1}{1 + y^2}.$$

Since these are all continuous everywhere, the open set $U$ in the statement of Green's theorem can be taken to be $\mathbb{R}^2$. We can therefore apply Green's theorem. We get

$$\begin{aligned}\int_C \left(e^{4x} \sin x + 2y\right) dx + \left(x^2 + \arctan y\right) dy &= \iint_D (2x - 2)\, dA = \int_1^5 \int_2^4 (2x - 2)\, dydx \\ &= \int_1^5 2(2x - 2)\, dx = 2\left[x^2 - 2x\right]_1^5 = 32.\end{aligned}$$

Here, $D$ denotes the rectangle $[1, 5] \times [2, 4]$. □

Example 23.2. Evaluate

$$\int_\gamma \left(e^{4x} \sin x + 2y\right) dx + \left(x^2 + \arctan y\right) dy,$$

where $\gamma$ is the boundary of the semiannular region shown on Figure 23.3 (the radii of the two semicircles are 2 and 3).

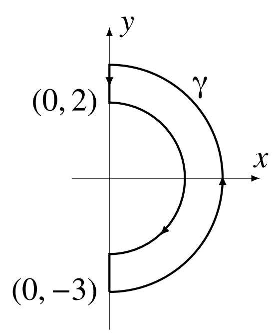



Figure 23.3. The region $D$

Solution. The integrand is the same as in the previous problem, but the curve is different. We can appeal to Green's theorem for the same reasons as in the previous example. We get

$$\int_\gamma \left(e^{4x} \sin x + 2y\right) dx + \left(x^2 + \arctan y\right) dy = \iint_D (2x - 2)\, dA,$$

where $D$ is the semiannular region shown on Figure 23.3. Note that this region is conveniently described in polar coordinates:

$$D = \left\{(r, \theta) : 2 \leq r \leq 3,\ -\pi/2 \leq \theta \leq \pi/2\right\}.$$

Thus, by converting the double integral to polar coordinates, we obtain

$$\begin{aligned}
\int_\gamma \left(e^{4x}\sin x + 2y\right) dx + \left(x^2 + \arctan y\right) dy &= \int_2^3 \int_{-\pi/2}^{\pi/2} (2r\cos\theta - 2) r\, d\theta dr \\
&= \int_2^3 \left[(2r\sin\theta - 2\theta) r\right]_{-\pi/2}^{\pi/2} dr \\
&= \int_2^3 \left(4r^2 - 2\pi r\right) dr = 25\tfrac{1}{3} - 5\pi.
\end{aligned}$$

□

In the above examples, we used Green's theorem to evaluate line integrals by means of double integrals. However, Green's theorem is a relation between the line integral and the double integral appearing in (23.1) that can be used in either direction. Sometimes, we want to use Green's theorem to evaluate a double integral by means of a line integral. For example, under the assumption of Theorem 23.1, we have

$$\int_\gamma x\, dy = \iint_D \left(\partial_x(x) - \partial_y(0)\right) dA = \iint_D 1\, dA = \operatorname{area}(D),$$

$$\int_\gamma (-y)\, dx = \iint_D \left(\partial_x(0) - \partial_y(-y)\right) dA = \iint_D 1\, dA = \operatorname{area}(D).$$

In other words, we have the following new formulas for the area enclosed by a simple, closed curve.

Theorem 23.2. *Let $D$ be a plane region, bounded by a piecewise smooth, simple, closed curve $\gamma$. Then*

$$\operatorname{area}(D) = \int_\gamma x\, dy = \int_\gamma (-y)\, dx = \frac{1}{2}\int_\gamma x\, dy - y\, dx. \tag{23.2}$$

Example 23.3. Find the area enclosed by the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.

Solution. We will use formula (23.2) and the parametric representation of the ellipse:

$$x = a\cos t, \qquad y = b\sin t \qquad (0 \leq \theta \leq 2\pi).$$

By (23.2),

$$\begin{aligned}
\operatorname{area}(D) &= \frac{1}{2}\int_C x\, dy - y\, dx = \frac{1}{2}\int_0^{2\pi} \left[(a\cos t)(b\cos t) - (b\sin t)(-a\sin t)\right] dt \\
&= \frac{1}{2}\int_0^{2\pi} ab\left(\cos^2 t + \sin^2 t\right) dt = \frac{1}{2}\int_0^{2\pi} ab\, dt = \pi ab.
\end{aligned}$$

□

### 23.3. Vector forms of Green's theorem*

We can use the curl and divergence of a vector field to give alternative formulations of Green's theorem. Suppose that $\gamma$ is a positively oriented, piecewise smooth, simple, closed curve with parametrization

$$\mathbf{r}(t) = x(t)\mathbf{e}_1 + y(t)\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (a \le t \le b),$$

that is, $\gamma$ is a plane curve viewed as a space curve. Furhter, consider the vector field

$$\mathbf{F}(x, y, z) = P(x, y)\mathbf{e}_1 + Q(x, y)\mathbf{e}_2 + 0\mathbf{e}_3.$$

By (21.12),

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_\gamma P(x, y)\, dx + Q(x, y)\, dy.$$

Also, by the definition of curl,

$$\operatorname{curl} \mathbf{F} = \left(\frac{\partial 0}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{e}_1 + \left(\frac{\partial P}{\partial z} - \frac{\partial 0}{\partial x}\right)\mathbf{e}_2 + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{e}_3 = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{e}_3,$$

because $P(x, y)$ and $Q(x, y)$ do not depend on $z$. Thus, we can rewrite (23.1) as

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \iint_D (\operatorname{curl} \mathbf{F}) \cdot \mathbf{e}_3\, dA. \tag{23.3}$$

Another formulation of Green's theorem involves the vector field

$$\mathbf{G}(x, y, z) = Q(x, y)\mathbf{e}_1 - P(x, y)\mathbf{e}_2 + 0\mathbf{e}_3,$$

which is orthogonal to the vector field $\mathbf{F}$ that we used in (23.3):

$$\mathbf{F} \cdot \mathbf{G} = P(x, y)Q(x, y) + Q(x, y)(-P(x, y)) + 0 = 0.$$

Note that

$$\operatorname{div} \mathbf{F} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Also, by (21.7),

$$\int_\gamma \mathbf{F} \cdot \mathbf{r} = \int_\gamma (\mathbf{F} \cdot \mathbf{T})\, ds,$$

where $\mathbf{T}$ is the unit tangent vector to $\gamma$. It is a nice exercise to try to use the properties of the dot product to convince yourself that

$$\mathbf{n}(t) = \frac{y'(t)\mathbf{e}_1 - x'(t)\mathbf{e}_2}{\sqrt{x'(t)^2 + y'(t)^2}}$$

is a unit vector that is normal to the curve $\gamma$ and points towards the exterior of the curve; it is called the *outward normal* vector of $\gamma$. Since $\mathbf{F} \cdot \mathbf{T} = \mathbf{G} \cdot \mathbf{n}$, we get

$$\int_\gamma \mathbf{F} \cdot \mathbf{r} = \int_\gamma (\mathbf{G} \cdot \mathbf{n})\, ds,$$

and we can restate (23.1) as

$$\int_\gamma (\mathbf{G} \cdot \mathbf{n})\, ds = \iint_D \operatorname{div} \mathbf{G}\, dA. \tag{23.4}$$

## Exercises

23.1. (a) Let $\gamma$ be the square with vertices $(1,1)$, $(2,1)$, $(2,2)$, and $(1,2)$. Parametrize $\gamma$ and use the formulas from Lecture #21 to evaluate

$$\int_\gamma x^2 y\,dx + xy^2\,dy.$$

(b) Evaluate the integral from part (a) by means of Green's theorem.

Use Green's theorem to evaluate the given line integral (all the curves are assumed positively oriented).

23.2. $\int_\gamma y^3\,dx - x^3\,dy$; $\gamma$ is the circle $x^2 + y^2 = 4$

23.3. $\int_\gamma (xy + \tan x)\,dx + \left(x^2 + e^{\cos y}\right) dy$; $\gamma$ is the boundary of the region bounded by $y = x$ and $y = x^2$

23.4. $\int_\gamma xy^3\,dx + 4x^2y^2\,dy$; $\gamma$ consists of the semicircle $x = \sqrt{9 - y^2}$ and the line segment from $(0,3)$ to $(0,-3)$

23.5. $\int_\gamma \left(xy^2 + \sin x^2\right) dx + \left((x^2 + y^2)^{3/2} + \arctan y\right) dy$; $\gamma$ is the boundary of region bounded by the semicircle $y = \sqrt{4 - x^2}$ and the lines $y = x$ and $y = -\sqrt{3}x$

23.6. $\int_\gamma e^x \sin y\,dx + e^x \cos y\,dy$; $\gamma$ is the ellipse $(x+1)^2 + 5(y-2)^2 = 7$

Use Green's theorem to evaluate the line integral $\int_\gamma \mathbf{F} \cdot d\mathbf{r}$ for the given field $\mathbf{F}(x,y)$ and the given curve $\gamma$.

23.7. $\mathbf{F}(x,y) = \left(y^{3/2} + x\right)\mathbf{e}_1 + \left(x^2 + y\right)\mathbf{e}_2$; $\gamma$ consists of the graph of $y = \sin^2 x$ from $(0,0)$ to $(\pi,0)$ and the line segment from $(\pi,0)$ to $(0,0)$

23.8. $\mathbf{F}(x,y) = \left(x^3 + y\right)\mathbf{e}_1 + \left(2xy + y^4\right)\mathbf{e}_2$; $\gamma$ is the cardioid $r = 2 + \cos\theta$, positively oriented

23.9. $\mathbf{F}(x,y) = x\mathbf{e}_1 + \left(x^3 + 3xy^2\right)\mathbf{e}_2$; $\gamma$ consists of the graph of $y = 4 - x^2$ from $(-2,0)$ to $(2,0)$ and the line segment from $(2,0)$ to $(-2,0)$

23.10. When the circle $x^2 + (y-1)^2 = 1$ is rolled along the $x$-axis in the positive direction, the point whose original position is the origin traces a curve called a *cycloid* (see Figure 23.4). This curve can be parametrized by the vector function

$$\mathbf{r}(t) = (t - \sin t)\mathbf{e}_1 + (1 - \cos t)\mathbf{e}_2 \qquad (t \geq 0).$$

Use Theorem 23.2 to find the area bounded by the $x$-axis and the first arc of the cycloid (that is, the part of the cycloid with $0 \leq t \leq 2\pi$).

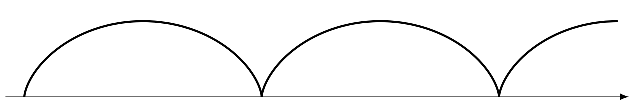


FIGURE 23.4. Cycloid

23.11. When a circle of radius 1 is rolled along the outside of the circle $x^2 + y^2 = 4$ a fixed point on the smaller circle traces a curve called an *epicycloid* (see Figure 23.5). This curve can be parametrized by the vector function

$$\mathbf{r}(t) = (3\cos t - \cos 3t)\mathbf{e}_1 + (3\sin t - \sin 3t)\mathbf{e}_2 \qquad (0 \leq t \leq 2\pi).$$

Use Theorem 23.2 to find the area enclosed by this epicycloid.

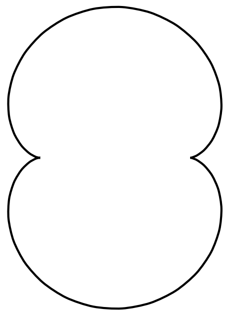


FIGURE 23.5. Epicycloid

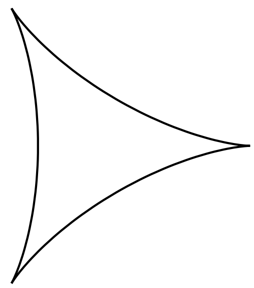


FIGURE 23.6. Hypocycloid

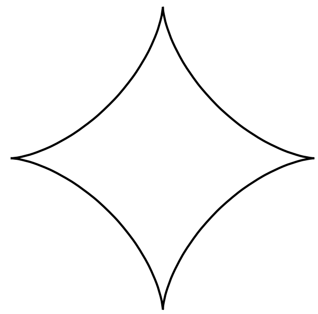


FIGURE 23.7. Astroid

23.12. When a circle of radius 1 is rolled along the inside of the circle $x^2 + y^2 = 9$ a fixed point on the smaller circle traces a curve called a *hypocycloid* (see Figure 23.6). This curve can be parametrized by the vector function

$$\mathbf{r}(t) = (2\cos t + \cos 2t)\mathbf{e}_1 + (2\sin t - \sin 2t)\mathbf{e}_2 \qquad (0 \le t \le 2\pi).$$

Use Theorem 23.2 to find the area enclosed by this hypocycloid.

23.13. When a circle of radius 1 is rolled along the inside of the circle $x^2 + y^2 = 16$ a fixed point on the smaller circle traces a different kind of hypocycloid called also an *astroid* (see Figure 23.7). This curve can be parametrized by the vector function

$$\mathbf{r}(t) = (3\cos t + \cos 3t)\mathbf{e}_1 + (3\sin t - \sin 3t)\mathbf{e}_2 \qquad (0 \le t \le 2\pi).$$

Use Theorem 23.2 to find the area enclosed by this astroid.

LECTURE 24

# Surface Integrals

## 24.1. Surface integrals of scalar functions

We start with a smooth parametric surface $\Sigma$, given by the vector function

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v) \in D).$$

Recall that the smoothness of $\Sigma$ means that $\mathbf{r}_u(u,v)$ and $\mathbf{r}_v(u,v)$ are continuous and $(\mathbf{r}_u \times \mathbf{r}_v)(u,v) \neq \mathbf{0}$ in $D$ (see §10.3). Further, suppose that $f(x,y,z)$ is a continuous function on a set $S$ in $\mathbb{R}^3$ that contains the surface $\Sigma$. In particular, $f(x,y,z)$ is defined at each point $(x,y,z)$ on $\Sigma$.

For simplicity, we shall assume that $D$ is a rectangle in the $uv$-plane: $D = [a,b] \times [c,d]$. We can treat the general case similarly, but there are more technical details to pay attention to and little extra insight to gain. For each $n \geq 1$, let $a = u_0 < u_1 < \cdots < u_n = b$ be the points that divide the interval $[a,b]$ into $n$ subintervals of equal lengths; let $c = v_0 < v_1 < \cdots < v_n = d$ be the points that divide the interval $[c,d]$ into $n$ subintervals of equal lengths. The points

$$\begin{array}{cccccc}
(u_n,v_0) & (u_n,v_1) & \cdots & (u_n,v_j) & \cdots & (u_n,v_n), \\
\vdots & \vdots & & \vdots & & \vdots \\
(u_i,v_0) & (u_i,v_1) & \cdots & (u_i,v_j) & \cdots & (u_i,v_n), \\
\vdots & \vdots & & \vdots & & \vdots \\
(u_1,v_0) & (u_1,v_1) & \cdots & (u_1,v_j) & \cdots & (u_1,v_n), \\
(u_0,v_0) & (u_0,v_1) & \cdots & (u_0,v_j) & \cdots & (u_0,v_n),
\end{array}$$

define a partition of the rectangle $D$ into $n^2$ congruent subrectangles. The vector function $\mathbf{r}$ maps each of those small rectangles to a small "patch" on $\Sigma$ and each grid point $(u_i,v_j)$ to a point $P_{ij}$ on the surface. The patches form a partition of $\Sigma$ into $n^2$ pieces, which we label as $\Sigma_{ij}$, $1 \leq i,j \leq n$, so that $\Sigma_{ij}$ is the image of the rectangle that has $(u_i,v_j)$ as its upper right corner point. Further, we pick $n^2$ sample points $(u_{ij}^*, v_{ij}^*)$ so that

$$u_{i-1} \leq u_{ij}^* \leq u_i, \quad v_{j-1} \leq v_{ij}^* \leq v_j.$$

These yield $n^2$ *sample points* $P_{ij}^*$ on the surface, with coordinates $x_{ij}^* = x(u_{ij}^*, v_{ij}^*)$, $y_{ij}^* = y(u_{ij}^*, v_{ij}^*)$, $z_{ij}^* = z(u_{ij}^*, v_{ij}^*)$. Then the *surface integral of $f$ over* $\Sigma$ is defined by

$$\iint_\Sigma f(x,y,z)\,dS = \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n f(x_{ij}^*, y_{ij}^*, z_{ij}^*)\Delta S_{ij}, \qquad \Delta S_{ij} = \operatorname{area}(\Sigma_{ij}), \tag{24.1}$$

provided that the limit exists.

Recall that, by Theorem 18.1,

$$\Delta S_{ij} = \iint_{D_{ij}} \|\mathbf{r}_u \times \mathbf{r}_v\|\,dA,$$

where $D_{ij}$ is the rectangle in the $uv$-plane whose image is $\Sigma_{ij}$. Using this formula, the above definition, the properties of continuous functions, and the definition of the double integral, it can be shown that for continuous functions $f$ the limit (24.1) always exists and the surface integral can be expressed as a double integral.

THEOREM 24.1. *Let $\Sigma$ be a smooth parametric surface, given by the vector function*

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v) \in D).$$

*If $f(x,y,z)$ is continuous on $\Sigma$, then*

$$\iint_\Sigma f(x,y,z)\,dS = \iint_D f(\mathbf{r}(u,v))\|\mathbf{r}_u \times \mathbf{r}_v\|\,dA. \tag{24.2}$$

Similarly to line integrals, we can extend the above definition of the surface integral to piecewise smooth surfaces. If $\Sigma$ is a piecewise smooth surface that is the union of the smooth surfaces $\Sigma_1, \Sigma_2, \dots, \Sigma_n$, $\Sigma = \Sigma_1 \cup \Sigma_2 \cup \cdots \cup \Sigma_n$, then we define

$$\iint_\Sigma f(x,y,z)\,dS = \iint_{\Sigma_1} f(x,y,z)\,dS + \iint_{\Sigma_2} f(x,y,z)\,dS + \cdots + \iint_{\Sigma_n} f(x,y,z)\,dS.$$

REMARK. Like all other types of integrals we have encountered so far, the surface integral can be interpreted in terms of averages. In this case, we are dealing with the average value of the function on the surface:

$$\operatorname*{average}_{(x,y,z)\in\Sigma} f(x,y,z) = \frac{1}{\text{area}(\Sigma)} \iint_\Sigma f(x,y,z)\,dS. \tag{24.3}$$

EXAMPLE 24.1. Evaluate

$$\iint_\Sigma xy\,dS,$$

where $\Sigma$ is the graph of $z = \frac{2}{3}\left(x^{3/2} + y^{3/2}\right)$ over the rectangle $0 \le x, y \le 1$.

