# Contents

## 1 Introduction to Three Dimensions

1.1 Describing Points in 3-Space ............................................. 3  
1.2 Surfaces from Graphs .................................................. 18  
1.3 Surfaces from Equations ............................................. 34  
1.4 Parametric Curves ..................................................... 48  
  1.4.1 Curves in $\mathbb{R}^2$ ............................................. 49  
  1.4.2 Curves in $\mathbb{R}^3$ ............................................. 52  
1.5 Parametric Surfaces .................................................... 64  
1.6 Describing Regions .................................................... 77

## 2 Introduction to Vectors

2.1 Geometry and Algebra of Vectors ...................................... 85  
2.2 The Dot Product ....................................................... 100  
2.3 The Cross Product .................................................... 110  
2.4 Calculus with Parametric Curves ..................................... 122  
2.5 Vector Fields—a first glance ......................................... 132

## 3 Differentiation

3.1 Functions ................................................................. 139  
3.2 Limits ................................................................. 142  
3.3 Partial Derivatives ..................................................... 149  
3.4 Tangent Planes to Surfaces .......................................... 156  
3.5 The Chain Rule ....................................................... 163  
3.6 Applications of the Gradient ....................................... 176  
3.7 The Hessian Test ..................................................... 183  
3.8 Constrained Optimization .......................................... 190

## 4 Integration

4.1 Double Integrals—The Definition .................................... 199  
4.2 Calculating Double Integrals ....................................... 205  
4.3 Triple Integrals ....................................................... 217  
4.4 Integration in Different Coordinate Systems ....................... 233  
4.5 Applications of Integration ....................................... 249  
4.6 Curve and Surface Integrals ....................................... 258
5 Vector Analysis 269
  5.1 Line Integrals .............................................. 269
  5.2 Conservative Vector Fields .................................. 281
  5.3 Green’s Theorem .............................................. 289
  5.4 Curl and Divergence ......................................... 298
  5.5 Surface Integrals and Stokes’ Theorem ....................... 303
  5.6 The Divergence Theorem ..................................... 321
Chapter 1

Introduction to Three Dimensions

In this chapter we begin our foray into multivariable calculus by getting comfortable with three dimensions. Section one introduces coordinate systems for describing points in three space. We see that Cartesian and polar coordinates in the plane extend naturally to Cartesian and cylindrical coordinates for \( \mathbb{R}^3 \), our notation for three-dimensional Euclidean space. A third coordinate system, spherical coordinates, is also introduced, rounding out those systems most used in this text. Sections 1.2, 1.3 and 1.5 study three different methods of describing surfaces in \( \mathbb{R}^3 \). Section 1.2 focuses on surfaces arising from graphs of functions of two variables, using level curves to aid understanding. Level curves turn out to be the equivalent of a topographic map for the graph of \( z = f(x, y) \). Solution sets of equations in three variables also give rise to surfaces in \( \mathbb{R}^3 \), and these are considered in Section 1.3. Parametric curves in the plane are reviewed in Section 1.4 and generalized to parametric curves and surfaces in \( \mathbb{R}^3 \). Each method of defining surfaces yields a different insight into the geometry of \( \mathbb{R}^3 \), and will be used regularly when discussing differentiation and integration. The chapter ends with a section on describing regions in space using systems of inequalities. This skill will be useful when determining limits of integration for multiple integrals.

1.1 Describing Points in 3-Space

Coordinates in \( \mathbb{R}^2 \):

Before describing coordinate systems in three dimensions we recall what we know about points in \( \mathbb{R}^2 \), the Cartesian plane. The common coordinate systems in \( \mathbb{R}^2 \) will extend naturally to coordinates for \( \mathbb{R}^3 \).

**Cartesian and Polar Coordinates**

In two dimensions there are two familiar methods for describing points. The most common coordinate system is Cartesian coordinates in which a point is
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

described by how far horizontally and vertically it is from the origin. To walk to the point \((x, y)\) from the origin \((0, 0)\) simply walk \(x\) units horizontally, then \(y\) units vertically.

Polar coordinates also describe points in the plane. Rather than using horizontal and vertical distances, polar coordinates tell you how far to walk and in what direction. To walk from \((0, 0)\) to the point with polar coordinates \((r, \theta)\), simply walk \(r\) units at an angle \(\theta\) with the positive \(x\)-axis. Thus in polar coordinates, \(r\) is the distance to the origin and \(\theta\) is the angle with the positive \(x\)-axis.

There is some ambiguity when using polar coordinates to describe points in the plane. Typically one chooses \(r \geq 0\), but we also make the convention that if \(r < 0\), go \(|r|\) units in the opposite direction from \(\theta\). Thus \((3, \frac{2\pi}{3})\) and \((-3, \frac{2\pi}{3})\) represent the same point in polar coordinates, namely, the point with Cartesian coordinates \((-\frac{3}{2}, \frac{3\sqrt{3}}{2})\). You’ll also remember that polar coordinates with the same \(r\) and angles that differ by a multiple of \(2\pi\) represent the same point. In general this ambiguity will not cause confusion.

**Constant Coordinate Curves**

Equations for curves in the plane can be given in either coordinate system, and we recall some particularly simple ones here. In particular, consider the curves obtained by fixing just one of the coordinates. Solution sets to the Cartesian equations \(x = c\) and \(y = d\) are vertical and horizontal lines, respectively (here \(c\) and \(d\) are constants). Moreover, since \(r\) denotes the distance to the origin, the polar equation \(r = c\) denotes a circle centered at \((0, 0)\) of radius \(c\). Finally, the polar equation \(\theta = d\) defines a ray emanating from the origin and making an angle of \(d\) with the positive \(x\)-axis (we assume \(r \geq 0\)). This has been a quick review of the curves we get by fixing one coordinate and letting the other vary.

![Figure 1.1.1: Review of Planar Coordinates](image)

Finally, the triangle pictured in Figure 1.1.1(c) verifies the relationships between Cartesian and polar coordinates. Trigonometry leads to the familiar change of coordinate formulas:
1.1. DESCRIBING POINTS IN 3-SPACE

<table>
<thead>
<tr>
<th>Change of Coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar to Cartesian</td>
</tr>
<tr>
<td>( x = r \cos \theta )</td>
</tr>
<tr>
<td>( y = r \sin \theta )</td>
</tr>
<tr>
<td>Cartesian to Polar</td>
</tr>
<tr>
<td>( r^2 = x^2 + y^2 )</td>
</tr>
<tr>
<td>( \theta = \tan^{-1} \left( \frac{y}{x} \right) )</td>
</tr>
</tbody>
</table>

The conversion \( \theta = \tan^{-1} \left( \frac{y}{x} \right) \) is valid as long as \( x > 0 \), since the range of the arctangent is \(-\pi/2 < \theta < \pi/2\). If \( x < 0 \), you must use \( \theta = \tan^{-1} \left( \frac{y}{x} \right) + \pi \).

Coordinates for \( \mathbb{R}^3 \):

To describe points in three dimensions, it stands to reason that a third coordinate is needed. The most direct way is to add a third coordinate to our two-dimensional coordinate systems just discussed. The result will be Cartesian and cylindrical coordinates for \( \mathbb{R}^3 \). We will describe a third coordinate system for \( \mathbb{R}^3 \), called spherical coordinates, which is useful as well.

**Cartesian Coordinates**

To obtain Cartesian coordinates for \( \mathbb{R}^3 \), start with an \( xy \)-plane and add a third axis through the origin which is perpendicular to both the \( x \)- and \( y \)-axes. Call the third axis the \( z \)-axis. Typically we think of the \( xy \)-plane as lying horizontally in space, and the \( z \)-axis as being the vertical direction. Cartesian coordinates \((x, y, z)\) of the point \( P \) in \( \mathbb{R}^3 \) mean the same as they did in two dimensions, with the \( z \)-coordinate giving the height of \( P \) above or below the \( xy \)-plane.

**Constant Coordinate Surfaces**

In three dimensions, then, the Cartesian equation \( z = c \) represents all points a fixed height \( c \) from the \( xy \)-plane. Thus \( z = c \) is an equation for a plane parallel to the \( xy \)-plane but \( c \) units from it. This is analogous to the two-dimensional situation, where the equation \( y = c \) describes a line parallel to, and \( c \) units from, the \( x \)-axis.

There are three coordinate planes in \( \mathbb{R}^3 \), the \( xy \)-, \( xz \)-, and \( yz \)-planes, which slice space into octants, pictured in Figure 1.1.2a. Can you think of equations for them (Hint: the equation for the \( y \)-axis in \( \mathbb{R}^2 \) is \( x = 0 \))? It should be clear that the three-dimensional Cartesian equation \( x = c \) describes a plane parallel to, and \( c \) units from, the \( yz \)-plane. See Figure 1.1.2b for the planes obtained by fixing a single Cartesian coordinate. Notice that the solution set to the equation \( x = c \) depends on what dimension you’re in. In \( \mathbb{R}^2 \) it is a line while in \( \mathbb{R}^3 \) its a plane. The context of the situation will dictate which interpretation you should give, so keep in mind what dimension you’re in!

Throughout this course it will be important to understand the relationship between English, analytic, and geometric descriptions of mathematical objects. We illustrate this with a few examples.

**Example 1.1.1. Describing a plane in \( \mathbb{R}^3 \)**

The phrase “the horizontal plane two units above the \( xy \)-plane” describes the horizontal plane in Figure 1.1.2 in English. Geometrically this surface is a
plane, and analytically it can be described as (the solution set of) the equation $z = 2$. We now know that equations like $z = c$ describe planes!

**Example 1.1.2. Describing a line in $\mathbb{R}^3$**

The simplest lines to describe in $\mathbb{R}^3$ are the coordinate axes. The $x$-axis can be described as the set of all points whose $y$- and $z$-coordinates are both zero, so the set of all points of the form $(x, 0, 0)$. Similarly, the $y$- and $z$-axes are all points of the form $(0, y, 0)$ and $(0, 0, z)$, respectively. We now consider lines parallel to the coordinate axes.

In $\mathbb{R}^3$, then, fixing one coordinate gives a plane. Fixing two coordinates, however, will give a line. For example, the solution set of the Cartesian system of equations $x = 1$, $y = -1$ is the set of all points $(1, -1, z)$ where $z$ is a variable. Thus it is a line parallel to the $z$-axis, and is the intersection of the planes $x = 1$ and $y = -1$ pictured in Figure 1.1.2(b). It is interesting that a single Cartesian equation yields a surface in $\mathbb{R}^3$ (e.g. $y = -1$ is a plane), while a system of two Cartesian equations yields a curve (e.g. $x = 1$, $y = -1$ describes a line). We call $C(t) = (1, -1, t), -\infty < t < \infty$ parametric equations for the line. ▲

**Example 1.1.3. Describing Regions in Space–Cartesian Coordinates**

We describe the portion of $\mathbb{R}^3$ defined by the system of inequalities

\[
\begin{align*}
0 &\leq x \leq 2 \\
0 &\leq y \leq 3 \\
0 &\leq z \leq 5
\end{align*}
\]
1.1. DESCRIBING POINTS IN 3-SPACE

The restrictions on $x$ indicate the solid is between the planes $x = 0$ and $x = 2$. Similarly, it is between the $xz$-plane and the plane $y = 3$, as well as between the $xy$-plane and $z = 5$. Thus, it is the rectangular box shown in Figure 1.1.3.

![Figure 1.1.3: A Rectangular Box](image)

We will usually think of the $yz$-plane as the plane of the paper, with the $x$-axis pointing out of the paper toward you. This is helpful to keep in mind when viewing static pictures, but software allows more flexibility.

We also mention that the distance formula for $\mathbb{R}^3$ is a natural generalization of the two-dimensional one. Let $(x_1, y_1, z_1)$ and $(x_2, y_2, z_2)$ be points in space, then the distance $d$ between them is given by

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}.$$  

**Example 1.1.4. Equations for Spheres**

This distance formula gives rise to Cartesian equations for spheres in $\mathbb{R}^3$. Indeed, a sphere is all points a fixed distance from a given point. The equation for a sphere radius $r$ and centered at $(a, b, c)$ is then

$$(x - a)^2 + (y - b)^2 + (z - c)^2 = r^2,$$

where we squared the distance formula to simplify the equation. Note the similarity between equations for spheres in three dimensions and those of circles in two.

**Cylindrical Coordinates**

Cylindrical coordinates for $\mathbb{R}^3$ are obtained by adding the $z$-coordinate to polar coordinates for the $xy$-plane in space. To get from the origin to the point $(r, \theta, z)$, first walk in the $xy$-plane $r$ units at an angle of $\theta$ with the $x$-axis. Then jump $z$ units vertically. Note that, while $r$ represented distance to the origin in polar coordinates, it represents distance to the $z$-axis in cylindrical coordinates. The right triangle pictured in Figure 1.1.4(a) indicates that the distance from $(r, \theta, z)$ to the origin in $\mathbb{R}^3$ is given by $\rho = \sqrt{r^2 + z^2}$.

**Constant Coordinate Surfaces**

As in Cartesian coordinates, let’s analyze what we get by fixing one cylindrical coordinate. As in Cartesian coordinates, fixing $z$ yields a horizontal plane.
Letting \( r \) be constant describes the set of all points a fixed distance from the \( z \)-axis. To get a feel for what this is, recall that in polar coordinates fixing \( r \) gave a circle. In three dimensions, that circle can be translated up and down the \( z \)-axis without changing the distance to the \( z \)-axis. In other words, without changing \( r \). Thus fixing \( r \) yields a cylinder in \( \mathbb{R}^3 \) whose axis of symmetry is the \( z \)-axis. Hence the name “Cylindrical” coordinates. Finally, if you fix \( \theta \) in polar coordinates, you get a ray emanating from the origin. As in the case of the cylinder, translate this up and down the \( z \)-axis to find what you get in space. The result is a half plane that makes an angle of \( \theta \) with the positive \( x \)-axis, and whose boundary is the \( z \)-axis.

(a) Geometry of Cylindrical Coords  
(b) Constant Coordinate Surfaces

Figure 1.1.4: Cylindrical Coordinates

**Math App 1.1.1. Cylindrical Constant Coordinate Surfaces**

Throughout this text Math Apps will be embedded hyperlinks to Maple Math Apps that enhance geometric understanding. Simply click on the link to have an interactive Maple window open, and explore the concepts therein. Try it now with the link pictured below.

Since both Cartesian and Cylindrical coordinates for \( \mathbb{R}^3 \) extend coordinate systems for \( \mathbb{R}^2 \), converting between them is the same as between Cartesian and
1.1. DESCRIPTING POINTS IN 3-SPACE

polar. There is the obvious addition that the $z$-coordinates are the same. Thus the Cartesian coordinates for the cylindrical coordinates $(r, \theta, z)$ are

$$(x, y, z) = (r \cos \theta, r \sin \theta, z).$$  

(1.1.1)

Example 1.1.5. Constant Coordinate Surfaces

We again emphasize the English, geometric, and analytic descriptions of a surface. The infinite cylinder with radius one and $z$-axis as core is pictured in Figure 1.1.4(b). The given cylinder is the solution set of the cylindrical equation $r = 1$. ▲

Cylinders with $z$-axis as core are constant coordinate surfaces when using cylindrical coordinates, as are half-planes with $z$-axis as boundary. The half of the $xz$-plane that contains the positive $y$-axis is given by the equation $\theta = \pi/2$.

Example 1.1.6. Curves of intersection in cylindrical coordinates

In Cartesian coordinates we found the solution set to the system $x = 1, y = -1$ was the line of intersection of the corresponding planes. Interesting curves also arise from intersecting constant coordinate surfaces in cylindrical coordinates.

Figure 1.1.4(b) indicates that the system of cylindrical equations $r = 1, \theta = \pi/2$ describes the intersection of the cylinder $r = 1$ and the half-plane $\theta = \pi/2$. The result is a vertical line. ▲

Remark: While discussing cylindrical coordinates, we should mention that there is nothing special about using the $z$-axis as the third coordinate. We could have just as easily used polar coordinates in the $yz$-plane, and the $x$-axis as the third coordinate. Then $r$ would be the distance to the $x$-axis, $\theta$ the angle with the positive $y$-axis and $x$ would just be $x$. See the figure below. Unless otherwise stated, however, cylindrical coordinates will mean $(r, \theta, z)$.

We finish our initial discussion on cylindrical coordinates by looking at the solution set to a system of inequalities. The idea of describing regions in space using a system of inequalities will be useful when setting up limits of integration in triple integrals.

Example 1.1.7. Describing Regions in Space—Cylindrical Coordinates

We describe the portion of $\mathbb{R}^3$ defined by the system of inequalities

$$0 \leq r \leq 2$$

$$0 \leq \theta \leq \frac{\pi}{2}$$

$$0 \leq z \leq 5$$

The restrictions on $r$, which is the distance to the $z$-axis, describe an infinite solid cylinder with core along the $z$-axis and radius two. The restrictions on $z$
cut it down to a solid cylinder radius 2 and height 5, with base in the $xy$-plane. Finally, the restriction on $\theta$ reduces it to that portion which is in the first octant (See Figure 1.1.5).

**Spherical Coordinates**

In cylindrical coordinates, two coordinates describe distances and one describes a direction. We now introduce spherical coordinates, in which two describe directions and only one is a distance. Spherical coordinates are denoted $(\rho, \theta, \phi)$, where $\rho$ is the distance to the origin, $\theta$ is our old friend from polar and cylindrical coordinates, and $\phi$ is the angle with the positive $z$-axis. See Figure 1.1.6.

**Constant Coordinate Surfaces**

Fixing $\theta$ just gives us a half-plane with boundary on the $z$-axis and making an angle of $\theta$ with the positive $x$-axis, as before. Fixing $\rho$ focuses on all points a fixed distance from the origin, take a guess at what shape that might be. The set of all points in $\mathbb{R}^3$ satisfying $\phi = c$ is the set of all points making a fixed angle with the positive $z$-axis. This is actually a cone with vertex at the origin and making an angle of $c$ with the positive $z$-axis.

**Math App 1.1.2. Spherical Constant Coordinate Surfaces**

Click the following hyperlink to view and manipulate a Math App illustrating constant coordinate surfaces $\phi = c$. 
Now consider curves of intersection of constant coordinate surfaces. Fixing both \( \theta \) and \( \phi \) results in a ray through the origin. Indeed, fixing \( \theta \) results in a half-plane while fixing \( \phi \) yields a cone. Fixing both is equivalent to taking the intersection of the half-plane and cone (convince yourself I’m not lying), which is a ray in the half-plane that makes the given angle with the \( z \)-axis.

**Example 1.1.8. Spherical equation from geometric description**

Find a spherical equation for the cone with vertex at the origin and that makes an angle of \( \pi/3 \) with the positive \( z \)-axis. Since the angle with the positive \( z \)-axis is the coordinate \( \phi \) in spherical coordinates, the spherical equation is \( \phi = \pi/3 \).

**Example 1.1.9. Curve of intersection between constant coordinate surfaces**

Describe, as carefully as possible, the curve of intersection of the surfaces \( \phi = 3\pi/4 \) and \( \rho = 1 \). The surfaces are pictured in Figure 1.1.6(b), and the intersection of the cone and sphere will be a circle. Since the cone is pointing straight down, it will be a horizontal circle. Further, using the trigonometry of the triangle corresponding to that Figure 1.1.6(a) we can determine the radius and height below the \( xy \)-plane. In the triangle we have hypoteneuse 1, since \( \rho = 1 \), and an angle of \( 3\pi/4 \) with the positive \( z \)-axis. Trigonometry implies that \( r = \sqrt{2}/2 \) and \( z = -\sqrt{2}/2 \).

In summary, then, the intersection of \( \phi = 3\pi/4 \) and \( \rho = 1 \) is a horizontal circle at height \( z = -\sqrt{2}/2 \) with radius \( r = \sqrt{2}/2 \).

**Example 1.1.10. Describing Regions in Space–Spherical Coordinates**
We describe the portion of $\mathbb{R}^3$ defined by the system of inequalities

\[
2 \leq \rho \leq 3 \\
0 \leq \phi \leq \pi/2 \\
0 \leq \theta \leq 3\pi/2
\]

The restrictions on $\rho$ indicate that the solid is between spheres of radius 2 and 3 centered at the origin. The region of $\mathbb{R}^3$ described by $0 \leq \phi \leq \pi/2$ is the top half of space, while the restriction $0 \leq \theta \leq 3\pi/2$ lets you go three-quarters of the way around the $z$-axis. Thus it is the solid pictured in Figure 1.1.7.

Coordinate Conversion

Converting between spherical and other coordinates is easily achieved using trigonometry and the right triangle illustrated in Figure 1.1.6(a). Notice that $r = \rho \sin \phi$ and $z = \rho \cos \phi$, giving the conversion from spherical to cylindrical. A simple substitution then gives $(x, y, z) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$. We summarize converting between the different coordinate systems in the following table.

<table>
<thead>
<tr>
<th>Converting Between Coordinate Systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cartesian</td>
</tr>
<tr>
<td>$x$</td>
</tr>
<tr>
<td>$y$</td>
</tr>
<tr>
<td>$z$</td>
</tr>
</tbody>
</table>

Other useful conversions

\[
r^2 = x^2 + y^2, \quad \rho^2 = x^2 + y^2 + z^2, \quad r = \rho \sin \phi
\]

It will frequently be helpful to translate between coordinate systems, and these conversions facilitate that translation.

Example 1.1.11. Verifying a conversion analytically
By direct substitution, and simplification, we verify \( x^2 + y^2 + z^2 = \rho^2 \).

\[
\begin{align*}
  x^2 + y^2 + z^2 &= (\rho \sin \phi \cos \theta)^2 + (\rho \sin \phi \sin \theta)^2 + (\rho \cos \phi)^2 \\
  &= \rho^2 \sin^2 \phi \cos^2 \theta + \rho^2 \sin^2 \phi \sin^2 \theta + \rho^2 \cos^2 \phi \\
  &= \rho^2 \sin^2 \phi \left( \cos^2 \theta + \sin^2 \theta \right) + \rho^2 \cos^2 \phi = \rho^2 \sin^2 \phi + \rho^2 \cos^2 \phi \\
  &= \rho^2. \quad \blacktriangle
\end{align*}
\]

**Example 1.1.12. Equations in different coordinate systems**

The conversions in the above table allow one to translate an equation from one coordinate system to another. For example, the Cartesian equation for a sphere centered at the origin radius 2 is \( x^2 + y^2 + z^2 = 4 \). Using the appropriate above conversion, we replace \( x^2 + y^2 + z^2 \) with \( \rho^2 \) to get the spherical equation \( \rho^2 = 4 \), which simplifies to \( \rho = 2 \).

Similarly, we can derive cylindrical and spherical equations for the Cartesian equation \( x = 5 \). Replacing \( x \) with \( r \cos \theta \) we obtain the cylindrical equation \( r \cos \theta = 5 \) for the plane \( x = 5 \). Using the appropriate spherical substitution we find the spherical equation is \( \rho \sin \phi \cos \phi = 5 \).

**Example 1.1.13. Equations for surface described geometrically**

Find an equation in each coordinate system for the plane parallel to the \( xz \)-plane, and 7 units to the right of it.

Since such planes have Cartesian equations \( y = c \) for some constant \( c \), we have \( y = 7 \) as the Cartesian equation of the plane. Substitutions yield:

\[
\begin{align*}
  \text{Cartesian} & \quad y = 7 \\
  \text{Cylindrical} & \quad r \sin \theta = 7 \\
  \text{Spherical} & \quad \rho \sin \phi \sin \theta = 7 \quad \blacktriangle
\end{align*}
\]

In this section we’ve introduced coordinate systems in three dimensions. They will all be used extensively in what follows. We summarize the important topics now.

**Things to know/Skills to have**

- The interpretation of each coordinate in Cartesian, cylindrical and spherical coordinates (e.g. \( z \) is the height above the \( xy \)-plane, while \( r \) is the distance to the \( z \)-axis).
- Converting between different coordinate systems, and why they work using trig.
- Be able to sketch and describe in English constant coordinate surfaces in each coordinate system.
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

- Be able to describe in English and sketch the intersection of two constant coordinate surfaces.
- Be able to give equations in each coordinate system for constant coordinate surfaces.
- Be able to use a system of inequalities to describe solid regions in $\mathbb{R}^3$.

Exercises

1. Sketch the following constant coordinate surfaces, include the coordinate planes in your sketch.
   (a) $z = -2$.
   (b) $x = 4$.
   (c) $y = -5$.

2. Sketch the following constant coordinate surfaces, include the coordinate axes in your sketch.
   (a) $z = -2$.
   (b) $r = 4$.
   (c) $\theta = -3\pi/4$.

3. Sketch the following constant coordinate surfaces, include the coordinate axes in your sketch.
   (a) $\rho = 2$.
   (b) $\phi = \pi/3$.
   (c) $\theta = -3\pi/4$.

4. Sketch the constant coordinate surfaces $x = -1$, $x = 0$, and $x = 3$ on the same set of axes.

5. Sketch the constant coordinate surfaces $r = 1$, $r = 3$, and $r = 5$ on the same set of axes.

6. Sketch the constant coordinate surfaces $\theta = 0$, $\theta = \pi/4$, and $\theta = 3\pi/4$ on the same set of axes.

7. Sketch the constant coordinate surfaces $\phi = \pi/6$, $\phi = \pi/2$, and $\phi = 3\pi/4$ on the same set of axes.

8. Sketch the constant coordinate surfaces $\rho = 1$, $\rho = 2$, and $\rho = 5$ on the same set of axes.

9. Give a one-sentence English description of the following constant coordinate surfaces:
1.1. DESCRIBING POINTS IN 3-SPACE

(a) \( z = 5. \)
(b) \( r = 4. \)
(c) \( \rho = 2. \)
(d) \( \phi = \pi/3. \)
(e) \( \theta = -3\pi/4. \)

10. Find Cartesian equations for the following surfaces.

(a) \( \rho = 2. \)
(b) \( \phi = \pi/2. \)
(c) \( r = 3. \)

11. Find cylindrical equations for the following surfaces.

(a) \( x^2 + y^2 = 9. \)
(b) \( z = 3. \)
(c) \( \rho \sin \phi = 6. \)

12. Find spherical equations for the following surfaces.

(a) \( r^2 + z^2 = 4. \)
(b) \( z = 3. \)
(c) \( x^2 + y^2 + z^2 = 9. \)

13. Find a Cartesian equation for the set \( S \) of all points 6 units above the \( xy \)-plane. Now find cylindrical and spherical equations for \( S \).

14. Find a Cartesian equation for the set \( S \) of all points 3 units to the left of the \( yz \)-plane. Now find cylindrical and spherical equations for \( S \).

15. Find a cylindrical equation for the set of all points in \( \mathbb{R}^3 \) which are 4 units from the \( z \)-axis.

16. Find a spherical equation for the set \( S \) of all points 5 units from the origin. Now find Cartesian and cylindrical equations for \( S \).

17. Find a Cartesian equation for the sphere centered at (0, 0, 1) with radius 1. Now find a polar equation for it.

18. Describe, as carefully as you can, the intersection of the constant coordinate surfaces given below. Include what geometric shape it is (e.g. a line, ray, circle, etc.), and how it sits in \( \mathbb{R}^3 \) (e.g. horizontally, parallel to the \( y \)-axis, etc.).

(a) \( y = 3, z = -2. \)
(b) \( z = 5, r = 2. \)
(c) $z = -2, \theta = \frac{3\pi}{4}$.
(d) $r = 5, \theta = -\frac{\pi}{3}$.
(e) $\rho = 3, \phi = \frac{\pi}{4}$.
(f) $\rho = 3, \phi = \frac{\pi}{2}$.
(g) $\rho = 5, \theta = \frac{4\pi}{3}$.
(h) $\phi = \frac{\pi}{6}, \theta = \frac{\pi}{6}$.

19. Use the Pythagorean Theorem to prove the distance formula in $\mathbb{R}^3$.
20. Justify in English the conversion $z = \rho \cos \phi$. (Hint: use Figure 1.1.6)
21. Justify in English the conversion $r = \rho \sin \phi$. (Hint: use Figure 1.1.6)
22. Justify, analytically and in English, the conversion $r^2 = x^2 + y^2$.
23. Justify, in English, the conversion $\rho^2 = x^2 + y^2 + z^2$.
24. Sketch the solids determined by the system of inequalities:

\[
0 \leq x \leq 3 \\
0 \leq y \leq 5 \\
0 \leq z \leq 1
\]

25. Sketch the solids determined by the system of inequalities:

\[
0 \leq x \leq 4 \\
-2 \leq y \leq 0 \\
0 \leq z \leq 3
\]

26. Sketch the solids determined by the system of inequalities:

\[
0 \leq r \leq 4 \\
0 \leq \theta \leq \pi \\
0 \leq z \leq 3
\]

27. Sketch the solids determined by the system of inequalities:

\[
2 \leq r \leq 3 \\
0 \leq \theta \leq \pi/2 \\
0 \leq z \leq 4
\]
28. Sketch the solids determined by the system of inequalities:

\[
0 \leq \rho \leq 2 \\
0 \leq \phi \leq \pi \\
0 \leq \theta \leq \pi/2
\]

29. Sketch the solids determined by the system of inequalities:

\[
1 \leq \rho \leq 2 \\
0 \leq \phi \leq \pi/2 \\
0 \leq \theta \leq \pi
\]

30. A cylindrical can has radius 5 and height 2. A coordinate system is introduced so that the center of mass of the can is at the origin, and its axis is the \( z \)-axis. What system of inequalities on cylindrical coordinates describes the region of space occupied by the can?

31. A spherical shell is centered at the origin. Its inner radius is 2 and it is half a unit thick. What system of inequalities in spherical coordinates describes the region of space occupied by the shell?

32. Define a different set of cylindrical coordinates, where \( r \) is the distance to the \( x \)-axis and \( \theta \) is the angle made with the positive \( y \)-axis. What are the change-of-coordinate functions from this system to Cartesian coordinates?

33. Let \( T \) be rotation of space counterclockwise around the \( z \)-axis through an angle of \( \frac{\pi}{2} \), and let \((\rho, \theta, \phi) = (2, -\frac{\pi}{4}, \frac{\pi}{4})\) be the spherical coordinates of the point \( P \). Find the spherical coordinates of the rotated point \( T(P) \).
1.2 Surfaces from Graphs

In single-variable calculus considerable effort is spent on studying curves defined as graphs of functions \( y = f(x) \). The derivative \( f'(x) \) is the instantaneous rate of change of \( f \), and can be used to determine when the graph of \( f \) is increasing or decreasing, the concavity of \( f \), and extreme values of \( f \). The integral of \( f \) can represent area or distance traveled, and be used to find physical quantities like arclength and centers of mass. Indeed, much of single-variable calculus is concerned with analyzing properties of functions \( f(x) \) and their graphs. In multivariable calculus we will be concerned with functions of several variables, and we begin our study in this section with analyzing graphs of functions of two variables.