SOLUTION. The surface can be parametrized by the vector function

$$\mathbf{r}(u,v) = u\mathbf{e}_1 + v\mathbf{e}_2 + \tfrac{2}{3}\left(u^{3/2} + v^{3/2}\right)\mathbf{e}_3 \qquad (0 \le u, v \le 1).$$

Hence,

$$\mathbf{r}_u(u,v) = \mathbf{e}_1 + 0\mathbf{e}_2 + u^{1/2}\mathbf{e}_3, \qquad \mathbf{r}_v(u,v) = 0\mathbf{e}_1 + \mathbf{e}_2 + v^{1/2}\mathbf{e}_3,$$

and

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ 1 & 0 & u^{1/2} \\ 0 & 1 & v^{1/2} \end{vmatrix} = \begin{vmatrix} 0 & u^{1/2} \\ 1 & v^{1/2} \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} 1 & u^{1/2} \\ 0 & v^{1/2} \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} \mathbf{e}_3 = -u^{1/2}\mathbf{e}_1 - v^{1/2}\mathbf{e}_2 + \mathbf{e}_3.$$

Applying (24.2), we get

$$\iint_S xy\,dS = \iint_{0\le u,v\le 1} uv\sqrt{u+v+1}\,dA = \int_0^1 \left[ v \int_0^1 u\sqrt{u+v+1}\,du \right] dv. \tag{24.4}$$

Using the integration by parts, we find that

$$\begin{aligned}\int_0^1 u\sqrt{u+v+1}\,du &= \tfrac{2}{3}\int_0^1 u\,d(u+v+1)^{3/2}\\ &= \left[\tfrac{2}{3}u(u+v+1)w^{3/2}\right]_0^1 - \tfrac{2}{3}\int_0^1 (u+v+1)^{3/2}\,du\\ &= \tfrac{2}{3}(v+2)^{3/2} - \tfrac{2}{3}\left[\tfrac{2}{5}(u+v+1)^{5/2}\right]_0^1\\ &= \tfrac{2}{3}(v+2)^{3/2} - \tfrac{4}{15}(v+2)^{5/2} + \tfrac{4}{15}(v+1)^{5/2}.\end{aligned}$$

Inserting the result back into (24.4) and using another integration by parts, we obtain

$$\begin{aligned}\iint_S xy\,dS &= \int_0^1 v\left[\tfrac{2}{3}(v+2)^{3/2} - \tfrac{4}{15}(v+2)^{5/2} + \tfrac{4}{15}(v+1)^{5/2}\right]dv\\ &= \int_0^1 vd\left(\tfrac{4}{15}(v+2)^{5/2} - \tfrac{8}{105}(v+2)^{7/2} + \tfrac{8}{105}(v+1)^{7/2}\right)\\ &= \left[v\left(\tfrac{4}{15}(v+2)^{5/2} - \tfrac{8}{105}(v+2)^{7/2} + \tfrac{8}{105}(v+1)^{7/2}\right)\right]_0^1\\ &\quad - \int_0^1 \left(\tfrac{4}{15}(v+2)^{5/2} - \tfrac{8}{105}(v+2)^{7/2} + \tfrac{8}{105}(v+1)^{7/2}\right)dv\\ &= \tfrac{4}{15}\left(3^{5/2} - \tfrac{2}{7}3^{7/2} + \tfrac{2}{7}2^{7/2}\right) - \tfrac{4}{15}\left[\tfrac{2}{7}(v+2)^{7/2} - \tfrac{4}{63}(v+2)^{9/2} + \tfrac{4}{63}(v+1)^{9/2}\right]_0^1\\ &= \tfrac{4}{105}\left(9\sqrt{3} + 16\sqrt{2}\right) - \tfrac{8}{105}\tfrac{1}{9}\left(81\sqrt{3} - 8\sqrt{2} - 2\right) \approx 0.372.\end{aligned}$$

□

## 24.2. Surface integrals of vector fields

Definition. Let $\mathbf{F}$ be a continuous vector field defined on an oriented surface $\Sigma$ with unit normal vector $\mathbf{n}$. Then the *surface integral of* $\mathbf{F}$ *over* $\Sigma$ is

$$\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} = \iint_\Sigma (\mathbf{F}\cdot\mathbf{n})\,dS. \tag{24.5}$$

Remark. Since $\mathbf{F}\cdot\mathbf{n}$ is the size of the projection of $\mathbf{F}$ in the direction of $\mathbf{n}$, we can interpret the surface integral of $\mathbf{F}$ as the average size of its projection onto the direction of orientation of $\Sigma$.

Suppose that the surface is given by the vector function

$$\mathbf{r}(u,v) = x(u,v)\mathbf{e}_1 + y(u,v)\mathbf{e}_2 + z(u,v)\mathbf{e}_3 \qquad ((u,v)\in D),$$

and that it is oriented using the unit normal vector

$$\mathbf{n}(u,v) = \frac{\mathbf{r}_u\times\mathbf{r}_v}{\|\mathbf{r}_u\times\mathbf{r}_v\|}.$$

Combining (24.5) and (24.2), we get

$$\begin{aligned}\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} = \iint_\Sigma (\mathbf{F}\cdot\mathbf{n})\,dS &= \iint_\Sigma \left(\mathbf{F}\cdot\frac{\mathbf{r}_u\times\mathbf{r}_v}{\|\mathbf{r}_u\times\mathbf{r}_v\|}\right)dS\\ &= \iint_D \left(\mathbf{F}\cdot\frac{\mathbf{r}_u\times\mathbf{r}_v}{\|\mathbf{r}_u\times\mathbf{r}_v\|}\right)\|\mathbf{r}_u\times\mathbf{r}_v\|\,dA = \iint_D \left(\mathbf{F}\cdot(\mathbf{r}_u\times\mathbf{r}_v)\right)dA.\end{aligned} \tag{24.6}$$

In particular, when $\Sigma$ is the graph $z = f(x, y)$, $(x, y) \in D$, with the upward orientation (recall Example 10.8), we obtain

$$\iint_{\Sigma} (P\mathbf{e}_1 + Q\mathbf{e}_2 + R\mathbf{e}_3) \cdot d\mathbf{S} = \iint_D \big( - P(x, y) f_x(x, y) - Q(x, y) f_y(x, y) + R(x, y) \big)\, dA. \tag{24.7}$$

Example 24.2. Evaluate

$$\iint_{\Sigma} (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3) \cdot d\mathbf{S},$$

where $\Sigma$ is the elliptic paraboloid $z = 4 - 2x^2 - 2y^2$, $z \geq 0$, with the upward orientation.

Solution. Condition $z \geq 0$ restricts $x, y$ to the circle

$$4 - 2x^2 - 2y^2 \geq 0 \quad \iff \quad x^2 + y^2 \leq 2.$$

Since we are using the upward orientation, we can appeal to (24.7). We have

$$\partial_x \big(4 - 2x^2 - 2y^2\big) = -4x, \quad \partial_y \big(4 - 2x^2 - 2y^2\big) = -4y,$$

so (24.7) gives

$$\begin{aligned}
\iint_{\Sigma} (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3) \cdot d\mathbf{S} &= \iint\limits_{x^2+y^2 \leq 2} \big(4x^2 - 8y^2 + z(x, y)\big)\, dA \\
&= \iint\limits_{x^2+y^2 \leq 2} \big(4x^2 - 8y^2 + 4 - 2x^2 - 2y^2\big)\, dA \\
&= \iint\limits_{x^2+y^2 \leq 2} \big(4 + 2x^2 - 10y^2\big)\, dA.
\end{aligned}$$

Converting the double integral to polar coordinates, we obtain

$$\begin{aligned}
\iint_{\Sigma} (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3) \cdot d\mathbf{S} &= \int_0^{2\pi} \int_0^{\sqrt{2}} \big(4 + 2r^2 \cos^2 \theta - 10r^2 \sin^2 \theta\big)\, r\, dr d\theta \\
&= \int_0^{2\pi} \Big[2r^2 + \tfrac{1}{2} r^4 \cos^2 \theta - \tfrac{5}{2} r^4 \sin^2 \theta\Big]_0^{\sqrt{2}}\, d\theta \\
&= \int_0^{2\pi} \big(4 + 2\cos^2 \theta - 10 \sin^2 \theta\big)\, d\theta \\
&= \int_0^{2\pi} \big(4 + (1 + \cos 2\theta) - 5(1 - \cos 2\theta)\big)\, d\theta = 0.
\end{aligned}$$

□

Example 24.3. Evaluate

$$\iint_{\Sigma} (xz\mathbf{e}_1 + x\mathbf{e}_2 + y\mathbf{e}_3) \cdot d\mathbf{S},$$

where $\Sigma$ is the hemisphere $x^2 + y^2 + z^2 = 25$, $y \geq 0$, oriented in the direction of the positive $y$-axis.

Solution. It is best to represent the hemisphere in spherical coordinates:

$$\mathbf{r}(\theta,\phi) = 5\cos\theta\sin\phi\mathbf{e}_1 + 5\sin\theta\sin\phi\mathbf{e}_2 + 5\cos\phi\mathbf{e}_3 \qquad (0 \le \theta,\phi \le \pi).$$

Then

$$\begin{aligned}
\mathbf{r}_\theta &= -5\sin\theta\sin\phi\mathbf{e}_1 + 5\cos\theta\sin\phi\mathbf{e}_2,\\
\mathbf{r}_\phi &= 5\cos\theta\cos\phi\mathbf{e}_1 + 5\sin\theta\cos\phi\mathbf{e}_2 - 5\sin\phi\mathbf{e}_3;\\
\mathbf{r}_\theta\times\mathbf{r}_\phi &= -25\sin^2\theta\sin\phi\cos\phi(\mathbf{e}_1\times\mathbf{e}_2) + 25\sin\theta\sin^2\phi(\mathbf{e}_1\times\mathbf{e}_3)\\
&\quad + 25\cos^2\theta\sin\phi\cos\phi(\mathbf{e}_2\times\mathbf{e}_1) - 25\cos\theta\sin^2\phi(\mathbf{e}_2\times\mathbf{e}_3)\\
&= -25\cos\theta\sin^2\phi\mathbf{e}_1 - 25\sin\theta\sin^2\phi\mathbf{e}_2 - 25\sin\phi\cos\phi\mathbf{e}_3\\
&= -25\sin\phi(\cos\theta\sin\phi\mathbf{e}_1 + \sin\theta\sin\phi\mathbf{e}_2 + \cos\phi\mathbf{e}_3).
\end{aligned}$$

Note that the vector $\mathbf{r}_\theta\times\mathbf{r}_\phi$ has a negative second component for all $\theta,\phi$ with $0 \le \theta,\phi \le \pi$. Therefore, we can orient $\Sigma$ in the direction of the positive $y$-axis by taking

$$\mathbf{n} = \frac{-(\mathbf{r}_\theta\times\mathbf{r}_\phi)}{\|\mathbf{r}_\theta\times\mathbf{r}_\phi\|} = \frac{\mathbf{r}_\phi\times\mathbf{r}_\theta}{\|\mathbf{r}_\phi\times\mathbf{r}_\theta\|}.$$

With this choice, we have

$$\begin{aligned}
\mathbf{F}(\mathbf{r}(\theta,\phi)) &= 25\cos\theta\sin\phi\cos\phi\mathbf{e}_1 + 5\cos\theta\sin\phi\mathbf{e}_2 + 5\sin\theta\sin\phi\mathbf{e}_3\\
&= 5\sin\phi(5\cos\theta\cos\phi\mathbf{e}_1 + \cos\theta\mathbf{e}_2 + \sin\theta\mathbf{e}_3),\\
\mathbf{F}(\mathbf{r}(\theta,\phi))\cdot\left(\mathbf{r}_\phi\times\mathbf{r}_\theta\right) &= 125\sin^2\phi\left(5\cos^2\theta\sin\phi\cos\phi + \sin\theta\cos\theta\sin\phi + \sin\theta\cos\phi\right).
\end{aligned}$$

Thus, (24.6) yields

$$\begin{aligned}
\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} &= \iint_{0\le\theta,\phi\le\pi} 125\sin^2\phi(5\cos^2\theta\sin\phi\cos\phi + \sin\theta\cos\theta\sin\phi + \sin\theta\cos\phi)\,dA\\
&= 125\int_0^\pi\int_0^\pi \left(5\cos^2\theta\sin^3\phi\cos\phi + \sin\theta\cos\theta\sin^3\phi + \sin\theta\sin^2\phi\cos\phi\right)d\theta d\phi.
\end{aligned}$$

We now use that

$$\int_0^\pi \cos^2\theta\,d\theta = \frac{\pi}{2}, \qquad \int_0^\pi \sin\theta\cos\theta\,d\theta = 0, \qquad \int_0^\pi \sin\theta\,d\theta = 2.$$

Substituting these values in the iterated integral above, we obtain

$$\begin{aligned}
\iint_S \mathbf{F}\cdot d\mathbf{S} &= 125\int_0^\pi (2.5\pi\sin^3\phi\cos\phi + 2\sin^2\phi\cos\phi)\,d\phi\\
&= 125\int_0^0 (2.5\pi u^3 + 2u^2)\,du = 0.
\end{aligned}$$

□

Example 24.4. The surface $\Sigma$ with vector equation

$$\mathbf{r}(u,v) = u\cos v\mathbf{e}_1 + u\sin v\mathbf{e}_2 + v\mathbf{e}_3,$$

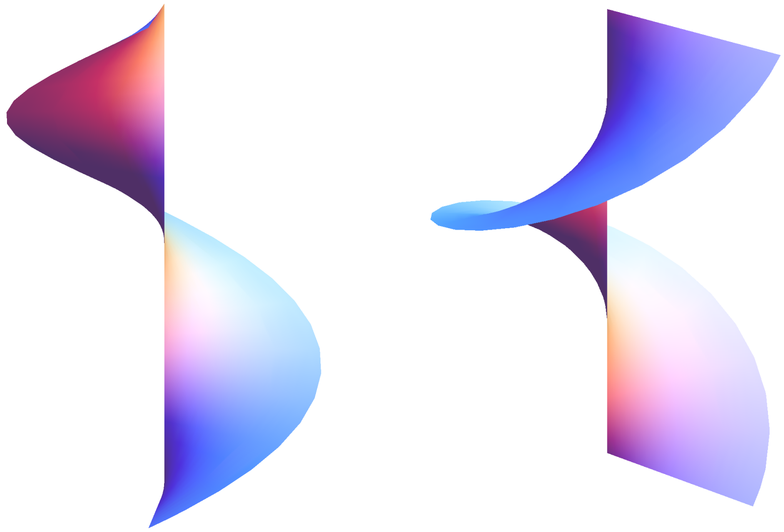


FIGURE 24.1. Two views of the helicoid

where $0 \le u \le 1$, $0 \le v \le \pi$, is called *helicoid* (see Figure 24.1). Evaluate the surface integral of

$$\mathbf{F}(x,y,z) = x\sqrt{1+x^2+y^2}\,\mathbf{e}_1 + y\sqrt{1+x^2+y^2}\,\mathbf{e}_2 + (x+y)z\,\mathbf{e}_3$$

over $\Sigma$ with upward orientation.

SOLUTION. We have

$$\begin{aligned}
\mathbf{r}_u &= \cos v\mathbf{e}_1 + \sin v\mathbf{e}_2, \quad \mathbf{r}_v = -u\sin v\mathbf{e}_1 + u\cos v\mathbf{e}_2 + \mathbf{e}_3;\\
\mathbf{r}_u \times \mathbf{r}_v &= u\cos^2 v(\mathbf{e}_1\times\mathbf{e}_2) + \cos v(\mathbf{e}_1\times\mathbf{e}_3) - u\sin^2 v(\mathbf{e}_2\times\mathbf{e}_1) + \sin v(\mathbf{e}_2\times\mathbf{e}_3)\\
&= \sin v\mathbf{e}_1 - \cos v\mathbf{e}_2 + u\mathbf{e}_3.
\end{aligned}$$

Note that the vector $\mathbf{r}_u \times \mathbf{r}_v$ has a positive third component for all $u$ with $0 \le u \le 1$. Therefore, we can orient $\Sigma$ in the upward direction by the unit normal vector

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}.$$

With this choice, we have

$$\mathbf{F}(\mathbf{r}(u,v)) = u\cos v\sqrt{1+u^2}\mathbf{e}_1 + u\sin v\sqrt{1+u^2}\mathbf{e}_2 + uv(\cos v + \sin v)\mathbf{e}_3$$

$$\mathbf{F}(\mathbf{r}(u,v))\cdot(\mathbf{r}_u \times \mathbf{r}_v) = u^2 v(\cos v + \sin v).$$

Thus, (24.6) yields

$$\begin{aligned}
\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} &= \int_0^\pi\int_0^1 u^2 v(\cos v + \sin v)\,du dv\\
&= \frac{1}{3}\int_0^\pi v(\cos v + \sin v)\,dv = \frac{1}{3}\int_0^\pi v\,d(\sin v - \cos v)\\
&= \frac{1}{3}\big[v(\sin v - \cos v)\big]_0^\pi - \frac{1}{3}\int_0^\pi (\sin v - \cos v)\,dv = \frac{\pi}{3} - \frac{2}{3}.
\end{aligned}$$

□

## Exercises

24.1. Evaluate the surface integral $\iint_\Sigma x^2yz\,dS$, where $\Sigma$ is the surface

$$z = 1 + 2x + 3y, \quad 0 \le x \le 3,\ 0 \le y \le 2,$$

that is, the part of the plane $z = 1 + 2x + 3y$ that lies above the rectangle $[0, 3] \times [0, 2]$.

24.2. Evaluate the surface integral $\iint_\Sigma x^2z^2\,dS$, where $\Sigma$ is the surface

$$z = \sqrt{x^2 + y^2}, \quad 1 \le x^2 + y^2 \le 4,$$

that is, the part of the cone $z^2 = x^2 + y^2$ that lies between the planes $z = 1$ and $z = 2$.

24.3. Evaluate the surface integral $\iint_\Sigma \left(x^2z + y^2z\right) dS$, where $\Sigma$ is the hemisphere $x^2 + y^2 + z^2 = 4$, $z \ge 0$, that is, the parametric surface given by

$$\mathbf{r}(u, v) = 2\cos u \sin v\mathbf{e}_1 + 2\sin u \sin v\mathbf{e}_2 + 2\cos v\mathbf{e}_3, \qquad 0 \le u \le 2\pi,\ 0 \le v \le \pi/2.$$

24.4. Evaluate the surface integral $\iint_\Sigma \left(x^2y + z^2\right) dS$, where $\Sigma$ is the part of the cylinder $x^2 + y^2 = 9$ that lies between the planes $z = 0$ and $z = 2$, that is, the parametric surface given by

$$\mathbf{r}(u, v) = 3\cos u\mathbf{e}_1 + 3\sin u\mathbf{e}_2 + v\mathbf{e}_3, \qquad 0 \le u \le 2\pi,\ 0 \le v \le 2.$$

24.5. Evaluate the surface integral $\iint_\Sigma \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F}(x, y, z) = xy\mathbf{e}_1 + yz\mathbf{e}_2 + zx\mathbf{e}_3$ and $\Sigma$ is the part of the paraboloid $z = 4 - x^2 - y^2$ that lies above the square $0 \le x, y \le 1$ and is oriented upward.

24.6. Evaluate the surface integral $\iint_\Sigma \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F}(x, y, z) = xze^y\mathbf{e}_1 - xze^y\mathbf{e}_2 + z\mathbf{e}_3$ and $\Sigma$ is the part of the plane $x + y + z = 1$ in the first octant with the upward orientation.

24.7. Evaluate the surface integral $\iint_\Sigma \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F}(x, y, z) = xz\mathbf{e}_1 + 2yz\mathbf{e}_3$ and $\Sigma$ is the paraboloid $y = x^2 + z^2$, $0 \le y \le 1$, oriented in the direction of the positive $y$-axis.

24.8. Evaluate the surface integral $\iint_\Sigma \mathbf{F} \cdot d\mathbf{S}$, where $\mathbf{F}(x, y, z) = x\mathbf{e}_1 + y\mathbf{e}_2 + xyz\mathbf{e}_3$ and $\Sigma$ is the helicoid given by

$$\mathbf{r}(u, v) = u\cos v\mathbf{e}_1 + u\sin v\mathbf{e}_2 + v\mathbf{e}_3, \qquad 0 \le u \le 1,\ 0 \le v \le \pi,$$

and oriented by $\mathbf{n} = \mathbf{r}_u \times \mathbf{r}_v$.

LECTURE 25

# Stokes' Theorem

In the last two lectures, we discuss the other two major integral theorems of vector calculus: Stokes' theorem and Gauss' divergence theorem. Stokes' theorem is a generalization of Green's theorem to closed curves on surfaces in space. Before we state it, we introduce the notion of orientation of curves on an oriented surface.

Definition. Suppose that $\Sigma$ is an oriented surface with unit normal vector $\mathbf{n}$. We say that a closed curve $\gamma$ on $\Sigma$ is *positively oriented*, if the vector $\mathbf{n} \times \mathbf{T}$ points towards the interior of the curve; here $\mathbf{T}$ is the unit tangent vector to $\gamma$.

Theorem 25.1 (Stokes). *Let* $\mathbf{F}$ *be a vector field that has continuous partial derivatives on an open set* $U$ *in* $\mathbb{R}^3$. *Let* $\Sigma$ *be an oriented, piecewise smooth surface that is contained in* $U$ *and is bounded by a simple, closed, piecewise smooth, positively oriented curve* $\gamma$. *Then*

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \iint_\Sigma \operatorname{curl} \mathbf{F} \cdot d\mathbf{S}. \tag{25.1}$$

We said above that Stokes' theorem is a generalization of Green's. It would be nice to corroborate that. Consider the parametric surface $\Sigma$ given by

$$\mathbf{r}(x, y) = x\mathbf{e}_1 + y\mathbf{e}_2 + 0\mathbf{e}_3 \qquad ((x, y) \in D),$$

that is, $\Sigma$ is the region $D$ in the $xy$-plane turned into a "three-dimensional surface". If we give this surface an upward orientation, then the unit normal vector $\mathbf{n}$ is $\mathbf{e}_3$, and Stokes' theorem gives

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \iint_\Sigma \operatorname{curl} \mathbf{F} \cdot d\mathbf{S} = \iint_\Sigma (\operatorname{curl} \mathbf{F}) \cdot \mathbf{e}_3 \, dS.$$

Converting the latter surface integral into a double integral, we find

$$\mathbf{r}_x(x, y) = \mathbf{e}_1, \quad \mathbf{r}_y(x, y) = \mathbf{e}_2, \quad \mathbf{r}_x \times \mathbf{r}_y = \mathbf{e}_3,$$

and hence,

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \iint_\Sigma (\operatorname{curl} \mathbf{F}) \cdot \mathbf{e}_3 \, dS = \iint_D (\operatorname{curl} \mathbf{F}) \cdot \mathbf{e}_3 \, dA.$$

This is exactly one of the vector forms of Green's theorem (recall (23.3)).

We now move to describe the main applications of Stokes' theorem. They are similar to the applications of Green's theorem. First, we can apply Stokes's theorem to convert a line integral over a closed space curve $\gamma$ to a surface integral over a surface $\Sigma$ that has $\gamma$ as its boundary.

Example 25.1. Evaluate

$$\int_\gamma \left[ \left(e^{x^3+x} + z^2\right)\mathbf{e}_1 + \left( \arctan y + x^2 \right)\mathbf{e}_2 + \left( \ln(z^2 + 1) + y^2 \right)\mathbf{e}_3 \right] \cdot d\mathbf{r},$$

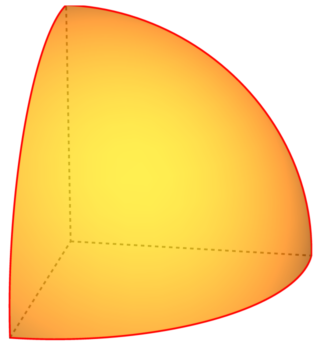


Figure 25.1. The spherical patch $x^2 + y^2 + z^2 = 9$, $x, y, z \geq 0$, and its boundary

where $\gamma$ is the positively oriented boundary of the part of the sphere $x^2 + y^2 + z^2 = 9$ that lies in the first octant, with downward orientation (see Figure 25.1.

Solution. All the partials of $\mathbf{F}$ will be continuous everywhere, so we can apply Stokes' theorem. We get

$$\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \iint_\Sigma \operatorname{curl} \mathbf{F} \cdot d\mathbf{S},$$

where $\Sigma$ is the given surface and

$$\mathbf{F}(x, y, z) = \left(e^{x^3+x} + z^2\right)\mathbf{e}_1 + \left(\arctan y + x^2\right)\mathbf{e}_2 + \left(\ln(z^2 + 1) + y^2\right)\mathbf{e}_3.$$

As we do usually when dealing with spheres, we parametrize $\Sigma$ using spherical coordinates:

$$\mathbf{r}(\theta, \phi) = 3\cos\theta \sin\phi \mathbf{e}_1 + 3\sin\theta\sin\phi\mathbf{e}_2 + 3\cos\phi\mathbf{e}_3 \qquad (0 \leq \theta, \phi \leq \pi/2).$$

Then, similarly to Example 24.3,

$$\begin{aligned} \mathbf{r}_\theta &= -3\sin\theta\sin\phi\mathbf{e}_1 + 3\cos\theta\sin\phi\mathbf{e}_2, \\ \mathbf{r}_\phi &= 3\cos\theta\cos\phi\mathbf{e}_1 + 3\sin\theta\cos\phi\mathbf{e}_2 - 3\sin\phi\mathbf{e}_3; \\ \mathbf{r}_\theta \times \mathbf{r}_\phi &= -9\sin\phi(\cos\theta\sin\phi\mathbf{e}_1 + \sin\theta\sin\phi\mathbf{e}_2 + \cos\phi\mathbf{e}_3). \end{aligned}$$

Since $-9\sin\phi\cos\phi \leq 0$ for all $\phi$, $0 \leq \phi \leq \pi/2$, the downward orientation of $\Sigma$ is given by

$$\mathbf{n} = \frac{\mathbf{r}_\theta \times \mathbf{r}_\phi}{\|\mathbf{r}_\theta \times \mathbf{r}_\phi\|}.$$

Furthermore, we have

$$\begin{aligned} \operatorname{curl} \mathbf{F} &= (2y - 0)\mathbf{e}_1 - (0 - 2z)\mathbf{e}_2 + (2x - 0)\mathbf{e}_3 = 2y\mathbf{e}_1 + 2z\mathbf{e}_2 + 2x\mathbf{e}_1; \\ \operatorname{curl} \mathbf{F}(\mathbf{r}(\theta, \phi)) &= 6\sin\theta\sin\phi\mathbf{e}_1 + 6\cos\phi\mathbf{e}_2 + 6\cos\theta\sin\phi\mathbf{e}_3, \\ \left(\operatorname{curl} \mathbf{F}(\mathbf{r}(\theta, \phi))\right) \cdot (\mathbf{r}_\theta \times \mathbf{r}_\phi) &= -54\sin^2\phi\left(\sin\theta\cos\theta\sin\phi + (\sin\theta + \cos\theta)\cos\phi\right). \end{aligned}$$

We conclude that

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \iint_{0\le\theta,\phi\le\pi/2} \big(\operatorname{curl}\mathbf{F}(\mathbf{r}(\theta,\phi))\big)\cdot(\mathbf{r}_\theta\times\mathbf{r}_\phi)\,dA$$
$$= -54\int_0^{\pi/2}\int_0^{\pi/2} \sin^2\phi\big(\sin\theta\cos\theta\sin\phi + (\sin\theta+\cos\theta)\cos\phi\big)\,d\theta d\phi$$
$$= -54\int_0^{\pi/2} \sin^2\phi\big[\tfrac{1}{2}\sin^2\theta\sin\phi + (-\cos\theta+\sin\theta)\cos\phi\big]_0^{\pi/2}\,d\phi$$
$$= -54\int_0^{\pi/2} \sin^2\phi\big(\tfrac{1}{2}\sin\phi + 2\cos\phi\big)\,d\phi$$
$$= -27\int_0^{\pi/2}\big(1-\cos^2\phi\big)\sin\phi\,d\phi - 108\int_0^{\pi/2}\sin^2\phi\cos\phi\,d\phi$$
$$= -27\int_0^1\big(1-u^2\big)\,du - 108\int_0^1 u^2\,du = -54.$$

□

Example 25.2. Evaluate

$$\int_\gamma (yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3)\cdot d\mathbf{r},$$

where $\gamma$ is the intersection curve of the ellipsoid $x^2+2y^2+z^2=4$ and the plane $x+z=2$, oriented positively relative to the outward orientation of the ellipsoid.

Solution. The intersection curve $\gamma$ of the ellipsoid and the plane is displayed on Figure 25.2. It can be seen from the figure that the positive orientation of $\gamma$ relative to the outward orientation of the ellipsoid is in the counterclockwise direction. Note that this is also the positive orientation of $\gamma$ relative to the upward orientation of the plane $x+z=2$. The equations of $\gamma$ are

$$x^2+2y^2+z^2=4,\quad x+z=2 \quad\Longleftrightarrow\quad x^2+2y^2+(2-x)^2=4,\quad z=2-x.$$

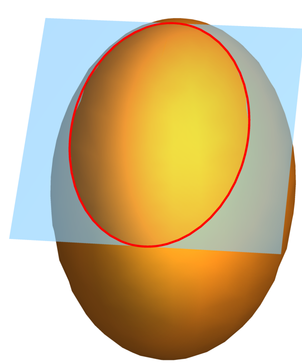


Figure 25.2. The intersection curve of $x^2+2y^2+z^2=4$ and $x+z=2$

Some basic algebra shows that

$$\begin{aligned} x^2 + 2y^2 + (2-x)^2 = 4 \quad &\iff \quad x^2 + 2y^2 + 4 - 4x + x^2 = 4 \quad \iff \quad 2x^2 - 4x + 2y^2 = 0 \\ &\iff \quad x^2 - 2x + y^2 = 0 \quad \iff \quad (x-1)^2 + y^2 = 1, \end{aligned}$$

so $\gamma$ is described by the equations

$$(x-1)^2 + y^2 = 1, \quad z = 2 - x.$$

That is, $\gamma$ contains the points $(x, y, z)$ on the plane $z = 2 - x$ for which the point $(x, y)$ lies on the circle $(x-1)^2 + y^2 = 1$ in the $xy$-plane. It is possible to parametrize $\gamma$ and to evaluate the given line integral directly (see the exercises), but we want to use Stokes' theorem.

Let $\Sigma$ be the portion of the plane $z = 2 - x$ that lies within $\gamma$, with the upward orientation. Then, by Stokes' theorem,

$$\int_\gamma (yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3) \cdot d\mathbf{r} = \iint_\Sigma \operatorname{curl}(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3) \cdot d\mathbf{S}.$$

We have

$$\begin{aligned} \operatorname{curl}(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3) &= \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_x & \partial_y & \partial_z \\ yz & 2xz & 3xy \end{vmatrix} = \begin{vmatrix} \partial_y & \partial_z \\ 2xz & 3xy \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} \partial_x & \partial_z \\ yz & 3xy \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} \partial_x & \partial_y \\ yz & 2xz \end{vmatrix} \mathbf{e}_3 \\ &= (3x - 2x)\mathbf{e}_1 - (3y - y)\mathbf{e}_2 + (2z - z)\mathbf{e}_3 = x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3. \end{aligned}$$

Further, $\Sigma$ is the graph of $z = 2 - x$ for $(x, y)$ in the disk $D$ with equation $(x-1)^2 + y^2 \le 1$. Thus, (24.7) gives

$$\begin{aligned} \iint_\Sigma \operatorname{curl}(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3) \cdot d\mathbf{S} &= \iint_\Sigma (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3) \cdot d\mathbf{S} \\ &= \iint_D (-x\partial_x(2-x) - (-2y)\partial_y(2-x) + z(x,y))\, dA \\ &= \iint_D (x + 0 + (2-x))\, dA = \iint_D 2\, dA. \end{aligned}$$

The last double integral represents twice the area of $D$, which is a disk of radius 1. Consequently,

$$\int_\gamma (yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3) \cdot d\mathbf{r} = \iint_D 2\, dA = 2\operatorname{area}(D) = 2\pi. \qquad \square$$

In the last two examples, we used Stokes' theorem to reduce a line integral of a vector field to a surface integral of its curl. In the next couple of examples, we describe the other main application of Stokes' theorem: to simplify the evaluation of surface integrals. To illustrate the first flavor of such applications, we take a look at a surface integral that we evaluated in Example 24.2 by writing it as a double integral. We now give an alternative solution, which uses Stokes' theorem.

Example 25.3. Evaluate

$$\iint_\Sigma (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3) \cdot d\mathbf{S},$$

where $\Sigma$ is the part of the paraboloid $z = 4 - 2x^2 - 2y^2$ that lies above the $xy$-plane, oriented upward.

FIGURE 25.3. The paraboloid $z = 4 - 2x^2 - 2y^2$, $z \geq 0$ and its boundary

SOLUTION. The main idea is to replace the given surface integral (which is a two-dimensional object) by a line integral along the surface boundary (which is a one-dimensional object). Recall from the previous example that

$$x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3 = \operatorname{curl}\mathbf{F}, \qquad \mathbf{F} = yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3.$$

Hence, when we apply Stokes' theorem to $\mathbf{F}$ on $\Sigma$, we get

$$\iint_\Sigma (x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3)\cdot d\mathbf{S} = \iint_\Sigma \operatorname{curl}\mathbf{F}\cdot d\mathbf{S} = \int_\gamma \mathbf{F}\cdot d\mathbf{r},$$

where $\gamma$ is the boundary of the given surface, positively oriented. The curve $\gamma$ is the intersection curve of the paraboloid and the $xy$-plane, that is, the circle $x^2 + y^2 = 2$, $z = 0$. We can parametrize $\gamma$ using

$$\mathbf{r}(t) = \sqrt{2}\cos t\mathbf{e}_1 + \sqrt{2}\sin t\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (0 \leq t \leq 2\pi).$$

Note that with this parametrization, $\gamma$ is positively oriented relative to the upward orientation of $\Sigma$ (see Figure 25.3). Thus,

$$\begin{aligned}\int_\gamma \mathbf{F}\cdot d\mathbf{r} &= \int_0^{2\pi} \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt\\ &= \int_0^{2\pi} \big(0\mathbf{e}_1 + 0\mathbf{e}_2 + 6\cos t\sin t\mathbf{e}_3\big)\cdot\big(-\sqrt{2}\sin t\mathbf{e}_1 + \sqrt{2}\cos t\mathbf{e}_2\big)\,dt = 0. \qquad \square\end{aligned}$$

In the above solution, the vector field $\mathbf{F}$ resulted from a "happy coincidence". Had we not just worked out Example 25.2, we would have started the above solution of Example 25.3 by "observing" that

$$x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3 = \operatorname{curl}\big(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\big).$$

Such an "observation" falls just short of reliance upon "divine intervention". In general, given a surface integral $\iint_\Sigma \mathbf{F}\cdot d\mathbf{S}$, in order to argue similarly to above, we need to find a vector field $\mathbf{G}$ such that $\operatorname{curl}\mathbf{G} = \mathbf{F}$ is some given vector field. This is a difficult problem, which limits the applicability of the method. Still, the above example is not a complete waste of time: sometimes, one has to calculate the surface integral of the curl of a given field, and in those situations one can argue as above.

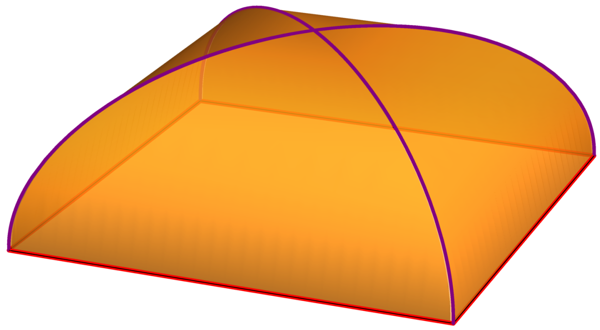


FIGURE 25.4. The surface Σ from Example 25.4

EXAMPLE 25.4. Evaluate

$$\iint_{\Sigma} \operatorname{curl}\left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{S},$$

where Σ is the piecewise smooth surface displayed on Figure 25.4[1], oriented upward (the boundary curve of Σ is the square formed by the lines $x = \pm 2$ and $y = \pm 2$ in the $xy$-plane).

SOLUTION. The main idea of this solution is to replace the given surface integral by another surface integral that is easier to evaluate. By Stokes' theorem,

$$\iint_{\Sigma} \operatorname{curl}\left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{S} = \int_{\gamma} \left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{r},$$

where γ is the boundary square of Σ, oriented positively relative to the upward orientation of the surface (that is, γ is oriented counterclockwise in the $xy$-plane). We now observe that γ is also the positively oriented boundary of the surface $\Sigma_0$ with parametrization

$$\mathbf{r}(x, y) = x\mathbf{e}_1 + y\mathbf{e}_2 + 0\mathbf{e}_3 \qquad (-2 \le x, y \le 2)$$

and with the upward orientation. Hence, another application of Stokes' theorem gives

$$\int_{\gamma} \left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{r} = \iint_{\Sigma_0} \operatorname{curl}\left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{S}.$$

We have

$$\operatorname{curl}\left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) = x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3,$$
$$\mathbf{r}_x \times \mathbf{r}_y = \mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3,$$

Hence, by (24.7),

$$\iint_{\Sigma} \operatorname{curl}\left(yz\mathbf{e}_1 + 2xz\mathbf{e}_2 + 3xy\mathbf{e}_3\right) \cdot d\mathbf{S} = \iint_{\Sigma_0} \left(x\mathbf{e}_1 - 2y\mathbf{e}_2 + z\mathbf{e}_3\right) \cdot d\mathbf{S}$$
$$= \int\int_{-2 \le x, y \le 2} \left(x\mathbf{e}_1 - 2y\mathbf{e}_2 + 0\mathbf{e}_3\right) \cdot \mathbf{e}_3 \, dA = 0. \qquad \square$$

[1]In fact, the surface is the graph of the function $z = \min\left(\sqrt{4 - x^2}, \sqrt{4 - y^2}\right)$, but that is unimportant for the purposes of the example.

## Exercises

Use Stokes' theorem to evaluate $\int_\gamma \mathbf{F} \cdot d\mathbf{r}$ for the given field $\mathbf{F}(x, y, z)$ and the given curve $\gamma$.

25.1. $\mathbf{F}(x, y, z) = 2z\mathbf{e}_1 + x\mathbf{e}_2 + 3y\mathbf{e}_3$; $\gamma$ is the ellipse that is the intersection of the plane $z = x$ and the cylinder $x^2 + y^2 = 4$, oriented clockwise

25.2. $\mathbf{F}(x, y, z) = (z - y)\mathbf{e}_1 + (z - x)\mathbf{e}_2 + (x + y)\mathbf{e}_3$; $\gamma$ is the triangle with vertices $(0, 1, 0)$, $(0, 2, 2)$ and $(1, 2, 2)$, oriented positively as viewed from above (i.e., relative to the upward orientation of its plane)

25.3. $\mathbf{F}(x, y, z) = x^2\mathbf{e}_1 - 2xy\mathbf{e}_2 + yz^2\mathbf{e}_3$; $\gamma$ is the boundary of the surface $z = x^2y$, $0 \leq x, y \leq 1$, oriented counterclockwise

25.4. $\mathbf{F}(x, y, z) = x^2y\mathbf{e}_1 + \frac{1}{3}x^3\mathbf{e}_2 + xy\mathbf{e}_3$; $\gamma$ is the curve of intersection of the hyperbolic paraboloid $z = y^2 - x^2$ and the cylinder $x^2 + y^2 = 1$, oriented counterclockwise as viewed from above

Use Stokes' theorem to evaluate $\iint_\Sigma \operatorname{curl} \mathbf{F} \cdot d\mathbf{S}$ for the given field $\mathbf{F}(x, y, z)$ and the given surface $\Sigma$.