Recall that the graph of a function, say \( f(x) = \sinh x = \frac{e^x - e^{-x}}{2} \), is the set of all points of the form \((x, f(x))\) as pictured below.

\[
\begin{align*}
\text{Figure 1.2.1: The hyperbolic sine function}
\end{align*}
\]

In this context the domain of the function is the horizontal axis, which is a subset of the plane containing the graph. The range consists of \( y \)-values, which are on the vertical axis. Combining the domain and range in the same space is so familiar, we rarely think about it. We take our cue from the two-dimensional case, and make the following definition:

**Definition 1.2.1.** The graph of a function \( f(x, y) \) is the set of all points in \( \mathbb{R}^3 \) of the form \((x, y, f(x, y))\).

To sketch the graph of \( z = f(x, y) \), consider the domain to be the \( xy \)-plane in \( \mathbb{R}^3 \) and the function value \( f(x, y) \) determines the height above it. For example, if \( f(x, y) = x^2 - y^2 \), the point \((2, 1, f(2, 1)) = (2, 1, 3)\) is on the graph of \( f \). We illustrate the graphs of two functions below, then describe the level curve technique for understanding the graph of a function of two variables.
Example 1.2.1. The graphs of two functions

We give, without justification, the graphs of two functions. These surfaces represent all points in \( \mathbb{R}^3 \) of the form \( (x, y, f(x, y)) \) for the respective functions.

\[
\begin{align*}
(a) \quad & f(x, y) = x^2 - y^2 \\
(b) \quad & f(x, y) = x^2 + y^2
\end{align*}
\]

Figure 1.2.2: The graphs of two functions

Our current task is to develop a method for determining the pictures of Figure 1.2.2 from the formulas. This method involves the use of level curves, which we will want to understand analytically, geometrically, and conversationally. We motivate level curves by considering topographic maps.

Figure 1.2.3: Longs Peak area

A topographic map is one that also illustrates the topography of a region using what are called contour lines. Contour lines are constant altitude curves, or curves that connect points of the same altitude. Figure 1.2.3 is a topographic
map of the area around Longs Peak in Rocky Mountain National Park, Colorado. The contour lines, together with shading and colors, help give us an idea of what the terrain is like. For example, consecutive darker curves represent an altitude gain of 250 feet. The closer they are together, the steeper the terrain in that area. The label on a curve tells you its altitude, and we can use them to determine which peak is higher, Longs Peak in the lower center of the region, or Mount Meeker, South East of Longs. One could also use contour lines to determine the direction a hiker at the middle of the map should walk to get downhill the fastest. A lot of information is encoded in a topographic map.

We wish to use similar techniques to understand graphs of functions \( z = f(x, y) \). The “mathematical mountain” in Figure 1.2.4(a), for example, is represented by the topographic map in Figure 1.2.4(b). Notice that the contour lines are labeled with their corresponding altitudes. If we had a big can of paint and a lot of time we could paint the curves on the mountain that correspond to the contour lines on the map, as in Figure 1.2.4(c). The contour lines of the map are not actually on the mountain, they live in two dimensions. However, looking only at the topographic map, the contour lines do give us an idea of what the mountain looks like.

![Figure 1.2.4: Topographic maps motivate level curves](image)

Our goal is to learn how to construct topographic maps from formulas to aide our understanding of graphs of functions of two variables. To do so, note that the constant altitude curves on the mountain in Figure 1.2.4(c) can be thought of as the curve of intersection of the mountain and a horizontal plane. For example, the highest curve is 630 feet above sea level, and can be thought of as the intersection of the mountain with the plane \( z = 630 \). Thus constant altitude curves on the mountain \( z = f(x, y) \) at height \( c \) can be thought of as the solution set in \( \mathbb{R}^3 \) of the system of equations

\[
\begin{align*}
    z &= f(x, y) \\
    z &= c.
\end{align*}
\]

The contour lines of Figure 1.2.4(b) are obtained by dropping the curves on the mountain down into the \( xy \)-plane.
This “dropping” is more precisely called projecting into the $xy$-plane, and is accomplished analytically by eliminating the $z$-coordinate. To eliminate the $z$-coordinate from the system of equations 1.2.1, merely turn them into a single equation by substituting $c$ for $z$, obtaining $c = f(x,y)$. Thus we have accomplished our goal. We have determined how to find equations for contour lines from a formula for the function $f$. Let’s look at a concrete example before going further.

**Example 1.2.2. A contour line, or level curve**

Let $f(x,y) = x^2 + y^2$, and find a contour line at height 4 for the graph of

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

As described above, the curve on the “mountain” $z = x^2 + y^2$ at height 4 is the solution set to the system of equations

\[
\begin{align*}
  z &= x^2 + y^2 \\
  z &= 4
\end{align*}
\]

To find the contour line, merely substitute 4 for $z$ in the first equation yielding

\[x^2 + y^2 = 4.
\]

The curve in the $xy$-plane defined by this equation is the desired contour line, and we will call it the **level curve** of the function $z = x^2 + y^2$ at level 4. It is obtained by projecting into the $xy$-plane the curve of intersection of the surface $z = x^2 + y^2$ and the plane $z = 4$. Figure 1.2.5 illustrates what’s going on geometrically.

![Figure 1.2.5: Geometric understanding of the $\mathbb{R}^3$ problem](image)

After this concrete example, we are ready to make a general definition. We remark that, although the term “contour line” is used in the context of topographic maps, we will transition to using the term **level curve**.
**Definition 1.2.2.** The level curve of \( f(x, y) \) at level \( c \) is the curve in the \( xy \)-plane given by the equation \( c = f(x, y) \).

We found that the level curve of \( f(x, y) = x^2 + y^2 \) at level \( c = 4 \) is the circle \( x^2 + y^2 = 4 \) (see Figure 1.2.5). The equation for the level curve is found by setting \( f(x, y) \) equal to the given level. Equivalently, we substitute the level \( c \) for \( z \) in the equation \( z = f(x, y) \). Let’s investigate this further with some examples.

**Example 1.2.3.** Level curves of \( f(x, y) = x^2 + y^2 \)

Using the above strategy we find the level curve at level \( c = 0 \) is the solution set of
\[
x^2 + y^2 = 0,
\]
which is a single point—the origin. Thus a level curve need not be a curve at all, but can be a single point. In fact, the situation can be more extreme. Level curves can be intersecting lines, multiple curves, or even fail to exist. For example, the level curve of \( f(x, y) = x^2 + y^2 \) at level \( c = -1 \) does not exist since it is the solution set of the equation
\[
x^2 + y^2 = -1.
\]

If one is considering several level curves simultaneously, it is sometimes convenient to summarize the equations in tabular form. For example, we see the following equations corresponding to different levels:

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c = -1 )</td>
<td>( x^2 + y^2 = -1 )</td>
</tr>
<tr>
<td>( c = 0 )</td>
<td>( x^2 + y^2 = 0 )</td>
</tr>
<tr>
<td>( c = 0.5 )</td>
<td>( x^2 + y^2 = 0.5 )</td>
</tr>
<tr>
<td>( c = 1 )</td>
<td>( x^2 + y^2 = 1 )</td>
</tr>
<tr>
<td>( c = 2 )</td>
<td>( x^2 + y^2 = 2 )</td>
</tr>
<tr>
<td>( c = 4 )</td>
<td>( x^2 + y^2 = 4 )</td>
</tr>
</tbody>
</table>

Analytically, then, it is easy to find equations for level curves—just let the function equal the level. To develop a greater geometric understanding of the graph of \( z = f(x, y) \), we plot the level curves in the \( xy \)-plane. When plotting level curves in the plane, it is customary to label each curve with its corresponding level, as in a topographic map (see Figure 1.2.6). It is also common to consider several level curves at the same time. Doing this allows us to analyze the surface \( z = f(x, y) \). For example, the level curves in Figure 1.2.6 are concentric circles centered at the point at level zero. Take a moment to compare the surface in Figure 1.2.5(b) with its corresponding level curves in Figure 1.2.6▲

**Math App 1.2.1.** Visualizing level curves

Click the hyperlink below to visualize the process of slicing the surface \( z = f(x, y) \) and graphing the level curve side by side. Use the slider to change the level, and see how the level curves change.
1.2. SURFACES FROM GRAPHS

Example 1.2.4. Level curves of \( f(x, y) = x^2 - y^2 \)

Let's now analyze the surface in Figure 1.2.2. The level curve of \( f(x, y) = x^2 - y^2 \) at level \( c = 1 \) is the hyperbola \( x^2 - y^2 = 1 \). It has asymptotes \( y = \pm x \), vertices \((\pm 1, 0)\), and opens sideways. The level curves of \( f(x, y) = x^2 - y^2 \) at levels \( c = -1, 0, 1, 2 \) are the curves in the \( xy \)-plane given by the equation \( f(x, y) = c \). They are

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c = -1 )</td>
<td>( x^2 - y^2 = -1 )</td>
</tr>
<tr>
<td>( c = 0 )</td>
<td>( x^2 - y^2 = 0 )</td>
</tr>
<tr>
<td>( c = 1 )</td>
<td>( x^2 - y^2 = 1 )</td>
</tr>
<tr>
<td>( c = 2 )</td>
<td>( x^2 - y^2 = 2 )</td>
</tr>
</tbody>
</table>

Most of these curves are hyperbolas in the plane, with one exception. At level \( c = 0 \), then, the level curve is given by \( x^2 - y^2 = 0 \). The left hand side factors to \((x + y)(x - y) = 0\), which holds when either \( y = x \) or \( y = -x \). Thus the level curve at level zero is really a pair of intersecting lines. Plotting the level curves in the \( xy \)-plane yields Figure 1.2.7.

Math App 1.2.2. Level curves of quadratic functions

The previous two examples have both involved quadratic functions. Click the hyperlink below to explore level curves of the function \( f(x, y) = Ax^2 + Bxy + Cy^2 \)
for different values of $A$, $B$, $C$. It turns out that the discriminant $B^2 - 4AC$ is a magic number! See if you can tell what the level curves look like when $B^2 - 4AC$ is positive, negative and zero.

You now have two examples of level curves under your belt, and it’s time to see how to use them to visualize the surface $z = f(x, y)$. Recall that Figures 1.2.6 and 1.2.7 are topographic maps of their corresponding surfaces. To use these maps to visualize your surface, you just need to add the third dimension. We now illustrate how level curves can tell us when a surface is above or below the $xy$-plane.

**Example 1.2.5. Interpreting level curves of $f(x, y) = x^2 - y^2$.**

In Figure 1.2.7 the level curves at level 0 cut the plane into four pieces: the left, right, top and bottom. On each piece the function $f(x, y) = x^2 - y^2$ is either always positive or always negative, telling us if the graph of $f$ is above or below that portion of its domain. The levels are positive on the left and right pieces, indicating the graph is above the $xy$-plane over those regions. Similarly, the graph of $f$ is below the $xy$-plane in the top and bottom regions. ▲

The surface of Figure 1.2.2(a) is called a saddle (it kinda looks like one, doesn’t it: the sides go down but the front and back go up), and the origin is a saddle point of the surface. Now if a circus wanted to design a saddle for a monkey who rode horses in the show, they’d have to consider its tail. A monkey
1.2. SURFACES FROM GRAPHS

A saddle would have to go down in the back so its tail could relax and the monkey would be comfortable. The surface given by \( f(x, y) = x(y - x)(x + y) \) is a monkey saddle. What are the level curves at level \( c = 0 \) for the Monkey Saddle? These curves divide the \( xy \)-plane into regions on which the graph of \( f \) is always above or always below the \( xy \)-plane. Over which regions is it above? below?

![Figure 1.2.8: A Monkey Saddle](image)

We now include an example where a little algebra is helpful before sketching level curves.

**Example 1.2.6. Level Curves of \( f(x, y) = x^2 - 2x + y^2 + 4y + 1 \)**

Sketch the level curves of \( f(x, y) = x^2 - 2x + y^2 + 4y + 1 \) at levels \( c = -1, 0, 1 \).

First, we complete the square on \( x \) and \( y \) to see that \( f(x, y) = (x - 1)^2 + (y + 2)^2 - 4 \). Setting the function equal to each level gives the equation for the level curve in the \( xy \)-plane. We have

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Curve</th>
<th>Simplified Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( c = -1 )</td>
<td>(-1 = (x - 1)^2 + (y + 2)^2 - 4)</td>
<td>(3 = (x - 1)^2 + (y + 2)^2)</td>
</tr>
<tr>
<td>( c = 0 )</td>
<td>(0 = (x - 1)^2 + (y + 2)^2 - 4)</td>
<td>(4 = (x - 1)^2 + (y + 2)^2)</td>
</tr>
<tr>
<td>( c = 1 )</td>
<td>(1 = (x - 1)^2 + (y + 2)^2 - 4)</td>
<td>(5 = (x - 1)^2 + (y + 2)^2)</td>
</tr>
</tbody>
</table>

Thus the level curves are all circles centered at the point \((1, -2)\) with radii \(\sqrt{3}, 2, \sqrt{5}\), respectively. See Figure 1.2.9(a) for a sketch of the level curves. It turns out that this surface is again a paraboloid, seen in Figure 1.2.9(b), but be forewarned that not every surface with circular level curves is a paraboloid! ▲

**Level curves and critical points**

One can look at a topographic map and tell where the mountain peaks are. Similarly, level curves can tell us where relative maxima and minima of our function are, as well as saddle points. More precisely, level curves near extreme values look like concentric ellipses (for nice functions \( f(x, y) \)), while near a saddle they look more like hyperbolas. We illustrate with two examples.

**Example 1.2.7. Extreme values from level curves**
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

(a) Level curves (b) The surface \( z = x^2 - 2x + y^2 + 4y + 1 \)

Figure 1.2.9: A shifted paraboloid

Some level curves of the function \( f(x, y) = e^{-x^2}(y^3 - y) \) are given in Figure 1.2.10 and are labeled with their corresponding heights. The points \( P \) and \( Q \) in this example are degenerate level curves. Since the level curves are elliptical in shape around the points \( P \) and \( Q \), those points are relative extreme points for the surface. Since the levels are decreasing toward \( P \), we know \( P \) is a relative minimum of the surface. Similarly, the levels are increasing toward \( Q \) so it is a relative maximum.

Figure 1.2.10: Some level curves of \( f(x, y) = e^{-x^2}(y^3 - y) \)

In the previous example we used the fact that if a point is a level curve (like \( P \) or \( Q \) above), and the curves around them are elliptical in shape, the points correspond to relative extrema of \( f(x, y) \) (some additional assumptions on the function \( f \) are necessary, but we ignore the details for now). We’ve also seen that near a saddle point the level curves look like hyperbolas (see Figure 1.2.7). We use these facts to interpret the following level curve diagram.

Example 1.2.8. Interpreting level curves, again
The level curves in Figure 1.2.11 are for the function \( f(x, y) = x^3 + 6xy^2 - 6x \). Notice that they are not connected. For example, in Figure 1.2.11(a) the level curves at level 5 appear in the upper and lower right, then close to the point \( P \) on the left. There are actually only two components of the level curve \( 5 = x^3 + 6xy^2 - 6x \), as in Figure 1.2.11(b), the display just wasn’t wide enough to show it. In any case, this illustrates that level curves can have several pieces.

![Figure 1.2.11: Some level curves of \( f(x, y) = x^3 + 6xy^2 - 6x \)](image)

Reasoning as in the previous example and discussion, we wish to classify the points \( P, Q, R, S \) as relative maxima, minima, or saddles. Since levels are increasing toward \( P \), \( f(x, y) \) attains a relative maximum there. Similarly, \( f(x, y) \) has a relative minimum at \( Q \). The level curve at level 0 intersects itself at \( R \) and \( S \), and look like hyperbolas nearby, so \( R \) and \( S \) are saddles of \( f \).

### Level Curves and Composition

From experience we know that composing functions is a powerful way to create new functions from old ones. With a little care we can talk about composing functions of several variables. For example, the function \( f(x, y) = x^2 - y^2 \) takes two variables as input and gives a real number as output. We say that \( f \) maps \( \mathbb{R}^2 \) to \( \mathbb{R} \), and denote it by

\[
f: \mathbb{R}^2 \to \mathbb{R}.
\]

Since \( x^2 - y^2 \) is a real number, the composition \( g \circ f(x, y) = g(f(x, y)) = g(x^2 - y^2) \) only makes sense if the domain of \( g \) is the real numbers. More generally, the composition \( g \circ f \) is only defined when the domain of \( g \) contains the range of \( f \). Let’s look at some examples

#### Example 1.2.9. Composition of functions

Let \( f(x, y) = x^2 - y^2 \), \( g(u) = e^u \), and \( h(v) = v^2 + 1 \). Form \( g \circ f(x, y) \) and \( h \circ f(x, y) \).

By direct substitution we see that

\[
g \circ f(x, y) = g(x^2 - y^2) = e^{x^2 - y^2}
\]

\[
h \circ f(x, y) = h(x^2 - y^2) = (x^2 - y^2)^2 + 1
\]
We note that $f \circ g(u) = f(e^u)$ makes no sense since $e^u$ is a real number and $f$ needs ordered pairs as input. Similarly, $f \circ h$ is not defined. ▲

Given a function $f(x, y)$ of two variables, and a single-variable function $g(u)$, we wish to see how composition affects level curves. More precisely, we’d like to relate the level curves of $g \circ f(x, y)$ to those of $f(x, y)$. We do so in the following lemma.

**Lemma 1.2.1.** With $f$ and $g$ as above, let $C$ be the level curve of $f(x, y)$ at level $c$. Then $C$ is (part of) a level curve of $g \circ f(x, y)$ at $g(c)$.

**Proof.** Let $(a, b)$ be any point on $C$. Since $C$ is the level curve of $f$ at level $c$, we have $f(a, b) = c$. To prove the claim we must show that $g \circ f(a, b) = g(c)$. Indeed, we have

$$g \circ f(a, b) = g(f(a, b)) = g(c).$$

Intuitively this says that level curves of $f(x, y)$ and $g \circ f(x, y)$ are the same sets, only with different levels. The situation is more complicated than that, depending on $g$, but the intuition is worth remembering. We will see that even though $f$ and $g \circ f$ share the same level curves, the graphs of the functions can be quite different.

**Example 1.2.10.** Exponentiating a saddle

The function $h(x, y) = e^{x^2 - y^2}$ can be thought of as the composition $g \circ f(x, y)$ where $f(x, y) = x^2 - y^2$ and $g(u) = e^u$. By Lemma 1.2.1, the level curves of $h$ are those of $f$ with different levels attached. Recall that the level curves of $f(x, y) = x^2 - y^2$ are the hyperbolas of Figure 1.2.7, so these are also level curves for $h$. Since $g(u) = e^u$, exponentiating the levels in Figure 1.2.7 gives the $h$-levels. See Figure 1.2.12.

(a) Level curves of $f(x, y) = x^2 - y^2$  (b) Level curves of $h(x, y) = e^{x^2 - y^2}$

**Figure 1.2.12: Level curves of compositions**

Before continuing, we show analytically that the level curve of $f(x, y)$ at level $-1$ is the level curve of $h(x, y)$ at level $e^{-1}$. The level curve of $h(x, y)$ at level $e^{-1}$ is the solution set to the equation $e^{-1} = h(x, y)$, which is

$$e^{-1} = e^{x^2 - y^2}.$$
Taking the natural logarithm of both sides gives

\[-1 = x^2 - y^2,\]

which is the level curve of \( f(x, y) \) at level \(-1\). ▲

In the previous example \( g(u) \) was a one-to-one function, and that made the level curves of \( f \) and \( g \circ f \) the same sets. If \( g \) is not one-to-one, more complicated things can happen. We illustrate with another example.

**Example 1.2.11. Composing with many-to-one functions**

The function \( h(x, y) = \sin(x^2 - y^2) \) can be thought of as the composition of \( f(x, y) = x^2 - y^2 \) and the function \( g(u) = \sin u \). The level curve of \( h \) at level \( c = 1 \) is the solution set to the equation

\[ \sin(x^2 - y^2) = 1. \]

As \( \sin \alpha = 1 \) when \( \alpha = \pi/2 + 2\pi n \) for any integer \( n \), this set is equivalent to the solutions to the infinite system of equations

\[ x^2 - y^2 = \frac{\pi}{2} + 2\pi n. \]

Thus one “level curve” of \( h(x, y) = g \circ f(x, y) \) is the union of infinitely many level curves of \( f(x, y) \). These can be seen in Figure 1.2.13 as the hyperbolic ridges of the rippled surface \( z = h(x, y) \). ▲

![Figure 1.2.13: The surface \( z = \sin(x^2 - y^2) \)](image)

**Example 1.2.12. Domain restrictions**

Recall that for \( u \geq 1 \) the inverse secant, \( g(u) = \sec^{-1} u \), is the angle between 0 and \( \pi/2 \) whose secant is \( u \). Thus for any \( f(x, y) \) whose range is at least one we can form \( g \circ f(x, y) \). For example, \( f(x, y) = x^2 + y^2 + 1 \) is the paraboloid of Figure 1.2.2(b) shifted up one unit, so we can form

\[ h(x, y) = g \circ f(x, y) = \sec^{-1} \left( x^2 + y^2 + 1 \right). \]
By Lemma 1.2.1 we know that the level curves of $h$ are the same as those of $f$ with different levels. Hence they are concentric circles, centered at the origin. More analysis shows that the level curve of $h$ at level $0 \leq c < \pi/2$ is

$$\sec^{-1}\left(x^2 + y^2 + 1\right) = c$$

$$x^2 + y^2 + 1 = \sec(c)$$

$$x^2 + y^2 = \sec(c) - 1.$$

Thus the graph of $h$ intersects the plane $z = c$ in a circle radius $\sqrt{\sec(c) - 1}$. The surface $z = h(x, y)$ is pictured in Figure 1.2.14.

In this section we’ve been introduced to surfaces that arise as graphs of functions, and the level curve approach toward understanding them. Level curves sketched in the same plane are the mathematical analogue of topographic maps. We were also introduced to composition of functions, and how level curves of compositions are related.

**Things to know/Skills to have**

- Know the definition of the graph of a function of two variables.
- Be able to sketch level curves for a given function and level.
- Be able to determine maxima, minima, and saddle points for a function from its level curves.
1.2. SURFACES FROM GRAPHS

- Determine if a composition of functions makes sense.
- Determine the level curves of a composition of functions.

Exercises

1. Decide whether each point $P$ is a relative maximum, relative minimum, or a saddle point. Give a one-sentence justification for your answer.

2. Sketch the level curves of $f(x, y) = xy$ at levels $c = -1, 0, 1, 2$. What kind of surface is this?

3. Sketch the level curves of $f(x, y) = x^2 + 4y^2$ at levels $c = 0, 4, 16$. For what levels $c$ does the level curve $f(x, y) = c$ not exist? What does this imply about the intersection of the surface $z = f(x, y)$ and the plane $z = c$ for these levels?

4. Some level curves of the function $f(x, y) = x^3 - 3x + y^2$ are pictured below. Guess whether the points $P$ and $Q$ represent maxima, minima, or saddles. What keeps you from being definite in your answer?

5. Sketch the level curves of $f(x, y) = 3x + 2y + 7$ for levels $c = -2, 1, 8$. Can you guess at the shape of the surface?

6. Sketch the level curves of $f(x, y) = x^2 - y^2 - 4x - 2y$ at levels $c = -4, -3, -2, 0$. What kind of surface do you get?

7. Sketch the level curves of $f(x, y) = \sqrt{16 - x^2 - y^2}$ for levels $c = 0, 1, 2, 3, 4$. Can you describe the surface?
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

8. Sketch the level curves of \( f(x, y) = x^2 + y^2 - 2xy \) for levels \( c = 0, 1, 4 \). Can you guess at the shape of the surface? (Hint: factor \( f(x, y) \) first)

9. Maximize \( f(x, y) = x - \sqrt{3}y \) subject to the constraint \( x^2 + y^2 = 1 \).

10. Sketch the level curves for \( f(x, y) = 3x - 3y + 4 \) at levels \( c = 0, 4, 16 \).
An ant is crawling on the surface \( z = f(x, y) \) above the unit circle in the \( xy \)-plane. What are the highest and lowest elevations the ant attains? Sketch the surface in \( \mathbb{R}^3 \).

11. Sketch the level curves for \( f(x, y) = x^2 + y^2 - 6x + 16y \) at levels \( c = 0, 4, 16 \).
An ant is crawling on the surface \( z = f(x, y) \) above the unit circle in the \( xy \)-plane. What are the highest and lowest elevations the ant attains? Sketch the surface in \( \mathbb{R}^3 \).

12. Different surfaces can have very similar level curves!

(a) Sketch the level curves of \( f(x, y) = x^2 + y^2 \) at levels \( c = 0, 1, 2 \).

(b) Sketch the level curves of \( f(x, y) = \sqrt{x^2 + y^2} \) at levels \( c = 0, 1, 2 \).

(c) The level curves for both surfaces are concentric circles centered at the origin. How can you tell the surfaces apart from their level curves?

13. Describe the curves of intersection of the surface \( z = x^3 - y + 1 \) with the planes \( x = -1, x = 0, \) and \( x = 1 \). Sketch the curves and connect them to form sketch of the surface.

14. Describe the curves of intersection of the surface \( z = x^3 - y + 1 \) with the planes \( y = -1, y = 0, \) and \( y = 1 \). Sketch the curves and connect them to form sketch of the surface.

15. Describe the curves of intersection of the surface \( z = e^x + y^2 \) with the planes \( y = -1, y = 0, \) and \( y = 1 \). Sketch the curves and connect them to form sketch of the surface.

16. The saddle \( f(x, y) = (x - y)(x + y) \) has two portions sloping down and two sloping up. The Monkey Saddle \( f(x, y) = x(x - y)(x + y) \) has three portions of the surface sloping down (two for the legs and one for the tail) and three up. Can you guess a formula for a surface with four upward and four downward sloping portions? Can you generalize?

17. Which of the following compositions are defined on the \( xy \)-plane?

(a) \( g \circ f \) where \( f(x, y) = x^2 - y^2 + 1, g(u) = u^2 + 1 \).

(b) \( g \circ f \) where \( f(x, y) = x^2 - y^2 + 1, g(u) = \ln u \).

(c) \( g \circ f \) where \( f(u) = \ln u, g(x, y) = x^2 + y^2 + 1 \).

(d) \( g \circ f \) where \( f(x, y) = x^2 - y^2 + 1, g(u) = \frac{1}{1 + u^2} \).
1.2. **SURFACES FROM GRAPHS**

18. Sketch some level curves of \( h(x, y) = \ln(x^2 + y^2 + 1) \). Describe the surface in a sentence or two.

19. Sketch some level curves of \( h(x, y) = \cos(x^2 - y^2) \). Describe the surface in a sentence or two.

20. Compare the level curves of \( h(x, y) = \cos(x^2 + y^2) \) and \( h(x, y) = \cos(x^2 + y^2)/(1 + x^2 + y^2) \). Include a description of the surfaces.

21. Sketch some level curves of \( h(x, y) = 1/(1 + x^2 + y^2) \), and describe what the surface does near infinity.
1.3 Surfaces from Equations

In this section we present a potpourri of surfaces arising as the solution sets of equations, and techniques to study them. We study slicing surfaces with planes, quadric surfaces, generalized cylinders, planes, level surfaces, and equations in other coordinate systems. The basic technique for studying quadric surfaces and level surfaces will be to slice the surfaces with planes parallel to coordinate planes. Generalized cylinders turn out to be easy to recognize from their equations, and easy to sketch! Geometric understanding will help analyze surfaces given to us using equations in other coordinate systems. Enjoy the variety!

Slicing Surfaces

We set out to understand surfaces arising from the graphs of functions from their formulas, and have found that level curves are one helpful tool in that regard. The basic idea is to use two-dimensional slices (the level curves) to understand a surface in three dimensions. In the level curve setting, we always slice the surface with horizontal planes. In this section we generalize Section 1.2 in two ways. One generalization is that we consider surfaces which need not be graphs of functions. This is analogous to considering curves in the plane defined as solution sets to equations which aren’t necessarily solved for one variable. For example, the unit circle \( x^2 + y^2 = 1 \) is not the graph of a function. The other generalization is that the planes we slice them with need not be horizontal. Allowing for this flexibility can give greater insight to the structure of some surfaces. Which planes give you greatest insight into a surface is usually dictated by the equation that defines it.

We begin with some examples slicing graphs of functions with other planes, then proceed to surfaces whose defining equations are not functions of one of the variables.

Example 1.3.1. Slicing Surfaces for Geometric Understanding

Level curves of the function \( f(x, y) = x^3 - x + y^2 \) are curves in the \( xy \)-plane defined by

\[ x^3 - x + y^2 = c, \]

for some constant \( c \). Since these curves are hard to visualize, it may be easier to find the intersections with other planes instead. Let’s investigate the graph of \( f(x, y) = x^3 - x + y^2 \) by slicing it with planes parallel to the \( xz \)-plane.

First, the curve of intersection of the surface \( z = x^3 - x + y^2 \) with the plane \( y = 0 \) is the solution to the system of equations

\[
\begin{align*}
z &= x^3 - x + y^2 \\
y &= 0.
\end{align*}
\]

An equation for the curve is obtained by substituting 0 for \( y \) in the first equation, yielding \( z = x^3 - x \). One can plot this curve in the \( xz \)-plane, as in Figure 1.3.1(b). The intersection with the plane \( y = -3 \) has equation \( z = x^3 - x + (-3)^2 = x^3 - x + 9 \), and is pictured in 1.3.1(a). Similarly the \( y = 2 \) cross-section is in 1.3.1(c), and they are put together in Figure 1.3.2.
1.3. SURFACES FROM EQUATIONS

![Figure 1.3.1: Planar cross-sections of \( f(x,y) = x^3 - x + y^2 \)]

More generally, to describe the intersection of the surface with the plane \( y = c \) analytically we substitute \( c \) for \( y \), obtaining \( z = x^3 - x + c^2 \). This is a vertical translation of the curve \( z = x^3 - x \) by \( c^2 \) units. So as you move back and forth along the \( y \)-axis, the curve \( z = x^3 - x \) gets translated up by the appropriate amount, sweeping out the surface as in Figure 1.3.2. The grid lines on the surface that are “parallel” to the highlighted one in the middle are all intersections with planes \( y = c \). This technique is convenient because the \( x \)'s and \( y \)'s in \( f(x,y) = x^3 - x + y^2 \) are added together, making the vertical translation easy to see.