25.5. $\mathbf{F}(x, y, z) = 3y\mathbf{e}_1 - xz\mathbf{e}_2 - yz^2\mathbf{e}_3$; $\Sigma$ is the part of the elliptic paraboloid $2z = x^2 + y^2$ below $z = 2$, oriented upward

25.6. $\mathbf{F}(x, y, z) = yz\mathbf{e}_1 + xz\mathbf{e}_2 + xy\mathbf{e}_3$; $\Sigma$ is the part of the paraboloid $z = 9 - x^2 - y^2$ above $z = 8$, oriented downward

25.7. $\mathbf{F}(x, y, z) = x^2\mathbf{e}_1 - 2xy\mathbf{e}_2 + yz^2\mathbf{e}_3$; $\Sigma$ is the the surface $z = x^2y$, $0 \leq x, y \leq 1$, oriented upward

25.8. $\mathbf{F}(x, y, z) = y\mathbf{e}_1 + xz\mathbf{e}_2 + 2x\mathbf{e}_3$; $\Sigma$ is the hemisphere $x^2 + y^2 + z^2 = 1$, $y \geq 0$, oriented in the direction of the positive $y$-axis

LECTURE 26

# Gauss' Divergence Theorem

We now turn to another generalization of Green's theorem, known as *Gauss' divergence theorem*. It relates the surface integral over the boundary of a simply-connected solid to a triple integral over the entire solid, similarly to how Green's theorem relates a line integral over the boundary of a simply-connected plane region to the double integral over the entire region.

Like Green's theorem, the divergence theorem requires that we pay attention to terminology. Recall the definition of a simply-connected region from §20.3. By a *simply-connected solid*, we mean a simply-connected region in space that is closed (contains all its boundary points). For such a solid, its *boundary* consists of the points on its "outer" (visible) surface. The boundary of a simply-connected solid is the rightful three-dimensional analog of the simple, closed curve appearing in Green's theorem. We also define the *positive orientation* of the boundary of a solid (assuming that it is orientable) to be the outward orientation.

THEOREM 26.1 (Gauss' divergence theorem). *Let $S$ be a bounded, simply-connected solid in $\mathbb{R}^3$, whose boundary is the piecewise smooth, positively oriented surface $\Sigma$. Let $\mathbf{F}$ be a vector field whose first-order partials are continuous on an open set $U$ containing $S$. Then*

$$\iint_\Sigma \mathbf{F} \cdot d\mathbf{S} = \iiint_S \operatorname{div} \mathbf{F} \, dV. \tag{26.1}$$

EXAMPLE 26.1. Evaluate

$$\iint_\Sigma \big((x^3 + y \sin z)\mathbf{e}_1 + (y^3 + z \sin x)\mathbf{e}_2 + 3z\mathbf{e}_3\big) \cdot d\mathbf{S},$$

where $\Sigma$ is the positively oriented boundary of the solid $S$ bounded by the $xy$-plane and by the hemispheres

$$z = \sqrt{1 - x^2 - y^2}, \quad z = \sqrt{4 - x^2 - y^2}.$$

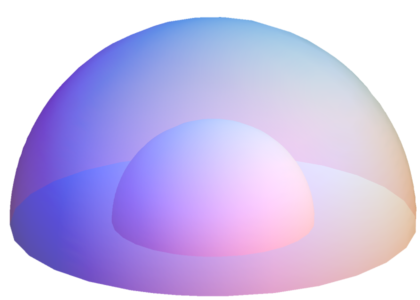


FIGURE 26.1. The solid $1 \le x^2 + y^2 + z^2 \le 4$, $z \ge 0$

SOLUTION. We write $\mathbf{F} = (x^3 + y\sin z)\mathbf{e}_1 + (y^3 + z\sin x)\mathbf{e}_2 + 3z\mathbf{e}_3$. Then, by Gauss' theorem,

$$\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} = \iiint_S \operatorname{div}\mathbf{F}\, dV = \iiint_S \left(3x^2 + 3y^2 + 3\right) dV.$$

The solid $S$ is displayed on Figure 26.1. It is described by the inequalities

$$1 \le x^2 + y^2 + z^2 \le 4, \quad z \ge 0.$$

These inequalities are translated in spherical coordinates as

$$1 \le \rho \le 2, \quad 0 \le \theta \le 2\pi, \quad 0 \le \phi \le \pi/2.$$

Hence, converting the triple integral over $S$ to spherical coordinates, we obtain

$$\begin{aligned}
\iint_\Sigma \mathbf{F}\cdot d\mathbf{S} &= \int_1^2 \int_0^{\pi/2} \int_0^{2\pi} (3\rho^2\sin^2\phi + 3)\rho^2\sin\phi\, d\theta d\phi d\rho \\
&= 6\pi \int_1^2 \int_0^{\pi/2} (\rho^4\sin^2\phi + \rho^2)\sin\phi\, d\phi d\rho \\
&= 6\pi \int_1^2 \left[\int_0^{\pi/2} \left(\rho^4(1-\cos^2\phi) + \rho^2\right)\sin\phi\, d\phi\right] d\rho \\
&= 6\pi \int_1^2 \left[\int_0^1 \left(\rho^4(1-u^2) + \rho^2\right) du\right] d\rho \\
&= 6\pi \int_1^2 \left(\tfrac{2}{3}\rho^4 + \rho^2\right) d\rho = \tfrac{194}{5}\pi.
\end{aligned}$$

□

EXAMPLE 26.2. Evaluate

$$\iint_\Sigma \left(x^2\mathbf{e}_1 + y^2\mathbf{e}_2 + z^2\mathbf{e}_3\right)\cdot d\mathbf{S},$$

where $\Sigma$ is the positively oriented surface of the solid $S$ shown on Figure 26.2 (a cube of side length 2 with a corner cube of side length 1 removed).

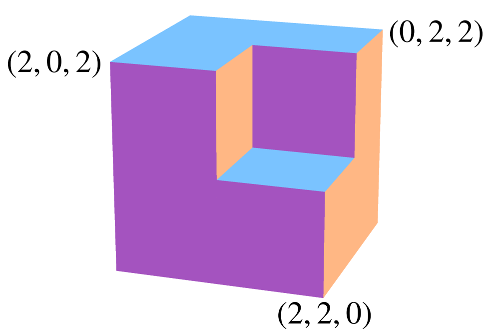



FIGURE 26.2. The solid $E$

SOLUTION. It is possible to evaluate the surface integral using double integrals. However, were we to do that, we would have to parametrize each face separately, and there are nine faces. Furthermore, we would have to split the three double integrals corresponding to the $L$-shaped faces in two, so we would have to evaluate twelve double integrals altogether. On the other hand, $S$ is the difference of two cubes, so we can calculate a triple integral over $S$ by writing it as the difference of two triple integrals over those cubes. This suggests that we use the divergence theorem.

By Gauss' theorem,

$$\iint_\Sigma \left(x^2\mathbf{e}_1 + y^2\mathbf{e}_2 + z^2\mathbf{e}_3\right)\cdot d\mathbf{S} = \iiint_S (2x+2y+2z)\,dV.$$

Since $E$ is the difference of the cubes $[0,2]^3$ and $[1,2]^3$, it is easier to calculate the triple integral over $E$ as the difference of the triple integrals over those two cubes:

$$\begin{aligned}
\iiint_S (2x+2y+2z)\,dV &= \iiint_{[0,2]^3}(2x+2y+2z)\,dV - \iiint_{[1,2]^3}(2x+2y+2z)\,dV\\
&= \int_0^2\int_0^2\int_0^2 (2x+2y+2z)\,dzdydx - \int_1^2\int_1^2\int_1^2 (2x+2y+2z)\,dzdydx\\
&= \int_0^2\int_0^2 \left[2xz+2yz+z^2\right]_0^2 dydx - \int_1^2\int_1^2 \left[2xz+2yz+z^2\right]_1^2 dydx\\
&= \int_0^2\int_0^2 (4x+4y+4)\,dydx - \int_1^2\int_1^2 (2x+2y+3)\,dydx\\
&= \int_0^2 \left[4xy+2y^2+4y\right]_0^2 dx - \int_1^2 \left[2xy+y^2+3y\right]_1^2 dx\\
&= \int_0^2 (8x+16)\,dx - \int_1^2 (2x+6)\,dx = \left[4x^2+16x\right]_0^2 - \left[x^2+6x\right]_1^2 = 39.
\end{aligned}$$

□

With a little bit of extra effort, we can also use the divergence theorem to evaluate surface integrals of scalar functions.

Example 26.3. Evaluate

$$\iint_\Sigma \left(x^2z + y^2z\right)dS,$$

where $\Sigma$ is the sphere $x^2+y^2+z^2=4$.

Solution. Note that we have to deal with the surface integral of a scalar function and not of a vector field. Thus, Gauss' theorem is not directly applicable. However, we know that the sphere $\Sigma$ can be parametrized using spherical coordinates by the vector function

$$\mathbf{r}(\theta,\phi) = 2\cos\theta\sin\phi\mathbf{e}_1 + 2\sin\theta\sin\phi\mathbf{e}_2 + 2\cos\phi\mathbf{e}_3,$$

where $0 \le \theta \le 2\pi$ and $0 \le \phi \le \pi$. We also know (we went through this calculation several times already) that the (outward) unit normal vector is then

$$\mathbf{n} = \mathbf{r}_\phi \times \mathbf{r}_\theta = \cos\theta\sin\phi\mathbf{e}_1 + \sin\theta\sin\phi\mathbf{e}_2 + \cos\phi\mathbf{e}_3.$$

Note that in Cartesian coordinates $\mathbf{n} = \frac{1}{2}\left(x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3\right)$. We can use this observation to represent the integrand in the given integral in the form $\mathbf{F}\cdot\mathbf{n}$, with $\mathbf{F}$ some nice vector field; we can then appeal to Gauss' theorem. For example, if we choose $\mathbf{F} = 2(x^2+y^2)\mathbf{e}_3$, we obtain

$$\mathbf{F}\cdot\mathbf{n} = 0\left(\tfrac{1}{2}x\right) + 0\left(\tfrac{1}{2}y\right) + 2(x^2+y^2)\left(\tfrac{1}{2}z\right) = x^2z + y^2z,$$

whence

$$\iint_\Sigma \left(x^2z + y^2z\right) dS = \iint_\Sigma (\mathbf{F}\cdot\mathbf{n})\, dS = \iint_\Sigma \mathbf{F}\cdot d\mathbf{S} = \iiint_B \operatorname{div}\mathbf{F}\, dV = 0.$$

Here, $B$ is the ball $x^2 + y^2 + z^2 \leq 4$.

Note that the above choice of $\mathbf{F}$ is not unique. We could have chosen also

$$\mathbf{F} = 2xz\mathbf{e}_1 + 2yz\mathbf{e}_2, \qquad \mathbf{F} = xz\mathbf{e}_1 + yz\mathbf{e}_2 + (x^2 + y^2)\mathbf{e}_3,$$

or any other vector field $\mathbf{F}$ that satisfies

$$\mathbf{F}\cdot\mathbf{n} = x^2z + y^2z.$$

However, in most cases, those choices would yield triple integrals that equal 0 for less obvious reasons than the one we encountered. For example, the choice $\mathbf{F} = 2xz\mathbf{e}_1 + 2yz\mathbf{e}_2$ yields

$$\iint_\Sigma \left(x^2z + y^2z\right) dS = \iiint_B \operatorname{div}\mathbf{F}\, dV = \iiint_B 4z\, dV.$$

The last integral is still 0, but this requires some thought (try to convince yourself by interpreting it in terms of the average value of $f(x, y, z) = 4z$). □

We can also use the divergence theorem to calculate volumes by means of surface integrals similarly to how we earlier used Green's theorem to calculate areas by means of line integrals. For example, if $S$ is a simply-connected solid and $\Sigma$ is its boundary, then

$$\iint_\Sigma x\mathbf{e}_1\cdot d\mathbf{S} = \iiint_S \operatorname{div}(x\mathbf{e}_1)\, dV = \iiint_S 1\, dV = \text{vol.}(S).$$

Similarly, we obtain formulas for the volume of $S$ that use the vector fields $y\mathbf{e}_2$ and $z\mathbf{e}_3$. Here are some of the surface integrals representing vol.$(S)$:

$$\text{vol.}(E) = \iint_\Sigma x\mathbf{e}_1\cdot d\mathbf{S} = \iint_\Sigma y\mathbf{e}_2\cdot d\mathbf{S} = \iint_\Sigma z\mathbf{e}_3\cdot d\mathbf{S} = \frac{1}{3}\iint_\Sigma (x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3)\cdot d\mathbf{S}. \tag{26.2}$$

Example 26.4. Find the volume of the spiral pipe shown on Figure 26.3. The outer surface of the pipe is parametrized by the vector function

$$\mathbf{r}(u, v) = (2 + \sin v)\cos 2u\mathbf{e}_1 + (2 + \sin v)\sin 2u\mathbf{e}_2 + (u + \cos v)\mathbf{e}_3 \qquad (0 \leq u, v \leq 2\pi),$$

and its two openings are the disks $(x - 2)^2 + z^2 \leq 1$ and $(x - 2)^2 + (z - 2\pi)^2 \leq 1$ in the $xz$-plane.

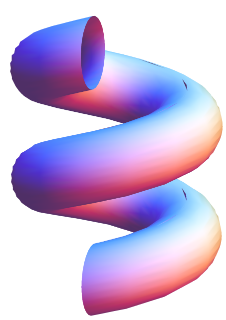


Figure 26.3. The surface $\mathbf{r}(u, v) = (2 + \sin v)\cos 2u\mathbf{e}_1 + (2 + \sin v)\sin 2u\mathbf{e}_2 + (u + \cos v)\mathbf{e}_3$

Solution. Let $\Sigma$ be the positively oriented, piecewise smooth surface consisting of the parametric surface $\Sigma_0$ that is the outer surface of the pipe and the two disks representing its two openings. We denote the disks $\Sigma_1$ and $\Sigma_2$. Then $\Sigma$ is the boundary of the simply-connected solid $S$ whose volume we want to find. By (26.2),

$$\text{vol.}(S) = \iint_{\Sigma} z\mathbf{e}_3 \cdot d\mathbf{S} = \left( \iint_{\Sigma_0} + \iint_{\Sigma_1} + \iint_{\Sigma_2} \right) z\mathbf{e}_3 \cdot d\mathbf{S}.$$

The unit normal vectors to $\Sigma_1$ are $\mathbf{e}_2$ and $-\mathbf{e}_2$, and of those two, it is $-\mathbf{e}_2$ that points out of the solid $S$. Thus, by (26.2),

$$\iint_{\Sigma_1} z\mathbf{e}_3 \cdot d\mathbf{S} = \iint_{\Sigma_1} \big((z\mathbf{e}_3) \cdot (-\mathbf{e}_2)\big)\, dS = 0.$$

Similarly,

$$\iint_{\Sigma_2} z\mathbf{e}_3 \cdot d\mathbf{S} = \iint_{\Sigma_2} \big((z\mathbf{e}_3) \cdot (\mathbf{e}_2)\big)\, dS = 0.$$

Hence,

$$\text{vol.}(S) = \iint_{\Sigma_0} z\mathbf{e}_3 \cdot d\mathbf{S}.$$

We evaluate the last surface integral using the orientation

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{\|\mathbf{r}_u \times \mathbf{r}_v\|}.$$

Thus, by (24.6),

$$\iint_{\Sigma_0} z\mathbf{e}_3 \cdot d\mathbf{S} = \iint\limits_{0 \le u,v \le 2\pi} \big((z(u,v)\mathbf{e}_3) \cdot (\mathbf{r}_u \times \mathbf{r}_v)\big)\, dA.$$

Note that the dot product depends only on the third component of $\mathbf{r}_u \times \mathbf{r}_v$. We have

$$\begin{aligned}
\mathbf{r}_u(u,v) &= -2(2+\sin v)\sin 2u\mathbf{e}_1 + 2(2+\sin v)\cos 2u\mathbf{e}_2 + \mathbf{e}_3, \\
\mathbf{r}_v(u,v) &= \cos v\cos 2u\mathbf{e}_1 + \cos v\sin 2u\mathbf{e}_2 - \sin v\mathbf{e}_3.
\end{aligned}$$

Thus, the cross product $\mathbf{r}_u \times \mathbf{r}_v$ is the sum of six nonzero terms:

$$\begin{aligned}
\mathbf{r}_u \times \mathbf{r}_v &= -2\cos v(2+\sin v)\sin^2 2u(\mathbf{e}_1 \times \mathbf{e}_2) + 2\cos v(2+\sin v)\cos^2 2u(\mathbf{e}_2 \times \mathbf{e}_1) + \text{other terms} \\
&= -2\cos v(2+\sin v)\big(\sin^2 2u + \cos^2 2u\big)\mathbf{e}_3 + \text{other terms} \\
&= -2\cos v(2+\sin v)\mathbf{e}_3 + \text{other terms}.
\end{aligned}$$

The omitted terms are cross products involving $\mathbf{e}_1 \times \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3$, etc., which do not contribute to the coefficient of $\mathbf{e}_3$ in $\mathbf{r}_u \times \mathbf{r}_v$. Thus,

$$(z(u,v)\mathbf{e}_3) \cdot (\mathbf{r}_u \times \mathbf{r}_v) = ((u+\cos v)\mathbf{e}_3) \cdot (\mathbf{r}_u \times \mathbf{r}_v) = -2\cos v(u+\cos v)(2+\sin v).$$

We conclude that

$$\begin{aligned}
\iint_{\Sigma_0} z\mathbf{e}_3 \cdot d\mathbf{S} &= \iint\limits_{0 \le u,v \le 2\pi} \big(-2\cos v(u+\cos v)(2+\sin v)\big)\, dA \\
&= \int_0^{2\pi} \left[ \int_0^{2\pi} \big(-2u\cos v(2+\sin v) - 2\cos^2 v\sin v - 4\cos^2 v\big)\, dv \right] du.
\end{aligned}$$

We have

$$\int_0^{2\pi} (2+\sin v)\cos v\,dv = 0, \qquad \int_0^{2\pi} \cos^2 v \sin v\,dv = 0, \qquad \int_0^{2\pi} \cos^2 v\,dv = \pi,$$

so

$$\iint_{\Sigma_0} z\mathbf{e}_3 \cdot d\mathbf{S} = \int_0^{2\pi} \big( -2u(0) - 2(0) - 4\pi \big)\,du = -8\pi^2.$$

We obtained a negative volume, because we did not check whether the unit vector **n** gives the inward or the outward orientation of the surface. It turns out that it gives the inward (negative) orientation; hence, the negative volume. Consequently, the volume of the pipe is $8\pi^2$. □

## Exercises

Use Gauss' theorem to evaluate $\iint_\Sigma \mathbf{F} \cdot d\mathbf{S}$ for the given field $\mathbf{F}(x,y,z)$ and the given surface $\Sigma$ oriented outward.

26.1. $\mathbf{F}(x,y,z) = y^3\mathbf{e}_1 + z^3\mathbf{e}_2 + x^3\mathbf{e}_3$; $\Sigma$ is the boundary of the solid $S$ defined by $x^2+y^2+z^2 \le 3$, $x \ge 0$, $y \ge 0$

26.2. $\mathbf{F}(x,y,z) = x^3\mathbf{e}_1 + y^3\mathbf{e}_2 + z^3\mathbf{e}_3$; $\Sigma$ as in Exercise 26.1

26.3. $\mathbf{F}(x,y,z) = x^2z^3\mathbf{e}_1 + x^2y\mathbf{e}_2 + y^2z\mathbf{e}_3$; $\Sigma$ is the boundary of the solid $S$ enclosed by the sphere $x^2+y^2+z^2 = 4$ inside the cylinder $x^2+y^2 = 1$

26.4. $\mathbf{F}(x,y,z) = xy^2\mathbf{e}_1 + yz^2\mathbf{e}_2 + zx^2\mathbf{e}_3$; $\Sigma$ is the boundary of the solid shown on Figure 26.4, that is, the part of the solid sphere $x^2+y^2+z^2 \le 4$ outside the first octant

26.5. $\mathbf{F}(x,y,z) = y^3e^z\mathbf{e}_1 + y\sqrt{4-x^2}\mathbf{e}_2 + e^y\cos x\mathbf{e}_3$; $\Sigma$ is the boundary of the solid $S$ that lies below the surface $z = 2 - x^3 - y^3$, $-1 \le x, y \le 1$, and above the $xy$-plane

Figure 26.4. Seven eights of the sphere $x^2+y^2+z^2 \le 4$

26.6. Use Gauss' theorem to evaluate $\iint_\Sigma \left(x^2+y^2+z^3\right) dS$, where $\Sigma$ is the sphere $x^2+y^2+z^2 = 9$.

Use the volume formulas (26.2) to calculate the volume of the given solid.

26.7. The solid enclosed by the ellipsoid $x^2+2y^2+4z^2 = 4$, parametrized by

$$\mathbf{r}(\theta,\phi) = 2\cos\theta\sin\phi\mathbf{e}_1 + \sqrt{2}\sin\theta\sin\phi\mathbf{e}_2 + \cos\phi\mathbf{e}_3 \qquad (0 \le \theta \le 2\pi, 0 \le \phi \le \pi)$$

26.8. The solid bounded by the torus

$$\mathbf{r}(u,v) = \cos u(4+2\cos v)\mathbf{e}_1 + \sin u(4+2\cos v)\mathbf{e}_2 + 2\sin v\mathbf{e}_3 \qquad (0 \le u, v \le 2\pi)$$

APPENDIX A

# Determinants

An $n \times n$ *matrix* is an array of real numbers with $n$ rows and $n$ columns. For example, $A$ below is a $2 \times 2$ matrix and $B$ is a $3 \times 3$ matrix:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & -3 & 1 \\ 0 & 2 & -1 \\ 1 & 1 & -2 \end{bmatrix}.$$

The *determinant of a* $2 \times 2$ *matrix*

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is the following number (the product of its diagonal entries minus the product of its off-diagonal entries):

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$$

The *determinant of a* $3 \times 3$ *matrix*

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

is given by *expansion along the first row*, a process which requires the computation of three $2 \times 2$ determinants:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

Example A.1. Compute the determinant of the matrix $B$ above.

Solution. We have

$$\begin{aligned} \det(B) &= 1 \begin{vmatrix} 2 & -1 \\ 1 & -2 \end{vmatrix} - (-3) \begin{vmatrix} 0 & -1 \\ 1 & -2 \end{vmatrix} + 1 \begin{vmatrix} 0 & 2 \\ 1 & 1 \end{vmatrix} \\ &= (2 \cdot (-2) - 1 \cdot (-1)) + 3(0 \cdot (-2) - 1 \cdot (-1)) + (0 \cdot 1 - 2 \cdot 1) = -2. \end{aligned}$$ □

APPENDIX B

# Definite Integrals of Single-Variable Functions

## B.1. Definition of the definite integral

Definition. Let $f(x)$ be a bounded function on $[a, b]$. For each $n = 1, 2, \ldots$, set $\Delta_n = (b - a)/n$ and define the numbers

$$x_0 = a, \quad x_1 = a + \Delta_n, \quad x_2 = a + 2\Delta_n, \quad \ldots, \quad x_n = a + n\Delta_n = b.$$

Also, choose *sample points* $x_1^*, x_2^*, \ldots, x_n^*$, with

$$x_0 \leq x_1^* \leq x_1, \quad x_1 \leq x_2^* \leq x_2, \quad \ldots, \quad x_{n-1} \leq x_n^* \leq x_n.$$

The *definite integral of $f$ from $a$ to $b$* is

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta_n, \tag{B.1}$$

provided that the limit exists. The sums on the right side of (B.1) are called *Riemann sums of $f$ on* $[a, b]$.

Remarks. 1. In the above definition, the limit is supposed to exist and be independent of the choice of the sample points. This is not true for any old function, but it does hold for functions that are continuous on $[a, b]$ or have only a finite number of **jump** discontinuities inside $[a, b]$. For such functions, we may evaluate the integral using any choice of the sample points (e.g., we can take each $x_i^*$ to be the midpoint of the respective subinterval).

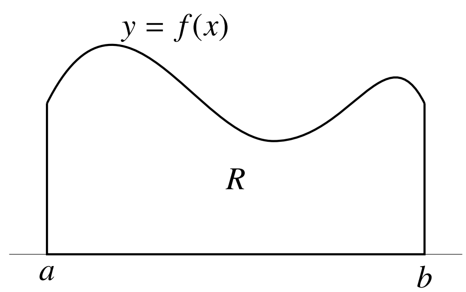



Figure B.1. The definite integral as area

2. In single-variable calculus, the definite integral of a continuous function is usually motivated by the *area problem*: If $f(x)$ is a continuous function on $[a, b]$, with $f(x) \geq 0$, calculate the area of the region (see Figure B.1)

$$R = \big\{(x, y) : a \leq x \leq b,\ 0 \leq y \leq f(x)\big\}.$$

The answer turns out to be the definite integral of $f(x)$:

$$\text{area}(R) = \int_a^b f(x)\,dx.$$

3. A different interpretation of the definite integral arises, if we divide both sides of (B.1) by $b - a$. Since $\Delta_n = (b-a)/n$, we have

$$\frac{1}{b-a}\sum_{i=1}^{n} f(x_i^*)\Delta_n = \frac{\Delta_n}{b-a}\sum_{i=1}^{n} f(x_i^*) = \frac{1}{n}\sum_{i=1}^{n} f(x_i^*).$$

Hence, dividing both sides of (B.1) by $b - a$, we derive

$$\frac{1}{b-a}\int_a^b f(x)\,dx = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n} f(x_i^*). \tag{B.2}$$

Note that the expression $\frac{1}{n}\sum_i f(x_i^*)$ is the average of the sample values $f(x_1^*), f(x_2^*), \dots, f(x_n^*)$, which are about equally spaced. Thus, the limit on the right side of (B.2) represents the average of "infinitely many, equally spaced" sample values of $f(x)$. Consequently, we may argue that the left side of (B.2) is the average value of $f(x)$:

$$\operatorname*{average}_{x\in[a,b]} f(x) = \frac{1}{b-a}\int_a^b f(x)\,dx. \tag{B.3}$$

This alternative interpretation of the definite integral is rarely emphasized in single-variable calculus, but is much more appropriate for the purpose of these notes, because we want to draw parallels between the properties of definite integrals and the properties of other types of integrals (double, triple, line, surface) whose geometric meaning may be less obvious.

## B.2. Properties of the definite integral

The following is a list of some general properties of definite integrals:

i) $\displaystyle\int_a^b [f(x) + g(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$

ii) $\displaystyle\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx$

iii) $\displaystyle\int_a^c f(x)\,dx + \int_c^b f(x)\,dx = \int_a^b f(x)\,dx$

iv) If $f(x) \le g(x)$, then $\displaystyle\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$

v) If $m \le f(x) \le M$, then $\displaystyle m(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$.

vi) $\displaystyle\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx, \qquad \int_a^a f(x)\,dx = 0.$

Of course, the most important property of definite integrals is the Fundamental Theorem of Calculus, also known as the Newton–Leibniz theorem.

Theorem B.1 (Fundamental theorem of calculus). *Suppose that $f(x)$ is a continuous function on $[a, b]$.*

i) *We have*

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

*that is, $\int_a^x f(t)\,dt$ is an antiderivative of $f(x)$.*

ii) *If $F(x)$ is any antiderivative of $f(x)$, then*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

APPENDIX C

# Polar Coordinates

## C.1. Definition

Consider a point $P$ in the plane. Recall that its *Cartesian coordinates* $(x, y)$ are the signed distances from $P$ to two perpendicular axes (the $x$- and $y$-axes). In this appendix, we recall another set of coordinates: the *polar coordinates* of $P$. The *polar coordinate system* consists of a point $O$ called the *pole* (or the *origin*) and a ray starting at $O$ called the *polar axis*. The polar coordinates of a point $P \neq O$ are a pair of real numbers $(r, \theta)$, where $r$ is the distance from $O$ to $P$ and $\theta \in [0, 2\pi)$ is the oriented angle between the polar axis and the line $OP$ (see Figure C.1); the polar coordinates of the pole $O$ are $(0, \theta)$, where $\theta$ is any number in $[0, 2\pi)$.

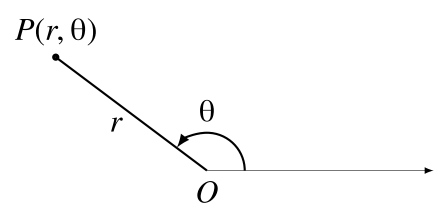



FIGURE C.1. Polar coordinates

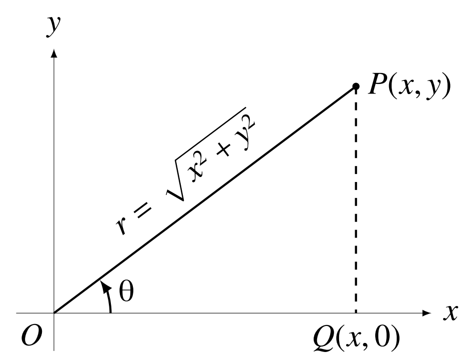



FIGURE C.2. Polar and Cartesian coordinates

## C.2. Relations between polar and Cartesian coordinates

We want to be able to convert polar coordinates to Cartesian and vice versa. In order to do so, we need to set a Cartesian and a polar coordinate systems in the same copy of the plane. We choose the two coordinate systems so that they share the same origin and the polar axis coincides with the positive $x$-axis (see Figure C.2). Since the distance between the origin and $P(x, y)$ is the same no matter what coordinate system we use, we have

$$r = |OP| = \sqrt{x^2 + y^2}.$$

The relation between $\theta$ and $(x, y)$ is more technical, but not much more complicated. First, consider a point $P(x, y)$ in the first quadrant. From $\triangle OPQ$ on Figure C.2,

$$x = r\cos\theta, \quad y = r\sin\theta, \quad \tan\theta = \frac{y}{x}.$$

The first two of these equations hold for all points $P(x, y)$, even if they do not lie in the first quadrant, that is, the Cartesian coordinates of a point $P(r, \theta)$ are obtained from the formulas

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{C.1}$$

On the other hand, the relation $\tan\theta = y/x$ extends only to the points with $x \neq 0$ and determines $\theta$ uniquely only on the half-planes $x > 0$ or $x < 0$. Thus, when we write the formulas for converting Cartesian into polar coordinates

$$r = \sqrt{x^2 + y^2}, \quad \tan\theta = \frac{y}{x}, \tag{C.2}$$

we must remember that the latter formula must be applied cautiously.

Example C.1. The point $P$ whose polar coordinates are $(2, \frac{2}{3}\pi)$ has Cartesian coordinates

$$x = 2\cos\left(\tfrac{2}{3}\pi\right) = -1, \quad y = 2\sin\left(\tfrac{2}{3}\pi\right) = \sqrt{3}.$$

□

Example C.2. The Cartesian point $P(2, 2)$ has polar coordinates

$$r = \sqrt{2^2 + 2^2} = 2\sqrt{2}, \quad \theta = \arctan\left(\frac{2}{2}\right) = \frac{\pi}{4}.$$

Here, we used that since $P$ lies in the first quadrant, the solution of the second equation in (C.2) is given by $\theta = \arctan(y/x)$. The Cartesian point $P(2, -2)$ has polar coordinates $(r, \theta)$, where

$$r = \sqrt{2^2 + (-2)^2} = 2\sqrt{2}$$

and $\theta$ is such that

$$\tan\theta = -1, \quad \theta \in \left(\tfrac{3}{2}\pi, 2\pi\right),$$

that is, $P(2\sqrt{2}, \frac{7}{4}\pi)$. □

Example C.3. The Cartesian point $P(-3, 4)$ has polar coordinates$(r, \theta)$, where

$$r = \sqrt{(-3)^2 + 4^2} = 5$$

and $\theta$ is such that

$$\tan\theta = -\frac{4}{3}, \quad \theta \in \left(\tfrac{1}{2}\pi, \pi\right).$$

Thus, $P(5, \arctan(-\frac{4}{3}) + \pi)$. □

## C.3. Polar curves

A *polar curve* is the set of points in the plane whose (polar) coordinates $(r, \theta)$ satisfy an equation $r = f(\theta)$, where $f$ is some function. Note that using equations (C.1), we can write every polar curve in parametric form:

$$x = r\cos\theta = f(\theta)\cos\theta, \quad y = r\sin\theta = f(\theta)\sin\theta.$$

Example C.4. The curve $r = 3$ is the set of points at distance 3 from the origin, that is, a circle of radius 3 centered at $O$ (whose Cartesian equation is $x^2 + y^2 = 9$, of course). Notice that we reach the same conclusion if we use (C.2) to convert the given polar equation into a Cartesian equation:

$$r = 3 \iff \sqrt{x^2 + y^2} = 3 \iff x^2 + y^2 = 9.$$

□

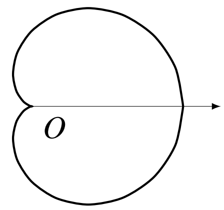


FIGURE C.3. The cardioid $r = 1 + \cos\theta$

EXAMPLE C.5. The curve $\theta = \frac{3}{4}\pi$ is the ray starting at the origin and intersecting the polar axis at an angle of 135°. □

EXAMPLE C.6. The curve $r = -2\cos\theta$ is a circle. To see this, we pass to Cartesian coordinates:

$$r = -2\cos\theta \quad \Longleftrightarrow \quad r^2 = -2r\cos\theta \quad \Longleftrightarrow \quad x^2 + y^2 = -2x.$$

The last equation is the Cartesian equation of a circle of radius 1 centered at $(-1, 0)$:

$$x^2 + y^2 = -2x \quad \Longleftrightarrow \quad x^2 + 2x + 1 + y^2 = 1 \quad \Longleftrightarrow \quad (x + 1)^2 + y^2 = 1.$$

□

EXAMPLE C.7. The curve $r = 1 + \cos\theta$ is known as a *cardioid*, because its graph is shaped like a heart (see Figure C.3). Unlike the curves in the previous examples, the cardioid does not have a simple Cartesian equation which we can use to sketch its graph. □

## C.4. Area and arc length in polar coordinates

THEOREM C.1. *Let $0 \le a < b \le 2\pi$ and let $f$ be a positive continuous function defined on $[a, b]$. Then the area $A$ of the region in the plane bounded by the curves*

$$\theta = a, \quad \theta = b, \quad r = f(\theta),$$

*is given by*

$$A = \frac{1}{2}\int_a^b [f(\theta)]^2\, d\theta.$$

EXAMPLE C.8. Find the area enclosed by the cardioid (see Example C.7).

SOLUTION. By Theorem C.1, the area is given by the integral

$$\begin{aligned}\frac{1}{2}\int_0^{2\pi}(1 + \cos\theta)^2\, d\theta &= \frac{1}{2}\int_0^{2\pi}\left(1 + 2\cos\theta + \cos^2\theta\right) d\theta \\ &= \frac{1}{2}\int_0^{2\pi}\left[1 + 2\cos\theta + \tfrac{1}{2}(1 + \cos 2\theta)\right] d\theta \\ &= \frac{1}{2}\left[\theta + 2\sin\theta + \tfrac{1}{2}\theta + \tfrac{1}{4}\sin 2\theta\right]_0^{2\pi} = \frac{3\pi}{2}.\end{aligned}$$

□

Theorem C.2. *Let $0 \leq a < b \leq 2\pi$ and let $f$ be a positive continuous function defined on $[a, b]$. Then the length $L$ of the curve $r = f(\theta)$ is given by*

$$L = \int_a^b \sqrt{f(\theta)^2 + f'(\theta)^2}\, d\theta.$$

Example C.9. Find the length $L$ of the cardioid.

Solution. By Theorem C.2,

$$\begin{aligned} L &= \int_0^{2\pi} \sqrt{(1 + \cos\theta)^2 + (-\sin\theta)^2}\, d\theta = \int_0^{2\pi} \sqrt{2 + 2\cos\theta}\, d\theta \\ &= \int_0^{2\pi} \sqrt{4\cos^2(\theta/2)}\, d\theta = \int_0^{2\pi} 2|\cos(\theta/2)|\, d\theta = 4\int_0^{\pi} |\cos u|\, du \\ &= 4\int_0^{\pi/2} \cos u\, du - 4\int_{\pi/2}^{\pi} \cos u\, du = 8. \end{aligned}$$

□

APPENDIX D

# Introduction to *Mathematica*

## D.1. *Mathematica* notebooks

A standard *Mathematica* file is called a notebook and has extension `.nb`. To create a new notebook, select `File→New→Notebook` from the main menu or press `Ctrl+N`. That will result in a new window filled with white space and named *Untitled-1*, *Untitled-2*, or similar.

Notebooks consist of one or more cells. Each cell represents a logical unit of the notebook and is shown by a blue square bracket on the right side of the screen. There are several different types of cells, but the three main types are: input cells, output cells, and text cells. In addition to these, there are types of cells that provide special formatting: title cells, subtitle cells, section cells, etc.

**D.1.1. Selecting and editing cells.** To select an existing cell, click on its cell bracket. A selected cell is indicated by a thick black bracket. To select more than one consecutive cell, click on one cell bracket and then drag over the others or click on the desired cells while pressing the `Ctrl` key down. To select cells that are not next to each other, click on the first one then, while holding the `Ctrl` key down, point and click on the next one and so on. This is particularly useful in printing a selection of cells. One can also copy and paste selected cells.

To edit the content of a cell, simply click on the desired location and start editing. One can also change an entire cell by selecting it and choosing a different style type from the pull-down menu in the toolbar.

**D.1.2. Creating a new cell.** If you position the mouse between two existing cells or after the last cell and then click, a horizontal line (called the cell insertion point or line) will be shown across the screen. This indicates where the next new cell will be inserted. Now, if you begin to type, *Mathematica* will create a new cell containing the text that you type. If you type several lines (using the `Enter` key to begin new lines), the cell bracket will grow to enclose everything you type. Note that, by default, every new cell you type in is created as an input cell. That means that *Mathematica* treats the symbols inside that cell as *Mathematica* code and may try to execute it. That is not always desirable, so you must always convert the style of cells you consider text from *Input* to *Text*. For example, a cell containing your name should have type *Text*, *Subtitle* or *Subsubtitle*, not *Input*.

**D.1.3. Executing cells.** To execute a cell (to get an answer to an arithmetic operation, for example), placing the cursor anywhere within the cell and then press `Shift+Enter` simultaneously or press the `Enter` key on the numeric keypad. As soon as a cell is executed, it will be labeled with `In[1]` for the first cell. The result will be shown in a new cell labeled `Out[1]`. The next cell executed will be labeled `In[2]` regardless of its actual location relative to the first cell. *Mathematica* pays attention only to the order in which the cell is executed rather than to its relative position with respect to other cells. Note that *Mathematica* will execute only input cells.