![Figure 1.3.2: The Surface \( z = x^3 - x + y^2 \)]

**Example 1.3.2. Fun with trigonometric functions**

Equations for the level curves of \( f(x,y) = \cos x + \sin y \) are of the form

\[
\cos(x) + \sin(y) = c,
\]

which are definitely not fun to work with. If we want to have fun, we’ll choose other planes to get some two-dimensional slices with. If we choose to slice the surface \( z = \cos x + \sin y \) with planes parallel to the \( yz \)-plane, that amounts to...
replacing $x$ with a constant and analyzing the curves. The curve of intersection with the $yz$-plane (or $x = 0$), is given by $z = \cos \theta + \sin x = 1 + \sin y$, which is a translation of the sine curve one unit up. In fact, slicing with the plane $x = c$ gives the curve

$$z = \cos c + \sin y,$$

which is a vertical translation of $z = \sin y$ by the value $\cos c$. The surface $z = \cos x + \sin y$ can be thought of, then, as taking the sine curve in the $y$-direction, and having it ride along a roller coaster in the $x$-direction. The roller coaster is the cosine curve.

Figure 1.3.3: The Surface $z = \cos x + \sin y$

The preceeding two examples involved surfaces that were graphs of functions, but didn’t have particularly nice level curves. The strategy was to get two-dimensional slices by cutting with planes parallel to one of the coordinate planes, other than the $xy$-plane. We now consider some surfaces arising from equations that can’t be solved for one of the variables. We will find that slicing with a variety of planes gives geometric insight to the surfaces.

**Quadric Surfaces**

A quadric surface is the solution to a quadratic equation in three variables. The quadratic equations we will consider have the form

$$Ax^2 + By^2 + Cz^2 + Dx + Ey + Fz = G$$

(we avoid equations with mixed terms $xy$, $xz$ or $yz$). Our basic strategy for understanding such surfaces will be to sketch a “skeleton” by intersecting the surface with coordinate planes, then “connect the curves”. We illustrate this with two examples.

**Example 1.3.3. An ellipsoid**

The solution set to $\frac{x^2}{a^2} + y^2 = 1$ is an ellipse. Analogously, the solution set to the three-variable equation $\frac{x^2}{a^2} + y^2 + \frac{z^2}{b^2} = 1$ is an ellipsoid. To get a feel for what it looks like, consider the intersections with coordinate planes. When first sketching these surfaces, it might help to sketch the intersections in separate planes, as in Figure 1.3.4 then piece them together. To find equations for these curves of intersection, set one of the variables to zero and sketch the resulting curve in the appropriate coordinate plane. The intersection of
1.3. SURFACES FROM EQUATIONS

Figure 1.3.4: Planar cross-sections of \( \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 1 \)

the surface \( \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 1 \) with the \( xy \)-plane is gotten by setting \( z = 0 \) (technically, we’re solving the system of equations consisting of both, and getting the intersection of the surfaces). Analogously, the intersection with the \( xz \)-and \( yz \)-planes are the ellipses \( \frac{x^2}{4} + \frac{z^2}{9} = 1 \) and \( y^2 + \frac{z^2}{9} = 1 \), respectively. The planes pieced together form a “skeleton” for the whole ellipsoid pictured in Figure 1.3.5.

▲

Example 1.3.4. The one-sheeted hyperboloid

The solution set to the equation \( x^2 + y^2 - z^2 = 1 \) is called a hyperboloid of one sheet. Intersecting it with the coordinate planes yields the circle \( x^2 + y^2 = 1 \) in the \( xy \)-plane, and the hyperbolas \( x^2 - z^2 = 1 \) and \( y^2 - z^2 = 1 \) in the \( xz \)- and \( yz \)-planes, respectively (see Figure 1.3.6).

The one sheeted hyperboloid can also be thought of as a surface of revolution. To see this, just revolve one component of \( x^2 - z^2 = 1 \) around the \( z \)-axis. You can also see this analytically by changing the equation to cylindrical coordinates. Recall that \( x^2 + y^2 = r^2 \), so the Cartesian equation for the surface \( x^2 + y^2 - z^2 = 1 \) becomes the cylindrical equation \( r^2 - z^2 = 1 \). You notice the equation is independent of \( \theta \), which implies that this is a surface of revolution. ▲

Generalized Cylinders

We have applied the strategy of slicing surfaces with planes to several examples at this point. There are some surfaces that arise from simple equations
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

Figure 1.3.6: A one-sheeted hyperboloid

which are even easier to visualize. They are generalized cylinders, and we discuss them and their equations now.

We have already seen that the cylindrical coordinate equation \( r = 2 \) describes a cylinder in space. Translating to Cartesian coordinates for \( \mathbb{R}^3 \) this corresponds to the equation \( \sqrt{x^2 + y^2} = 2 \), or \( x^2 + y^2 = 4 \), which is missing the \( z \)-coordinate entirely. One way to think of \( r = 2 \), then, is that it is the curve \( x^2 + y^2 = 4 \) in the \( xy \)-plane translated up and down along the \( z \)-axis. More generally, we’ll say that any surface whose Cartesian equation is missing one variable is a cylinder.

To visualize such surfaces:

1. Sketch the curve in the appropriate plane.
2. Translate the curve along the axis of the missing variable.

**Example 1.3.5. A hyperbolic cylinder**

Sketch the surface \( x^2 - 4z^2 = 1 \). We know this is a generalized cylinder since the \( y \)-coordinate is missing from the equation. The first step is to sketch the curve \( x^2 - 4z^2 = 1 \) in the \( xz \)-plane. This is a hyperbola with vertices \((\pm1, 0, 0)\), and asymptotes \( z = \pm x/2 \) (see the dark curve of intersection in Figure 1.3.7).

To get the surface, simply translate the hyperbola back and forth along the \( y \)-axis, as in Figure 1.3.7.

**Example 1.3.6. A washboard**

You can sketch the surface \( y = \sin z \) in \( \mathbb{R}^3 \) in two steps. First sketch the curve \( y = \sin z \) in the \( yz \)-plane, then translate the curve back and forth along the \( x \)-axis. This process is illustrated in Figure 1.3.8.

**Planes in \( \mathbb{R}^3 \)**

The general form of an equation for a line in \( \mathbb{R}^2 \) is \( Ax + By = C \). Analogously, the general form for an equation of a plane in \( \mathbb{R}^3 \) is \( Ax + By + Cz = D \), where \( A \) through \( D \) are constants and at least one of \( A \), \( B \), or \( C \) is not zero. The constants in the equation \( Ax + By + Cz = D \) turn out to have significant
1.3. SURFACES FROM EQUATIONS

The generalized cylinder $x^2 - 4z^2 = 1$

An easy way to visualize the plane $5x + 2y + z = 10$ is to plot the intersections with the coordinate axes, connect the dots to form a triangle, then extend it to a plane. The $y$- and $z$-coordinates of points on the $x$-axis in $\mathbb{R}^3$ are both zero. This implies that letting $y = z = 0$ in the equation for the plane gives the $x$-coordinate, and we get the intercept $(2, 0, 0)$. Similarly, we get $y$-intercept $(0, 5, 0)$ and $z$-intercept $(0, 0, 10)$. Plotting these points and connecting the dots gives a triangle that lives in the plane (see Figure 1.3.9(a)). Extending the triangle in all directions gives the plane itself, as in Figure 1.3.9(b).

**Example 1.3.7. The plane** $5x + 2y + z = 10$

A level curve is a curve in the plane given by $f(x, y) = c$ for some function of two variables $f(x, y)$. Analogously, a level surface is a surface in space obtained by setting a function of three variables equal to a constant, in symbols $f(x, y, z) = c$. Since the result is a three variable equation, level surfaces are examples of surfaces given as the solution set to an equation. We illustrate them with several examples.
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

Figure 1.3.9: Visualizing the plane $Ax + By + Cz = D$

Example 1.3.8. Planar level surfaces

Let $f(x, y, z) = x - y + 2z$, and sketch the level surfaces at levels $c = -1, 1, 4$. The level surface at level $c = -1$ is the plane $x - y + 2z = -1$. As above we can find the $x$-, $y$-, and $z$-intercepts, plot them and sketch. In this case they are $(-1, 0, 0), (0, 1, 0)$, and $(0, 0, -1/2)$. The resulting level surface is the lowest plane pictured in Figure 1.3.10. The remaining level surfaces are pictured as well.

Figure 1.3.10: Level surfaces of $f(x, y, z) = x - y + 2z$

As the level increases, then, the level surface remains planar but is translated up the $z$-axis. In this example, all level surfaces are the same geometric object—a
plane. We will soon see that this is not always the case. ▲

**Example 1.3.9. Quadric level surfaces**

In this example we sketch the level surfaces of \(g(x, y, z) = x^2 - y^2 - z^2\) at levels \(c = -1, 0, 1, 2\). We treat the cases separately, and it will be interesting to see how different levels can lead to very different surfaces.

The level \(c = -1\) yields the equation \(-1 = x^2 - y^2 - z^2\). Multiplying by negative one gives \(y^2 + z^2 - x^2 = 1\), which is the same equation as in Example 1.3.4 except the roles of the \(x\) and \(z\) variables are switched. As Example 1.3.4 yielded a one-sheeted hyperboloid with \(z\)-axis as core, the surface \(y^2 + z^2 - x^2 = 1\) is a one-sheeted hyperboloid with \(x\)-axis as core (see Figure 1.3.11(a)).

At level zero, the resulting equation simplifies to \(x^2 = y^2 + z^2\). We analyze this surface by first finding intersections with the coordinate planes. Letting \(z = 0\) gives \(x^2 = y^2\), which is the pair of lines \(y = \pm x\). Similarly, letting \(y = 0\) gives the pair of lines \(z = \pm x\) in the \(xz\)-plane. Further, one sees from the equation \(x^2 = y^2 + z^2\) that fixing \(x = d\) yields a circle in that plane of radius \(d\). Thus the surface is a cone with \(x\)-axis as core, as in Figure 1.3.11(b).

![Figure 1.3.11: Level surfaces of \(g(x, y, z) = x^2 - y^2 - z^2\)](image)

The level surface at level \(c = 1\) is the solution to the equation \(x^2 - y^2 - z^2 = 1\). Clearly if \(x = 0\) there are no solutions to the equation, so this surface misses the \(yz\)-plane. The intersection with the \(xy\)-plane is the hyperbola \(x^2 - y^2 = 1\), and with the \(xz\)-plane is \(x^2 - z^2 = 1\). These hyperbolas are highlighted in Figure 1.3.11(c). Notice that letting \(x = d > 1\) yields the curve \(y^2 + z^2 = d^2 - 1\). This is a circle in the plane \(x = d\) centered on the \(x\)-axis. Thus the surface is the two-sheeted hyperboloid of Figure 1.3.11(c).
The level surface at $c = 2$ is similar to $c = 1$, and the results of this discussion are summarized in the following table. The level surfaces are pictured on the same set of axes in Figure 1.3.11(d).

<table>
<thead>
<tr>
<th>Level</th>
<th>Level Surface</th>
<th>Surface Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c = -1$</td>
<td>$-1 = x^2 - y^2 - z^2$</td>
<td>One-sheeted hyperboloid</td>
</tr>
<tr>
<td>$c = 0$</td>
<td>$0 = x^2 - y^2 - z^2$</td>
<td>Cone</td>
</tr>
<tr>
<td>$c = 1$</td>
<td>$1 = x^2 - y^2 - z^2$</td>
<td>Two-sheeted hyperboloid</td>
</tr>
<tr>
<td>$c = 2$</td>
<td>$2 = x^2 - y^2 - z^2$</td>
<td>Two-sheeted hyperboloid</td>
</tr>
</tbody>
</table>

In summary, then, for $c < 0$ the level surface $x^2 - y^2 - z^2 = c$ will be a one-sheeted hyperboloid. As $c \to 0^-$, the “waist” of the hyperboloid shrinks. In the limit, then, one gets a cone when $c = 0$. The cone pinches the waist so that when $c > 0$ the level surfaces are two-sheeted hyperboloids. ▲

Notice that just as level curves of $f(x,y)$ fill up the plane, level surfaces of $g(x,y,z)$ fill up $\mathbb{R}^3$. We conclude the level surface portion with an application.

**Example 1.3.10. Temperature near a heat source**

Suppose the function $T(x,y,z) = 20 - (x^2 + y^2 + z^2)$ gives the temperature (in degrees Celcius) near a heat source at the origin. In this context, the level surface at level $c = 11$ is the surface of constant temperature $11^\circ C$ (this is sometimes called an isothermal surface). Thus, if our original function is temperature level surfaces represent surfaces of constant temperature. In our case, the surface is the solution set to the equation $11 = 20 - (x^2 + y^2 + z^2)$. Isolating the variables gives

$$11 = 20 - (x^2 + y^2 + z^2)$$

$$x^2 + y^2 + z^2 = 9,$$

which is the equation of a sphere centered at the origin of radius 3. More generally, the isothermal surfaces of this example are concentric spheres centered at the origin. ▲

**Equations in other coordinate systems**

Since we are studying surfaces given by solution sets of an equation, it makes sense to look at the solution set of cylindrical and spherical equations. Even simple equations can give quite interesting surfaces.

**Example 1.3.11. A familiar surface**

Using different coordinate systems can also lead to a better geometric understanding of surfaces. To examine the graph of $f(x,y) = x^2 + y^2$, we can write the equation $z = f(x,y)$ in cylindrical coordinates as follows:

$$z = x^2 + y^2 = (r \cos \theta)^2 + (r \sin \theta)^2 = r^2 \cos^2 \theta + r^2 \sin^2 \theta = r^2.$$ 

Since $z = r^2$ has no $\theta$ in it, points on the surface are independent of $\theta$. This really means the surface has rotational symmetry about the $z$-axis! To sketch it,
then, you can sketch the curve in the $yz$-plane, then rotate it about the $z$-axis. The $yz$-plane is given by $x = 0$, so the equation for the curve of intersection is $z = y^2$, a parabola. Revolving the parabola $z = y^2$ around the $z$-axis yields the surface.

**Example 1.3.12. A cone**

To understand the graph of $f(x, y) = \sqrt{x^2 + y^2}$, we can write the equation $z = \sqrt{x^2 + y^2}$ in spherical coordinates. Using the change of coordinates, we see that $z = \sqrt{x^2 + y^2}$ becomes

\[
\rho \cos \phi = \sqrt{(\rho \sin \phi \cos \theta)^2 + (\rho \sin \phi \sin \theta)^2}
\]

\[
= \sqrt{\rho^2 \sin^2 \phi}
\]

\[
= \rho \sin \phi.
\]

(1.3.1)

Now $\rho \cos \phi = \rho \sin \phi$ simplifies to $\tan \phi = 1$, which is satisfied by the spherical equation $\phi = \pi/4$. As before, this is a cone that makes an angle of $\pi/4$ with the positive $z$-axis.

**Example 1.3.13. A cylindrical approach**

To understand the surface $z = \sqrt{3r}$, it’s probably easiest to change to spherical coordinates, as in the previous example. Since $z = \rho \cos \phi$ and $r = \rho \sin \phi$, equation $z = \sqrt{3r}$ becomes the spherical equation $\rho \cos \phi = \sqrt{3} \rho \sin \phi$. Dividing by $\rho \cos \phi$ and solving gives $\tan \phi = 1/\sqrt{3}$, from which we get $\phi = \tan^{-1}(1/\sqrt{3}) = \pi/6$. Thus the surface $z = \sqrt{3r}$ is the cone that makes an angle of $\pi/6$ with the positive $z$-axis.
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

Figure 1.3.13: The Cone \( z = \sqrt{x^2 + y^2} \)

Of course, this generalizes to \( z = Ar \) for any constant \( A \). Translating to spherical coordinates and solving gives:

\[
\begin{align*}
  z &= Ar \\
  \rho \cos \phi &= A \rho \sin \phi \\
  \cot \phi &= A.
\end{align*}
\]

So the coefficient of \( r \) is the cotangent of the angle the cone \( z = Ar \) makes with the \( z \)-axis.

Notice that the equation \( z = Ar \) is independent of \( \theta \). Hence the surface has rotational symmetry. To sketch it, then, you can sketch the intersection with the \( yz \)-plane and rotate it about the \( z \)-axis. ▲

**Example 1.3.14. A surface of revolution in spherical coordinates**

Sketch the surface \( \rho = 1 + \cos \phi \), for \( 0 \leq \phi \leq \pi \).

To do so, sketch the portion in the \( yz \)-plane, where it looks like a cardiod pointing down (recall that \( \phi \) is the angle with the positive \( z \)-axis). Rotating gives the upside-down “apple” of Figure 1.3.14 ▲

**Example 1.3.15. A sea shell**

Let’s look at the surface described by the spherical equation \( \rho = \theta \), for \( \theta > 0 \).

Fixing \( \theta = \frac{\pi}{3} \) is equivalent to intersecting the surface with the half-plane. So in the half plane, we’re looking at all points of the form \( \rho = \frac{\pi}{3} \). Since \( \rho \) is the distance to the origin, the intersection of the surface \( \rho = \theta \) with the half plane \( \theta = \frac{\pi}{3} \) is a semicircle of radius \( \frac{\pi}{3} \). Generalizing this observation, we notice that as \( \theta \) increases the intersection with \( \rho = \theta \) is a semicircle with increasing radius. The result is \( \rho = \theta \) looks like a sea shell.

**Things to know/Skills to have**

In this section we studied surfaces resulting from the solution set of equations. You know you understand the material of this section when you are able to:
1.3. SURFACES FROM EQUATIONS

Figure 1.3.14: The surfaces $\rho = 1 + \cos \phi$ and $\rho = \theta$

- Sketch the plane given by a linear equation in three variables.
- Sketch certain quadric surfaces, using their intersection with coordinate planes.
- Sketch level surfaces for a function of three variables.
- Sketch generalized cylinders.
- Sketch solution sets to equations given in cylindrical and spherical coordinates.

Exercises

1. Sketch the plane $7x - 3y + z = 21$.
2. Sketch the plane $-x + 2y + z = 8$.
3. Sketch the plane $4x - 3y - 3z = 12$.
4. Find coefficients $A$, $B$, $C$ so that the plane $Ax + By + Cz = 15$ has intercepts $(1, 0, 0)$, $(0, -5, 0)$, and $(0, 0, 3)$.
5. Find coefficients $A$, $B$, $C$ so that the plane $Ax + By + Cz = 8$ has intercepts $(-4, 0, 0)$, $(0, 2, 0)$, and $(0, 0, -8)$.
6. Assume $D > 0$, and the plane $Ax + By + Cz = D$ intersects the positive $z$-axis. What can you say about the coefficient $C$?
7. Find an equation for the plane with intercepts $(2, 0, 0)$, $(0, -4, 0)$, and $(0, 0, 6)$.
8. Find an equation for the plane with intercepts \((-3, 0, 0), (0, 3, 0),\) and \((0, 0, -2)\).

9. Sketch the intersections of the surface \(x^2 + y^2 - z^2 = -1\) with the coordinate planes. Sketch the entire surface. Why might this be called a two-sheeted hyperboloid?

10. Sketch the intersections of the surface \(x^2 + 4y^2 + z^2 = 1\) with the coordinate planes. What type of surface is this?

11. Sketch the intersections of the surface \(x^2 + y^2 - z^2 = 0\) with the coordinate planes. What type of surface is this?

12. Sketch the intersections of the surface \(x^2 - y + z^2 = 1\) with the coordinate planes. What type of surface is this?

13. Let \(f(x, y, z) = x + y + z\), and sketch the level surfaces at levels \(-1, 1, 2\).

14. Sketch the level surfaces of \(f(x, y, z) = x - y^2 - z^2\) at levels \(-1, 0, 2\).

15. Sketch the level surfaces of \(f(x, y, z) = x^2 + y^2 - z^2\) at levels \(-1, 0, 2\).

16. Sketch the level surfaces of \(g(x, y, z) = x^2 + y^2 + z^2\) at levels \(-1, 0, 4\).

17. Sketch the level surfaces of \(f(x, y, z) = 4x^2 + y^2 + 9z^2\) at levels \(-1, 1, 9\).

18. A wire runs along the \(z\)-axis, and conducts heat so that the temperature near it is given by \(T(x, y, z) = 12 - x^2 - y^2\) (so temperature is independent of \(z\)). Find the isothermal surfaces at temperatures \(0^\circ, 6^\circ, 10^\circ\).

19. Level surfaces and electric potential: The electric potential \(V\) a distance \(\rho\) units from a point charge \(Q\) is given by

\[
V = \frac{Q}{4\pi\epsilon_0\rho},
\]

where \(\epsilon_0 \approx 8.854 \times 10^{-12}\) farads per meter is the permittivity of free space. Suppose a point charge of 8 picocoulombs \((8 \times 10^{-12}\) Coulombs\) is placed at the origin in \(\mathbb{R}^3\).

(a) Find a spherical equation for \(V\), assuming \(\epsilon_0 = 8.8 \times 10^{-12}\).

(b) Find a Cartesian equation for \(V\).

An equipotential surface is one on which the potential \(V\) is constant—i.e. a level surface for \(V!\)

(c) Find a Cartesian equation for the 10 volt equipotential surface in this problem.

(d) Describe the equipotential surfaces of \(V\) geometrically.
20. Level surfaces and gravitational force: Two point masses, with mass $m_1$ and $m_2$, attract each other with a certain gravitational force. According to Newton’s law of universal gravitation, the magnitude of force $F$ between them is given by

$$F = G\frac{m_1 m_2}{\rho^2},$$

where $G \approx 6.674 \times 10^{-11} \text{N} \text{(m/kg)}^2$ is the gravitational constant and $\rho$ is the distance between the masses. If a unit mass is placed at the origin, find a Cartesian equation for the gravitational force between it and another unit mass placed at $(x, y, z)$. Find a Cartesian equation for the level surface $F = 100$.

21. Sketch the surface whose polar equation is $z^2 + r^2 = 9$. Hint: change to spherical coordinates.

22. Sketch the surface $z = r$, assume the convention that $r \geq 0$.

23. Find a cylindrical equation for the surface $\phi = \pi/4$, and sketch.

24. Let $0 < \alpha < \pi$ be a constant. Find a cylindrical equation for the surface $\phi = \alpha$.

25. (a) Sketch the curve in $\mathbb{R}^2$ given by the polar equation $r = \theta$.

(b) Sketch the surface in $\mathbb{R}^3$ given by the cylindrical equation $r = \theta$.

26. Sketch the surface in $\mathbb{R}^3$ whose Cartesian equation is $x = y^2$. Hint: this is a generalized cylinder.

27. Sketch the surface in $\mathbb{R}^3$ whose Cartesian equation is $x^2 + z^2 = 4$.

28. Sketch the surface in $\mathbb{R}^3$ whose Cartesian equation is $z = 4 - x^2$.

29. What is the Cartesian equation for the right circular cylinder of radius 1, with the $x$-axis as its core?
1.4 Parametric Curves

Curves in the plane can be described as graphs of functions and solution sets to equations. In the previous two sections we saw how these methods of describing curves in \( \mathbb{R}^2 \) generalize naturally to describe surfaces in \( \mathbb{R}^3 \). There is a third way we commonly describe curves in the plane: parametrically. It turns out that parametric curves in \( \mathbb{R}^2 \) have two natural generalizations to three dimensions. One generalization leads to parametric curves in \( \mathbb{R}^3 \), and the other to parametric surfaces! In this section we recall the definition of planar parametric curves, and extend it to curves in \( \mathbb{R}^3 \) by adding a third coordinate. Then, in Section 1.5, we will introduce parametric surfaces by adding a second parameter.

Recall that parametric curves are given by so-called parametric equations \( x = f(t), \ y = g(t) \). We sometimes use the vector notation \( C(t) = (f(t), g(t)) \), combining the two parametric equations into a single ordered pair. In this context, both the horizontal and vertical coordinates are functions of a third variable \( t \), called the parameter.

Example 1.4.1. A Fermat spiral

The curve given by the parametric equations

\[
C(t) = (f(t), g(t)) = (\sqrt{t} \cos t, \sqrt{t} \sin t), \quad 0 \leq t \leq 8\pi
\]

is the Fermat spiral of Figure 1.4.1. Recall that parametric curves are implicitly endowed with a direction—that of increasing \( t \). For the Fermat spiral the direction of increasing \( t \) is counterclockwise. ▲

![Figure 1.4.1: The Fermat Spiral](image)

We describe two ways of interpreting parametric curves. One way to think of parametric curves is that they describe the position \( C(t) \) of an ant crawling on the plane at time \( t \). In the previous example, then, at time \( t = \frac{\pi}{2} \) we substitute into the parametric equations to find the ant is at position...
Another way to think of parametric curves is that they tell you how to bend, stretch, and shrink a piece of wire into the plane. In our example, we start with a straight wire from 0 to $8\pi$ on the $t$-number line. The parametric equations $x(t) = \sqrt{t} \cos t$ and $y(t) = \sqrt{t} \sin t$ are the instructions for how to bend and stretch the wire into the plane. Thus the domain of the parametric curve $C(t)$ is the $t$-axis, which is a separate space from the range of $C(t)$. Then $C(t)$ is viewed as a mapping from one to the other. The range of $C(t)$ is a subset of $\mathbb{R}^2$, namely the curve itself. Conceptually you can think of it as in Figure 1.4.2.

This is a different way of viewing curves than we might be used to. When working with the graphs of functions, the domain and range are in the same space; whereas parametric curves place the domain and range in separate spaces.

### 1.4.1 Curves in $\mathbb{R}^2$

In this subsection we review parameterizations of several familiar curves in the plane: lines, ellipses and hyperbolas. We also review techniques for parameterizing the graph of a function and a curve given by a polar equation. Both techniques generalize nicely to parametric surfaces.

**Parameterizing Lines**

It turns out that any line can parameterized using linear equations. For any constants $a, b, c, d$, the parametric equations

\[
    x = at + b \\
    y = ct + d
\]

for $-\infty < t < \infty$ yield a line. One way to see this is solve the first equation for $t$, giving $t = (x - b)/a$. Substituting into the parametric equation for $y$ gives

\[
    y = c \left( \frac{x - b}{a} \right) + d = \frac{c}{a} x + \frac{ad - bc}{a}.
\]
As long as \( a \neq 0 \), this gives an equation for a line. If \( a = 0 \), the result is a vertical line.

**Parametrizing Graphs of Functions**

It is easy to find parametric equations for the graph of a function. More precisely, to parameterize the curve \( y = f(x) \), simply let \( x = t \) and \( y = f(t) \). Parametric equations for our hyperbolic sine function are then

\[
\begin{align*}
x &= t \\
y &= e^t - e^{-t}
\end{align*}
\]

for \(-\infty < t < \infty\). This seems rather artificial, but it does represent a change of perspective which can sometimes be useful. Similarly, parametric equations for the graph of \( y = x^3 - x \) are \( C(t) = (t, t^3 - t) \), for \(-\infty < t < \infty\). The direction of these parameterizations is the direction of increasing \( t \), which is left to right.

**Example 1.4.2. Parametrizing Ellipses**

Recall that \( C(t) = (x(t), y(t)) = (\cos t, \sin t) \) for \( 0 \leq t \leq 2\pi \) parameterizes the unit circle. The ellipse

\[
\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1
\]

is obtained by stretching the unit circle by a factor of \( a \) units horizontally and \( b \) units vertically. The same rescaling of parametric equations parameterizes the ellipse. More precisely, \( C(t) = (x(t), y(t)) = (a\cos t, b\sin t) \) for \( 0 \leq t \leq 2\pi \) parameterizes the ellipse.

To see this analytically, a simple substitution verifies that the parametric equations satisfy the Cartesian equation of the ellipse:

\[
\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{(a\cos t)^2}{a^2} + \frac{(b\sin t)^2}{b^2} = \cos^2 t + \sin^2 t = 1. \tag{1.4.2}
\]

Notice that the given parameterization starts at \((a, 0)\) and traverses the ellipse counterclockwise. One can also parameterize the ellipse in a clockwise fashion. One way to do this is to negate the \( y \)-coordinate, giving \( C(t) = (a\cos t, -b\sin t) \). This starts at \((a, 0)\), goes clockwise, and lies on the ellipse.
Another clockwise parameterization comes from switching the sine and cosine functions, resulting in \( C(t) = (a \sin t, b \cos t) \) for \( 0 \leq t \leq 2\pi \). Indeed this starts at the point \( C(0) = (0, b) \), and at time \( t = \pi/2 \) we have \( C(\pi/2) = (a, 0) \). Thus the parameterization has a different starting point, and goes clockwise around the ellipse. A calculation as in Equation 1.4.2 verifies that \( C(t) \) is on the ellipse.

\[ \text{Example 1.4.3. Parameterizing Hyperbolas} \]

This section started with the graph of \( f(x) = \sinh x \). The hyperbolic cosine function is similar, and defined by \( \cosh x = \frac{e^x + e^{-x}}{2} \). Note that \( \cosh x \geq 1 \) for all \( x \). There is a hyperbolic Pythagorean identity, which is analogous to the familiar \( \sin^2 x + \cos^2 x = 1 \). It is not hard to verify that \( \cosh^2 x - \sinh^2 x = 1 \).

Using this, one can show that \( x = a \cosh t, y = b \sinh t \) parameterizes the right half of the hyperbola \( \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \) (we are assuming both \( a \) and \( b \) are positive).

To verify that points on the parametric curve \( C(t) = (a \cosh t, b \sinh t) \) are on the hyperbola \( \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \), substitute and simplify:

\[
\frac{x^2}{a^2} - \frac{y^2}{b^2} = \frac{(a \cosh t)^2}{a^2} - \frac{(b \sinh t)^2}{b^2} = \cosh^2 t - \sinh^2 t = 1.
\]

Since \( \cosh t \geq 1 \), the curve \( C(t) = (a \cosh t, b \sinh t) \) for \( -\infty < t < \infty \) is the right component of the hyperbola. Since \( \sinh t \) is negative for \( t < 0 \), the parameterization starts at the bottom right, goes to the vertex \((1, 0)\) at \( t = 0 \), and continues up as \( t \) increases.

\[ \text{Parameterizing Polar Equations} \]

Curves in \( \mathbb{R}^2 \) can also be described as solution sets to polar equations \( r = f(\theta) \). For example, the polar equation \( r = \sqrt{\theta} \) for \( \theta > 0 \) yields Fermat’s spiral of Figure 1.4.1. To find parametric equations for Fermat’s spiral, you need to find coordinate functions for Cartesian coordinates \( x \) and \( y \). This can be accomplished in two steps. First let \( \theta = t \) which implies \( r = f(t) \), and then use the change of coordinates \( x = r \cos \theta, \; y = r \sin \theta \). In our case, then, substituting \( \theta = t \) and \( r = f(t) = \sqrt{t} \) into the change of coordinates yields the parametric equations we had originally:

\[ \mathbf{C}(t) = (x, y) = \left( \sqrt{t} \cos t, \sqrt{t} \sin t \right), \text{ for } t \geq 0. \]

\[ \text{Parameterizing Polar Equations} \]

To parameterize a curve given by the polar function \( r = f(\theta) \), let \( \theta \) be the parameter \( t \) and use the change of coordinates formula. More precisely, a parameterization for the curve \( r = f(\theta), \; a \leq \theta \leq b \), is

\[ \mathbf{C}(t) = (f(t) \cos t, f(t) \sin t), \text{ for } a \leq \theta \leq b \]
Example 1.4.4. A three-leafed rose

Parameterize the three-leaf rose \( r = \cos 3\theta \) pictured in Figure 1.4.3.

Letting \( \theta = t \) we find \( r = \cos 3t \). Substituting into the coordinate conversion we get

\[
C(t) = (\cos 3t \cos t, \cos 3t \sin t), \quad 0 \leq t \leq \pi.
\]

The limits on \( t \) come from observing that for \( \pi \leq t \leq 2\pi \) the curve traces back over itself. We could also parameterize just the horizontal petal by determining when \( r = 0 \). These would be the \( t \)-values nearest 0 where \( \cos 3t = 0 \), or \( t = \pm \pi/6 \). Thus the limits \( -\pi/6 \leq t \leq \pi/6 \) parameterize the horizontal petal. ▲

This technique for parameterizing planar curves given a polar equation will generalize to parameterizing curves and surfaces in higher dimensions. For example, one can parameterize a surface given by an equation in spherical coordinates using the appropriate coordinate conversion.

1.4.2 Curves in \( \mathbb{R}^3 \)

We have seen that curves in \( \mathbb{R}^3 \) can be described as the intersection of two surfaces. Thus, they are the solution to a system of two equations in three variables. For example, the solution set of the system \( \rho = 3, \phi = \pi/4 \) is the circle of intersection of the sphere \( \rho = 3 \) and the cone \( \phi = \pi/4 \). Using two equations to describe curves, however, would make it difficult (impossible?) to compute arclength, centers of mass, circulation of fluid, and various other quantities related to curves. Using parametric equations, however, allows for such calculations, and is therefore our preferred method for working with curves in three space.

We have just reminded ourselves that a parametric curve in the plane is given by two coordinate functions \( C(t) = (x(t), y(t)) \). Recall that we had two intuitive interpretations of parametric curves to aide our understanding. If the parameter \( t \) represents time, one can think of \( x(t) \) and \( y(t) \) as giving the horizontal and vertical position of a bug crawling in the plane at time \( t \). Alternatively, parametric equations can be thought of as instructions for bending, stretching...
and shrinking a wire into the plane. The same intuitive interpretations are valid for curves in $\mathbb{R}^3$, the only additional information you need is a third coordinate function $z(t)$. We give a formal definition, and illustrate parametric curves with several examples.

**Definition 1.4.1.** A parametric curve is one given by coordinate functions $x(t), y(t),$ and $z(t)$, for $a \leq t \leq b$. We denote them using ordered triples as in:

$$C(t) = (x(t), y(t), z(t)), \ a \leq t \leq b.$$ 

**Example 1.4.5.** Lines

Lines in the plane have linear coordinate functions. Analogously, a line in space is given by parametric equations of the form:

$$x = At + x_0$$
$$y = Bt + y_0$$
$$z = Ct + z_0, \ -\infty < t < \infty$$

Letting $t = 0$ in these parametric equations shows that $(x_0, y_0, z_0)$ is a point on the line. After discussing vectors we will see that $(A, B, C)$ gives the direction of the line. Therefore lines in space are determined by two pieces of information: a point and a direction. We look at a couple examples that use parametric equations to determine some geometric properties of lines. ▲

**Example 1.4.6.** Lines and coordinate planes

Let $\ell$ be the line given by parametrically by $C(t) = (2t + 1, -t + 3, 4t - 2)$. Find the point of intersection of $\ell$ and the $yz$-plane.

First we find “when” $\ell$ intersects the $yz$-plane by solving $x(t) = 0$. In this case we solve $2t + 1 = 0$ to get $t = -1/2$. Now find the point on $\ell$ at that time by evaluating $C(t)$ at $t = -1/2$. The point where $\ell$ intersects the $yz$-plane is $C(-1/2) = (0, 3.5, -4)$. ▲

**Example 1.4.7.** Intersecting lines

Determine if the lines $C_1(t) = (3t, 2t - 1, -t + 2)$ and $C_2(t) = (t + 3, t, -t + 3)$ intersect.

If the lines intersect, the coordinate functions must all be equal. There is one subtlety here: if the lines intersect, it may not be at the same time. In other words, bugs flying along these two paths could visit the same point, but not hit each other because they arrive there at different times. To account for this, change one of the parameters to a different letter, say $C_2(s) = (s + 3, s, -s + 3)$. Now equating coordinate functions gives a system of three equations in two variables. A solution to the system corresponds to a point of intersection of the lines. Solving

$$3t = s + 3$$
$$2t - 1 = s$$
$$-t + 2 = -s + 3$$
gives \( t = 2, \ s = 3 \), which means the lines intersect.

Now substitute to find the point of intersection is \( C_1(2) = C_2(3) = (6, 3, 0) \).

\[ \begin{aligned}
\text{Parameterizing Surface Intersections:} & \quad \text{In general the intersection of two surfaces in } \mathbb{R}^3 \text{ will be a curve. In some cases there are effective strategies for parameterizing curves of intersection. We treat three such cases now: when one surface is a plane, when one surface is a generalized cylinder, and where the surfaces are given in cylindrical or spherical coordinates.}

\text{Example 1.4.8. \ Horizontal slices} \\
\text{Parameterize the curve of intersection of the elliptical paraboloid } z = x^2 + 4y^2 \text{ and the horizontal plane } z = 36.

\text{We know we want coordinate functions for the curve. Further, the } x- \text{ and } y-\text{coordinates of the curve of intersection satisfy the equation of the level curve } x^2 + 4y^2 = 36, \text{ which is an ellipse. Thus the } x- \text{ and } y-\text{coordinate functions parameterize the curve}

\[ \frac{x^2}{36} + \frac{y^2}{9} = 1, \]

\text{and are given by}

\[ \begin{aligned}
x &= 6 \cos t \\
y &= 3 \sin t, \quad 0 \leq t \leq 2\pi.
\end{aligned} \]

Since \( z \) is a constant, its coordinate function will be too, and we get the parameterization \( C(t) = (6 \cos t, 3 \sin t, 36), \quad 0 \leq t \leq 2\pi. \)

\text{Recall that a generalized cylinder is a surface whose defining Cartesian equation is missing a variable, like } x^2 + z^2 = 9. \text{ Given a second surface, which is the graph of a function } y = f(x, z), \text{ let } C \text{ be the curve of intersection with the}
cylinder. One can parameterize $C$ by first parameterizing the curve $x^2 + z^2 = 9$ in the $xz$-plane, then substituting the resulting coordinate functions in $f(x, z)$ to find $y(t)$. We illustrate this with some examples.

**Example 1.4.9. A cylinder and a plane**

Let $C$ be the curve of intersection of the cylinder $x^2 + z^2 = 9$ and the plane $x - y + 2z = 10$ as in Figure [1.4.5]. We can parameterize $C$ in two steps:

**Step 1:** Parameterize the curve $x^2 + z^2 = 9$ in the $xz$-plane.

**Step 2:** Use the equation $x - y + 2z = 10$ to find $y(t)$.

![Figure 1.4.5: Parameterizing surface intersections](image)

(a) Plane and Cylinder  
(b) Paraboloid and Generalized Cylinder

We now carry out the above steps. The curve $x^2 + z^2 = 9$ can be parameterized using coordinate functions $x(t) = 3 \cos t$, $z(t) = 3 \sin t$ for $0 \leq t \leq 2\pi$. To accomplish step 2, solve the plane equation for $y$, giving $y = x + 2z - 10$. Substituting we find the coordinate function $y(t) = 3 \cos t + 6 \sin t - 10$. Thus the curve $C$ is parameterized by

$$C(t) = (3 \cos t, 3 \cos t + 6 \sin t - 10, 3 \sin t), \text{ for } 0 \leq t \leq 2\pi. \square$$

**Example 1.4.10. Paraboloid and “washboard”**

We do a second example, employing the above technique. Parameterize the curve of intersection of the generalized cylinder $y = \cos x$ and the paraboloid $z = x^2 + y^2$.

The first step is to parameterize the graph of $y = \cos x$. Recall from Equation [1.4.1] that to parameterize the graph of a function, let the independent variable be the parameter and use the function to parameterized the dependent variable. In this case we get $x = t$, $y = \cos t$ for $-\infty < t < \infty$. This gives the coordinate functions for $x$ and $y$, and we use the equation of the surface $z = x^2 + y^2$ to find
the coordinate function for $z$. In this case $z = t^2 + \cos^2 t$, and we summarize:

$$C(t) = \left(t, \cos t, t^2 + \cos^2 t\right), \text{ for } -\infty < t < \infty. \uparrow$$

### Parameterizing Intersections of Cylinders and Surfaces

Let $g(x, y) = 0$ be a generalized cylinder in $\mathbb{R}^3$, $z = f(x, y)$, and $C$ be the intersection of the two surfaces. To parameterize $C$:

1. Parameterize the curve $g(x, y) = 0$ in the $xy$-plane. This gives coordinate functions $x(t)$ and $y(t)$.

2. Substitute the coordinate functions from step 1 into the function $f(x, y)$ to find the coordinate function $z(t)$.

Note that there is nothing special about the cylinder missing the variable $z$. The same strategy works for cylinders $g(x, z) = 0$ and surfaces $y = f(x, z)$, etc.

#### Concept Connection

In some applications the direction of a curve is important. In particular, both Green’s Theorem and Stokes’ Theorem require a certain orientation (or direction) on a curve. The reason for this is many physical quantities have a direction associated with them (velocity, force, etc.), and going the wrong direction gives the negative of the answer you want! We do one example where the direction of the parameterization is specified.

**Example 1.4.11. Paraboloid and cylinder–clockwise orientation**

Parameterize the curve of intersection of the paraboloid $z = -x^2 + 6x - y^2 - 8y$ and the cylinder $x^2 + y^2 = 25$, so that it is oriented clockwise when viewed from above.

The added feature of this example is the requirement that the parameterization be oriented clockwise. To do so, note that when viewing “from above” the $z$-coordinate disappears and we are looking at what’s called the projection in the $xy$-plane. Thus, the “clockwise from above” requirement just means that our parameterization of the curve $x^2 + y^2 = 25$ in the $xy$-plane be clockwise. The standard parameterization of the circle is counterclockwise, but recall that negating the $y$-coordinate function results in a clockwise parameterization.

Thus step 1 of the process gives $x(t) = 5\cos t$ and $y(t) = -5\sin t$ for $0 \leq t \leq 2\pi$. Substituting these equations in the formula for the paraboloid gives

$$z(t) = -(5\cos t)^2 + 30\cos t - (-5\sin t)^2 + 40\sin t$$
$$= 30\cos t + 40\sin t - 25.$$

Thus $C(t) = (5\cos t, -5\sin t, 30\cos t + 40\sin t - 25)$, for $0 \leq t \leq 2\pi$, is a clockwise parameterization of the curve. \uparrow
We have seen several examples of parameterizing curves of intersection when one of the surfaces is a generalized cylinder. Surfaces given by equations in spherical or cylindrical coordinates can also intersect in curves that are relatively easy to parameterize. The general strategy is to let the independent variable be the parameter and use the conversions between coordinate systems to find coordinate functions.

Example 1.4.12. Intersecting Constant Coordinate Surfaces

Parameterize the curve of intersection of the sphere $\rho = 5$ and the half-plane $\theta = \pi/3$.

Parametric curves are given by coordinate functions of a single variable, and there is nothing special about using the letter $t$ for the parameter. Thus substituting the given values for $\rho$ and $\theta$ into the conversions gives coordinate functions of the single variable $\phi$, which is a parameterization! More explicitly,

$$C(\phi) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$$

$$= (5 \sin \phi \cos(\pi/3), 5 \sin \phi \sin(\pi/3), 5 \cos \phi)$$

$$= \left(\frac{5}{2} \sin \phi, \frac{5\sqrt{3}}{2} \sin \phi, 5 \cos \phi\right), \ 0 \leq \phi \leq \pi.$$

If you aren’t comfortable using $\phi$ as a parameter, you can replace it with $t$ for now. However, you do want to recognize that any letter can be used as a parameter. ▲

Example 1.4.13. An example in cylindrical coordinates

Parameterize the curve of intersection of $z = r^2$ and $r = \theta$, for $\theta \geq 0$.

Analytically, this is easy to do. Let the independent variable, $r$, be the parameter. The given equations express all the cylindrical coordinates $(r, \theta, z)$ in
terms of \( r \), so that \((r, \theta, z) = (r, r, r^2)\). Now substitute these values into the coordinate conversions to get the coordinate functions \(x(r), y(r),\) and \(z(r)\). Since the result is coordinate functions of one variable, this is a parameterization.

\[
C(r) = (r \cos \theta, r \sin \theta, z) = \left( r \cos r, r \sin r, r^2 \right), \quad 0 \leq r
\]

(a) \( z = r^2 \) and \( \theta = r \) \hspace{1cm} (b) The Curve of Intersection

Figure 1.4.7: Cylindrical Coordinate Surfaces

It’s a little more difficult to see what is going on geometrically. First, the surface \( z = r^2 \) has Cartesian equation \( z = x^2 + y^2 \), so is our favorite paraboloid. Second, the curve \( \theta = r \) in the \( xy \)-plane is a spiral since as \( \theta \) increases the distance to the origin does too. Since \( z \) is missing from the equation, the surface is obtained by translating the spiral up and down along the \( z \)-axis. The resulting surface is a spiral cookie cutter, and the curve of intersection is pictured in Figure 1.4.7.

Parameterizing Curves Using Other Coordinate Systems

To parameterize the intersection of two surfaces given in cylindrical or spherical coordinates:

1. Parameterize the given coordinates (i.e. get them all in terms of one variable).
2. Use conversions to Cartesian coordinates to find the coordinate functions.

We conclude our introduction to parametric curves with the classic example of a helix, describing knots, and determining when a curve lies on a surface. These examples follow.
Example 1.4.14. Helicies

Everyone who has seen parametric curves has seen the helix, and then shown it to their brother because it was cool. The parametric equations for a helix are \( C(t) = (\cos t, \sin t, t) \), for \(-\infty < t < \infty\). Note that if you project into the \( xy \)-plane (i.e. cover up the \( z \)-coordinate), you get parametric equations for the unit circle in the plane. Thus as \( t \) increases the bug flying along \( C(t) \) keeps circling the \( z \)-axis in a counterclockwise fashion when viewed from above. The third coordinate function \( z(t) = t \) means that the bug is increasing in altitude all the while.

![Figure 1.4.8: A helix and Stevedore's knot](image)

More general helicies have parametric equations \( C(t) = (a \cos t, a \sin t, bt) \) for constants \( a \) and \( b \). These all live on the cylinders given by \( x^2 + y^2 = a^2 \). How might you parameterize a helix that lies on the elliptical cylinder \( \frac{x^2}{4} + y^2 = 1 \)?

Example 1.4.15. Knots

Knots are curves that don’t intersect themselves, and end where they started. Lissajous knots are knots with particularly easy parameterizations, they are always of the form:

\[
\begin{align*}
x &= \cos(At + x_0) \\
y &= \cos(Bt + y_0) \\
z &= \cos(Ct + z_0), \quad 0 < t < 2\pi.
\end{align*}
\]

The Lissajous knot \( C(t) = (\cos(3t + 1.5), \cos(2t + .2), \cos(5t)) \) is called a Stevedore knot and is pictured in Figure 1.4.8. You can change the constants in a Lissajous parameterization and try to get different knots.

Example 1.4.16. Analytically verifying curves on surfaces
To show that the parametric line \( C(t) = (1, t, t) \) lies on the one-sheeted hyperboloid \( x^2 + y^2 - z^2 = 1 \), just substitute the parametric equations for the line in for the Cartesian coordinates and verify that the equation is satisfied. In this case we see \( x^2 + y^2 - z^2 = 1 + t^2 - t^2 = 1 \) satisfies the equation. To see that the line \( C(t) = (t + 1, 0, -t) \) doesn’t lie on the surface, substitute and simplify to get \( x^2 + y^2 - z^2 = (1 + t)^2 + 0^2 - (-t)^2 = 1 + 2t + t^2 - t^2 = 1 + 2t \). Since the result is not identically 1, this does not lie on the surface. In fact, the only value of \( t \) which makes this one is \( t = 0 \). This means that the only point where the line intersects the surface is \( C(0) = (1, 0, 0) \). These lines and the surface are pictured in Figure 1.4.9.

![Figure 1.4.9: Curves on Surfaces](image)

To show that the curve \( C(t) = (\sin t \cos t, \sin^2 t, \cos t) \) lies on the unit sphere \( x^2 + y^2 + z^2 = 1 \), we substitute and simplify.

\[
\begin{align*}
  x^2 + y^2 + z^2 &= (\sin t \cos t)^2 + \sin^4 t + \cos^2 t \\
  &= \sin^2 t (\cos^2 t + \sin^2 t) + \cos^2 t \\
  &= \sin^2 t + \cos^2 t = 1.
\end{align*}
\]

**Things to know/Skills to have**

In this section we studied parametric curves in \( \mathbb{R}^2 \) and \( \mathbb{R}^3 \). You know you understand the material of this section when you are able to:

- Parameterize ellipses clockwise and counterclockwise.
- Parameterize graphs of \( y = f(x) \).
1.4. **PARAMETRIC CURVES**

- Parameterize graphs of polar equations \( r = f(\theta) \).
- Parameterize lines in \( \mathbb{R}^3 \).
- Parameterize curves of intersection of surfaces whether they are planes, generalized cylinders, or surfaces in other coordinate systems.
- Determine if a parametric curve lies on a given surface.

**Exercises**

1. Parameterize the circle \( x^2 + y^2 = 9 \) clockwise. Now parameterize it counterclockwise.

2. Parameterize the circle \((x - 2)^2 + y^2 = 4\). (Hint: how should you change the parameterization of \( x^2 + y^2 = 4 \) to get the desired result?)

3. Parameterize the circle \((x + 3)^2 + (y - 1)^2 = 25\).

4. Find parametric equations for the circle in the \( xy \)-plane with center \((h, k)\) and radius \( r \). (i.e. the circle \((x - h)^2 + (y - k)^2 = r^2\)).

5. Parameterize the ellipse \( \frac{x^2}{16} + \frac{y^2}{9} = 1 \) counterclockwise.

6. Parameterize the ellipse \( 4x^2 + 9y^2 = 1 \) clockwise.

7. Parameterize the ellipse \( 9x^2 + z^2 = 1 \) in the \( xz \)-plane.

8. Parameterize the curve in \( \mathbb{R}^2 \) given by the function \( y = x^4 + 8x \).

9. Parameterize the curve in \( \mathbb{R}^2 \) given by the function \( y = \tan^{-1} x \).

10. Parameterize \( y = \frac{x + 1}{x^2 + 4} \).

11. Parameterize the graph of \( y = \sqrt{1 - x^2} \), be sure to include limits on your parameter.

12. Parameterize one component of the hyperbola \( y^2 - x^2 = 1 \).

13. Parameterize one component of \( x^2 - y^2 = 4 \).

14. Parameterize the left component of \( \frac{x^2}{4} - \frac{y^2}{9} = 1 \).

15. Find parametric equations for the Cardiod \( r = 1 + \cos \theta \).

16. Parameterize the four-leaf rose \( r = \cos 2\theta \). Now parameterize just the petal pointing straight up.

17. Parameterize the spiral \( r = \frac{\theta}{\pi} \), for \( \theta \geq 0 \). Where does this curve intersect the positive \( x \)-axis? The negative \( x \)-axis?
18. Parameterize the infinite spiral \( r = e^\theta \).

19. Where do the following lines intersect the coordinate planes?

(a) \( \mathbf{C}(t) = (2t + 4, -t - 5, 3t + 1) \)

(b) \( \mathbf{C}(t) = (2t, -4t, 7t) \)

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

20. Determine if the lines \( \mathbf{C}_1(t) = (t + 1, -2t, t - 2) \) and \( \mathbf{C}_2(t) = (t, 2 - t, t - 3) \) intersect.

21. Determine if the lines \( \mathbf{C}_1(t) = (3t, 2t + 1, 2 - t) \) and \( \mathbf{C}_2(t) = (t + 2, 2 - t, 2t - 1) \) intersect.

22. Without doing any algebra, determine whether or not the lines \( \mathbf{C}_1(t) = (4t, 3, -t + 1) \) and \( \mathbf{C}_2(t) = (1 + t, -1, t + 2) \) intersect.

23. At what point does the line \( \mathbf{C}(t) = (4t + 1, 2t - 3, t + 4) \) intersect the plane \( x - y + 2z = 8 \)?

24. At what point does the helix \( \mathbf{C}(t) = (\cos t, \sin t, t) \) intersect the paraboloid \( z = x^2 + y^2 \)?

25. Parameterize the curve of intersection of the cylinder \( x^2 + y^2 = 4 \) and the plane \( 2x + y - 6z = 6 \) clockwise when viewed from above.

26. Parameterize the curve of intersection of the cylinder \( r = 4 \) and the paraboloid \( z = 4 - x^2 - 4x - y^2 \) clockwise when viewed from above.

27. Parameterize the intersection of the generalized cylinder \( z = x^2 \) and the plane \( 2x - y + 3z = 4 \).

28. Parameterize the intersection of the generalized cylinder \( 9y^2 + z^2 = 36 \) and the surface \( x = y^2 + z^2 + 2y \).

29. Parameterize the curve of intersection of the generalized cylinder \( x = \sin z \) and the surface \( y = x^2 - z^2 \).

30. Parameterize the curve of intersection of the plane \( z = 6 \) and the half-plane \( \theta = \pi/6 \). Careful with the limits on your parameterization.

31. Parameterize the curve of intersection of the cylinder \( r = 3 \) and the half-plane \( \theta = -\pi/4 \).

32. Parameterize the curve of intersection of the plane \( z = -3 \) and the cylinder \( r = 2 \), clockwise when viewed from above.

33. Parameterize, counterclockwise when viewed from above, the curve of intersection of the cone \( \phi = \pi/3 \) and the sphere \( \rho = 2 \).

34. Parameterize the curve of intersection of the sphere \( \rho = 3 \) and the half-plane \( \theta = 2\pi/3 \).
35. Parameterize the curve of intersection of the cone $\phi = \pi/4$ and the half-plane $\theta = 5\pi/4$.

36. Show that the curve $C(t) = (t \cos t, t \sin t, t)$, for $t \geq 0$, lies on the cone $z = \sqrt{x^2 + y^2}$.

37. Show that the curve $C(t) = (t \cos t, t^2, t \sin t)$, for $t \geq 0$, lies on the paraboloid $y = x^2 + z^2$. 
1.5 Parametric Surfaces

A curve in space is one-dimensional, so you need one parameter to describe it. A surface is two dimensional, so the parametric equations will be functions of two variables: \( S(s, t) = (x(s, t), y(s, t), z(s, t)) \). More formally we have

**Definition 1.5.1.** A parametric surface is one defined by two-variable coordinate functions \( S(s, t) = (x(s, t), y(s, t), z(s, t)) \), together with limits on the parameters \( a \leq s \leq b, \ c \leq t \leq d \).

It should be noted that the individual coordinate functions could be constant or have just one variable, as long as the coordinate functions combined involve two variables. Thus \( S(s, t) = (s^3 - 2t, t^2 + 1, 5) \), \( 0 \leq s, t \leq 1 \) is a parametric surface, but \( C(s, t) = (t^3 - 2t, \cos t, t^2) \) and \( x(s, t) = (t^2, s^2u^3, t - s + u) \) aren’t (one has too few variables, the other too many).

The “bug flying” interpretation that works for parametric curves doesn’t work so well for parametric surfaces. The bending wire interpretation, however, can be generalized nicely to this setting. Instead of wire, imagine starting with a sheet of aluminum foil. The parametric equations are then instructions for how to bend it into space (foil doesn’t stretch, so maybe elastic fabric is a better choice). More precisely, the domain of a parametric surface \( S(s, t) \) is a subset of the \( st \)-plane, while the range lives in \( \mathbb{R}^3 \). In this section we give several examples and strategies for parameterizing surfaces.

**Example 1.5.1. Planes**

As parametric equations for lines are linear functions in \( t \), parametric equations for planes are linear functions in two variables. They have the form:

\[
\begin{align*}
x &= As + Dt + x_0 \\
y &= Bs + Et + y_0 \\
z &= Cs + Ft + z_0, \quad -\infty < s, t < \infty
\end{align*}
\]

Letting \( s = t = 0 \) shows that \((x_0, y_0, z_0)\) is a point on the plane. We’ll see in Chapter 2 that the coefficients of \( s \) determine a direction in the plane, as do the coefficients of \( t \).

**Constant Coordinate Surfaces** Parameterizing constant coordinate surfaces is straightforward. An immediate substitution parameterizes the plane \( y = -3 \):

\[
S(x, z) = (x, -3, z), \quad -\infty < x, z < \infty.
\]

Parameterizations of the planes \( x = c \) and \( z = d \) are obtained similarly.

We will use parameters \( s \) and \( t \) for many parameterizations in this book, however that is not necessary as long as there are two variables. In this case it’s easier, and more illuminating, to use the variables \( x \) and \( z \) as parameters.
To parameterize constant coordinate surfaces in other coordinate systems, you again substitute the given constant. Now, however, you substitute it in the relevant coordinate system. For example, to parameterize the sphere \( \rho = 5 \) you get:

\[
S(\theta, \phi) = (5 \sin \phi \cos \theta, 5 \sin \phi \sin \theta, 5 \cos \phi), \quad 0 \leq \phi \leq \pi, \quad 0 \leq \theta \leq 2\pi.
\]

We point out that when \( \phi = 0 \) we have \( S(\theta, 0) = (0, 0, 5) \) for all \( \theta \). Similarly, \( S(\theta, \pi) = (0, 0, -5) \) for all \( \theta \). The fact that many points in the domain get mapped to the same point in the image is not desirable. In this case it makes little difference because they are isolated points on the sphere, but it is something to be aware of when determining limits of a parameterization. For example, if we let \( \phi \) range from 0 to \( 2\pi \) every point on the sphere would be hit twice since \( S(\theta, \phi) = S(\theta + \pi, 2\pi - \phi) \).

Parameterizing Constant Coordinate Surfaces

To parameterize a constant coordinate surface, substitute the constant into the conversion to Cartesian coordinates.

One last example will round out this paragraph: A parameterization of \( r = 7 \) is

\[
S(s, t) = (7 \cos t, 7 \sin t, s), \quad 0 \leq t \leq 2\pi, \quad -\infty < s < \infty.
\]

Parameterizing Graphs: In Section 1.2 we saw that the graph of a function of two variables is a surface in \( \mathbb{R}^3 \). Parametric equations for the surface \( z = f(x, y) \) are obtained in the same way that we parameterize the curve \( y = f(x) \) in the plane. Simply let the dependent variables be the parameters, and use the function \( f(x, y) \) to find the parameterization for \( z \).

Example 1.5.2. Parameterizing a Saddle

Using the above strategy we see that one way to parametrize the saddle \( z = x^2 - y^2 \) is \( S(s, t) = (s, t, s^2 - t^2) \). Again, this seems artificial, but represents thinking of the domain and surface as living in separate spaces. The idea is illustrated in Figure 1.5.1 for the restrictions \(-2 \leq s, t \leq 2\).

We note that the parametric equations for \( S(s, t) \) map curves in the domain to curves on the surface. For example, letting \( t = -1 \) in the domain gives a horizontal line in the \( st \)-plane which the parameterization \( S(s, t) \) maps to the thickened curve on the surface given parametrically by \( S(s, -1) = (s, -1, s^2 - 1) \).

Parameterizing Graphs

To parameterize the graph of \( z = f(x, y) \) let the dependent variables be the parameters and use the function to parameterize the independent variable.
Example 1.5.3. Using different coordinates

This approach to parameterizing surfaces works for functions of any two Cartesian coordinates, not just $x$ and $y$. We illustrate by parameterizing the surface $y = \sin(x^2 - 2x + z^2)$.

In this case, the independent variables are $x$ and $z$ so they become the parameters, and we have $S(x, z) = (x, \sin(x^2 - 2x + z^2), z)$ for $-\infty < x, z < \infty$. We also remark that the parameters can be any variables we choose, so long as the coordinate functions use two variables combined, thus we can leave them as $x$ and $z$. The surface is pictured in Figure 1.5.2 for $-2 \leq x, z \leq 2$. ▲

Math App 1.5.1. A parametric paraboloid
In this math app you view the paraboloid \( z = x^2 + y^2 \) parameterized by
\[
S(s, t) = (s \cos t, s \sin t, s^2), \quad 0 \leq s \leq 2, \quad 0 \leq t \leq 2\pi.
\]

Direct substitution verifies that the parametric equations satisfy the Cartesian equation:
\[
x^2 + y^2 = (s \cos t)^2 + (s \sin t)^2 \quad \text{substitution}
= s^2 \left( \cos^2 t + \sin^2 t \right) = s^2 \quad \text{algebra}
= z. \quad \text{un-substitution}
\]

To develop an intuition for how the parametric equations bend, twist and stretch the rectangle \( 0 \leq s \leq 2, \quad 0 \leq t \leq 2\pi \) onto the paraboloid \( z = x^2 + y^2 \), click on the hyperlink below.

**Concept Connection:** It turns out that thinking of graphs parametrically will be useful when using calculus to study surfaces. In particular, any formula we derive for parametric surfaces will yield a formula for the graph \( z = f(x, y) \) by thinking of it as the parametric surface \( S(x, y) = (x, y, f(x, y)) \). Such formulas include finding tangent planes to surfaces, determining surface area, and much more. This becomes a powerful tool for remembering the myriad of formulas we encounter in this course!

**Parameterizing Generalized Cylinders:** Recall that generalized cylinders are surfaces whose Cartesian equations are missing a variable. We can think of them as a planar curve that gets translated back and forth along a perpendicular axis. The only way you know it’s a surface, rather than a curve, is by the context. To parameterize a cylinder, first parameterize the planar curve using a single parameter, then let the variable that’s missing be the second parameter. We illustrate the process with an example.