Sometimes, one wants to execute more than one cell. That can be done by selecting the desired cells and then pressing `Shift+Enter`. One can also execute the entire notebook by selecting `Evaluation→Evaluate Notebook` from the main menu. This will make *Mathematica* run through the entire notebook and execute all input cells in the order of appearance. Note that if there is text contained in an input cell, *Mathematica* will try to execute that text. The results of the latter can be comical—for example, executing an input cell containing the text "Today is 10/10/2010." results in the following:

```
 In[1]:= Today is 10/10/2010.
Out[1]:= 0.000497512 is Today
```

## D.2. Basic syntax

**D.2.1. Arithmetic operations and numerical values.** *Mathematica* supports the basic arithmetic operations. The following symbols are used:

- `+` indicates addition
- `-` indicates subtraction
- `*` indicates multiplication
- `/` indicates division
- `ˆ` indicates exponentiation

Here are some examples of those rules in action:

```
 In[1]:= 2+2*3
Out[1]:= 8

 In[2]:= 2+2 3
Out[2]:= 8

 In[3]:= 5ˆ3
Out[3]:= 125

 In[4]:= 2/3-1/6
Out[4]:= 1/2
```

Note the second example in this list: the space between the numbers 2 and 3 is interpreted as a multiplication. A space between two expressions is another way to indicate multiplication in *Mathematica*. Sometimes, even the space is unnecessary: e.g., "`2*x`", "`2 x`" and "`2x`" are all interpreted by *Mathematica* as "$2x$".

Notice that *Mathematica* returns the sum of two rational numbers as a rational number. That is because by default, *Mathematica* uses exact arithmetic. This means that a rational number, such as $\frac{1}{2}$, will normally be displayed as a fraction and not as the decimal 0.5. In fact, the latter is considered an approximation having a limited precision. If we want the numerical value of a fraction, we can use the built-in command `N[]`. In general, the command

```
N[..., n]
```

returns the numerical value of the expression in place of the ..., with $n$ correct decimal digits. When the value of $n$ is missing, the default precision is used. For example, to get the numeric value of $\frac{7}{94}$, we can try one of the following:

```
 In[1]:= N[7/94]
Out[1]:= 0.0744681

 In[2]:= N[7/94, 30]
Out[2]:= 0.0744680851063829787234042553191
```

An alternative way to get the numeric value of an expression is to add a left apostrophe ‘ after it. For example,

```
 In[1]:= 7/94‘
Out[1]:= 0.0744681
```

It is important to remember that *Mathematica* performs arithmetic computations in the usual mathematical order: exponentiations are performed first; then multiplications and divisions, from left to right; finally, additions and subtractions, also from left to right. When this natural order is to be changed, we use parentheses. For example:

| | | |
|---|---|---|
| `2^3 - 1/3*4 + 13` | represents | $2^3 - \frac{1}{3} * 4 + 13 = 8 - \frac{4}{3} + 13 = 19\frac{2}{3}$; |
| `(2^3 - 1)/(3*4 + 13)` | represents | $\dfrac{2^3 - 1}{3^4 + 13} = \frac{7}{94}$. |

**D.2.2. Elementary built-in functions and constants.** Some of *Mathematica*'s built-in constants are:

| | |
|---|---|
| `Pi` | $\pi = 3.14159265358979\ldots$ |
| `E` | $e = 2.71828182845904\ldots$ |
| `I` | $i = \sqrt{-1}$ |

Note that *Mathematica* is case-sensitive, so `pi` is **not** the number $\pi$! All built-in constants and functions in *Mathematica* begin with an upper-case letter. Furthermore, function arguments in *Mathematica* are enclosed in square brackets `[]`, as opposed to the parentheses used in usual mathematical notation. Hence, the input `Sqrt[16]` returns the value 4, whereas `sqrt[16]`, `Sqrt(16)` and `sqrt(16)` do not work. Table D.1 below lists some of the elementary built-in functions that you will use often.

| *Mathematica* function | Usual notation/Description |
|---|---|
| `Sin[x],Cos[x],Tan[x],Sec[x]` | $\sin(x), \cos(x), \tan(x), \sec(x)$ |
| `ArcSin[x],ArcCos[x],ArcTan[x]` | $\arcsin(x), \arccos(x), \arctan(x)$ |
| `Sqrt[x]` | $\sqrt{x}$ |
| `Abs[x]` | $\|x\|$ |
| `Log[x]` | $\ln(x)$ |
| `Log10[x]` | $\log(x)$, $\log_{10}(x)$ |
| `Log[b,x]` | $\log_b(x)$ |
| `N[x,n]` | Approximate value of $x$ with $n$ digits |
| `Exp[x], E^x` | $e^x$, exponential function base $e$ |
| `n!` | Factorial of $n$: $n! = 1 \cdot 2 \cdot 3 \cdots n$ |

TABLE D.1. *Mathematica* syntax of some basic functions

*Mathematica* gives exact results whenever possible. So, typing `Pi` as input will return $\pi$. If it is a numerical approximation to $\pi$ that we want, we should ask for it by using the `N[]`-command, as shown in the following examples:

```
 In[1]:= N[Pi]
Out[1]:= 3.14159

 In[1]:= N[Pi,50]
Out[2]:= 3.1415926535897932384626433832795028841971693993751
```

Here are some more examples of how *Mathematica* tries to use exact calculations:

```
 In[1]:= Sin[Pi/2]
Out[1]:= 1

 In[2]:= Log10[1000]
Out[2]:= 3

 In[3]:= Log10[200]
Out[3]:= Log[200]/Log[10]

 In[4]:= N[Log10[1000]]
Out[4]:= 2.30103

 In[5]:= 5!
Out[5]:= 120
```

As we mentioned earlier, *Mathematica* considers decimals, such as 0.5, approximate numbers. Therefore, most numerical calculations using such decimals are limited in precision. For example, compare the outputs of the following two calculations:

```
N[E^(1/2),50]
1.6487212707001281468486507878141635716537761007101

N[E^0.5,50]
1.64872
```

In the first calculation, we ask *Mathematica* to evaluate $e^{1/2} = \sqrt{e}$, which is an exact number having an infinite precision. The second calculation, on the other hand, is limited in precision by the presence of the exponent 0.5, which *Mathematica* treats differently from $\frac{1}{2}$. For this reason, it is preferred to represent decimals by fractions in numerical calculations.

**D.2.3. Defining functions of one variable.** To define a new function $f(x)$ in *Mathematica*, we typically use syntax of the form

```
f[x_] := ****
```

where the stars should be replaced by the appropriate formula. The symbols `:=` (the *delayed assignment operator*) separate the definition in two. The left-hand side of the definition (`f[x_]` above) specifies the name and the variable of the function; the right-hand side describes the function, usually by a formula involving the variable. For example, the command

```
f[x_] := x^2
```

defines the function $f(x) = x^2$. Note that we enclose the argument between square brackets and that we use an underscore character after the argument's name on the left-hand side of the definition.

*Mathematica* becomes aware of the definition of a user-defined function not when one has typed in the definition of the function, but only after the cell containing that definition has been executed. A common mistake is to type the definition of a function one wants to use and then to move on to using that function without executing the definition first. If a function has not been previously defined, *Mathematica* is likely to simply repeat any input that uses the undefined function. For example,

```
 In[1]:= someNewFunction[5]
Out[1]:= someNewFunction[5]
```

Note that when *Mathematica* does not recognize the name of a function or a variable, it displays that name in blue: e.g., `someNewFunction[]`. Once the definition of the function has been loaded (so that *Mathematica* knows what that name means), the highlighting changes from blue to black and calls to the function result in its evaluation. For example,

```
 In[1]:= someNewFunction[x_] := Sin[x]

 In[2]:= someNewFunction[5]
Out[2]:= Sin[5]

 In[3]:= N[someNewFunction[5],10]
Out[3]:= -0.9589242747
```

Remark. There are two conventions that you should follow when defining your own functions:

- To avoid conflicts with *Mathematica*'s built-in commands, **always** start the name of your functions with a small letter. Since all built-in commands start with a capital letter, this convention avoids the accidental overwriting of a built-in command.
- To avoid conflicts with the definitions of namesake functions, clear the name of the function before you define it. *Mathematica* remembers old definitions even if we delete the cells containing them (or even close the notebook in which they were introduced), and this can occasionally cause problems. To force *Mathematica* to forget such old definitions and start afresh, we use of the built-in function `Clear`. It is best to precede each function's definition with a call to `Clear` to erase potential prior definitions. For example, to define the function $g(x) = x^3 + 2x$, we should use

```
Clear[g]
g[x_] := x^3+2x
```

**D.2.4. Lists.** Many of *Mathematica*'s more sophisticated commands have inputs and outputs in the form of lists. A list is a collection of expressions (numbers, variables, functions, other lists, etc.) separated by commas and enclosed in a pair of curly braces. For example, `{x,0,2}` is a list of three elements: the variable $x$ and the numbers 0 and 2. Here is another example:

```
L = {1,3,5,7,9,11}
```

This defines a list L whose elements are the odd integers $1, 3, \dots, 11$. When a list has been given a name, we can use that name followed by a number in double brackets (`[[ ]]`) to access a specific element of the list. For example, the fourth element of the list L above can be obtained by `L[[4]]`.

Thus, we have now encountered four different types of brackets in *Mathematica*, each with its own dedicated use. These types are:

| | |
|---|---|
| `()` | Round parentheses indicate the order of operations.<br>Do not use braces or brackets instead! |
| `[]` | Square brackets enclose the argument of a function.<br>Do not use round parentheses instead! |
| `{}` | Curly brackets (braces) enclose a list. |
| `[[]]` | Double square brackets refer to a particular element of a list. |

## D.3. Plotting commands

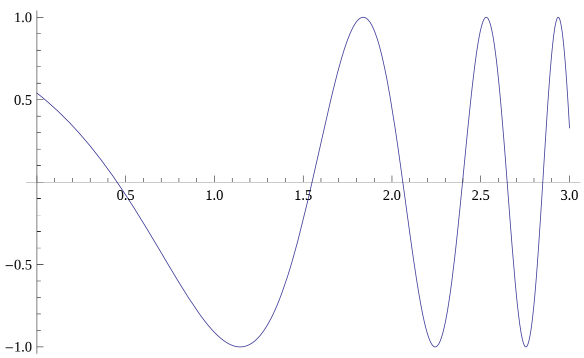
### D.3.1. Plotting functions of one variable.

D.3.1.1. *Basic plots.* To graph a function (or an expression) of one variable, we use the `Plot` command. The command takes at least two arguments, a function (or an expression), and a range. The range is a list of three elements: the variable in the expression, for example $x$, a minimum value for $x$, and a maximum value for $x$. For example, the command

```
Plot[Sin[x], {x,0,2Pi}]
```

will plot the sine function over the range $0 \le x \le 2\pi$. One can also graph a previously defined function by using its name as the first argument of the `Plot` command as shown in the following example.

```
Clear[f]
f[x_] := Cos[E^x]
Plot[f[x], {x,0,3}]
```



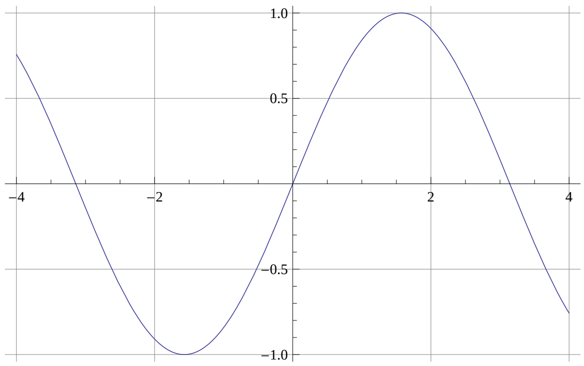

D.3.1.2. *Graphics options.* An option is a rule of the form `OptionName->OptionValue`, where `OptionName` is the name of the option and `OptionValue` is the value that it is assigned. If an option is not supplied explicitly, the default value is used. The following command will display all the possible plotting options in *Mathematica* together with their default settings.

```
Options[Plot]
{AlignmentPoint->Center, AspectRatio->1/GoldenRatio,
 Axes->True, AxesLabel->None, AxesOrigin->Automatic,
 ...
 Ticks->Automatic, TicksStyle->{},
 WorkingPrecision->MachinePrecision}
```

However, a much better way to learn about the various options of the `Plot` command is to read the documentation entry on it in *Mathematica*'s Documentation Center (under the *Help* menu). Here, we present only a few of the more commonly used options.

**`GridLines`.** We can use this option to simulate graphing paper:

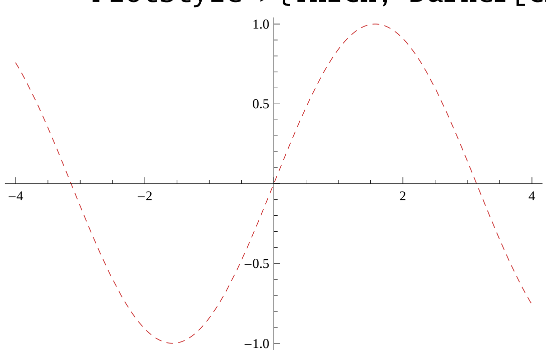
```
Plot[Sin[x], {x,-4,4}, GridLines->Automatic]
```

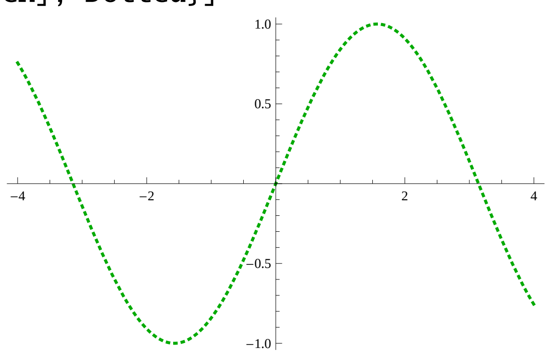



The two basic values for this option are `None` (the default) and `Automatic`, which makes *Mathematica* generate the most reasonable rectangular grid for the given graph. There are also more sophisticated alternatives that can be used to control the size and formatting of the grid—see the documentation entry on `GridLines`.

**`PlotStyle`.** We can use this option to change the appearance of the plot. The value of `PlotStyle` can be a single directive (e.g., `Blue` or `Thick`) or a list of such directives. Often used styles are: `Thin` or `Thick`, `Dashed` or `Dotted`, and various colors. In particular, there are several ways to specify the color of the plot. The most common is by the function `RGBColor`, which takes three arguments: the amount of red, the amount of green, and the amount of blue. These arguments must be numbers between 0 and 1, and they need not add up to 1. There are also several predefined colors that can be assigned to `PlotStyle`; those include: `Red, Green, Blue, Yellow, Cyan, Magenta, LightOrange, Lighter[Purple], Darker[Brown]`, and similar modifications. Here are a couple of examples using different styling options:

```
Plot[Sin[x], {x,-4,4},
     PlotStyle->{RGBColor[0.8,0.2,0.2], Dashed}]
Plot[Sin[x], {x,-4,4},
     PlotStyle->{Thick, Darker[Green], Dotted}]
```



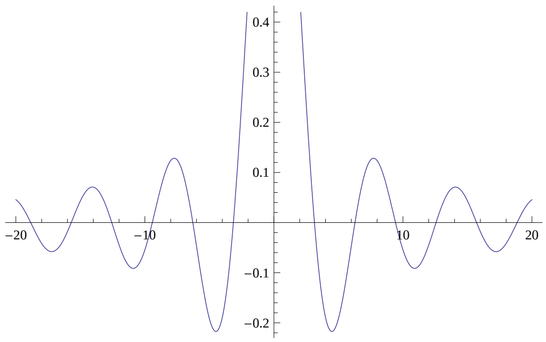



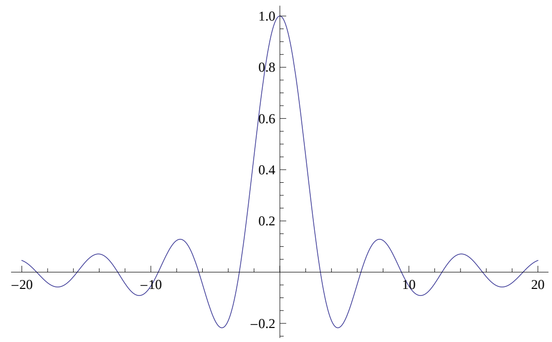
**PlotRange.** We can use this option to specify the range of $y$-coordinates to include in the plot. The most often used variants of this option are

```
PlotRange->All
PlotRange->{a,b}
```

When PlotRange is set to All, *Mathematica* plots all the points on the graph. When PlotRange is set to an actual range {a,b}, *Mathematica* plots only those points on the graph that have $y$-coordinates in the range $a \le y \le b$. When this option is not set, *Mathematica* displays only points where it thinks the function is "interesting". Sometimes, this does not work very well and *Mathematica* needs to be told to plot all points. For example, compare the outputs of the following two attempts to plot $(\sin x)/x$:

```
Plot[Sin[x]/x, {x,-20,20}]
Plot[Sin[x]/x, {x,-20,20}, PlotRange->All]
```



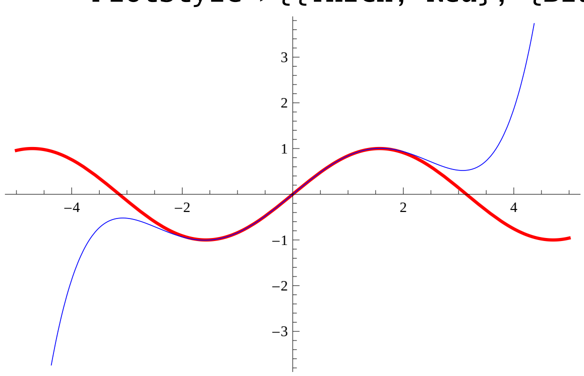



D.3.1.3. *Superimposing several graphs.* One can superimpose the graphs of several functions by supplying the functions as a list in the first argument of a Plot command. The following command will plot $\sin x$ and $x - x^3/6 + x^5/120$.

```
Plot[{Sin[x], x-x^3/6+x^5/120}, {x,-5,5}]
```

One can use the PlotStyle option to get more easily distinguishable curves as seen in the following example. Notice that the argument of PlotStyle is a list of two lists: the first list is the list of plotting options applied to the graph of $\sin x$ and the second list is the list of plotting options applied to the graph of $x - x^3/6 + x^5/120$.

```
Plot[{Sin[x], x-x^3/6+x^5/120}, {x,-5,5},
     PlotStyle->{{Thick, Red}, {Blue}}]
```

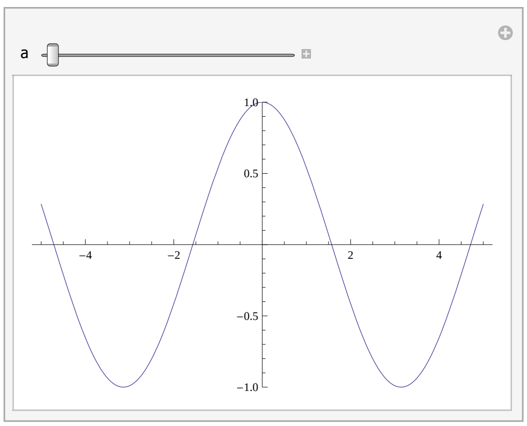


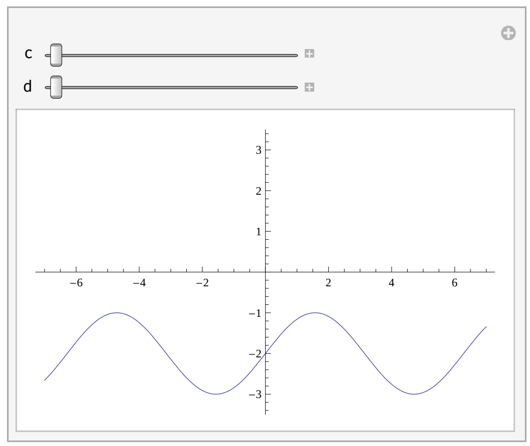
## D.4. Dynamic output

**D.4.1. `Manipulate`.** A major strength of *Mathematica* is its ability to handle dynamically updated variables. The basic command to access *Mathematica*'s dynamic features is `Manipulate`. It takes two or more inputs: the first input is the expression that we want to update dynamically and the subsequent inputs are lists indicating ranges of parameters that appear in the expression. Here are two examples of dynamic plots:

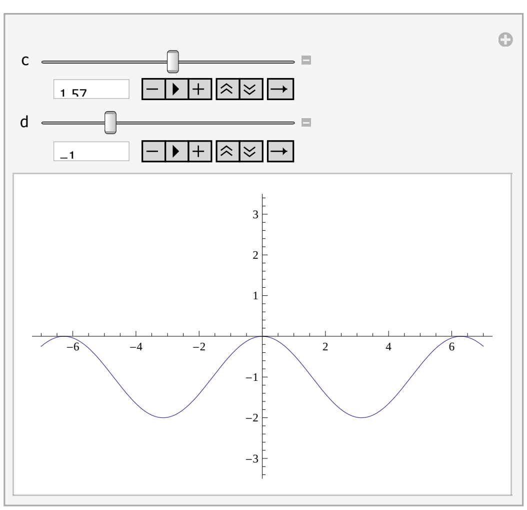
```
Manipulate[
  Plot[Cos[x], {x,-a,a}, PlotRange->{-1,1}], {a,5,0.001}]
```

```
Manipulate[
  Plot[Sin[x+c]+d, {x,-7,7}, PlotRange->{-3.5,3.5}],
  {c,0,3}, {d,-2,2}]
```



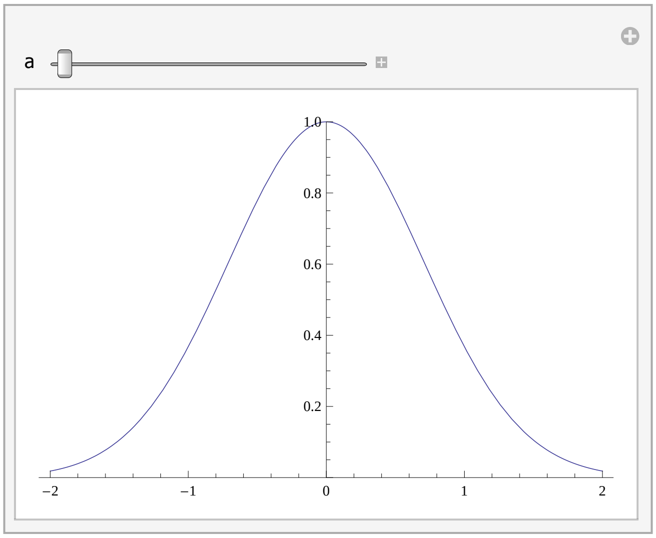




Note the little + buttons next to the sliders in the above dynamic outputs. Clicking on one of those opens a navigation menu for that particular slider. The navigation menu allows us to see animated versions of the output, to find the current value of the parameter, or to set the parameter equal to a particular value. For example, we can see that setting the two parameters to $c = 1.57$ and $d = -1$ in the second dynamic plot results in a sine curve that looks exactly as the graph $y = \cos x - 1$.