**Example 1.5.4. Parametric Generalized Cylinders**

Parameterize the cylinder \( z = x^2 - 3x - 10 \). First we parameterize the parabola in the \( xz \)-plane. Since \( z \) is a function of \( x \) in this case, we let \( x = s \) and \( z = s^2 - 3s - 10 \). To finish parameterizing the surface let \( y = t \), yielding the parametric surface \( S(s, t) = (s, t, s^2 - 3s - 10) \) for \( -\infty < s, t < \infty \).

Similarly, to parameterize the cylinder \( y^2 + z^2 = 1 \) start by parameterizing the unit circle in the \( yz \)-plane then let \( x \) be the second parameter. One parameterization is \( S(s, t) = (s \cos t, \sin t), \) with \( 0 \leq t \leq 2\pi, \quad -\infty < s < \infty \). ▲
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

Parameterizing Generalized Cylinders

To parameterize a generalized cylinder, parameterize the curve given by the equation and let the remaining variable be the second parameter.

Other Coordinate Systems

We have already seen that one can use the change from polar or spherical to Cartesian coordinates to parameterize constant coordinate surfaces. Similar techniques can be used to parameterize surfaces given by equations in other variables. We illustrate with several examples.

Example 1.5.5. A corkscrew

The surface given by the cylindrical equation \( z = \theta \) can be parameterized by replacing \( \theta \) with \( z \) in the change of coordinates formula to get \( \mathbf{S}(r, z) = (r \cos z, r \sin z, z) \). This surface looks better if we allow \( r \) to be negative, and is plotted in Figure 1.5.4(a) for \(-2 \leq r \leq 2 \) and \(-3 \leq z \leq 3\). ▲

Parameterizing Surfaces: Equations in other coordinate systems

Given an equation in cylindrical or spherical coordinates:

1. Use the equation to express one of the coordinates in terms of the other two, then

2. Use the conversion to Cartesian to parameterize the surface.

Example 1.5.6. Parametric Peach

The surface \( \rho = \phi \), for \( 0 \leq \phi \leq \pi \), looks somewhat like an apricot or peach. Using the equation we can replace \( \phi \) with \( \rho \), and let \( \theta \) be the other parameter. We get the parameterization

\[ \mathbf{S}(s, t) = (\phi \sin \phi \cos \theta, \phi \sin \phi \sin \theta, \phi \cos \phi), \quad 0 \leq \phi \leq \pi, \; 0 \leq \theta \leq 2\pi. \]
Recall that \( x = \rho \sin \phi \cos \theta, \ y = \rho \sin \phi \sin \theta, \) and \( z = \rho \cos \phi \). In Figure 1.5.4 we plotted the peach for \( 0 \leq \theta \leq 3\pi/2 \) to highlight the inside (and \( 0 \leq \phi \leq \pi \)).

**Surfaces of Revolution:** A surface of revolution is obtained by revolving a curve around an axis. It turns out that such surfaces have nice parameterizations. The first step is to parameterize the curve that is being revolved, which requires one parameter. The second parameter is used to carry out the revolution. Below is a procedure that works for the case where the \( z \)-axis is the axis of revolution.

**Parameterizing Surfaces of Revolution**

Let \( S \) be a surface of revolution about the \( z \)-axis. To parameterize \( S \):

1. Parameterize the curve \( C \) of \( S \) in the half-plane \( \theta = 0 \) by \((x(s), 0, z(s))\) for \( a \leq s \leq b \).

2. Parametric equations for \( S \) are

\[
S(s, t) = (x(s) \cos t, x(s) \sin t, z(s)), \ a \leq s \leq b, \ 0 \leq \theta \leq 2\pi.
\]
rotating $P$ through any angle gives another point on it, and the point $(r_p, t, z_p)$ is on $S$ for all values $0 \leq t \leq 2\pi$. Thus we let $\theta = t$ be one of the parameters.

To parameterize $r$ and $z$, note that the half-plane $\theta = 0$ intersects $S$ in a curve $C$. Since $C$ lies in $\theta = 0$, the $y$-coordinate of points on $C$ is 0 and one can use curve techniques to find coordinate functions $x(s)$ and $z(s)$. This obviously gives us $z(s)$, but we also get $r(s)$ for free! Recall that the cylindrical coordinate $r$ is the distance to the $z$-axis. In the half-plane $\theta = 0$ the $x$-coordinate gives the distance to the $z$-axis: so in this half-plane $r = x$. Therefore $r = x(s)$ is the desired parameterization for $r$. Now that we have parametric equations for $r$, $\theta$, and $z$, using coordinate conversion parameterizes the surface, as in step 2 above. This discussion justifies the above procedure.

We briefly note that this technique is more general than we’ve described. For example, if the half-plane $\theta = \pi/2$ is used in step one, the only difference is that the $y$-coordinate gives $r$. More precisely, suppose $C(s) = (0, y(s), z(s)), a \leq s \leq b$, parametrizes a curve in the $yz$-plane. One parameterization of the surface obtained by revolving $C(s)$ around the $z$-axis is

$$S(s, t) = (y(s) \cos t, y(s) \sin t, z(s)), a \leq s \leq b, 0 \leq \theta \leq 2\pi.$$ 

Additionally, if the $x$- or $y$-axes are used for the revolution, one uses similar techniques with “cylindrical” coordinates defined appropriately.

**Example 1.5.7. Revolving the hyperbolic cosine curve**

The hyperbolic cosine function is the even function defined by $\cosh x = (e^x + e^{-x})/2$. To get a surface, rotate the curve $z = \cosh x$, for $x \geq 0$, about the $z$-axis.

Our goal is to find parametric equations for this surface, and a first step is parameterizing the curve. To parameterize the function, let the independent variable be the parameter and use the function to parameterize the dependent
variable. Thus the given curve has parametric equations $(x, y, z) = (s, 0, \cosh s)$, for $s \geq 0$, and is the darkened curve in Figure 1.5.6(a). Applying step 2 of the process gives the parameterization

$$S(s, t) = (s \cos t, s \sin t, \cosh s), \quad s \geq 0, \quad 0 \leq t \leq 2\pi.$$ 

\[ \blacktriangle \]

(a) Hyperbolic Cosine Revolved

(b) A torus

Figure 1.5.6: Surfaces of Revolution

More generally, to parameterize a surface of revolution you start by parameterizing the curve. Then use a second parameter with sines and cosines to rotate about the given axis. We illustrate with a second example, the torus.

**Example 1.5.8. Parameterizing a Torus**

The curve $z^2 + (y - 2)^2 = 1$ is a circle in the $yz$-plane with center $(2, 0)$ and radius 1 (note that we are not thinking of this as a “generalized cylinder” here). It has parametric equations $y = \cos s + 2, z = \sin s$, for $0 \leq s \leq 2\pi$. To rotate it around the $z$-axis, use cylindrical coordinates as above with $r = \cos s + 2, \theta = t$, and $z = \sin s$. The change to Cartesian coordinates gives the parametric surface

$$S(s, t) = ((\cos s + 2) \cos t, (\cos s + 2) \sin t, \sin s), \quad 0 \leq s, t \leq 2\pi.$$ 

\[ \blacktriangle \]

The torus is interesting enough that we’ll take a closer look at it. The grid lines in Figure 1.5.6(b) come about by fixing one of the parameters, and letting the other run. In the case of the torus, the grid lines that go through the hole are called meridians and those going around the hole are longitudes. To test your understanding, determine whether the grid line $s = \pi$ is a meridian or longitude!

The torus in Figure 1.5.6(b) is the image of the square $0 \leq s, t \leq 2\pi$ in the $st$-plane. Why do the parametric equations imply that the left and right sides of the square map to the same grid line, where the left side is $s = 0$, and the right is $s = 2\pi$? They also show that the top and bottom of the square map
to the same longitude. You can get more interesting curves by letting \( t = 2u \), \( s = 3u \) for a new parameter \( u \) with \( 0 \leq u \leq 2\pi \). The result is a curve that goes twice around the hole and three times through it. This is equivalent to considering the image of the line segments below, each of which has slope \( 3/2 \) and the ends get glued when identifying top and bottom, and left and right. In fact, you can choose any two relatively prime integers \( p, q \) (not just 2 and 3), and do the same thing. The resulting curve is called a \( pq \)-torus knot because it lives on the surface of a torus.

Parameterizing Surfaces with Restricted Domain: In many applications we will be concerned with only portions of surfaces. For example, we may want to consider the portion of a plane that lies over a disk in the \( xy \)-plane, or that part of a paraboloid over a rectangle. The restrictions on the domain usually change the limits of your parameterization, leaving the coordinate functions the same. To find the limits on the parameters, one needs to describe regions in the plane using systems of inequalities. This topic is discussed more fully in section 1.6, but we illustrate here with a few examples.

**Example 1.5.9. Portions of a plane**

Parameterize the portion of the plane \( 3x - 2y - 4z = -8 \) over the rectangle \( 0 \leq x \leq 4, \ -2 \leq y \leq 3 \).

To parameterize the plane, we can solve for \( z \) and think of it as the graph of \( z = (3x - 2y + 8)/4 \). Letting \( x \) and \( y \) be parameters, and using the function to find the third coordinate function, we get the parameterization of the entire plane:

\[
S(x, y) = \left( x, y, \frac{3x - 2y + 8}{4} \right), \quad -\infty < x, y < \infty.
\]
1.5. PARAMETRIC SURFACES

The restriction to the given rectangle changes the limits of the parameterization, but not the coordinate functions, so we have (see Figure 1.5.8(a))

\[ S(x, y) = \left( x, y, \frac{3x - 2y + 8}{4} \right), \quad 0 \leq x \leq 4, \quad -2 \leq y \leq 3. \]

Figure 1.5.8: Restricted domain parameterization

Now suppose we want to parameterize the portion of the same plane that lies above the disk radius 3 in the \( xy \)-plane (i.e., all points satisfying \( x^2 + y^2 \leq 9 \)). One way to do this is to describe the disk using a system of inequalities involving \( x \) and \( y \). For example, the \( y \)-coordinate is always between the top and bottom semicircles, so \(-\sqrt{9-x^2} \leq y \leq \sqrt{9-x^2}\). Since \( x \) is between \( \pm 3 \), one parameterization is

\[ S(x, y) = \left( x, y, \frac{3x - 2y + 8}{4} \right), \]

with the limits on the parameterization being

\[-3 \leq x \leq 3, \quad -\sqrt{9-x^2} \leq y \leq \sqrt{9-x^2}.\]

The disadvantage to this parameterization is the limits on \( y \) are ugly (see Figure 1.5.8(b)). We can use polar coordinates instead to make the limits prettier.

In polar coordinates, the disk is described by the system of inequalities

\[ 0 \leq r \leq 3, \quad 0 \leq \theta \leq 2\pi. \]

If we use \( r \) and \( \theta \) as our parameters, the limits will be nice. To do so, we just use the conversions to Cartesian coordinates \( x = r \cos \theta \), \( y = r \sin \theta \). Making the substitution we get the alternate parameterization (see Figure 1.5.8(c))

\[ S(r, \theta) = \left( r \cos \theta, r \sin \theta, \frac{3r \cos \theta - 2r \sin \theta + 8}{4} \right), \]

with limits

\[ 0 \leq r \leq 3, \quad 0 \leq \theta \leq 2\pi. \]
Take a moment to compare the grid lines in Figures 1.5.8(b) and 1.5.8(c). Can you describe why they are different? Which is more aesthetically pleasing?

Example 1.5.10. Parameterizing a Paraboloid Portion

Parameterize the portion of the paraboloid \( z = x^2 + y^2 - 2x \) lying above the square \(-1 \leq x, y \leq 1\), and above the unit disk \( x^2 + y^2 \leq 1\). The different parameterizations are pictured in Figure 1.5.9.

As in the previous example we use Cartesian coordinates to parameterize the paraboloid over the square, and polar over the disk. Over the square we have

\[
S(x, y) = (x, y, x^2 + y^2 - 2x), \quad -1 \leq x, y \leq 1.
\]

Substituting coordinate conversions, and using appropriate limits, gives

\[
S(s, t) = (s \cos t, s \sin t, s^2 \cos^2 t + s^2 \sin^2 t - 2s \cos t)
= (s \cos t, s \sin t, 1 - 2 \cos t), \quad 0 \leq s \leq 1, \quad 0 \leq t \leq 2\pi.
\]

Things to Know/Skills to have

- A parametric surface is one given by 2-variable coordinate functions

\[
S(s, t) = (x(s, t), y(s, t), z(s, t)), \quad a \leq s \leq b, \quad c \leq t \leq d
\]

- **Constant coordinate surfaces** are parameterized using conversions to Cartesian coordinates.

- One parameterization of the graph \( z = f(x, y) \) is obtained by letting the independent variables be the parameters and using the function to parameterize the dependent variable.
1.5. PARAMETRIC SURFACES

- Parameterize **generalized cylinders** by parameterizing the corresponding curve with one parameter, and letting the other parameter be the missing variable.

- To parameterize surfaces given in **other coordinate systems**, first parameterize the given coordinates, then convert to Cartesian using coordinate conversions.

- To parameterize a **surface of revolution**: first parameterize the curve in a plane, then use rotational symmetry to finish the parameterization.

- Parameterize surfaces with a restricted domain.

**Exercises**

1. Parameterize the following constant-coordinate surfaces. Be sure to include limits on your parameters.
   
   (a) The cylinder $r = 3$.
   (b) The half-plane $\theta = 5\pi/6$.
   (c) The sphere $\rho = 5$.
   (d) The cone $\phi = \pi/4$.

2. Parameterize the ellipsoid $\frac{x^2}{4} + \frac{y^2}{2} + \frac{z^2}{9} = 1$. Hint: stretch the parameterization of the unit sphere, in the same way ellipses are stretched circles.

3. Parameterize the following generalized cylinders. Be sure to include limits on your parameters.
   
   (a) $z = e^y$.
   (b) $y = \sin x$.
   (c) $y^2 + z^2 = 1$.

4. Parameterize the surface obtained by revolving the line $z = 2x + 1$ around the $z$-axis.

5. Parameterize the surface obtained by revolving the curve $z = e^{y^2}$ for $y \geq 0$ about the $z$-axis.

6. Parameterize the torus obtained by revolving the circle $(y - 3)^2 + z^2 = 4$ around the $z$-axis.

7. Parameterize the one-sheeted hyperboloid $x^2 + y^2 - z^2 = 1$, by thinking of it as one component of the curve $x^2 - z^2 = 1$ revolved around the $z$-axis.

8. Parameterize the surface given by the cylindrical equation $z = \cos r$. Sketch it using computer software.
9. Parameterize the surface given by $\rho = 5 + \cos 10\phi$ for $0 \leq \phi \leq \pi$. Sketch it using computer software.

10. Parameterize the surface given by the spherical equation $\theta = 2\phi$. Sketch it using computer software.

11. Parameterize the portion of the plane $3x + 2y + 5z = 30$ above the disk $x^2 + y^2 \leq 9$.

12. Parameterize the portion of the paraboloid $z = x^2 + 4x + y^2$ that lies above the unit square with vertices $(0,0)$, $(1,0)$, $(1,1)$, and $(0,1)$.

13. Parameterize the portion of the plane $x - 2y - 3z = 4$ that lies inside the cylinder $r = 2$.

14. Parameterize the portion of the sphere $x^2 + y^2 + z^2 = 9$ that lies above the unit disc.

15. Parameterize the top portion of the sphere $\rho = 2$ that lies inside the cone $z = r$. 
1.6  Describing Regions

In this section we develop the skill of describing regions using systems of inequalities, which will be necessary for setting up limits of integration in multiple integrals.

Example 1.6.1. A region in the plane

Find a system of inequalities that describes the region \( R \) in the first quadrant under \( y = 4 - x^2 \). There are two approaches to this problem, depending on how you view the region, and we illustrate both.

Thinking of the \( x \)-axis as the bottom of the region and the parabola as the top, we see that the \( y \)-coordinate of any point in the region satisfies \( 0 \leq y \leq 4 - x^2 \). After making this observation, look at the shadow of the region in the \( x \)-axis to find limits on \( x \). The restrictions on \( y \) are valid for \( 0 \leq x \leq 2 \). Thus \( R \) can be described by the system of inequalities:

\[
\begin{align*}
0 &\leq y \leq 4 - x^2 \\
0 &\leq x \leq 2
\end{align*}
\]

\[\text{Figure 1.6.1: Parabolic Region}\]

The region \( R \) can also be thought of as being delimited by the \( y \)-axis on the left and the parabola on its right. This point of view restricts the \( x \)-coordinate of points in \( R \). To find the system of inequalities on \( x \) we have to solve the parabola’s equation for \( x \), giving \( x = \sqrt{4 - y} \). Now consider the shadow of the region on the \( y \)-axis to find restrictions on \( y \). We end up with the system

\[
\begin{align*}
0 &\leq x \leq \sqrt{4 - y} \\
0 &\leq y \leq 4.
\end{align*}
\]

\[\text{Figure 1.6.1: Parabolic Region}\]

This technique generalizes. If \( R \) is a region in the plane with top \( y = f(x) \) and bottom \( y = g(x) \), and its shadow on the \( x \)-axis is the interval \( a \leq x \leq b \), then \( R \) is defined by the system of inequalities

\[
\begin{align*}
0 &\leq x \leq \sqrt{4 - y} \\
0 &\leq y \leq 4.
\end{align*}
\]
CHAPTER 1. INTRODUCTION TO THREE DIMENSIONS

\[ g(x) \leq y \leq f(x) \]
\[ a \leq x \leq b. \]

If \( R \) is a region in the plane with right side \( x = f(y) \) and left side \( x = g(y) \), and the shadow on the \( y \)-axis is the interval \( c \leq y \leq d \), then \( R \) is defined by the system of inequalities

\[ g(y) \leq x \leq f(y) \]
\[ c \leq y \leq d. \]

**Example 1.6.2. A more complicated region**

Let \( R \) be the region in the first quadrant bounded by \( y = 2 \) and \( y = 4 - 2x \).

If we want to use the "Top/Bottom" approach, you notice that the top function changes from \( y = 2 \) to \( y = 4 - 2x \) at \( x = 1 \). To accurately describe \( R \) we have to use two systems of inequalities. The region \( R \) consists of two regions \( R_1 \) and \( R_2 \), described by the two systems of inequalities,

\[
\begin{align*}
\text{Region } R_1 & : \\
0 \leq y & \leq 2 \\
0 \leq x & \leq 1
\end{align*}
\]

\[
\begin{align*}
\text{Region } R_2 & : \\
0 \leq y & \leq 4 - 2x \\
1 \leq x & \leq 2.
\end{align*}
\]

![Figure 1.6.2: A Left/Right Region](image)

On the other hand, the region \( R \) has just one left and right side. Therefore it can be described using a single system of inequalities using a “Left/Right” approach. Solving the right side for \( x \) gives \( x = 2 - \frac{1}{2}y \), and we get the system

\[
\begin{align*}
0 \leq x & \leq 2 - \frac{1}{2}y \\
0 \leq y & \leq 2.
\end{align*}
\]

**Example 1.6.3. A disk**
1.6. DESCRIBING REGIONS

The unit disk in the plane is all points inside the unit circle, and is defined by \( \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq 1\} \). The top of the region is the upper semicircle, and the bottom is the lower. Solving the equation \( x^2 + y^2 = 1 \) of the boundary curve for \( y \) gives \( y = \pm \sqrt{1 - x^2} \). You could also use the left and right semicircle, solving the equation for \( x \). Therefore the disk is given by

\[
\begin{align*}
\text{Top/Bottom} & : -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2} \\
\text{Left/Right} & : -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2}
\end{align*}
\]

Both of these descriptions are cumbersome, and the unit disk is much more easily described in polar coordinates. We see that \( r \) is between 0 and 1, and there is no restriction on \( \theta \), therefore the unit disk is described by the system of inequalities

\[
\begin{align*}
0 \leq r & \leq 1 \\
0 \leq \theta & \leq 2\pi.
\end{align*}
\]

Example 1.6.4. A tetrahedron

A tetrahedron is the three dimensional analogue of a triangle. To construct a triangle, connect three vertices with line segments and take all the points inside. The vertices can’t lie on a line, or it doesn’t work. For a tetrahedron, connect four non-coplanar points in space and take all the points inside (as with the triangle, we also mean the points on the edge, or boundary).

Consider the tetrahedron \( T \) determined by the origin and the intercepts of the plane \( 2x + y - 3z = 6 \). The intercepts are \((3, 0, 0)\), \((0, 6, 0)\), and \((0, 0, -2)\), so the triangle they determine lies below the \( xy \)-plane, and is the bottom of the tetrahedron \( T \). The top of the tetrahedron is the \( xy \)-plane. Therefore the \( z \)-coordinate of any point in \( T \) is between the equations for those surfaces solved for \( z \). Since the \( xy \)-plane is \( z = 0 \), we have \((2x + y - 6)/3 \leq z \leq 0\). To get inequalities on \( x \) and \( y \), we reduce to a two-dimensional picture by taking the shadow of \( T \) in the \( xy \)-plane (the technical term is the projection of \( T \) into the \( xy \)-plane). Analytically, that amounts to letting \( z = 0 \) in the boundary...
surfaces. In this case you get the triangle in the first quadrant bounded by the line $2x+y = 6$. Now find limits on $x$ and $y$ using the two-dimensional techniques described above. This gives that $T$ is described by the system

$$
\frac{(2x+y-6)}{3} \leq z \leq 0 \\
0 \leq y \leq 2x - 6 \\
0 \leq x \leq 3
$$

Figure 1.6.4: A tetrahedron

Alternatively you can consider the plane as describing the right side of $T$, and the triangle in the $xz$-plane as the left side. Solving equations for $y$ gives limits on $y$, then project into the $xz$-plane and use two-dimensional techniques to find the remaining limits. Thinking of the plane as the front, solving equations for $x$, and projecting into the $yz$-plane yields the final system of inequalities describing $T$.

<table>
<thead>
<tr>
<th>Right/Left</th>
<th>Front/Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0 \leq y \leq 6 - 2x + 3z$</td>
<td>$0 \leq x \leq (6 - y + 3z)/2$</td>
</tr>
<tr>
<td>$\frac{(2x-6)}{3} \leq z \leq 0$</td>
<td>$0 \leq y \leq 3z + 6$</td>
</tr>
<tr>
<td>$0 \leq x \leq 3$</td>
<td>$-2 \leq z \leq 0$.</td>
</tr>
</tbody>
</table>

To describe a region in $\mathbb{R}^3$ using a system of inequalities in Cartesian coordinates, first decide if it’s a Top/Bottom, Right/Left, or Front/Back solid region. If it’s a Front/Back solid, solve all equations for $x$ getting the front surface is $x = f(y, z)$ and the back $x = g(y, z)$. The limits on $x$ are then $g(y, z) \leq x \leq f(y, z)$. Now project the solid into the $yz$-plane and use two-dimensional techniques to find the remaining limits. For a Top/Bottom approach:

1. Solve equations of surfaces for $z$: suppose the top surface is $z = f(x, y)$ and the bottom $z = g(x, y)$.
2. The limits on $z$ are $g(x, y) \leq z \leq f(x, y)$.
3. Project the solid into the $xy$-plane, and use two-dimensional techniques.
Example 1.6.5. A snowcone

In this example we describe a region in space as a system of inequalities in Cartesian, cylindrical, and spherical coordinates. The region $R$ is above the cone $z = r$ and inside the unit sphere, which looks like a snowcone.

Figure 1.6.5: A snowcone

Cylindrical Coordinates

We begin by describing all surfaces involved in cylindrical coordinates. The equation for the cone is given in cylindrical coordinates, while the easiest equation for the sphere is $\rho = 1$. Since $\rho^2 = z^2 + r^2$, we have that $R$ is above $z = r$ and below $z = \sqrt{1 - r^2}$. This determines the restrictions on $z$, and we must consider the shadow in the $xy$-plane to get limits on $r$ and $\theta$.

It is clear that this shadow is a disk centered at the origin. To find the radius of the disk, set the cylindrical equations equal to each other and solve for $r$. This gives $r = \sqrt{1 - r^2}$, which implies $r = \sqrt{2}/2$. The limits on $r$ and $\theta$ are now clear, and we see that the solid region $R$ is described in cylindrical coordinates by

$$
\begin{align*}
0 & \leq \rho \leq 1 \\
0 & \leq \theta \leq 2\pi \\
0 & \leq r \leq \sqrt{2}/2.
\end{align*}
$$

Spherical Coordinates

In spherical coordinates, the equations of the boundary surfaces are particularly straightforward. The sphere is $\rho = 1$ and the cone $\phi = \pi/4$. To get limits in spherical coordinates requires a slightly different approach than we’ve encountered so far. The lower limit on $\rho$ is the equation of the surface where a ray from the origin enters the region. Any ray entering $R$ starts in $R$, since the origin is in $R$, so the lower limit is $0$. The upper limit is the equation of the surface where the ray leaves the region, which is $\rho = 1$ in our case. Thus we have $0 \leq \rho \leq 1$. Notice that $\theta$ is unrestricted, and $\phi$ is between $0$ and $\pi/4$, so the snow cone is
0 \leq \rho \leq 1
0 \leq \theta \leq 2\pi
0 \leq \phi \leq \pi/4.

**Cartesian Coordinates**

In Cartesian coordinates the sphere is $x^2 + y^2 + z^2 = 1$, and the cone $z = \sqrt{x^2 + y^2}$. Since the sphere is the top and the cone the bottom of the solid we’ll use a Top/Bottom approach. The shadow in the $xy$-plane is the disk with boundary $x^2 + y^2 = 1/2$. Using the same strategy as for the unit disk above gives

\[
\sqrt{x^2 + y^2} \leq z \leq \sqrt{1 - x^2 - y^2}
\]
\[
-\sqrt{1/2 - x^2} \leq y \leq \sqrt{1/2 - x^2}
\]
\[
-1 \leq x \leq 1.
\]

**Exercises**

1. Find a system of inequalities that describes the region in the first quadrant bounded by the line $y = 6 - 2x$.

2. Find a system of inequalities that describes the region in $\mathbb{R}^2$ between the parabolas $y = 1 - x^2$ and $y = x^2 - 1$.

3. Describe the region of $\mathbb{R}^2$ between $x = 0$ and $x = 2$, above $y = x$ and below $y = 4$ using a system of inequalities.

4. Let $R$ be the triangle in the plane with vertices $(-1, 0)$, $(0, 1)$, and $(1, 0)$. Describe $R$.

5. Use inequalities to describe the finite regions between $y = x^3 - x$ and the $x$-axis (ignore the unbounded regions).

6. Use polar coordinates to describe the quarter of the unit circle in the second quadrant.

7. Find a system of inequalities in polar coordinates defining the region between circles of radii 3 and 5 centered at the origin.

8. Determine a system of inequalities describing the tetrahedron in the first octant bounded by the plane $2x + y + 3z = 6$.

9. Determine a system of inequalities describing the tetrahedron bounded by coordinate planes and the plane $x - y + 3z = 9$.

10. Use a system of inequalities to describe the solid under the cone $z = 5 - r$ and above the $xy$-plane.
11. Describe the solid inside the cylinder \( r = 1 \) and under the unit sphere using a system of inequalities.

12. Describe the region in space between spheres of radius 1 and 2 using inequalities in spherical coordinates.

13. The region \( R \) is inside the cylinder \( y^2 + z^2 = 1 \) and above the \( xy \)-plane, for \(-1 \leq x \leq 2\). Describe \( R \) using a system of inequalities on Cartesian coordinates.
Chapter 2

Introduction to Vectors

In this chapter we are introduced to vectors, which are directed line segments, or arrows. Vectors are a very powerful tool for studying geometric properties of surfaces, and have many applications in physics. For these reasons, this chapter is fundamental for the rest of the book.

Vectors are defined in Section 2.1 along with vector addition and scalar multiplication. Both of these operations on vectors are defined geometrically, then a concise algebraic description is determined. An application to the calculus of parametric curves is given to illustrate the utility of vector analysis. The next sections introduce two vector products: the dot- and cross-products. Both are easily defined algebraically, and both have considerable geometric significance. The dot product, defined in Section 2.3 is related to angles between vectors, and can be used to project one vector onto another. The cross product is studied in Section 2.3 and can be used to find areas of parallelograms and normal vectors to planes. While vectors are fresh in our minds we introduce vector fields in Section 2.5. These will be used extensively in Chapter 5.

2.1 Geometry and Algebra of Vectors

The Geometry of Vectors

We begin by defining what we mean by a vector, and when two vectors are the same. This is followed by geometric definitions of vector addition and scalar multiplication which seem quite natural. Decomposing vectors into components will connect the geometric definitions to algebraic manipulation, facilitating vector analysis.

Definition 2.1.1. A vector is a quantity that has a magnitude and direction. Thus a vector is a directed line segment.

A vector can be thought of as an arrow, as in 2.1.1. Vectors will be denoted by bold face letters, like \( \mathbf{v} \). The magnitude of the vector will be denoted \( ||\mathbf{v}|| \). The number \( ||\mathbf{v}|| \) is also called the norm, or length, of \( \mathbf{v} \). There is one special vector,
the zero vector $\mathbf{0}$. It is the vector with magnitude zero, which symbolically looks like $\|\mathbf{0}\| = 0$. The bold-faced zero represents a vector, while the zero on the right represents a real number.

Note that polar coordinates are similar in flavor to vectors since $r$ gives the length and $\theta$ the direction. In spherical coordinates, $\rho$ gives the length and $\theta$ and $\phi$ combine to provide the direction. Thus we’ve already been thinking of points in vector fashion and we’ll now formalize that a little more.

Geometrically, the vector $\mathbf{v}$ is represented by an arrow pointing in the direction of $\mathbf{v}$ with length $\|\mathbf{v}\|$. When thinking of them in this way, we call the tail of the arrow the initial point and the terminal point is its tip.

**Example 2.1.1. Vectors**

The initial point of vector $\mathbf{v}$ in Figure 2.1.1 is $(0, 2)$ and its terminal point is $(3, -1)$. The length $\|\mathbf{v}\|$ of the vector $\mathbf{v}$ is the distance from its initial to its terminal point. Using the distance formula in the plane we have

$$\|\mathbf{v}\| = \sqrt{(3-0)^2 + (-1-2)^2} = 3\sqrt{2}. \quad \blacktriangle$$

**Definition 2.1.2.** Two vectors are the same vector if they have the same magnitude and direction.

This means it doesn’t matter where the vectors start: if they have the same magnitude and direction, they are the same vector! So all the vectors pictured in Figure 2.1.1 are the same. This sometimes causes confusion at first, but just remember that you can move vectors around and they don’t change (provided you keep the same length and direction).

A position vector for the point $P$ will be the vector $\mathbf{P}$ with its initial point at the origin and its terminal point at $P$. In Figure 2.1.1 the position vector representing $\mathbf{v}$ is labeled $\mathbf{P}$, and is the position vector of the point $P = (3, -3)$. We’ll denote position vectors using angled parentheses, as in $\mathbf{P} = (3, -3)$. We will
2.1. GEOMETRY AND ALGEBRA OF VECTORS

frequently need to think of points as position vectors, and then as points again. It is important to be versatile enough to switch between the two interpretations with ease.

Example 2.1.2. Length of a position vector

Find the length of the vector \( \mathbf{P} = \langle 3, 7, -2 \rangle \).

Since the initial point of \( \mathbf{P} \) is the origin and the terminal point is \((3, 7, -2)\), the length of \( \mathbf{P} \) is just the distance from the point \( P \) to the origin. We have \( \| \mathbf{P} \| = \sqrt{3^2 + 7^2 + (-2)^2} = \sqrt{62}. \)

We now give geometric descriptions of vector addition and scalar multiplication.

**Vector Addition** Let \( \mathbf{v}, \mathbf{w} \) be vectors, we define their sum \( \mathbf{v} + \mathbf{w} \) as follows. The vector \( \mathbf{v} + \mathbf{w} \) is obtained by concatenating \( \mathbf{v} \) and \( \mathbf{w} \). In other words, placing \( \mathbf{w} \) on the terminal point of \( \mathbf{v} \), then going from tail to tip (see Figure 2.1.2). This moving \( \mathbf{w} \) around makes sense, since vectors are not tied to initial points. Notice that vector addition is commutative! Technically the sum \( \mathbf{v} + \mathbf{w} \) is obtained by placing \( \mathbf{w} \) on the tip of \( \mathbf{v} \), rather than placing \( \mathbf{v} \) on the tip of \( \mathbf{w} \). Observe that the two constructions form a parallelogram, and both \( \mathbf{w} + \mathbf{v} \) and \( \mathbf{v} + \mathbf{w} \) are the diagonal of that parallelogram. Algebraically, then \( \mathbf{w} + \mathbf{v} = \mathbf{v} + \mathbf{w} \). One can also see geometrically that vector addition is associative, but the picture is not very enlightening.

![Figure 2.1.2: Vector Addition and Scalar Multiplication](image)

**Example 2.1.3. Vector Addition and Position Vectors**

Let \( \mathbf{v} \) be the position vector \( \langle 3, 1 \rangle \), and \( \mathbf{w} = \langle 1, 2 \rangle \). Then \( \mathbf{v} + \mathbf{w} \) is obtained by placing \( \mathbf{w} \) on the tip of \( \mathbf{v} \), and going from tail to tip (see Figure 2.1.2(a)).
Since \( \mathbf{v} \) is a position vector, the initial point of \( \mathbf{v} + \mathbf{w} \) is the origin. To find the terminal point start at the origin, walk along \( \mathbf{v} \) then walk along \( \mathbf{w} \). Thus the total horizontal distance traveled is the horizontal distance from \( \mathbf{v} \) plus the horizontal distance from \( \mathbf{w} \). Similarly the total vertical distance for \( \mathbf{v} + \mathbf{w} \) is the sum of the vertical distances for \( \mathbf{v} \) and \( \mathbf{w} \). The tip of \( \mathbf{v} + \mathbf{w} \), then, is the point whose coordinates are the sums \((3 + 1, 1 + 2)\), and \( \mathbf{v} + \mathbf{w} \) is the position vector \( \langle 4, 3 \rangle \). Take a moment to notice what this equality looks like using position vector notation:

\[
\langle 3, 1 \rangle + \langle 1, 2 \rangle = \langle 3 + 1, 1 + 2 \rangle = \langle 4, 3 \rangle.
\]

This is not a coincidence, and provides an algebraic method for summing two position vectors—add them coordinate-wise! ▲

This observation generalizes to vector addition in arbitrary dimensions, so we see

<table>
<thead>
<tr>
<th>Adding Position Vectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let ( \mathbf{v} = \langle a, b, c \rangle ) and ( \mathbf{w} = \langle d, e, f \rangle ) be position vectors, then</td>
</tr>
<tr>
<td>( \mathbf{v} + \mathbf{w} = \langle a + d, b + e, c + f \rangle )</td>
</tr>
<tr>
<td>(2.1.2)</td>
</tr>
</tbody>
</table>

**Scalar Multiplication**  We can multiply vectors by real numbers. Intuitively, multiplying the vector \( \mathbf{v} \) by the number \( \alpha \) just stretches it by a factor of \( \alpha \). Since this multiplication results in a rescaling, real numbers will also be called *scalars*. Moreover, this multiplication will be called scalar multiplication. If \( \alpha \) is negative, then \( \alpha \mathbf{v} \) goes in the opposite direction from \( \mathbf{v} \). Symbolically we see that \( \|\alpha \mathbf{v}\| = |\alpha|\|\mathbf{v}\| \), where \( |\alpha| \) is the absolute value of a number while both \( \|\alpha \mathbf{v}\| \) and \( \|\mathbf{v}\| \) represent lengths of a vectors. If \( \alpha = 0 \), we have that \( 0 \cdot \mathbf{v} = \mathbf{0} \) is the zero vector.

One should convince themselves that \( \alpha \mathbf{v} + \beta \mathbf{v} = (\alpha + \beta)\mathbf{v} \) for scalars \( \alpha, \beta \in \mathbb{R} \) and any vector \( \mathbf{v} \).

**Example 2.1.4. Scalar Multiplication of Position Vectors**

In Figure 2.1.2(c) we see \( \mathbf{v} = \langle 1, 3/4 \rangle \) and \( \|\mathbf{v}\| = 5/4 \). Then \( 4\mathbf{v} \) should be in the same direction but four times as long. The initial point of \( \mathbf{v} \) is the origin, so multiplying the coordinates of the terminal point of \( \mathbf{v} \) by 4 gives the terminal point for \( 4\mathbf{v} \). Algebraically we see that

\[
4 \langle 1, 3/4 \rangle = \langle 4 \cdot 1, 4 \cdot 3/4 \rangle = \langle 4, 3 \rangle. \quad \blacktriangleleft
\]

Thus scalar multiplication with position vectors can be done coordinate-wise as well! We summarize with
2.1. GEOMETRY AND ALGEBRA OF VECTORS

Scalar Multiplication of Position Vectors

Let $\alpha \in \mathbb{R}$ be a scalar, and let $v = \langle a, b, c \rangle$ be a position vector, then

$$\alpha v = \langle \alpha a, \alpha b, \alpha c \rangle.$$  \hspace{1cm} (2.1.3)

Algebra with Vectors: The operations of vector addition and scalar multiplication have convenient algebraic interpretations when we work with position vectors. Sometimes, however, vectors are not given as position vectors; rather they are given in terms of their initial and terminal points, or by a length and direction. Before we can manipulate such vectors algebraically we have to decompose them into their components, which is a fancy way of saying we find a position vector representative. Given an arbitrary vector, then, we now define its components and show how to find them. We can then use the rules of Equations 2.1.2 and 2.1.3 to engage in vector analysis.

To begin with, we introduce special vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$. The vector $i$ is has magnitude one unit in the positive $x$-direction. Another way to say this is that $i$ is the position vector of the point $P = (1,0,0)$, so $i = \langle 1,0,0 \rangle$. Similarly, $j = \langle 0,1,0 \rangle$ and $k = \langle 0,0,1 \rangle$, are the position vectors of $(0,1,0)$ and $(0,0,1)$, respectively. When considering vectors in $\mathbb{R}^2$, we still use $i$ and $j$ to represent the position vectors of $(1,0)$ and $(0,1)$. These vectors are so special, they get a name.

Definition 2.1.3. The vectors $i$, $j$, $k$ are called the standard basis vectors for $\mathbb{R}^3$.

The standard basis vectors will give us an easy way to describe all vectors in $\mathbb{R}^3$, as well as calculate with them. Since scalar multiplication is just stretching, the vector $ai$ can be described as the vector with initial point $(0,0,0)$ and terminal point $(a,0,0)$ (i.e. $ai = \langle a,0,0 \rangle$). Similarly we obtain the vectors $bj = \langle 0,b,0 \rangle$ and $ck = \langle 0,0,c \rangle$. Moreover, when we start at the origin concatenate the vectors $ai$, $bj$, and $ck$, we travel $a$ units in the $x$-direction, $b$ in the $y$-direction, and $c$ in the $z$-direction. At the end of it all we are at the terminal point $(a,b,c) \in \mathbb{R}^3$. Algebraically we have $ai + bj + ck = \langle a,b,c \rangle$ (see Figure 2.1.3).

Definition 2.1.4. The components of the vector $v = ai + bj + ck = \langle a,b,c \rangle$ are the scalars $a$, $b$, $c$.

More precisely $a$ is the component of $v$ in the $i$ direction, $b$ is the component of $v$ in the $j$ direction, and $c$ is the $k$-component of $v$.

Writing a given vector as a sum of scalar multiples of other vectors, like $i$, $j$, and $k$, is an important skill. How you do it depends on how the vector is given to you initially, and we illustrate it with a couple examples below. First we define a linear combination of vectors.

Definition 2.1.5. A linear combination of vectors is any sum of scalar multiples of them.
Example 2.1.5. A linear combination

The vector $3\mathbf{i} + 2\mathbf{j} + 4\mathbf{k}$ is a linear combination of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$. It’s $\mathbf{i}$-component is 3, and 4 is its component in the $\mathbf{k}$ direction. ▲

Finding Vector Components: We have mentioned that vectors aren’t always given as a linear combination of the standard basis vectors. We may be given initial and terminal points, or a distance and direction. In the next few examples we will illustrate writing vectors as linear combinations of the standard basis vectors. The strategies for doing so depend on how the vector is initially given to you.

Example 2.1.6. Given Terminal and Initial Points

The vector $\mathbf{v}$ has initial point $(-2, -1, 1)$, and terminal point $(3, 5, -4)$. Write $\mathbf{v}$ as a linear combination of the standard basis vectors.

To do so, note that the coefficient of $\mathbf{i}$ is the distance in the $x$-direction from initial to terminal point of $\mathbf{v}$. In this case, that is from $-2$ to $3$, so the component is $3 - (-2) = 5$. Using similar reasoning in the other coordinate directions, we have

$$
\mathbf{v} = (3 - (-2))\mathbf{i} + (5 - (-1))\mathbf{j} + (-4 - 1)\mathbf{k} = 5\mathbf{i} + 6\mathbf{j} - 5\mathbf{k}.
$$

We see that the vector $\mathbf{v}$, which is given as having initial point $(-2, -1, 1)$ and terminal point $(3, 5, -4)$, is the same as the position vector $\langle 5, 6, -5 \rangle$. ▲

This example generalizes, and we state it in an intuitive way: To write $\mathbf{v}$ as a linear combination of the standard basis vectors, take the terminal point minus the initial point. Technically we aren’t “subtracting points”. We are subtracting the position vectors of the points. Analytically the two processes are equivalent.
2.1. GEOMETRY AND ALGEBRA OF VECTORS

The vector $v$ from $P(a, b, c)$ to $Q(d, e, f)$ is given algebraically by

$$v = Q - P = (d - a, e - b, f - c).$$

Geometrically we see that $v$ and $Q - P$ are opposite sides of a parallelogram, and are therefore the same vector.

Example 2.1.7. Given Magnitude and Direction

Let $v$ be the position vector in $\mathbb{R}^2$ that makes an angle of $\frac{2\pi}{3}$ with the positive $x$-axis and is 3 units long. Write $v$ as a linear combination of standard basis vectors.

Notice that this vector is defined by specifying a direction and magnitude. In polar coordinates, the endpoint of the position vector is at $r = 3$ and $\theta = \frac{2\pi}{3}$. Using the change of coordinates from polar to Cartesian, the endpoint is $(3 \cos \frac{2\pi}{3}, 3 \sin \frac{2\pi}{3}) = (-\frac{3}{2}, \frac{3\sqrt{3}}{2})$. Since the coordinates of the terminal point of a position vector are the components, we have

$$v = -\frac{3}{2}i + \frac{3\sqrt{3}}{2}j = \left\langle -\frac{3}{2}, \frac{3\sqrt{3}}{2} \right\rangle. \blacklozenge$$

Example 2.1.8. Spherical coordinates and vectors in $\mathbb{R}^3$

In $\mathbb{R}^2$ you can specify a direction by fixing $\theta$. In $\mathbb{R}^3$ you must fix both $\theta$ and $\phi$ to determine a direction. The coordinate transformations $(x, y, z) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$ will facilitate determining components for a given vector. Let $v$ be the vector 2 units long in the direction determined by $\theta = \frac{2\pi}{3} \text{ and } \phi = \frac{\pi}{4}$. Think of $v$ with its initial point at the origin. Then the length of $v$ is $\rho$, and the specifications determine the terminal point: $(x, y, z) = (2 \sin \frac{\pi}{4} \cos \frac{2\pi}{3}, 2 \sin \frac{\pi}{4} \sin \frac{2\pi}{3}, 2 \cos \frac{\pi}{4}) = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}, \sqrt{2})$. In vector notation we have
\[
v = \frac{\sqrt{2}}{2} \mathbf{i} + \frac{\sqrt{6}}{2} \mathbf{j} + \sqrt{2} \mathbf{k} = \left\langle -\frac{\sqrt{2}}{2}, \frac{\sqrt{6}}{2}, \sqrt{2} \right\rangle.
\]

Our original definitions of vector addition and scalar multiplication involved geometric interpretations of vectors. Expressing vectors as linear combinations of the standard basis makes working with them much easier algebraically. We summarize the algebra of vectors in terms of standard basis vectors as well as position vectors.

**Algebra with position vectors**

\[
\langle a, b, c \rangle + \langle d, e, f \rangle = \langle a + d, b + e, c + f \rangle \\
t \langle a, b, c \rangle = \langle ta, tb, tc \rangle \\
\| \langle a, b, c \rangle \| = \sqrt{a^2 + b^2 + c^2}
\]

**Algebra using standard basis vectors**

If \( v = \langle a, b, c \rangle \) and \( w = \langle d, e, f \rangle \) and \( t \in \mathbb{R} \), we have

\[
v + w = \langle ai + bj + ck \rangle + \langle di + ej + fk \rangle \\
= \langle a + di + (b + e)j + (c + f)k \rangle \\
tv = t\langle ai + bj + ck \rangle = t\mathbf{i} + t\mathbf{j} + t\mathbf{k} \\
\|v\| = \|ai + bj + ck\| = \sqrt{a^2 + b^2 + c^2},
\]

**Unit Vectors:** Vectors that are one unit long enjoy many nice properties, so we give them a name. A **unit vector** is any vector of length one.

**Example 2.1.9. Unit Vectors**

The vector \( \mathbf{u} = \frac{1}{\sqrt{3}} \mathbf{i} - \frac{1}{\sqrt{3}} \mathbf{j} + \frac{1}{\sqrt{3}} \mathbf{k} \) is a unit vector since

\[
\|\mathbf{u}\| = \left\| \frac{1}{\sqrt{3}} \mathbf{i} - \frac{1}{\sqrt{3}} \mathbf{j} + \frac{1}{\sqrt{3}} \mathbf{k} \right\| \\
= \sqrt{\left( \frac{1}{\sqrt{3}} \right)^2 + \left( -\frac{1}{\sqrt{3}} \right)^2 + \left( \frac{1}{\sqrt{3}} \right)^2} = 1. \quad \blacksquare
\]

We now use what we know about the unit circle in \( \mathbb{R}^2 \) and unit sphere in \( \mathbb{R}^3 \) to find forms of unit vectors more generally. Any vector of the form \( \mathbf{u} = \cos \theta \mathbf{i} + \sin \theta \mathbf{j} \) is a unit vector in \( \mathbb{R}^2 \). Moreover, any unit length position vector in \( \mathbb{R}^3 \) has its terminal point on the unit sphere \( \rho = 1 \). Letting \( \rho = 1 \) in the change-of-coordinates shows that any unit vector in \( \mathbb{R}^3 \) is of the form \( \mathbf{u} = \sin \phi \cos \theta \mathbf{i} + \sin \phi \sin \theta \mathbf{j} + \cos \phi \mathbf{k} \) for some values of \( \theta \) and \( \phi \).

**Example 2.1.10. Normalizing Vectors**
2.1. GEOMETRY AND ALGEBRA OF VECTORS

On occasion we will be given a vector \( \mathbf{v} \) and want to find a unit vector in the same direction. To do so, merely divide \( \mathbf{v} \) by its length \( \| \mathbf{v} \| \). This process is called normalizing the vector \( \mathbf{v} \). In this example, we wish to normalize the vector \( \mathbf{v} = 2\mathbf{i} + 1\mathbf{j} - 3\mathbf{k} \). First note that the length is \( \| \mathbf{v} \| = \sqrt{2^2 + 1^2 + (-3)^2} = \sqrt{14} \). Dividing the vector \( \mathbf{v} \) by this scalar gives

\[
\mathbf{u} = \frac{\mathbf{v}}{\| \mathbf{v} \|} = \frac{2}{\sqrt{14}} \mathbf{i} + \frac{1}{\sqrt{14}} \mathbf{j} - \frac{3}{\sqrt{14}} \mathbf{k}.
\]

Force: Force in physics is a push acting on an object as a result of its interaction with another object. This could be the force of gravity on us, resulting from our interaction with the earth, or the force of friction a table applies to a glass of water sliding across it. Force is a vector quantity, having both magnitude and direction. Its magnitude is measured in Newtons, which is the force needed to cause a 1m/s/s acceleration on a 1kg object. If several forces are acting on an object, the net force is just the vector sum of them. We illustrate with an example.

Example 2.1.11. Calculating Net Force

A simple pendulum has a 10kg mass, so that the constant force due to gravity is \( \mathbf{F}_{\text{grav}} = 10(0, -9.8) \) with magnitude \( \| \mathbf{F}_{\text{grav}} \| = 98 \text{N} \). There is a second force acting on the pendulum, that of tension due to the string (we ignore air resistance). The force due to tension \( \mathbf{F}_{\text{tens}} \) is not constant, but lets consider a particular moment in time.

Assume that when the pendulum is at its maximum height makes an angle of 60° below the horizontal. At this height we have \( \mathbf{F}_{\text{tens}} = \frac{1}{2} \langle -49\sqrt{3}, 147 \rangle \). Find the net force \( \mathbf{F}_{\text{net}} \) on the pendulum at that moment, together with its magnitude.

The net force is simply the vector sum of the forces acting on the pendulum,
so

\[ F_{\text{net}} = F_{\text{grav}} + F_{\text{tens}} = 10 \langle 0, -9.8 \rangle + \frac{1}{2} \left\langle -49\sqrt{3}, 147 \right\rangle \]

\[ = \langle 0, -98 \rangle + \left\langle -\frac{49\sqrt{3}}{2}, \frac{147}{2} \right\rangle = \left\langle -\frac{49\sqrt{3}}{2}, -\frac{49}{2} \right\rangle. \]

Further, we compute the magnitude of the net force to be

\[ \|F_{\text{net}}\| = \sqrt{\left( -\frac{49\sqrt{3}}{2} \right)^2 + \left( -\frac{49}{2} \right)^2} = 49. \]

**Parametric Lines** The operations of vector addition and scalar multiplication shed some light on parameterizations of lines. In fact, they demonstrate that lines (even in \( \mathbb{R}^3 \)) are determined by two things: a point \( P \) on the line and a vector \( v \) in the direction of the line. We will derive this by first considering lines through the origin.

Suppose \( \ell \) is a line through the origin in \( \mathbb{R}^3 \) and \( v = \langle A, B, C \rangle \) is a position vector that lies on \( \ell \). Then any scalar multiple \( tv \) also lies on \( \ell \), since it’s just a rescaling of \( v \). In fact, since the set of all scalar multiples of \( v \) form a line, it must be true that \( C(t) = tv, \ -\infty < t < \infty \), parameterizes \( \ell \)!

We now consider lines that don’t go through the origin. Let \( \ell \) be a line through the point \( P \), and in the direction of the vector \( v \). Then \( C(t) = tv \) parameterizes the line through the origin parallel to \( \ell \). To get a parameterization for \( \ell \) we just have to translate it so that it goes through \( P \). Adding the position vector \( P \) to each point on the line \( tv \) moves the origin to \( P \), and the rest of the line \( tv \) to \( \ell \). Thus parametric equations for \( \ell \) are \( C(t) = P + tv \). Notice that the right-hand side of this equality is a linear combination of the vectors \( P \) and \( v \), which is another vector. However, the left-hand side we are considering at point on the line, not a vector. Thus we are blurring the distinction between position vectors and points here.

### Parameterization Lines

One parameterization of the line \( \ell \) through the point \( P \) and in the direction of the vector \( v \) is

\[ C(t) = P + tv, \ -\infty < t < \infty. \]

**Example 2.1.12.** *Given point and direction*

Parameterize the line through \((1, -2, 1)\) parallel to the vector \( v = 2\mathbf{i} + 3\mathbf{j} - \mathbf{k} \).

Using the above strategy we get

\[ C(t) = \langle 1, -2, 1 \rangle + t \langle 2, 3, -1 \rangle = \langle 1, -2, 1 \rangle + \langle 2t, 3t, -t \rangle = \langle 1 + 2t, -2 + 3t, 1 - t \rangle, \text{ for } -\infty < t < \infty. \]

**Example 2.1.13.** *Line given two points*
2.1. GEOMETRY AND ALGEBRA OF VECTORS

Figure 2.1.5: Parametric Lines and Planes

Parameterize the line through the points $(3, -1, 3)$ and $(2, 4, -1)$.

We need a point on the line and a vector in the line. Notice that the vector from $(3, -1, 3)$ to $(2, 4, -1)$ lies in the line. Using the “terminal minus initial” idea we see that a vector in the line is $v = (2 - 3)i + (4 - (-1))j + (-1 - 3)k = -1i + 5j - 4k$. Using this we get the parametric equations

$$C(t) = (3, -1, 3) + t(-1, 5, -4) = (3 - t, -1 + 5t, 3 - 4t),$$

for $-\infty < t < \infty$. ▶

In many applications we are interested in parameterizing the line segment between two points. This occurs frequently enough that we highlight one way to do it. Suppose we want to parameterize the segment from point $P$ to point $Q$. As in Example 2.1.13, we can parameterize the line by noticing that the vector $Q - P$ is in the direction of the line. By our strategy for parameterizing lines the entire line $\ell$ is given by $C(t) = P + t(Q - P), -\infty < t < \infty$, and we have to determine the range of $t$ values that gives the segment from $P$ to $Q$. Notice that when $t = 0$ we have $C(0) = P$. Similarly $C(1) = P + Q - P = Q$. Therefore the segment is given by $C(t) = P + t(Q - P)$ for $0 \leq t \leq 1$. Combining like terms on $C(t)$ we get the parameterization $C(t) = P + t(Q - P) = (1-t)P + tQ$ for $0 \leq t \leq 1$.

Parameterizing Line Segments

One parameterization of the line segment from $P$ to $Q$ is

$$C(t) = (1 - t)P + tQ, \text{ for } 0 \leq t \leq 1. \quad (2.1.4)$$

Example 2.1.14. Line Segments

Parameterize the line segment from $(3, 1, 3)$ to $(-2, 3, 5)$. According to Equa-
We have
\[ C(t) = (1 - t) \langle 3, 1, 3 \rangle + t \langle -2, 3, 5 \rangle \]
\[ = \langle (1 - t)3 + t(-2), (1 - t) + 3t, (1 - t)3 + 5t \rangle \]
\[ = \langle 3 - 5t, 1 + 2t, 3 + 2t \rangle, \quad 0 \leq t \leq 1. \]

**Example 2.1.15. Piecewise Linear Curves**

Parameterize the triangle with vertices \( P(1, 3, -4) \), \( Q(5, -1, 6) \), and \( R(-2, 1, 3) \) counterclockwise when viewed from above.

In order to parameterize a triangle, you parameterize each side separately. Each side is a line segment, so we can use Equation 2.1.4. We now consider how to parameterize counterclockwise when viewed from above. When viewing from above one cannot distinguish heights, so we project \( P, Q \), and \( R \) into the \( xy \)-plane as in Figure 2.1.6. To ensure a counterclockwise orientation, we go from \( P \) to \( R \) to \( Q \), and back to \( P \) again.

Let \( C_1 \) be the line segment from \( P \) to \( R \), \( C_2 \) from \( R \) to \( Q \), and \( C_3 \) from \( Q \) to \( P \). Using Equation 2.1.4 we get the parameterizations:

\[ C_1(t) = (1 - t) \langle 1, 3, -4 \rangle + t \langle -2, 1, 3 \rangle = \langle 1 - 3t, 3 - 2t, -4 + 7t \rangle \]
\[ C_2(t) = (1 - t) \langle -2, 1, 3 \rangle + t \langle 5, -1, 6 \rangle = \langle -2 + 7t, 1 - 2t, 3 + 3t \rangle \]
\[ C_3(t) = (1 - t) \langle 5, -1, 6 \rangle + t \langle 1, 3, -4 \rangle = \langle 5 - 4t, -1 + 4t, 6 - 10t \rangle \]

where the limits on each parameterization is \( 0 \leq t \leq 1 \). 

**Parametric Planes:** Parametric planes are given by \( S(t) = (x_0 + As + Dt, y_0 + Bs + Et, z_0 + Cs + Ft) \), which goes through the point \( P = (x_0, y_0, z_0) \). Let \( u = \langle A, B, C \rangle \), and \( v = \langle D, E, F \rangle \). Then the plane can be rewritten in vector format as
2.1. GEOMETRY AND ALGEBRA OF VECTORS

\[ S(s, t) = (x_0 + As + Dt)i + (y_0 + Bs + Et)j + (z_0 + Cs + Ft)k \]
\[ = x_0i + y_0j + z_0k + s(Al + Bj + Ck) + t(Di + Ej + Fk) \]
\[ = P + su + tv. \]

In this context, \( P \) is a point on the plane and \( u, v \) are vectors in the plane.
We can use these observations to parameterize planes.

**Example 2.1.16. Parametric Planes**

Find parametric equations for the plane containing points \((3, -1, 2), (-3, 2, 2)\) and \((1, 4, 3)\).

First we find two vectors in the plane, thinking of \((3, -1, 2)\) as the initial point of both and the remaining points as terminal points. We have

\[ u = (-3 - 3)i + (2 - (-1))j + (2 - 2)k = -6i + 3j = \langle -6, 3, 0 \rangle \]
\[ v = (1 - 3)i + (4 - (-1))j + (3 - 2)k = -2i + 5j + j = \langle -2, 5, 1 \rangle \]

Using Equation 2.1.6 we get

\[ S(s, t) = (3, -1, 2) + s\langle -6, 3, 0 \rangle + t\langle 2, 5, 1 \rangle, \quad -\infty < s, t < \infty. \]

By restricting the limits on the parameters, we can parameterize portions of planes as well. In particular, we can easily parameterize the parallelogram determined by two position vectors. Let \( u = \langle a, b, c \rangle \) and \( v = \langle d, e, f \rangle \) be two position vectors in \( \mathbb{R}^3 \). As long as \( u \) and \( v \) are not parallel they determine a parallelogram \( \mathcal{P} \), with vertices \((0, 0, 0), (a, b, c), (d, e, f)\), and whose final vertex is the endpoint of \( u + v = (a + d, b + e, c + f) \). Any point \( P \) in \( \mathcal{P} \) has a position vector that is a linear combination of \( u \) and \( v \), so that \( P = su + tv \) (see Figure 2.1.7). In fact, the scalars \( s \) and \( t \) are both between 0 and 1 since \( P \) goes only partway in both the \( u \) and \( v \) directions. Thus a parameterization of the parallelogram \( \mathcal{P} \) is

\[ S(s, t) = su + tv, \quad 0 \leq s, t \leq 1. \]

**Example 2.1.17. Parameterizing a Parallelogram**

Parameterize the parallelogram determined by the position vectors \( \langle 3, -4, 2 \rangle \) and \( \langle 2, 2, 1 \rangle \).

By the above remarks, one parameterization is

\[ S(s, t) = (3s + 2t, -4s + 2t, 2s + t), \quad 0 \leq s, t \leq 1. \]

**Exercises**
1. Which of the following vectors are the same?
   (a) \( \mathbf{v} = (3, -3) \)
   (b) \( \mathbf{v} \) has initial point: \((3, -2)\) and terminal point: \((2, -3)\)
   (c) \( \|\mathbf{v}\| = 3\sqrt{2} \) and \( \mathbf{v} \) makes an angle of \(-\pi/4\) with the positive x-axis.
   (d) \( \|\mathbf{v}\| = \sqrt{2} \) and \( \mathbf{v} \) makes an angle of \(5\pi/4\) with the positive x-axis.
   (e) \( \mathbf{v} = (-1, -1) \)

2. Which of the following vectors are the same?
   (a) \( \mathbf{v} = (4, 3) \)
   (b) \( \mathbf{v} \) has initial point: \((5, -3)\) and terminal point: \((9, 0)\)
   (c) \( \|\mathbf{v}\| = 5 \) and \( \mathbf{v} \) makes an angle of \(-\pi/2\) with the positive x-axis.
   (d) \( \|\mathbf{v}\| = 5 \) and \( \mathbf{v} \) makes an angle of \(\tan^{-1}(3/4)\) with the positive x-axis.
   (e) \( \mathbf{v} = (0, -5) \)

3. Find the components of a vector in the plane that is 4 units long and makes an angle of \(\frac{4\pi}{3}\) with the positive x-axis.

4. Normalize the vector \( \langle 3, 4, 7 \rangle \).

5. Find a unit vector that points in the same direction as the vector from \((2, 1, 5)\) to \((3, -1, 2)\).

6. Normalize the vector in \(\mathbb{R}^2\) given by \( \|\mathbf{v}\| = 7 \) and \( \mathbf{v} \) makes an angle of \(5\pi/6\) with the positive x-axis.

7. Find the components of the unit vector in the direction \( \phi = \frac{\pi}{4}, \theta = \frac{3\pi}{4} \).

8. Find the vector length 5 in the direction \( \phi = \frac{3\pi}{4}, \theta = -\frac{\pi}{3} \).
9. Normalize the vector $\langle 13, -1, 4 \rangle$.