Although dynamic evaluations are a wonderful tool, they also require some care to avoid falling in certain traps. Consider the following dynamic plot.

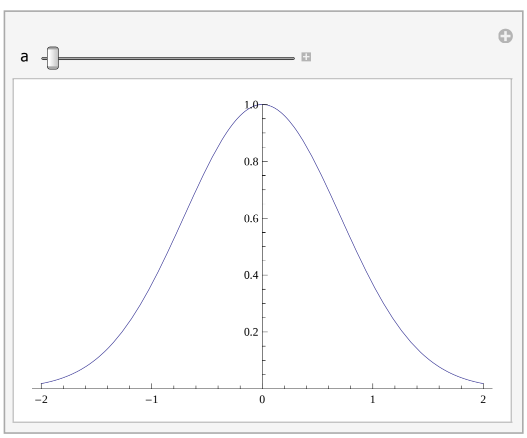
```
Clear[f]
f[x_] := Exp[-x^2]
Manipulate[Plot[f[x], {x,-a,a}, PlotRange->{0,1}], {a,2,4}]
```

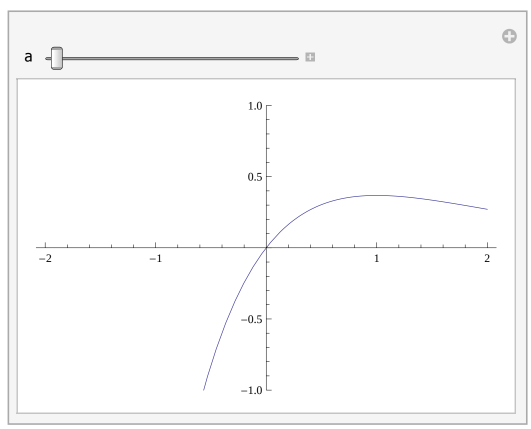



We see a nice bell-shaped curve. Now, let us take a look at another dynamic plot.

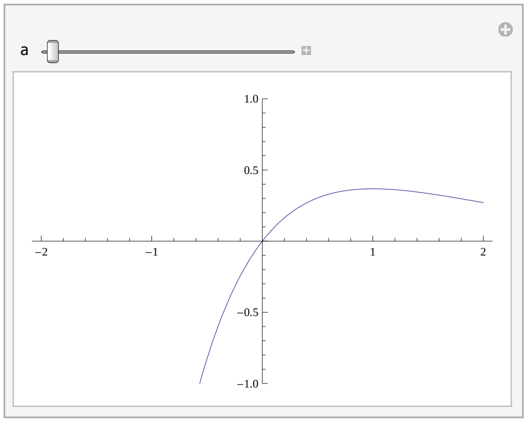
```
Clear[f]
f[x_] := x*Exp[-x]
Manipulate[Plot[f[x], {x,-a,a}, PlotRange->{-1,1}], {a,2,4}]
```

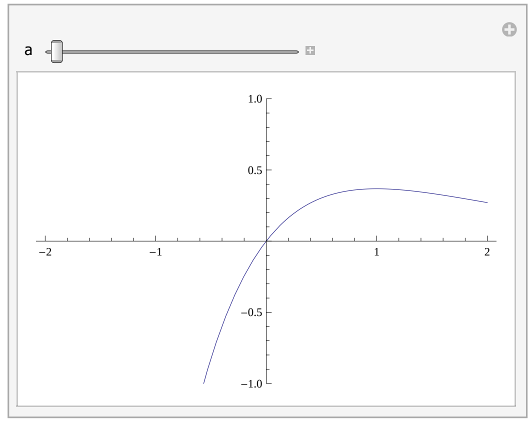
You may expect that after executing these commands, the pair of plots will look something like this:



However, the reality will look more like this:




This is not a bug! Since both plots are supposed to present the graph of the function $f(x)$ and are to be updated dynamically, changing the definition of $f(x)$ will result in both plots changing to match the new definition of $f(x)$. For example, if we execute the commands

```
Clear[f]
f[x_] := x^2
```

both plots above will change to dynamic plots of the parabola $y = x^2$. The safest way to avoid such confusion is not to reuse function names when using `Manipulate`. For example, the dynamic plots of the functions $e^{-x^2}$ and $xe^{-x}$ above should have been entered as

```
Clear[f1, f2]
f1[x_] := Exp[-x^2]
Manipulate[Plot[f1[x], {x,-a,a}, PlotRange->{0,1}], {a,2,4}]

f2[x_] := x*Exp[-x]
Manipulate[Plot[f2[x], {x,-a,a}, PlotRange->{-1,1}], {a,2,4}]
```

## D.5. Calculus-related *Mathematica* commands

*Mathematica* has built-in commands for many of the operations that you will learn in calculus. In this section, we list some of those commands.

**D.5.1. Limits.** To evaluate a limit in *Mathematica*, we use the built-in command `Limit` is easy. For example, to evaluate

$$\lim_{x\to 0} \frac{e^x - 1 - x - x^2/2}{x^3},$$

we just execute the following the command.

```
 In[1]:= Limit[(E^x-1-x-x^2/2)/x^3, x->0]
Out[1]:= 1/6
```

(Recall that *Mathematica* uses `E` for $e = 2.71828\ldots$)

One-sided limits can be evaluated using the option `Direction`. The value assigned to this option can be any nonzero number, but only the sign of the number really matters: if $a > 0$, then

`Direction->a` specifies a left limit; if $a < 0$, then `Direction->a` specifies a right limit. Here are several examples:

```
 In[1]:= Limit[Abs[x]/x, x->0, Direction->1]
Out[1]:= -1

 In[2]:= Limit[Abs[x]/x, x->0, Direction->-1]
Out[2]:= 1

 In[3]:= Limit[Exp[1/(x-1)], x->1, Direction->1]
Out[3]:= 0

 In[4]:= Limit[Exp[1/(x-1)], x->1, Direction->-1]
Out[4]:= Infinity
```

In particular, the last input/output pair demonstrates what happens when a limit is infinite. It should be noted that if no direction is specified and the two-sided limit does not exist, then *Mathematica* automatically reverts to evaluation of the right limit (`Direction->-1`). This explains, for example, the following evaluation.

```
 In[1]:= Limit[Abs[x]/x, x->0]
Out[1]:= 1
```

*Mathematica* can also handle limits at $\pm\infty$. To evaluate a limit as $x \to \infty$, we simply replace the number in the second argument of `Limit` by the word `Infinitiy` (or by `-Infinity`, for limits as $x \to -\infty$). For example,

```
 In[1]:= Limit[Log[x]/x, x->Infinity]
Out[1]:= 0

 In[2]:= Limit[\Sqrt[x^2+4x]/(3x - 7), x->-Infinity]
Out[2]:= -1/3
```

Finally, a few words about limits which do not exist. We already know what happens when a limit is infinite. What if the limit fails to exist for a different reason? For example, it is known from calculus that the limit $\lim_{x\to 0} \sin(1/x)$ does not exist. *Mathematica* reports that fact as follows:

```
 In[1]:= Limit[Sin[1/x], x->0]
Out[1]:= Interval[{-1,1}]
```

So, in this case *Mathematica* returns a whole range as the value of the limit instead of a single number. That means that the limit does not exist.

**D.5.2. Differentiation.** *Mathematica* has a built-in command to calculate the derivative of a function. To find $f'(x)$, given $f(x)$, we simply execute one of the following two commands

```
 In[1]:= f'[x]
Out[1]:= D[f[x],x]
```

For example, if $f(x) = x^3 + \cos(x^2)$, then

```
 In[1]:= f'[x]
Out[1]:= 2x - 2x Sin[x^2]
```

```
 In[2]:= D[f[x],x]
Out[2]:= 2x - 2x Sin[x^2]
```

Note that the latter version of the differentiation command works also for partial derivatives of multivariable functions. For example, if

$$f(x, y) = \sin(x + 2y) + (x^2 - y^2)\exp(x^2 + 3y^2),$$

then `D[f[x,y],x]` and `D[f[x,y],y]` calculate, respectively, the $x$- and $y$-partials of $f(x, y)$.

```
 Clear[f]
 In[1]:= f[x_,y_] := Sin[x+2y] + (x^2-y^2)Exp[x^2+3y^2]
         D[f[x,y], x]
Out[2]:= (2x^3-2xy^2+2x)E^(x^2+3y^2) + Cos[x+2y]

 In[3]:= D[Sin[x+2y] + (x^2-y^2)Exp[x^2+3y^2], y]
Out[3]:= (6yx^2-6y^3-2y)E^(x^2+3y^2) + 2Cos[x+2y]
```

**D.5.3. Integration.** *Mathematica*'s basic command for evaluation of integrals is `Integrate`. It can evaluate both definite and indefinite integrals. In its two basic forms, this command looks like this

| | |
|---|---|
| `Integrate[f[x], x]` | evaluates $\int f(x)\,dx$ |
| `Integrate[f[x], {x,a,b}]` | evaluates $\int_a^b f(x)\,dx$ |

The second version of `Intergate` tries to evaluate the given integral exactly in "closed form," but that is not always possible. In such cases, it may be preferable to settle for a very close numeric approximation to the value of the definite integral. *Mathematica* provides another command for this purpose. The command

```
NIntegrate[f[x], {x,a,b}]
```

uses sophisticated algorithms to approximate the definite integral $\int_a^b f(x)\,dx$ and the result is usually fairly close to the exact answer. By default, the answer that you obtain by executing this command is accurate to at least 6 digits; for higher precision, the options `PrecisionGoal` and `WorkingPrecision` can be used (see the documentation entry on `NIntegrate`). Here are some examples of integral evaluations:

```
 In[1]:= Integrate[ArcTan[2x+3], x]
Out[1]:= (3/2)ArcTan[3+2x] + x*ArcTan[3+2x] - (1/4)Log[5+6x+2x^2]

 In[2]:= Integrate[Exp[-x^2], {x,0,2}]
Out[2]:= (1/2) Sqrt[Pi] Erf[2]

 In[3]:= NIntegrate[Exp[-x^2], {x,0,2}]
Out[3]:= 0.882081

 In[4]:= NIntegrate[Exp[-x^2], {x,0,2}, PrecisionGoal->10]
Out[4]:= 0.882081

 In[5]:= NIntegrate[Exp[-x^2], {x,0,2}, WorkingPrecision->10]
Out[5]:= 0.8820813907
```

```
 In[6]:= NIntegrate[Exp[-x^2], {x,0,2},
   WorkingPrecision->10, PrecisionGoal->10]
Out[6]:= 0.8820813908
```

The last three examples in this series are meant as a warning that use of the high-precision options of `NIntegrate` requires some care. The safest solution is to set both `PrecisionGoal` and `WorkingPrecision` to the same number (as in the last example). What exactly goes wrong in the two earlier examples is beyond the reach of this introduction to *Mathematica*; suffice it to say that if you need to use these options, you should first study carefully the documentation.

## D.6. User-defined commands

*Mathematica*'s includes some programming features which allow the sophisticated user to define his/her own commands. User-defined commands share features with both built-in commands and user-defined functions. Similar to built-in commands, user-defined commands may have multiple inputs (of varying types) and may produce all kinds of output, including numeric answers, text, plots, etc. On the other hand, similar to user-defined functions, we must type and **execute** the definition of a user-defined command before we can use it. In the laboratories, we will make extensive use of such custom commands, so it is important to learn how to apply them properly, even if the actual "nuts and bolts" of the command seem mysterious.

The definition of the command `newCommand` always takes one of the following two forms:

```
newCommand[*list of arguments*] := Module[
  *some Mathematica code*
 ]
newCommand[*list of arguments*] := DynamicModule[
  *some Mathematica code*
 ]
```

Here, `*list of arguments*` and `*some Mathematica code*` are placeholders for the actual list of arguments and the definition of the command. In order to be able to use such a command, we must do two things:

- We must familiarize ourselves with the arguments the command accepts and the output that it returns. These are usually described in the manual, shortly before or after the definition of the command.
- We must **execute** the cell containing the definition of the command. Until then all occurrences of the name of the command will appear in blue (e.g., `newCommand`) and *Mathematica* will simply repeat any command containing a call to the new command.

The next three pages demonstrate these points. They present printouts of the definition of a module that appears in an actual TU calculus laboratory assignment and several calls to it. On the first printout, the definition of `newton` has not been executed and the "output" merely repeats the input. On the second printout, the definition has been executed but the arguments do not match the expected ones—in one case two inputs are missing, and in the other $n$ is not an integer. On the third printout, all the rules have been followed and the call to `newton` has resulted in an actual plot.

## 12.3. Geometric Interpretation

Newton's method for solving $f(x) = 0$ is equivalent to the following geometric process. Pick a starting number $x_0$ near the root $r$. Draw the tangent line to the graph of $f$ at the point $(x_0,\ f(x_0))$. The $x$-intercept, $x_1$, of the tangent line is generally closer to $r$ than $x_0$ was. The process is then repeated until some prescribed accuracy is attained.

The module below, called **newton**, repeats the above construction $n$ times, where $n$ is a given positive integer. It stores the values of $x_0, x_1, \ldots, x_n$ in a list named **data** and produces a plot of the iteration. The calling parameters of the module are:

$f$ = the function $f$;
$a$ = the left end of some interval that contains $r$;
$b$ = the right end of the interval;
$x_0$= the initial value to start the iteration;
$n$ = the number of iterations desired.

```
newton[f_, a_, b_, x0_, n_] := Module[{g}, g[x_] = x - f[x] / f'[x];
  data = NestList[g, x0, n]; data2 = Flatten[Table[
     {data[[i]], 0, data[[i]], f[data[[i]]]}, {i, Length[data]}]];
       data3 = Partition[data2, 2];
       data4 = Drop[data3, -1];
       p1 = Plot[f[x], {x, a, b}, PlotStyle -> {Red, Thick},
    AxesLabel -> {"x", "y"}, PlotRange -> All];
       p2 = Graphics[{Hue[.6], Line[data4]}];
       Show[{p1, p2}]]
```

We now apply this module to the function $f(x) = x^2 - 2$ with $x_0 = 0.5$ and the interval $[0,\ 2.5]$ in order to obtain a geometric illustration of the first iteration ($n = 1$).

```
Clear[f];
f[x_] := x^2 - 2
```

```
In[1]:= newton[f, 0, 2.5, .5, 1]
Out[1]= newton[f, 0, 2.5, 0.5, 1]
```

## 12.3. Geometric Interpretation

Newton's method for solving $f(x) = 0$ is equivalent to the following geometric process. Pick a starting number $x_0$ near the root $r$. Draw the tangent line to the graph of $f$ at the point $(x_0, f(x_0))$. The $x$-intercept, $x_1$, of the tangent line is generally closer to $r$ than $x_0$ was. The process is then repeated until some prescribed accuracy is attained.

The module below, called **newton**, repeats the above construction $n$ times, where $n$ is a given positive integer. It stores the values of $x_0, x_1, ..., x_n$ in a list named **data** and produces a plot of the iteration. The calling parameters of the module are:

$f$ = the function $f$;
$a$ = the left end of some interval that contains $r$;
$b$ = the right end of the interval;
$x_0$= the initial value to start the iteration;
$n$ = the number of iterations desired.

```
In[1]:= newton[f_, a_, b_, x0_, n_] := Module[{g}, g[x_] = x - f[x] / f'[x];
   data = NestList[g, x0, n]; data2 = Flatten[Table[
      {data[[i]], 0, data[[i]], f[data[[i]]]}, {i, Length[data]}]];
        data3 = Partition[data2, 2];
        data4 = Drop[data3, -1];
        p1 = Plot[f[x], {x, a, b}, PlotStyle -> {Red, Thick},
     AxesLabel -> {"x", "y"}, PlotRange -> All];
        p2 = Graphics[{Hue[.6], Line[data4]}];
        Show[{p1, p2}]]
```

We now apply this module to the function $f(x) = x^2 - 2$ with $x_0 = 0.5$ and the interval $[0, 2.5]$ in order to obtain a geometric illustration of the first iteration ($n = 1$).

```
In[2]:= Clear[f];
        f[x_] := x^2 - 2

In[4]:= newton[f, .5, 2]

Out[4]= newton[f, 0.5, 2]

In[5]:= newton[f, 0, 2.25, .5, 1.5]

        NestList::intnm : Non-negative machine-sized integer expected at position 3 in NestList[g$116, 0.5, 1.5]. ≫
```

## 12.3. Geometric Interpretation

Newton's method for solving $f(x) = 0$ is equivalent to the following geometric process. Pick a starting number $x_0$ near the root $r$. Draw the tangent line to the graph of $f$ at the point $(x_0, f(x_0))$. The $x$-intercept, $x_1$, of the tangent line is generally closer to $r$ than $x_0$ was. The process is then repeated until some prescribed accuracy is attained.