10. Find a vector 3 units long in the opposite direction from $v = \langle 3, 1, -4 \rangle$.

11. Describe, in English, why $\alpha v + \beta v = (\alpha + \beta)v$ for scalars $\alpha, \beta \in \mathbb{R}$ and any vector $v$.

12. A boy pulls a sled along a horizontal path with a force of 40 Newtons. If the rope makes an angle of $60^\circ$ with the horizontal, find the component of force in the direction of movement.

13. What scalar multiple of $v = \langle 3, 1, -2 \rangle$ must be added to the vector $u = \langle 4, 2, 1 \rangle$ to result in the vector $w = \langle -5, -1, 7 \rangle$?

14. Find parametric equations for the line through $(-3, 1, 4)$ and $(2, 5, 1)$.

15. Parameterize the line through $(2, -7, 3)$ and in the direction of $\langle -2, 3, 1 \rangle$.

16. Parameterize the line through $(0, 1, 4)$ and in the direction of $\langle 6, -2, -4 \rangle$.

17. Find a vector parallel to the line $C(t) = (2 - 3t, 3t - 7, 2t + 1)$.

18. Parameterize the line through $(2, 0, 5)$ and parallel to the line $C(t) = (2 + t, t + 5, 2t - 6)$.

19. Parameterize the line segment from $(2, 3, 2)$ to $(7, -6, 3)$.

20. Parameterize the line segment from $(-3, 4, 2)$ to $(6, 5, 7)$.

21. Parameterize the triangle with vertices $(2, -3, 2), (0, -2, -1),$ and $(-4, -7, 3)$ counterclockwise when viewed from above.

22. Parameterize the triangle with vertices $(2, 2, 5), (-5, 2, 1),$ and $(0, 7, -3)$ counterclockwise when viewed from above.

23. Parameterize the triangle with vertices $(-4, 5, 6), (3, 1, 2),$ and $(4, 7, 7)$ clockwise when viewed from above.

24. Find parametric equations for the plane through $(0, 1, 0), (-1, 2, 4),$ and $(3, -2, 5)$.

25. Parameterize the parallelogram determined by $u = \langle 2, 1, 3 \rangle$ and $v = \langle -1, 4, 4 \rangle$.

26. Find two vectors on the boundary of the parallelogram given by the parameterization $S(s, t) = \langle 3t + 4s, t - s, 4t + 2s \rangle$, $0 \leq s, t \leq 1$. 

2.1. GEOMETRY AND ALGEBRA OF VECTORS
2.2 The Dot Product

There are several ways to “multiply” vectors. One method is called the dot product, denoted \( \mathbf{v} \cdot \mathbf{w} \), and it multiplies two vectors to get a scalar. After giving an (unmotivated) algebraic definition of the dot product, we notice some of its algebraic properties in Theorem 2.2.1. These will be useful in applications, calculations, and proving further theorems. Then we prove Theorem 2.2.2, which gives a geometric interpretation of the dot product—relating it to the angle between the vectors. The geometric significance of the dot product provides a nice interpretation of Cartesian equations for planes, as well as a tool for projecting one vector onto another. Without further ado, we define the dot product.

**Definition 2.2.1.** The dot product of the vectors \( \mathbf{v} = \langle a, b, c \rangle \) and \( \mathbf{w} = \langle d, e, f \rangle \) is the scalar

\[
\mathbf{v} \cdot \mathbf{w} = ad + be + cf.
\]

**Example 2.2.1.** Calculating Dot Products

(a) Let \( \mathbf{v} = \langle 1, -2, 1 \rangle \), and \( \mathbf{w} = \langle -3, -5, 4 \rangle \). Then

\[
\mathbf{v} \cdot \mathbf{w} = \langle 1, -2, 1 \rangle \cdot \langle -3, -5, 4 \rangle = 1 \cdot (-3) + (-2) \cdot (-5) + 1 \cdot 4 = 11.
\]

(b) If \( \mathbf{v} = \langle 1, 2 \rangle \), find a vector \( \mathbf{w} = \langle a, b \rangle \) such that \( \mathbf{v} \cdot \mathbf{w} = 0 \). The thing that makes this example difficult is that there are infinitely many answers, and you have a choice. We have to solve the equation

\[
\mathbf{v} \cdot \mathbf{w} = 1 \cdot a + 2 \cdot b = 0.
\]

Perhaps the easiest solution is to interchange coordinates and negate one of them. Thus \( \mathbf{w} = \langle 2, -1 \rangle \) is a solution. Notice that any scalar multiple \( t \mathbf{w} = \langle 2t, -t \rangle \) is also a solution! In fact, this represents all solutions. ▲

The dot product has many convenient algebraic properties that follow directly from the definition. It will be useful to know how it interacts with vector addition and scalar multiplication. We list some properties in the following theorem.

**Theorem 2.2.1.** Let \( \mathbf{v} \), \( \mathbf{w} \), and \( \mathbf{x} \) be vectors, and \( \alpha \) a scalar. Then

1. \( \mathbf{v} \cdot \mathbf{w} = \mathbf{w} \cdot \mathbf{v} \)
2. \( (\alpha \mathbf{v}) \cdot \mathbf{w} = \alpha (\mathbf{v} \cdot \mathbf{w}) \)
3. \( (\mathbf{v} + \mathbf{w}) \cdot \mathbf{x} = \mathbf{v} \cdot \mathbf{x} + \mathbf{w} \cdot \mathbf{x} \)
4. \( \|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} \)
2.2. THE DOT PRODUCT

Proof. We prove the first statement, and leave the remaining verifications as exercises. Let \( \mathbf{v} = \langle a, b, c \rangle \) and \( \mathbf{w} = \langle d, e, f \rangle \) be vectors. A straightforward calculation shows
\[
\mathbf{v} \cdot \mathbf{w} = a \cdot d + b \cdot e + c \cdot f \\
= d \cdot a + e \cdot b + f \cdot c \\
= \mathbf{w} \cdot \mathbf{v}.
\]
The first equality follows from the definition of the dot product, the second is justified because multiplication in \( \mathbb{R} \) is commutative, and the last is another application of the definition of the dot product. \( \square \)

In addition to algebraic properties, the dot product has an important geometric interpretation. It comprises the next theorem:

**Theorem 2.2.2.** Let \( \mathbf{v} \) and \( \mathbf{w} \) be non-zero vectors, and let \( \theta \) be the angle between them. Then
\[
\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta.
\]

Proof. The result follows from an application of the law of cosines, which states that \( C^2 = A^2 + B^2 - 2AB \cos \theta \), where \( A, B, C \) are lengths of sides of a triangle and \( \theta \) is the angle opposite \( C \). If \( \mathbf{v} = \langle a, b, c \rangle \) and \( \mathbf{w} = \langle d, e, f \rangle \), then

\[
\mathbf{v} - \mathbf{w} = \langle a - d, b - e, c - f \rangle.
\]
Moreover, the vectors \( \mathbf{v}, \mathbf{w} \) and \( \mathbf{v} - \mathbf{w} \) form a triangle in which the angle \( \theta \) between \( \mathbf{v} \) and \( \mathbf{w} \) is opposite side \( \mathbf{v} - \mathbf{w} \). Applying the law of cosines, and part 4 of Theorem 2.2.1 to this situation gives
\[
\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\| \cos \theta \\
(a - d)^2 + (b - e)^2 + (c - f)^2 = (a^2 + b^2 + c^2) + (d^2 + e^2 + f^2) - 2\|\mathbf{v}\|\|\mathbf{w}\| \cos \theta \\
-2ad - 2be - 2cf = -2\|\mathbf{v}\|\|\mathbf{w}\| \cos \theta \\
\langle a, b, c \rangle \cdot \langle d, e, f \rangle = \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta.
\]

Proving the desired result. \( \square \)
CHAPTER 2. INTRODUCTION TO VECTORS

There are many consequences of Theorem 2.2.2 that we will encounter. We start by calculating an angle, then show how to find Cartesian equations of planes. We also see how to use the dot product to find projections of one vector onto another.

Example 2.2.2. Finding Angles Between Vectors

Find the angle between \( \mathbf{v} = \langle 5, 1, -2 \rangle \) and \( \mathbf{w} = \langle -3, 3, 4 \rangle \). By the above formula we have

\[
\theta = \cos^{-1}\left( \frac{\langle 5, 1, -2 \rangle \cdot \langle -3, 3, 4 \rangle}{\sqrt{5^2 + 1^2 + (-2)^2} \sqrt{(-3)^2 + 3^2 + 4^2}} \right) = \cos^{-1}\left( -\frac{20}{\sqrt{1020}} \right). \]

\[\blacktriangle\]

Theorem 2.2.2 also provides us with an easy way to check if two vectors are perpendicular. We take a moment to mention that perpendicular, orthogonal, and normal are three words for the same thing: meeting at 90°. We have the following corollary:

Corollary 2.2.1. The non-zero vectors \( \mathbf{v} \) and \( \mathbf{w} \) are orthogonal if and only if their dot product is zero, i.e. \( \mathbf{v} \cdot \mathbf{w} = 0 \).

Proof. Since \( \mathbf{v} \) and \( \mathbf{w} \) are non-zero vectors, their lengths are positive. This means the only way \( \|\mathbf{v}\| \|\mathbf{w}\| \cos \theta \) can be zero is if \( \cos \theta = 0 \). Thus, by Theorem 2.2.2 we have \( \mathbf{v} \cdot \mathbf{w} = 0 \) if and only if \( \cos \theta = 0 \). For us, \( \theta \) is always between 0 and \( \pi \), so this occurs if and only if \( \theta = \pi/2 \) and \( \mathbf{v} \) meets \( \mathbf{w} \) orthogonally. \( \square \)

Cartesian Equations for Planes: We introduced Cartesian equations for planes in Section 1.3. At that time we asked you to believe that solutions to linear equations in three variables (i.e. equations of the form \( Ax + By + Cz = D \)) yield planes in \( \mathbb{R}^3 \) by analogy with the general equation \( Ax + By = C \) of a line in \( \mathbb{R}^2 \). With the help of Theorem 2.2.2 we can now justify what you’ve taken for granted, and shed light on what information the constants \( A, B, C, \) and \( D \) give. We begin with an example.

Example 2.2.3. Orthogonal vectors

Describe all non-zero vectors which are orthogonal to \( \mathbf{v} = \langle 2, -3, 1 \rangle \).

To accomplish this, let \( \mathbf{w} = \langle x, y, z \rangle \) be any vector orthogonal to \( \mathbf{v} \). Corollary 2.2.1 implies that the set of all vectors orthogonal to \( \mathbf{v} = \langle 2, -3, 1 \rangle \) is all vectors \( \mathbf{w} = \langle x, y, z \rangle \) satisfying the equation

\[ \mathbf{v} \cdot \mathbf{w} = \langle 2, -3, 1 \rangle \cdot \langle x, y, z \rangle = 2x - 3y + z = 0. \]

Recall that this is the equation for a plane! \( \blacktriangle \)

Example 2.2.3 uses the fact that two non-zero vectors are orthogonal if and only if their dot product is zero. We extend this use to find Cartesian equations of planes. More precisely, let \( \mathcal{P} \) be a plane with \( \mathbf{n} = \langle A, B, C \rangle \) a normal vector.
to \( P \) and \((x_0, y_0, z_0)\) a point on \( P \). If \((x, y, z)\) is any other point in \( P \), then the vector \( v \) joining \((x_0, y_0, z_0)\) to \((x, y, z)\) is orthogonal to \( n \) (see Figure 2.2.2). This forces the dot product to be zero, and we get the equation

\[
0 = n \cdot v = (A, B, C) \cdot (\langle x, y, z \rangle - \langle x_0, y_0, z_0 \rangle)
\]

\[
= (A, B, C) \cdot \langle x, y, z \rangle - (A, B, C) \cdot \langle x_0, y_0, z_0 \rangle
\]

\[
= Ax + By + Cz - (Ax_0 + By_0 + Cz_0)
\]

In the final equation, the only variables are \( x, y, \) and \( z \), since everything else is a constant. Separating the variables from the constants we get the following:

**Equations for planes**

A Cartesian equation for the plane orthogonal to \( n = \langle A, B, C \rangle \) and through \((x_0, y_0, z_0)\) is

\[
Ax + By + Cz = D
\]

where \( D = Ax_0 + By_0 + Cz_0 \).

In vector notation, we have the formula

\[
\mathbf{n} \cdot \langle x, y, z \rangle = \mathbf{n} \cdot \langle x_0, y_0, z_0 \rangle
\]

Figure 2.2.2: The Vector derivation of Cartesian equations for planes

In Section 2.1 we showed that lines in \( \mathbb{R}^3 \) are determined by two pieces of information, a point on it and a vector lying in the line. Similarly we now see that planes are determined by two pieces of information: a point on it and a vector orthogonal to it.

**Example 2.2.4. Cartesian Equations for Planes**

(a) Find a Cartesian equation for the plane through \((2, -1, -3)\) with normal vector \( \mathbf{n} = \langle 3, 5, 2 \rangle \). Since the components of the normal vector \( \mathbf{n} \) are the
coefficients of $x$, $y$, and $z$, and the constant is the dot product of $n$ with the point, we have that

$$3x + 5y + 2z = \langle 2, -1, -3 \rangle \cdot \langle 3, 5, 2 \rangle$$
$$3x + 5y + 2z = -5$$

is an equation for the plane.

(b) Find a unit normal vector to the plane $2x + y - z = 4$. The coefficients tell us that $\langle 2, 1, -1 \rangle$ is a normal vector to the plane, so normalizing it gives the answer

$$u = \frac{n}{\|n\|} = \frac{\langle 2, 1, -1 \rangle}{\sqrt{6}} = \left\langle \frac{2}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}} \right\rangle.$$  ▲

Projections: We now investigate a second application of Theorem 2.2.2: vector projections. The projection of $v$ onto $w$ is obtained in the following way. Pick representatives of $v$ and $w$ with the same initial point, and drop a perpendicular line from the tip of $v$ to $w$. The resulting vector in the direction of $w$ is the projection of $v$ onto $w$, and is denoted $\text{proj}_w v$ (see Figure 2.2.3). The component of $v$ in the direction of $w$ is the signed length of $\text{proj}_w v$. It is denoted $\text{comp}_w v$, and is positive if $\text{proj}_w v$ points in the same direction as $w$ and negative if they point in opposite directions. Our goal is to see how the dot product allows us to easily calculate both $\text{proj}_w v$ and $\text{comp}_w v$.

![Figure 2.2.3: Projecting Vectors](image)

Letting $\theta$ be the angle between $v$ and $w$, and consider the right triangle created by $v$, $\text{proj}_w v$, and the dotted line in Figure 2.2.3. Recall that $\text{comp}_w v$ is the signed length of $\text{proj}_w v$. This implies that whether $\theta$ is acute or obtuse,
we have
\[ \cos \theta = \frac{\text{comp}_w v}{\|v\|}. \]

Solving for \( \text{comp}_w v \) we see from the triangle in Figure 2.2.3 that \( \text{comp}_w v = \|v\| \cos \theta \). Now use the result of Theorem 2.2.2 to compute
\[ \text{comp}_w v = \|v\| \cos \theta = \frac{v \cdot w}{\|v\| \|w\|} = \frac{v \cdot w}{\|w\|}. \]

So \( \text{proj}_w v \) is the vector that goes \( \frac{v \cdot w}{\|w\|} \) units in the direction of \( w \). Since the direction of \( w \) is the unit vector \( \frac{w}{\|w\|} \), we have
\[ \text{proj}_w v = \text{comp}_w v \frac{w}{\|w\|} = \frac{v \cdot w}{\|w\|} \frac{w}{\|w\|} = \frac{v \cdot w}{\|w\|^2} w = \frac{v \cdot w}{w \cdot w} w. \]

The projection \( \text{proj}_w v \) of the vector \( v \) onto the vector \( w \) is given by:
\[ \text{proj}_w v = \frac{v \cdot w}{w \cdot w} w. \]

The component of \( v \) in the \( w \) direction is the signed length of \( \text{proj}_w v \), and is given by:
\[ \text{comp}_w v = \frac{v \cdot w}{\|w\|}. \]

Remark: Some texts, in math or physics, will refer to what we’ve called the projection of \( v \) onto \( w \) as the component of \( v \) in the \( w \) direction. We’ll reserve the word component for a scalar quantity in this book.

Example 2.2.5. Projecting vectors

Find the projection \( \text{proj}_w v \) of \( v \) onto \( w \) where \( v = \langle 2, -3, 1 \rangle \) and \( w = \langle 1, 2, 1 \rangle \). By the above formula,
\[ \text{proj}_w v = \frac{v \cdot w}{w \cdot w} w \]
\[ = \frac{\langle 2, -3, 1 \rangle \cdot \langle 1, 2, 1 \rangle}{\langle 1, 2, 1 \rangle \cdot \langle 1, 2, 1 \rangle} \langle 1, 2, 1 \rangle \]
\[ = -\frac{1}{2} \langle 1, 2, 1 \rangle. \]

Example 2.2.6. Force on a pendulum

At a particular moment, a 10kg pendulum makes an angle of 45° with the horizontal as in Figure 2.2.4. Let \( F_{\text{grav}} = 10 \langle 0, -9.8 \rangle \) denote the force due to gravity. Find the projection of \( F_{\text{grav}} \) in the direction perpendicular to movement.
To do so, note that the pendulum travels in a circular arc, with the rope being a radius. Thus the rope is always perpendicular to the direction of movement. Since the force $F_{tens}$ due to tension is in the direction of the rope, the question asks us to project $F_{grav}$ into $F_{tens}$. From Figure 2.2.4 it is clear that $F_{tens}$ points in the same direction as $v = \langle -1, 1 \rangle$, so we find $\text{proj}_v F_{grav}$:

$$\text{proj}_v F_{grav} = \frac{\langle -1, 1 \rangle \cdot \langle 0, -98 \rangle}{\langle -1, 1 \rangle \cdot \langle -1, 1 \rangle} \langle -1, 1 \rangle = \frac{-98}{2} \langle -1, 1 \rangle = \langle 49, -49 \rangle.$$ \\n
An application: Work. In physics, the work done by a force $F$ in moving an object from one point to another is defined to be the component of force in the direction of movement times the distance traveled. In short, work is force times distance. This has a particularly nice description using vectors. Suppose an object moves in a straight line from point $A$ to point $B$. Its displacement can be described by the vector $d = B - A$ (here we’re thinking of the points $A$ and $B$ as position vectors, and taking terminal minus initial to find the vector from $A$ to $B$). The length $\|d\|$ is the distance traveled. Recall that the component of $F$ in the $d$ direction is the scalar $\frac{F}{\|d\|}$. Since work $w$ is the component of force in the direction of movement times the distance traveled, we have

$$w = \frac{F \cdot d}{\|d\|} \cdot \|d\| = F \cdot d.$$ \\n
So the work done by a constant force on an object moving in a straight line is the dot product of the force and displacement vectors!

Concept Connection: The above remarks apply to constant forces acting on objects moving linearly. We want to also consider work done by variable forces acting on objects moving along curves. This generalization is similar to the transition from calculating distance traveled from constant to variable velocity. As in the distance scenario, it turns out that to compute the work we’ll have to integrate, but that will have to wait until Chapter 5.

Example 2.2.7. Work by Constant Force
2.2. THE DOT PRODUCT

A boy pulls a sled horizontally 10 meters by applying a constant force of 16 Newtons at an angle of 60° from the horizontal. How much work did he do?

The displacement vector is \( \mathbf{d} = (10, 0) \) and the force vector is

\[
\mathbf{F} = (16 \cos 60°, 16 \sin 60°) = (8, 8\sqrt{3}).
\]

Thus the work done is

\[
w = \mathbf{F} \cdot \mathbf{d} = (10, 0) \cdot (8, 8\sqrt{3}) = 80 \text{Nm}.
\]

Figure 2.2.5: Work by the boy on the sled

In this case the work done is positive since the force is causing the sled to move. When multiple forces are involved, a given force may be hindering movement, resulting in a negative work done by that particular force. Geometrically, this occurs when the projection of the force vector points in the direction opposite the displacement vector. ▲

Example 2.2.8. Distance to a plane

Find the distance from the point \((2, 4, -3)\) to the plane \(3x + 3y + 4z = 1\).

To do this, pick a point in the plane and connect it to the point \((2, 4, -3)\) to get a vector \(\mathbf{v}\). In our case, the point \((-1, 0, 1)\) lies in the plane, so \(\mathbf{v} = \langle 2, 4, -3 \rangle - \langle -1, 0, 1 \rangle = \langle 3, 4, -4 \rangle\). Here is the key observation: the distance \(d\) from the tip of \(\mathbf{v}\) to the plane is equal to the length of the projection of \(\mathbf{v}\) onto the normal direction \(\mathbf{n} = \langle 3, 3, 4 \rangle\). This is just the absolute value of \(\text{comp}_n \mathbf{v}\), so we compute

\[
\text{comp}_n \mathbf{v} = \frac{\langle 3, 4, -4 \rangle \cdot \langle 3, 3, 4 \rangle}{\sqrt{\langle 3, 3, 4 \rangle \cdot \langle 3, 3, 4 \rangle}} = \frac{5}{\sqrt{34}}.\quad \text{▲}
\]

This example can be generalized to obtain

Distance to a Plane

The distance from the point \((d, e, f)\) to the plane \(Ax + By + Cz = D\) is the scalar

\[
\text{Distance} = \frac{|Ad + Be + Cf - D|}{\sqrt{A^2 + B^2 + C^2}}.
\]
In particular, the distance from the origin to the plane $Ax + By + Cz = D$ is seen to be
$$\frac{|D|}{\sqrt{A^2 + B^2 + C^2}}.$$ Thus we have geometric interpretations for each part of the equation $Ax + By + Cz = D$. The coefficients tell us a normal vector to the plane, while the constant can be used to find the distance to the origin. Further, if $\mathbf{n} = \langle A, B, C \rangle$ is a unit vector, the constant $D$ is the distance to the origin.

**Exercises**

1. Find the angle between the following pairs of vectors.
   (a) $\mathbf{v} = \langle \sqrt{3}, 1 \rangle$ and $\mathbf{w} = \langle 7, 0 \rangle$.
   (b) $\mathbf{v} = \langle 0, -1, 1 \rangle$ and $\mathbf{w} = \langle 0, 5, 0 \rangle$.
   (c) $\mathbf{v} = \langle -3, 4, 2 \rangle$ and $\mathbf{w} = \langle 2, 1, 1 \rangle$.

2. Let $\mathbf{v}$ have initial point $(2, 1, 3)$ and terminal point $(3, -2, 2)$, and let $\mathbf{w} = \langle 2, 1, 1 \rangle$. Find the angle between $\mathbf{v}$ and $\mathbf{w}$.

3. Find an equation for the line in $\mathbb{R}^2$ through the origin and orthogonal to $\mathbf{v} = \langle 3, 7 \rangle$. Sketch both the line and the vector.

4. Give a geometric interpretation of $\mathbf{v} \cdot \mathbf{w}$ when both $\mathbf{v}$ and $\mathbf{w}$ are unit vectors.
5. Let $\mathbf{u} = \langle a, b, c \rangle$, $\mathbf{v} = \langle d, e, f \rangle$, and $\mathbf{w} = \langle g, h, k \rangle$ be vectors, and $\alpha$ a scalar. Verify the remaining parts of Theorem 2.2.1.

(a) $(\alpha \mathbf{v}) \cdot \mathbf{w} = \alpha (\mathbf{v} \cdot \mathbf{w})$
(b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
(c) $|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}$

6. If $\mathbf{v} \cdot \mathbf{w} < 0$, what can you say about the angle between them?

7. Find an equation for the plane through $(4, -1, -3)$ with normal vector $\mathbf{n} = \langle 2, 1, 1 \rangle$.

8. Find both unit vectors orthogonal to the plane $5x - 3y + 2z = 6$.

9. Find the projection of $\langle -2, 4, 3 \rangle$ onto $\langle 1, 1, 2 \rangle$.

10. The force due to gravity acting on a $5kg$ pendulum is $\mathbf{F}_{grav} = \langle 0, -49 \rangle$. Find the projection of $\mathbf{F}_{grav}$ in the direction of $\mathbf{F}_{tens}$ when the pendulum makes an angle of $60^\circ$ with the horizontal.

11. Find the work done by the constant force $\mathbf{F} = \langle -3, 5 \rangle$ on a particle moving a displacement $\langle d \rangle = \langle 2, 0 \rangle$.

12. A boy pulls a sled 40 meters with a force of $20\text{N}$ at $45^\circ$ to the horizontal. Find the work done by the boy on the sled.

13. Find the component of $\mathbf{v} = \langle 2, 7, 4 \rangle$ in the direction of $\mathbf{w} = \langle 3, 1, -2 \rangle$.

14. Describe the set of all unit vectors $\mathbf{u}$ for which $\cos \theta > 0$.

15. Find the distance from $(3, -1, 2)$ to the plane $2x + y - 3z = 4$.

16. Find a formula for the set of all points 3 units from the plane $x + y + z = 3$, on the $\mathbf{n} = \langle 1, 1, 1 \rangle$ side of it.

17. Find a unit vector that makes an angle of $\pi/3$ with the $y$-axis and $\pi/4$ with the $x$-axis.

18. Find an equation for the plane parallel to $3x - 3y + 2z = 5$ and through the point $(4, 3, 5)$.

19. Find an equation for the plane orthogonal to the line $\mathbf{x}(t) = (2 + t, 3 - 2t, 1 - t)$ and through $(1, 0, -3)$. 
2.3 The Cross Product

In this section we introduce the cross product of two vectors. Whereas the dot product produces a scalar, the cross product of two vectors produces a third vector with several desirable geometric properties. Before defining the cross product, we discuss some preliminary topics in matrices and determinants.

**Matrix Preliminaries**

Matrices are an important mathematical tool that arise in a variety of contexts. We will not use them much in this course, but need some familiarity with them. What follows is the briefest of introductions to a subject deserving of much more. For instance, we do not cover determinants in general but introduce one method that is useful in taking cross products. We begin by defining matrices.

**Definition 2.3.1.** An \( m \)-by-\( n \) matrix is a rectangular array with \( m \) rows and \( n \) columns. If \( A \) is a matrix, then \( a_{ij} \) represents the entry in the \( i \)th row and \( j \)th column.

The singular form of matrices is *matrix* NOT *matrice* (there is no such word).

**Example 2.3.1. Matrices**

The following are matrices. The left has numeric entries, while the first row of the second is the standard basis vectors, and the entries in the third are functions. To clarify notation, note that in the first matrix we have \( a_{23} = -2 \), in the second \( a_{32} = 17 \), and \( a_{22} = 2x^2 \) in the third.

\[
\begin{pmatrix}
5 & -7 & 3 \\
-14 & 3 & -2 \\
2 & 4 & 6 \\
\end{pmatrix}
\begin{pmatrix}
i & j & k \\
3 & 2 & 1 \\
4 & 17 & 3 \\
\end{pmatrix}
\begin{pmatrix}
2y^2 & 2 + 3y^2 \\
2 + 3y^2 & 2x^2 \\
\end{pmatrix}
\]

We will want to calculate determinants and products of matrices.

**Definition 2.3.2.** The determinant of the 2 \( \times \) 2 matrix \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \) is

\[
\begin{vmatrix}
a & b \\
c & d \\
\end{vmatrix} = ad - bc.
\]

Thus, for the third matrix above we calculate its determinant:

\[
\begin{vmatrix}
2y^2 & 2 + 3y^2 \\
2 + 3y^2 & 2x^2 \\
\end{vmatrix} = (2y^2)(2x^2) - (2 + 3y^2)^2 = 4x^2y^2 - 9y^4 - 12y^2 - 4.
\]

Determinants of 3 \( \times \) 3 matrices can be defined using a technique called expansion by cofactors. We don’t go into the general technique, but show one method for calculating it here:
2.3. THE CROSS PRODUCT

\[
\begin{vmatrix}
  a & b & c \\
  d & e & f \\
  g & h & i \\
\end{vmatrix}
= \begin{vmatrix}
  e & f \\
  h & i \\
\end{vmatrix} a - \begin{vmatrix}
  d & f \\
  g & i \\
\end{vmatrix} b + \begin{vmatrix}
  d & e \\
  g & h \\
\end{vmatrix} c
\]
(2.3.1)

In this instance we are expanding across the first row. Notice the entries in the first row are multiplied by the determinants of the $2 \times 2$ matrices you get by deleting the row and column they are in. For example, since $b$ is in the first row, second column, the $2 \times 2$ matrix is the one obtained by deleting the first row and second column. Also notice that it's an alternating sum: the first and last terms are added while the middle term is subtracted. We illustrate with an example:

\[
\begin{vmatrix}
  i & j & k \\
  3 & 2 & 1 \\
  4 & 17 & 3 \\
\end{vmatrix}
= \begin{vmatrix}
  2 & 1 \\
  17 & 3 \\
\end{vmatrix} i - \begin{vmatrix}
  3 & 1 \\
  4 & 3 \\
\end{vmatrix} j + \begin{vmatrix}
  3 & 2 \\
  4 & 17 \\
\end{vmatrix} k
\]
(2.3.2)

\[
= -11i - 5j + 45k.
\]

Notice that when the entries in the first row are the standard basis vectors, the determinant of the matrix is itself a vector. This will turn out to be the cross product of the rows. This really isn’t fair, since determinants are defined for matrices with scalar valued entries, but we make an exception in this case.

We will also have occasion to multiply matrices. Let $A$ be an $m \times n$ matrix and $B$ an $n \times p$ matrix, so that the number of columns in $A$ is the same as the number of rows in $B$. We define the product $C = AB$ to be the $m \times p$ matrix whose $ij^{th}$ entry is the dot product of the $i^{th}$ row of $A$ with the $j^{th}$ column of $B$. We illustrate with an example:

\[
\begin{pmatrix}
  3 & 4 \\
  -2 & 5 \\
\end{pmatrix}
\begin{pmatrix}
  2 & 3 & 4 \\
  5 & 6 & 7 \\
\end{pmatrix}
= \begin{pmatrix}
  (3,4) \cdot (2,5) & (3,4) \cdot (3,6) & (3,4) \cdot (4,7) \\
  (-2,5) \cdot (2,5) & (-2,5) \cdot (3,6) & (-2,5) \cdot (4,7) \\
\end{pmatrix}
\]

\[
= \begin{pmatrix}
  3 \cdot 2 + 4 \cdot 5 & 3 \cdot 3 + 4 \cdot 6 & 3 \cdot 4 + 4 \cdot 7 \\
  -2 \cdot 2 + 5 \cdot 5 & -2 \cdot 3 + 5 \cdot 6 & -2 \cdot 4 + 5 \cdot 7 \\
\end{pmatrix}
\]
\[
= \begin{pmatrix}
  26 & 33 & 40 \\
  21 & 24 & 27 \\
\end{pmatrix}.
\]

We remark that the requirement that the number of rows of $A$ equal the number of columns of $B$ is necessary. If that doesn’t hold, the product $AB$ is not defined. The reason is that without this requirement it would be impossible to take the dot product of a row of $A$ with a column of $B$.