The module below, called **newton**, repeats the above construction $n$ times, where $n$ is a given positive integer. It stores the values of $x_0, x_1, \ldots, x_n$ in a list named **data** and produces a plot of the iteration. The calling parameters of the module are:

$f$ = the function $f$;
$a$ = the left end of some interval that contains $r$;
$b$ = the right end of the interval;
$x_0$= the initial value to start the iteration;
$n$ = the number of iterations desired.

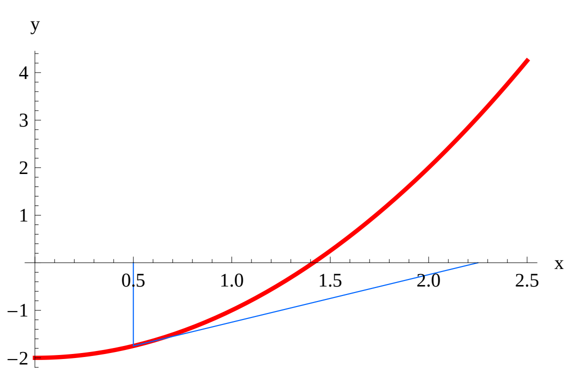
```
In[2]:= newton[f_, a_, b_, x0_, n_] := Module[{g}, g[x_] = x - f[x] / f'[x];
          data = NestList[g, x0, n]; data2 = Flatten[Table[
              {data[[i]], 0, data[[i]], f[data[[i]]]}, {i, Length[data]}]];
                 data3 = Partition[data2, 2];
                 data4 = Drop[data3, -1];
                 p1 = Plot[f[x], {x, a, b}, PlotStyle -> {Red, Thick},
            AxesLabel -> {"x", "y"}, PlotRange -> All];
                 p2 = Graphics[{Hue[.6], Line[data4]}];
                 Show[{p1, p2}]]
```

We now apply this module to the function $f(x) = x^2 - 2$ with $x_0 = 0.5$ and the interval $[0, 2.5]$ in order to obtain a geometric illustration of the first iteration ($n = 1$).

```
In[3]:= Clear[f];
        f[x_] := x^2 - 2
```

```
In[5]:= newton[f, 0, 2.5, .5, 1]
```

Out[5]=



APPENDIX E

# Answers to Selected Exercises

§1. **1.** $Q$; none; $Q$; $R$. **3.** $\sqrt{14}$, $\sqrt{14}$, $\sqrt{14}$; equilateral. **4.** $(x-1)^2+(y-1)^2+(z-1)^2 = 10$. **6.** A plane parallel to the $yz$-plane and 3 units behind it. **8.** All points on or between the planes $y = -1$ and $y = 2$. **10.** All points between the spheres centered at $O$ that have radii 2 and 3. **13.** $x^2+y^2+z^2 < 4$, $z > 0$. **15.** $(\frac{3}{2}\sqrt{2}, \frac{3}{2}\sqrt{2}, -3)$. **17.** $(0, -3, -1)$. **18.** $(\sqrt{2}, 7\pi/4, 4)$. **20.** $(2, \pi/6, -2)$. **24.** $(-\frac{3}{2}\sqrt{3}, -\frac{3}{2}, 0)$. **25.** $(0, \sqrt{2}, \sqrt{2})$. **26.** $(\sqrt{6}, \pi/4, \arccos(2/\sqrt{6}))$. **27.** $(2, 3\pi/4, 3\pi/4)$. **30.** $r^2 = 2z$, $\rho^2 \sin^2\phi = 2\rho\cos\phi$. **32.** $z = r$, $\phi = \pi/4$. **33.** $z = -2$, a horizontal plane. **35.** $x^2+y^2+z^2 = 4$, a sphere with radius 2 and center $O$. **37.** $x^2+y^2-2z = 0$. **40.** The sphere with radius 2 and center $O$. **42.** All points above the $xy$-plane that lie on or between the spheres with center $O$ and radii 2 and 4. **44.** All points below the $xy$-plane that lie on or inside the infinite cylinder with radius 1 and axis the $z$-axis.

§2. **2.** $2, \langle -5, -2\rangle, 2\sqrt{2}, \langle 4, -7\rangle$. **4.** $\sqrt{3}, \langle 5, -3, -4\rangle, \frac{1}{2}\sqrt{11}, \langle 3, -2, -1\rangle$. **8.** 0. **10.** 3. **12.** $\pi/3$. **14.** $\arccos(-1/\sqrt{28}) \approx 101°$. **15.** $\pi/2, \pi/3, \pi/6$. **18.** Neither. **19.** $\langle -1, -2, 2\rangle$. **21.** $\langle -7, -3, 5\rangle$. **23.** $\langle 2, -1, 3\rangle$. **26.** $2\sqrt{1329}$. **28.** 80. **30.** Undefined. **32.** $\frac{1}{2}\sqrt{110}$. **35.** $ab\mathbf{e}_1 - b\cos\phi\mathbf{e}_2 - b\sin\phi\mathbf{e}_3$.

§3. **2.** $x = t$, $y = -2+3t$, $z = 3+2t$ $(t \in \mathbb{R})$; $x = \frac{y+2}{3} = \frac{z-3}{2}$. **3.** Skew. **5.** $\arccos\left(-\frac{14}{3\sqrt{38}}\right) \approx 139°$. **7.** $3x-2z+3 = 0$. **9.** $-x+2y+z+5 = 0$. **10.** There is no such plane. **11.** $4x+5y-3z+1 = 0$. **13.** $\frac{x-2.4}{2} = \frac{y+0.4}{3} = \frac{z}{5}$. **16.** $\pi/2$. **17.** $\frac{8}{9}$.

§4. **1.** (a) $\langle \ln 2, \sin 2, e\rangle$; (b) $\langle 1, 2, 1\rangle$. **2.** $\left\langle \frac{2t}{t^2+1}, \frac{2}{\sqrt{1-4t^2}}, e^t\right\rangle$. **4.** $\left\langle 3t^2, 0, -6e^{2t}\right\rangle$. **7.** $\left\langle 0, \frac{1}{2}\left(e^2-e^{-2}\right), 0\right\rangle$.

§5. **1.** (a) Yes, $t = 0$; ; (c) no. **2.** Yes, $\mathbf{r}_1(0) = \mathbf{r}_2(2) = \langle 0, 0, 0\rangle$. **4.** $\left\langle \frac{3}{\sqrt{10}}, 0, \frac{1}{\sqrt{10}}\right\rangle$. **5.** $x = -1$, $y = 1$. **7.** $x = -5$, $z = 2$. **8.** $(0, \pm 5, -2)$. **9.** Smooth. **11.** Piecewise smooth. **14.** $\frac{16}{3\sqrt{53}}$. **16.** $1+\ln 2$. **17.** (a) $s(a) = 0$, $s(b) = L$; (e) $\mathbf{f}(s) = \left\langle \left(1+\frac{3s}{5}\right)\sin\left(4\ln(1+\frac{3s}{5})\right), \left(1+\frac{3s}{5}\right)\cos\left(4\ln(1+\frac{3s}{5})\right), 5\right\rangle$, $0 \le t \le 5$.

§6. **1.** Domain: all points $(x, y)$; range: $[0, \infty)$. **3.** Domain: the points $(x, y, z)$ for which $(x, y)$ is not on the hyperbola $x^2 - y^2 = 1$; range: all real numbers. **5.** The points on and above the line $y = -x/2$, except for those that lie on the circle $x^2+y^2 = 1$. **6.** Levels $-1$ and 0: no level curves; level 1: the line $y = 0$; level 2: the lines $y = 1$ and $y = -1$; level 3: the lines $y = \sqrt{2}$ and $y = -\sqrt{2}$. **8.** Level $-1$: no level curve; level 0: the level curve consists of a single point, $(0, 0)$; for $k = 1, 2, 3$, the $k$-level curve is the circle $x^2+y^2 = k$. **9.** $-29$. **11.** 0. **12.** D.N.E. **14.** 2. **15.** All $(x, y)$. **16.** All $(x, y) \neq (0, 0)$. **17.** The points with $x^2+y^2+4z^2 < 4$.

§7. **1.** Parabolas, parabolas, hyperbolas, hyperbolic paraboloid. **3.** Ellipses, ellipses, ellipses, ellipsoid. **6.** $(x-2)^2+4y^2+(z+1)^2 = 9$, an ellipsoid centered at $(2, 0, -1)$. **7.** $\mathbf{r}(t) = \langle 3\cos t, 3\sin t, 3\cos t + 3\sin t + 2\rangle$, $0 \le t \le 2\pi$. **8.** The elliptic paraboloid $z = x^2+y^2$. **11.** $\mathbf{r}(y, z) = \left\langle -2\sqrt{y^2+z^2+4}, y, z\right\rangle$, $-\infty < y, z < \infty$. **13.** (a) $\mathbf{r}(x, y) = \left\langle x, y, 4-x^2-y^2\right\rangle$, $x^2+y^2 \le 4$; (b) $\mathbf{r}(u, v) = \left\langle u\cos v, u\sin v, 4-u^2\right\rangle$, $0 \le u \le 2, 0 \le v \le 2\pi$. **14.** $\mathbf{r}(u, v) = \langle (b+a\cos v)\cos u, (b+a\cos v)\sin u, a\sin v\rangle$, $0 \le u, v \le 2\pi$.

§8. **1.** $4x^3 - 8xy^{3/2}$, $-6x^2y^{1/2}$. **3.** $3(3x+7yz)^{-1}$, $7z(3x+7yz)^{-1}$, $7y(3x+7yz)^{-1}$. **4.** $\left(2x^2+xy\right)\left(x^2+xy+y^2\right)^{x-1} + \left(x^2+xy+y^2\right)^x \ln\left(x^2+xy+y^2\right)$, $\left(2x^2+xy\right)\left(x^2+xy+y^2\right)^{x-1}$. **7.** $12x^2-8y^{3/2}$, $-12xy^{1/2}$, $-3x^2y^{-1/2}$. **8.** $-(2x+y)^2\cos\left(x^2+xy+y^2\right) - 2\sin\left(x^2+xy+y^2\right)$, $-(2x+y)(x+2y)\cos\left(x^2+xy+y^2\right) - \sin\left(x^2+xy+y^2\right)$, $-(x+2y)^2\cos\left(x^2+xy+y^2\right) - 2\sin\left(x^2+xy+y^2\right)$. **10.** $-12y^{1/2}$. **13.** $\left(6z^2+4yz^3\right)e^{2yz}$, $\left(12xz+24xyz^2+8xy^2z^3\right)e^{2yz}$. **14.** $2y(2t+3)+(2x+3y^2)(2t)$. **17.** $\dfrac{2(y-3z)e^t - 3(x+2z)e^{-t} + 2(2y-3y)e^{2t}}{xy+2yz-3xz}$. **18.** $(4xy-y^2)6s + (2x^2-2xy)3t$, $(4xy-y^2)2t + (2x^2-2xy)3s$. **21.** $(-y\sin xy + y3^x\ln 3)2s + (-x\sin xy + 3^x) + 2z(t+1)$,

$(-y\sin xy + y3^x \ln 3) + (-x\sin xy + 3^x)(-2t) + 2z(sg+1)$. **23.** $-6e^{-24}$. **24.** $\frac{9}{7}$. **27.** $\ln 7 + \frac{3}{7}(x-2) + \frac{5}{7}(y-1)$. **29.** $\epsilon_1(x,y) = (y-2)$, $\epsilon_2(x,y) = 0$. **30.** (a) 0, 0. **31.** (a) $a$, $d$.

**§9.** **1.** $\frac{2}{\sqrt{5}}e^{2xy}(x + 2y + x^2y + 2xy^2 + y^3 + 2x^3)$. **4.** $\frac{1}{60}\sqrt{10}$. **6.** 7, in the direction of $\langle -3, 2, -6\rangle$.

**§10.** **1.** $3x + 2y + z - 2 = 0$ and $\dfrac{x-2}{3} = \dfrac{y+1}{2} = z+2$. **3.** $x + y - 2z - 5 = 0$ and $x - 2 = y + 1 = \dfrac{z+2}{-2}$. **5.** (a) $x = x_0 + tx_0$, $y = y_0 + ty_0$, $z = z_0 + tz_0$. **9.** $\left(\frac{3}{16}, -\frac{1}{2}, \frac{1}{8}\right)$. **10.** The tangent plane to both is $x - 2y + z = 2$. **11.** $-x + \frac{\pi}{2}y - z = 0$. **14.** $-2y + z + 1 = 0$. **17.** $(\frac{2}{3}, \frac{10}{9}, \frac{26}{27})$. **18.** Smooth. **20.** Piecewise smooth. **21.** Not even piecewise smooth: it is non-smooth on each curve $xy = -\frac{\pi}{2} + 2k\pi$.

**§11.** **1.** A saddle point at $(\frac{1}{2}, -2)$. **3.** Local minima at $(\pm 1, \pm 1)$ and a saddle point at $(0,0)$. **6.** Local maxima at $(\frac{\pm 1}{\sqrt{2}}, \frac{\pm 1}{\sqrt{2}})$, local minima at $(\frac{\pm 1}{\sqrt{2}}, \frac{\mp 1}{\sqrt{2}})$, and a saddle point at $(0,0)$. **8.** A local minimum at $(1, \frac{1}{2})$ and a saddle point at $(0,0)$. **10.** 8, 20. **12.** $-4\sqrt{2}, 4\sqrt{2}$. **14.** $\sqrt{2/3}$. **16.** The cube of side length 5.

**§12.** **1.** $\pm\frac{4}{5}5^{-1/4}$. **4.** The minimum is 7; there is no maximum. **6.** $S/3, S/3, S/3$.

**§13.** **1.** (a) $-\frac{8}{27}$. **3.** 2.5. **4.** $20\frac{2}{3}$. **6.** $8(\ln 3)^{-2}$. **8.** $\frac{3}{4}\ln 3 - \ln 2$. **10.** $\frac{25}{8}\pi$. **12.** $-1$.

**§14.** **1.** $-21\frac{1}{3}$. **3.** $\sin 1 - \cos 1$. **4.** 257.2. **6.** $\frac{1}{2}\ln 2$. **8.** $\frac{1}{2}(1 - \sin 1)$. **10.** 7. **13.** $4\ln 6$. **14.** $\int_0^1 \int_0^{y^2} f(x,y)\,dx\,dy$. **16.** $\int_0^2 \int_0^{\sqrt{2y}} f(x,y)\,dx\,dy + \int_2^4 \int_0^{4-y} f(x,y)\,dx\,dy$. **18.** $\int_0^{\ln 2}\int_{e^x}^2 f(x,y)\,dy\,dx$. **20.** $\frac{1}{4}\left(1 - e^{-1}\right)$. **21.** $\frac{1}{4}\ln 17$. **23.** 4.9. **24.** 6. **26.** $\frac{3}{2}\pi + \frac{56}{15}$. **27.** 243.2. **29.** $\frac{16}{3}$. **30.** $\frac{2}{3}\pi$.

**§15.** **2.** $(e-1)\frac{\pi}{2}$. **3.** $\left(15\sqrt{15} - 7\sqrt{7}\right)\frac{\pi}{6}$. **5.** $\frac{1}{8}\left(1 - 5e^{-4}\right)$. **6.** $8\pi$. **8.** $\left(36 - \frac{20}{3}\sqrt{5}\right)\pi$. **10.** $\frac{8}{3}\pi - \frac{32}{9}$. **11.** $\frac{\pi}{2}$.

**§16.** **2.** $\frac{1}{2}\left(1 - e^{-16}\right)$. **3.** 0.7. **5.** $\frac{64}{15}$. **7.** $\frac{1}{2}(1 - \sin 1)$. **9.** $\frac{16}{3}$.

**§17.** **1.** $\frac{34}{15}\pi$. **3.** $\frac{2\pi}{3}\left(2\sqrt{2} - 1\right) - \pi$. **4.** $2\pi^2$. **6.** $8\pi$. **8.** $\frac{2\pi}{35}$. **11.** $\frac{81}{2}\pi$. **13.** $\frac{14}{3}\pi$.

**§18.** **1.** $\frac{9}{4}\sqrt{6}$. **3.** $8\pi\left(2 - \sqrt{2}\right)$. **5.** $\pi\left(2\sqrt{6} - \frac{8}{3}\right)$. **8.** $\frac{64}{9}$, $(\frac{6}{5}, 0)$. **9.** 32, $\left(\frac{7}{6}, \frac{7}{6}, \frac{7}{6}\right)$. **12.** $I_n = \frac{1}{2}\frac{3}{2}\cdots\frac{2n-1}{2}\frac{\sqrt{\pi}}{2}$.

**§19.** **1.** $R : x^2 + 4y^2 \le 4$. **3.** $R : x^2 - 2xy + 2y^2 \le 2$. **5.** $R$ is a parallelogram with vertices $(0,0)$, $(1,1)$, $(1,-3)$, and $(2,-2)$. **6.** (c) $u$. **7.** $-6$. **9.** $\frac{1}{2}\ln 2 - \frac{5}{16}$. **11.** (a) $D : [1,4]\times[1,6]$; (c) $-4x^2 - 2y^2$; (d) 7.5. **13.** (a) $T$ is not one-to-one on the planes $u = 0$ and $v = 0$; (b) $u^2v$; (c) $p!q!r!s!/(p+q+r+s+3)!$.

**§20.** **2.** $\left\langle ye^{2y}\cos xy, xe^{2y}\cos xy + 2e^{2y}\sin xy\right\rangle$. **4.** $\left\langle \dfrac{2x}{x^2+yz}, \dfrac{z}{x^2+yz}, \dfrac{y}{x^2+yz}\right\rangle$. **5.** Open and simply-connected. **6.** Simply-connected, but not open. **7.** Open and connected, but not simply-connected. **9.** Open, but not connected. **11.** No. **13.** Yes, $f(x,y) = xe^y + c$. **14.** Yes, $f(x,y) = y^2\arctan x + c$. **16.** $\operatorname{curl}\mathbf{F} = \left\langle -4xz^2, x^2y + 4yz^2, -x^2z\right\rangle$, $\operatorname{div}\mathbf{F} = -6xyz$. **18.** $\operatorname{curl}\mathbf{F} = \left\langle 2xze^{2xy}, -2yze^{2xy} - x\sin xz, 2y\cos xy\right\rangle$, $\operatorname{div}\mathbf{F} = e^{2xy} + 2x\cos xy - z\sin xz$. **20.** Yes, $f(x,y,z) = xy^2\cos z + c$. **21.** No. **23.** No, because $\operatorname{div}(xy^2\mathbf{e}_1 + yz^2\mathbf{e}_2 + zx^2\mathbf{e}_3) \neq 0$.

**§21.** **1.** $\frac{1}{840}\left(125\sqrt{5} - 1\right)$. **3.** $\frac{7}{24}$. **5.** $3\pi^3\sqrt{13}$. **7.** 14. **9.** $15\pi$. **10.** $\frac{2}{3} - \cos 1 + \sin 2$.

**§22.** **1.** $-1 + 2e^{-1}$. **3.** $-4$. **5.** $256\cos 8$. **6.** $4\pi$.

**§23.** **2.** $-24\pi$. **4.** 162. **6.** 0. **7.** $\frac{4}{3} - \frac{1}{2}\pi^2$. **9.** $-142\frac{22}{35}$. **10.** $3\pi$. **12.** $2\pi$.

**§24.** **1.** $171\sqrt{14}$. **3.** $16\pi$. **5.** $713/180$. **7.** $-\frac{2}{3}\pi$.

**§25.** **1.** $8\pi$. **3.** $-\frac{13}{12}$. **5.** $-20\pi$. **7.** $-\frac{13}{12}$.

**§26.** **1.** 0. **2.** $\frac{27\sqrt{3}}{5}\pi$. **5.** $\frac{8}{3}\pi + 4\sqrt{3}$. **6.** $216\pi$. **7.** $\frac{8\sqrt{2}}{3}\pi$. **8.** $32\pi^2$.