**Algebraic Definition of $v \times w$:** With the matrix preliminaries in hand, we can define the cross product of two vectors.
CHAPTER 2. INTRODUCTION TO VECTORS

Definition 2.3.3. The cross product \( \mathbf{v} \times \mathbf{w} \) of the vectors \( \mathbf{v} = (a, b, c) \) and \( \mathbf{w} = (d, e, f) \) is the vector

\[
\mathbf{v} \times \mathbf{w} = \begin{vmatrix} i & j & k \\ a & b & c \\ d & e & f \end{vmatrix} = \begin{vmatrix} b & c \\ e & f \end{vmatrix} i - \begin{vmatrix} a & c \\ d & f \end{vmatrix} j + \begin{vmatrix} a & b \\ d & e \end{vmatrix} k 
\]

(2.3.3)

\[= (bf - ec)i - (af - dc)j + (ae - db)k.\]

Let’s interpret our determinant calculation in Equation 2.3.3 in cross product language. The calculation shows that if \( \mathbf{v} = (3, 2, 1) \) and \( \mathbf{w} = (4, 17, 3) \), then \( \mathbf{v} \times \mathbf{w} = (-11, -5, 45) \).

When we introduced the dot product, we took note of some of its algebraic properties in Theorem 2.2.1. For example, we saw that dot products distribute across vector sums, and scalars can be “factored out.” Further, dot products are commutative. We take a moment to highlight some algebraic properties of the cross product.

Theorem 2.3.1. Let \( \mathbf{v}, \mathbf{w}, \) and \( \mathbf{x} \) be vectors, and let \( \alpha \) be a scalar. Then

1. \( \mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v} \)
2. \((\alpha \mathbf{v}) \times \mathbf{w} = \alpha (\mathbf{v} \times \mathbf{w})\)
3. \((\mathbf{v} + \mathbf{w}) \times \mathbf{x} = \mathbf{v} \times \mathbf{x} + \mathbf{w} \times \mathbf{x}\)

Proof. We prove the first property. Let \( \mathbf{v} = (a, b, c) \) and \( \mathbf{w} = (d, e, f) \). Using the definition of the cross product, we see that

\[
\mathbf{w} \times \mathbf{v} = \begin{vmatrix} i & j & k \\ d & e & f \\ a & b & c \end{vmatrix} = \begin{vmatrix} e & f \\ a & c \end{vmatrix} i - \begin{vmatrix} d & f \\ a & e \end{vmatrix} j + \begin{vmatrix} d & e \\ a & b \end{vmatrix} k
\]

\[= (ec - bf)i - (dc - af)j + (db - ae)k.\]

Comparing this with Equation 2.3.3 shows that \( \mathbf{v} \times \mathbf{w} = -\mathbf{w} \times \mathbf{v} \). □

The remaining facts are left as exercises for the reader. One should observe that the cross product does not commute! Since switching the order introduces a negative sign, we say it anti-commutes. The cross product, however, does behave as you might expect with respect to scalar multiplication and vector addition. Scalars can be “factored” out, and the cross product “distributes” across vector addition.

Example 2.3.2. Algebraic properties of \( \mathbf{v} \times \mathbf{w} \)

Use the properties of Theorem 2.3.1 to show that \((\alpha \mathbf{v}) \times (\beta \mathbf{w}) = \alpha \beta (\mathbf{v} \times \mathbf{w})\).

Notice that property 2 of Theorem 2.3.1 only states that you can factor scalars out of the first vector in a cross product, while this example states you
can factor them from either/both vectors. This result follows from repeated applications of properties 1 and 2. The right hand column justifies each equality.

\[
(\alpha \mathbf{v}) \times (\beta \mathbf{w}) = \alpha (\mathbf{v} \times (\beta \mathbf{w})) \quad \text{Property 2}
\]
\[
= \alpha (- (\beta \mathbf{w}) \times \mathbf{v}) \quad \text{Property 1}
\]
\[
= \alpha \beta (\mathbf{w} \times \mathbf{v}) \quad \text{Property 2}
\]
\[
= \alpha \beta (\mathbf{v} \times \mathbf{w}). \quad \text{Property 1}
\]

**Geometric Properties of \( \mathbf{v} \times \mathbf{w} \):** It turns out that the cross product has geometric significance, which involves the notion of the right-hand rule for vectors. Let two vectors \( \mathbf{v}, \mathbf{w} \) have the same initial point. Let \( \mathbf{x} \) be orthogonal to both \( \mathbf{v} \) and \( \mathbf{w} \), then \( \mathbf{v}, \mathbf{w} \) and \( \mathbf{x} \) satisfy the right-hand rule if when you stretch the fingers of your right hand along \( \mathbf{v} \) and curl them toward \( \mathbf{w} \) through the acute angle between them, your thumb points in the direction of \( \mathbf{x} \). For example, the vectors \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) satisfy the right-hand rule, but \( \mathbf{j}, \mathbf{i}, \mathbf{k} \) don’t. There is an easy algebraic check to see if an ordered list of three vectors satisfy the right-hand rule. The non-zero vectors \( \mathbf{v}, \mathbf{w} \) and \( \mathbf{x} \) satisfy the right hand rule if and only if the determinant of the matrix with the vectors as rows is positive. In symbols: the non-zero vectors \( \mathbf{v}, \mathbf{w} \) and \( \mathbf{x} \) satisfy the right hand rule if and only if

\[
\begin{vmatrix}
  a & b & c \\
  d & e & f \\
  g & h & i \\
\end{vmatrix} > 0.
\]

The proof of this fact is beyond the scope of this text, and we omit it. We are now ready to state the following theorem.

**Theorem 2.3.2.** Let \( \mathbf{v} \) and \( \mathbf{w} \) be non-zero vectors. Then

1. \( \mathbf{v} \times \mathbf{w} \) is orthogonal to both \( \mathbf{v} \) and \( \mathbf{w} \), and the ordered list of vectors \( \mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w} \) satisfies the right-hand rule.

2. \( \|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\|\|\mathbf{w}\| \sin \theta \), where \( \theta \) is the angle between \( \mathbf{v} \) and \( \mathbf{w} \).

**Proof.** The fact that \( \mathbf{v} \times \mathbf{w} \) is orthogonal to both \( \mathbf{v} \) and \( \mathbf{w} \) is proven by showing the corresponding dot products are zero. To see that \( \mathbf{v}, \mathbf{w}, \mathbf{v} \times \mathbf{w} \) satisfies the right hand rule, one can take the determinant of the matrix with these vectors as rows to see it is positive.

We’ll prove part 2 letting \( \mathbf{v} = \langle a, b, c \rangle \) and \( \mathbf{w} = \langle d, e, f \rangle \). Let \( T \) be the area of the triangle determined by placing \( \mathbf{v} \) and \( \mathbf{w} \) at the same initial point. Then \( 2T \) is the area of the parallelogram spanned by \( \mathbf{v} \) and \( \mathbf{w} \), which is \( 2T = \|\mathbf{v}\|\|\mathbf{w}\| \sin \theta \). To see this, consider Figure 2.2.3 of the last section. Note that \( \|\mathbf{v}\| \sin \theta \) is the height of the dotted line and \( \|\mathbf{w}\| \) is the length of the base of the parallelogram, so \( 2T = \|\mathbf{v}\|\|\mathbf{w}\| \sin \theta \). Since \( \sin^2 \theta = 1 - \cos^2 \theta \), and \( \cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \), squaring the area of the parallelogram we get
(2T)^2 = (\|v\|\|w\| \sin \theta)^2 = \|v\|^2 \|w\|^2 (1 - \cos^2 \theta)
= \|v\|^2 \|w\|^2 \left(1 - \left(\frac{v \cdot w}{\|v\| \|w\|}\right)^2\right)
= \|v\|^2 \|w\|^2 - (v \cdot w)^2 = (a^2 + b^2 + c^2)(d^2 + e^2 + f^2) - (ad + be + cf)^2
= (ac - bd)^2 + (af - dc)^2 + (bf - ce)^2 = \|v \times w\|^2.

Therefore \|v \times w\| = 2T = \|v\|\|w\| \sin \theta$, as desired. $\square$

Remark: Implicit in the proof of Theorem 2.3.2 is the fact that \(\|v \times w\|\) is the area of the parallelogram determined by \(v\) and \(w\).

Corollary 2.3.1. Two non-zero vectors \(v, w\) are parallel if and only if
\[v \times w = 0.\]

Proof. The non-zero vectors \(v\) and \(w\) are parallel if and only if \(\theta = 0\) or \(\theta = \pi\). By Theorem 2.3.2(2), this is true if and only if \(v \times w = 0\). $\square$

Concept Connection: Since \(\|v\|\|w\| \sin \theta\) is the area of the parallelogram determined by \(v\) and \(w\), Theorem 2.3.2 has significant geometric consequences. In particular, this will be useful when developing integral formulas for surface area and, more generally, for integrating functions on surfaces. We will approximate the surface area using parallelogram-shaped tiles, much like the tiles on a roof (see Figure 4.5.2). The cross product will allow us to calculate areas in our approximation. The area of the surface, of course, will be the integral obtained by taking the limit as our approximation gets finer.

We interpret Theorem 2.3.2 when considering the cross product of two vectors in \(\mathbb{R}^2\). The planar vectors \(v = (a, b)\) and \(w = (c, d)\) can be thought of as vectors in \(\mathbb{R}^3\) with zero in the \(z\)-coordinate. In this case they are \(v = (a, b, 0)\) and \(w = (c, d, 0)\), and have cross product \(v \times w = (ad - bc)\mathbf{k}\). This means that the (absolute value of the) \(2 \times 2\)-determinant \[\begin{vmatrix} a & b \\ c & d \end{vmatrix} = (ad - bc)\] is the area of the parallelogram spanned by the rows.

We now consider several applications of Theorem 2.3.2. First, we’ll see that while the dot product determines when non-zero vectors are perpendicular, the cross product determines when they are parallel. Further, since a plane is determined by a point and a normal direction, Theorem 2.3.2(1) can be used to find equations for planes. We can also use it to parameterize lines of intersections of planes. We’ll also look at the triple scalar product, \((v \times w) \cdot x\), and its geometric meaning.

Example 2.3.3. Parametric to Cartesian equations of planes
Find a Cartesian equation for the plane given by the parametric equations \( x(s, t) = (2 - s + 2t, 3 + 2s - t, 4 - s + t) \).

We need a normal vector and a point on the plane. A point on the plane is \( x(0, 0) = (2, 3, 4) \). To find a normal vector, we find two vectors in the plane and take their cross product. Isolating the parameters gives

\[
x(s, t) = (2 - s + 2t, 3 + 2s - t, 4 - s + t) = (2, 3, 4) + s(-1, 2, -1) + t(2, -1, 1).
\]

Therefore the cross product of \(-1, 2, -1\) and \(2, -1, 1\) gives a normal vector. We get

\[
\begin{vmatrix}
i & j & k \\
-1 & 2 & -1 \\
2 & -1 & 1 \\
\end{vmatrix} = \begin{vmatrix} i & -1 & -1 \\
-1 & 2 & 1 \\
5 & -3 & 2 \\
\end{vmatrix} \cdot \begin{vmatrix} j & -1 & -2 \\
-1 & 1 & 2 \\
-3 & -1 & 1 \\
\end{vmatrix} k = i - j - 3k.
\]

Putting these together we get the equation for the plane:

\[
x - y - 3z = (1, -1, -3) \cdot (2, 3, 4) = -13.
\]

**Cartesian Equations of Planes given vectors in the plane**

To find a Cartesian equation for a plane containing the point \( P \) and the vectors \( v \) and \( w \):

1. Find the normal vector \( n = v \times w \).
2. Use the equation \( n \cdot \langle x, y, z \rangle = n \cdot P \)

**Example 2.3.4. A plane through three points**

Find a Cartesian equation for the plane through \( P = (-2, 3, 1), Q = (3, 4, -2) \), and \( R = (5, -3, 2) \).

First we find two vectors in the plane, then take their cross product to find a normal vector to the plane. Then we’ll have the information needed to write down a Cartesian equation. Two vectors in the plane are \( v = Q - P = \langle 3, 4, -2 \rangle - \langle -2, 3, 1 \rangle = \langle 5, 1, -3 \rangle \) and \( w = R - P = \langle 5, -3, 2 \rangle - \langle -2, 3, 1 \rangle = \langle 7, -6, 1 \rangle \). Their cross product is

\[
\begin{vmatrix}
i & j & k \\
5 & 1 & -3 \\
7 & -6 & 1 \\
\end{vmatrix} = \begin{vmatrix} i & -3 \\
-6 & 1 \\
5 & 7 \\
\end{vmatrix} + \begin{vmatrix} j & -3 \\
-6 & 1 \\
5 & 7 \\
\end{vmatrix} k = -17i - 26j - 37k
\]
Thus \( \mathbf{n} = \langle 17, 26, 37 \rangle \) is normal to the plane (Pop quiz: what happened to the negative sign?), \((-2, 3, 1)\) is on the plane, and an equation for the plane is

\[
17x + 26y + 37z = \langle 17, 26, 37 \rangle \cdot (-2, 3, 1) = 81.
\]

**Example 2.3.5. Lines of intersection of planes**

Parameterize the line of intersection of \(2x - 3y + z = 5\) and \(x + y - 2z = 3\).

To parameterize a line we need to know a point on the line and a vector in the line. To find a point on the line, we’ll decide to find where it intersects the \(xz\)-plane. Then \(y = 0\), and we’ve reduced the problem to finding the point of intersection of the lines \(2x - z = 5\) and \(x - 2z = 3\). Your favorite method of solving systems yields \(x = 7/3, z = -1/3\), so the point \((7/3, 0, -1/3)\) is on the line.

To find the direction of the line, we make a key observation. If a line lies in a plane, then its direction is orthogonal to the normal vector! Since our line lies on both planes, its direction is orthogonal to both normal vectors. Thus the cross product of the normal vectors gives the direction of the line. We calculate:

\[
\begin{vmatrix}
  \mathbf{i} & \mathbf{j} & \mathbf{k} \\
  2 & -3 & 1 \\
  1 & 1 & -2 \\
\end{vmatrix} = \begin{vmatrix}
  -3 & 1 \\
  1 & -2 \\
\end{vmatrix} \mathbf{i} + \begin{vmatrix}
  2 & 1 \\
  1 & 1 \\
\end{vmatrix} \mathbf{j} + \begin{vmatrix}
  2 & -3 \\
  1 & 1 \\
\end{vmatrix} \mathbf{k}
\]

\[= 5\mathbf{i} + 5\mathbf{j} + 5\mathbf{k},\]

which means that the vector \(\mathbf{v} = (1, 1, 1)\) lies in the direction of the line (Note that if \(\mathbf{v}\) lies on the line, so does any scalar multiple of it. Thus we have the equation

\[
\mathbf{x}(t) = \left(\frac{7}{3} + t, t, \frac{1}{3} + t\right).
\]

**The Triple Scalar Product:** We now combine the cross and dot products to define the triple scalar product of three vectors \(\mathbf{v}, \mathbf{w},\) and \(\mathbf{x}\). It turns out that this product gives the volume of the parallelepiped determined by the vectors. This interpretation is central to the development of a formula for finding the flux of a vector field across a surface in Chapter 5.

**Definition 2.3.4.** The triple scalar product of the ordered triple of vectors \(\mathbf{v}, \mathbf{w},\) and \(\mathbf{x}\) is \((\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x}\).

When calculating the triple scalar product, order of operations is important. First compute the cross product \(\mathbf{v} \times \mathbf{w}\), then dot the resulting vector with \(\mathbf{x}\). That’s why we say an “ordered” list. There is a quick way to calculate the triple scalar product using determinants.
Lemma 2.3.1. The triple scalar product of the vectors $\mathbf{v} = \langle a, b, c \rangle$, $\mathbf{w} = \langle d, e, f \rangle$, and $\mathbf{x} = \langle g, h, i \rangle$ is the determinant:

$$(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x} = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}.$$ 

Proof. To compute the triple scalar product we first calculate the cross product:

$$\mathbf{v} \times \mathbf{w} = (bf - ec)i - (af - dc)j + (ae - db)k.$$ 

Using this, we calculate the dot product

$$(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x} = ((bf - ec)i - (af - dc)j + (ae - db)k) \cdot (gi + hj + ik).$$

Expanding this calculation, and comparing it with Equation 2.3.1 completes the proof. □

Example 2.3.6. Triple scalar product

Find the triple scalar product of the vectors $\mathbf{v} = \langle -3, -2, 3 \rangle$, $\mathbf{w} = \langle 1, -3, 2 \rangle$, and $\mathbf{x} = \langle 2, 2, 3 \rangle$. By the above remarks, this amounts to calculating the determinant:

$$(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x} = \begin{vmatrix} -3 & -2 & 3 \\ 1 & -3 & 2 \\ 2 & 2 & 3 \end{vmatrix} = -3 \cdot \begin{vmatrix} -3 & 2 \\ 2 & 3 \end{vmatrix} - (-2) \cdot \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & -3 \\ 2 & 2 \end{vmatrix} = 61.$$ 

Any three vectors $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{x}$ determine a parallelepiped, as in Figure 2.3.1, which has a volume. When the vectors are ordered, we can define a signed volume of the parallelepiped. Take the volume to be positive if $\mathbf{v} \times \mathbf{w}$ and $\mathbf{x}$ point “in the same direction”, and negative otherwise. Of course, by pointing “in the same direction” we really mean $\text{comp}_{\mathbf{v} \times \mathbf{w}} \mathbf{x} > 0$. Since $\text{comp}_{\mathbf{v} \times \mathbf{w}} \mathbf{x} = (\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x} / \|\mathbf{v} \times \mathbf{w}\|$, and since $\|\mathbf{v} \times \mathbf{w}\| \geq 0$, we see that $\mathbf{v} \times \mathbf{w}$ and $\mathbf{x}$ point in the same direction if the numerator $(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x}$ is positive.

Theorem 2.3.3. The triple scalar product $(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x}$ is the signed volume of the parallelepiped determined by the ordered list of vectors $\mathbf{v}$, $\mathbf{w}$, and $\mathbf{x}$.

Proof. The volume of a parallelepiped is the area of the base parallelogram times the height. By Theorem 2.3.2, we know that $\|\mathbf{v} \times \mathbf{w}\|$ is the area of the base parallelogram. Using Theorem 2.2.2 we see

$$(\mathbf{v} \times \mathbf{w}) \cdot \mathbf{x} = \|\mathbf{v} \times \mathbf{w}\| \cdot \|\mathbf{x}\| \cos \psi,$$
where \( \psi \) is the angle between \( \mathbf{v} \times \mathbf{w} \) and \( \mathbf{x} \). If we can show that \( \|\mathbf{x}\| \cos \psi \) is the signed height of the parallelepiped, we are done. Since \( \mathbf{v} \times \mathbf{w} \) is orthogonal to both \( \mathbf{v} \) and \( \mathbf{w} \), \( \text{comp}_{\mathbf{v}\times\mathbf{w}} \mathbf{x} = \|\mathbf{x}\| \cos \psi \) is the signed height of the parallelepiped (see Figure 2.3.1). \( \square \)

At this point it may seem strange to discuss signed volume. In applications, however, the sign can determine whether net flow is going into or out of a solid, or whether the net charge of an electric field is positive or negative.

**Physical Motivation for \( \mathbf{v} \times \mathbf{w} \):** The geometric properties of \( \mathbf{v} \times \mathbf{w} \) are often used to define it. More precisely, one can define \( \mathbf{v} \times \mathbf{w} \) to be the unique vector satisfying:

1. \( \|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \|\mathbf{w}\| \sin \theta \).
2. \( \mathbf{v} \times \mathbf{w} \) is orthogonal to both \( \mathbf{v} \) and \( \mathbf{w} \), and \( \mathbf{v} \), \( \mathbf{w} \), and \( \mathbf{v} \times \mathbf{w} \) satisfy the right hand rule.

Then one derives Equation 2.3.3 for calculating \( \mathbf{v} \times \mathbf{w} \) using these properties.

We now show that, given the right interpretation of vectors \( \mathbf{v} \) and \( \mathbf{w} \), the vector \( \mathbf{v} \times \mathbf{w} \) is the velocity vector of a point rotating about an axis. Suppose a fluid is rotating about an axis and that \( P \) is a point on the axis. Now suppose that \( \mathbf{v} \) has its initial point at \( P \), and points in the direction of the axis so that the fluid rotates according to the right hand rule. In other words, if you point your right thumb along \( \mathbf{v} \), your fingers curl around \( \mathbf{v} \) in the direction the fluid is rotating. Let \( \alpha \) measure the angle about the axis of rotation, and suppose that \( \|\mathbf{v}\| \) is the angular velocity of the fluid \( \omega = \frac{d\alpha}{dt} \). Thus the vector \( \mathbf{v} \) completely describes how the fluid is flowing: its direction determines how the fluid rotates and its length determines the speed of rotation since \( \|\mathbf{v}\| = \frac{d\alpha}{dt} \).

Now let \( Q \) be any point in the fluid, and let \( \mathbf{w} \) be the vector from \( P \) to \( Q \). As the fluid flows around \( \mathbf{v} \), the point \( Q \) traces out a circle and there is a velocity vector \( \mathbf{v}_Q \) for \( Q \). Moreover, we claim the cross product \( \mathbf{v} \times \mathbf{w} \) is the velocity vector \( \mathbf{v}_Q \).

To see this we must show that the vectors \( \mathbf{v} \), \( \mathbf{w} \) and \( \mathbf{v}_Q \) satisfy the right hand rule and that the speed \( \|\mathbf{v}_Q\| \) is \( \|\mathbf{v}\| \|\mathbf{w}\| \sin \theta \). Then \( \mathbf{v}_Q \) satisfies the properties.
that determine $\mathbf{v} \times \mathbf{w}$, which implies $\mathbf{v}_Q = \mathbf{v} \times \mathbf{w}$. Placing the fingers of one’s right hand along $\mathbf{v}$ in Figure 2.3.2 and curling them toward $\mathbf{w}$ points the thumb in the direction of $\mathbf{v}_Q$. Thus $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{v}_Q$ satisfy the right-hand rule. As for $\|\mathbf{v}_Q\|$, note that $Q$ travels in a circle of radius $r = \|\mathbf{w}\| \sin \theta$ (see Figure 2.3.2). Further, since the radian measure of $\alpha$ is $\frac{s}{r}$ where $s$ is arc length and $r$ is radius, we see that $s = r \alpha$. Recall that $r$ is constant, so differentiating with respect to time gives $\frac{ds}{dt} = r \frac{d\alpha}{dt} = r \|\mathbf{v}\|$. Combining this with our expression for $r$ gives:

$$\|\mathbf{v}_Q\| = \frac{ds}{dt} = r \frac{d\alpha}{dt} = (\|\mathbf{w}\| \sin \theta) \|\mathbf{v}\| = \|\mathbf{v} \times \mathbf{w}\|.$$ 

Thus $\mathbf{v}_Q$ is the vector satisfying the right-hand rule with $\mathbf{v}$ and $\mathbf{w}$ with the appropriate length, implying that $\mathbf{v}_Q = \mathbf{v} \times \mathbf{w}$.

**Exercises**

1. Compute the determinants of the following matrices

   $$\begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ -2 & 3 \end{pmatrix}, \begin{pmatrix} 3x & -4xy \\ 6y^2 & -8x \end{pmatrix}, \begin{pmatrix} 1 & 3 \\ -2 & -6 \end{pmatrix}$$

2. Compute the determinants of the following matrices

   $$\begin{pmatrix} 3 & -4 & 2 \\ -2 & 3 & 1 \\ 0 & 2 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 4 & 1 \\ -2 & 3 & 2 \\ 0 & 7 & 3 \end{pmatrix}$$

3. Compare the determinants of the matrices with two rows interchanged:

   $$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \begin{vmatrix} a & b & c \\ g & h & i \\ d & e & f \end{vmatrix}$$
4. Compare the determinants of the matrices with rows interchanged:

\[
\begin{vmatrix}
  a & b & c \\
  d & e & f \\
  g & h & i \\
\end{vmatrix}
\quad \quad \quad
\begin{vmatrix}
  g & h & i \\
  a & b & c \\
  d & e & f \\
\end{vmatrix}
\]

5. For what values of \( x \) and \( y \) is the determinant \( \begin{vmatrix} 2y & 2x \\ 2x & 6y \end{vmatrix} \) positive?

6. For what values of \( x \) and \( y \) is the determinant \( \begin{vmatrix} 6xy & 3x^2 \\ 3x^2 & 0 \end{vmatrix} \) positive?

7. Perform the following matrix multiplications:

\[
(a) \begin{pmatrix} 2 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \quad \quad \quad (b) \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -2 & 3 \end{pmatrix}
\]

Is matrix multiplication commutative?

8. Multiply the following matrices, if possible:

\[
(a) \begin{pmatrix} -1 & 3 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & 4 \end{pmatrix} \quad \quad \quad (b) \begin{pmatrix} 4 & -3 & 2 \\ 6 & 7 & -2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}
\]

9. Calculate the cross product of the following pairs of vectors:

\[
(a) \mathbf{v} = \langle 2, 1, 0 \rangle, \quad \mathbf{w} = \langle 4, 3, 0 \rangle \quad \quad \quad (b) \mathbf{v} = \langle -3, 2, 0 \rangle, \quad \mathbf{w} = \langle 6, -4, 0 \rangle \\
(c) \mathbf{v} = \langle 3, 0, -2 \rangle, \quad \mathbf{w} = \langle 0, 4, -1 \rangle \quad \quad \quad (d) \mathbf{v} = \langle 1, 2, -2 \rangle, \quad \mathbf{w} = \langle -\frac{1}{2}, -1, 1 \rangle
\]

10. Find the area of the parallelogram determined by the vectors \( \mathbf{v} = \langle 3, 1 \rangle \) and \( \mathbf{w} = \langle -2, 3 \rangle \), and sketch it.

11. Find the area of the parallelogram determined by the vectors \( \mathbf{v} = \langle 5, -2, 3 \rangle \) and \( \mathbf{w} = \langle 2, -1, -3 \rangle \).

12. This problem involves scalar multiples of the vectors \( \mathbf{v} = \langle 3, -2, 1 \rangle \) and \( \mathbf{w} = \langle 1, 3, -2 \rangle \).
   
   (a) Calculate the area of the parallelogram determined by \( \mathbf{v} \) and \( \mathbf{w} \).
   
   (b) Calculate the area of the parallelogram determined by the vectors \( \mathbf{v} = \Delta s \langle 3, -2, 1 \rangle \) and \( \mathbf{w} = \Delta t \langle 1, 3, -2 \rangle \), where \( \Delta s \) and \( \Delta t \) are scalars.

   How is it related to your answer in part (a)?

13. Find a Cartesian equation for the plane through the point \( (3, 3, 2) \) and containing the vectors \( \mathbf{v} = \langle 2, -1, 2 \rangle \) and \( \mathbf{w} = \langle 1, 3, 1 \rangle \).

14. Find a Cartesian equation for the plane through the origin and containing the vectors \( \mathbf{v} = \langle 3, 1, 5 \rangle \) and \( \mathbf{w} = \langle -2, 1, 1 \rangle \).
15. Find a Cartesian equation for the plane given parametrically by $\mathbf{x}(s, t) = (3 - s + 2t, 1 + 2s + t, -2 + s - 3t)$, $-\infty < s, t < \infty$.

16. Find a Cartesian equation for the plane given parametrically by $\mathbf{x}(s, t) = (4 + 3s + 2t, -3 - 2s + t, 5s + t)$, $-\infty < s, t < \infty$.

17. Parameterize the line of intersection of the planes $3y - 2z = 4$ and $6x - 2y + 3z = 2$.

18. Parameterize the line of intersection of the planes $x + y + 3z = 6$ and $2x - y + z = 4$.

19. Find a Cartesian equation for the plane containing the line $\mathbf{x}(t) = (2 - t, 3 + 2t, t)$, $-\infty < t < \infty$ and perpendicular to the plane $x - y + z = 3$.

20. Find a Cartesian equation for the plane through the points $(-2, 3, 4)$, $(1, -1, 2)$ and $(3, 5, -1)$.

21. Find the volume of the parallelepiped spanned by the vectors $(2, 1, 3)$, $(4, -2, 2)$, and $(-1, 1, 1)$. 
2.4 Calculus with Parametric Curves

In this section we take a look at calculus with parametric curves. Finding the velocity vector for a curve follows from what we know about vector addition, scalar multiplication, and taking limits. Thus it is a nice application of the vector topics we’ve been discussing in this chapter. Arclength is also shown to be the integral of the speed of a curve. Moreover, tangents to curves provide another context in which we encounter parametric lines and planes, as well as applications of work. Without further ado, we consider the problem of finding velocity vectors of curves.

**Tangents to Curves:** You can think of a parametric curve \( C(t) = \langle x(t), y(t), z(t) \rangle \) as describing the position at time \( t \) of a bug flying through space. With this interpretation it makes sense to talk about the velocity of the bug. Now velocity is a vector quantity, since it has both magnitude and direction. Moreover, it should still be the rate of change of position with respect to time. In this context, the change in position is the vector \( C(t + h) - C(t) \) (see figure 2.4.1).

So dividing by the change in time and taking a limit we get:

\[
C'(t) = \lim_{h \to 0} \frac{C(t + h) - C(t)}{h} = \lim_{h \to 0} \frac{1}{h} \left( C(t + h) - C(t) \right)
\]

\[
= \lim_{h \to 0} \frac{1}{h} \left( \langle x(t + h), y(t + h), z(t + h) \rangle - \langle x(t), y(t), z(t) \rangle \right)
\]

\[
= \lim_{h \to 0} \frac{1}{h} \left( \langle x(t + h) - x(t), y(t + h) - y(t), z(t + h) - z(t) \rangle \right)
\]

\[
= \lim_{h \to 0} \left( \frac{x(t + h) - x(t)}{h}, \frac{y(t + h) - y(t)}{h}, \frac{z(t + h) - z(t)}{h} \right)
\]

\[
= \langle x'(t), y'(t), z'(t) \rangle. \quad (2.4.1)
\]

Therefore, to find the tangent vector \( C'(t) \) you differentiate component-wise! The basic idea is that differentiation involves vector substraction \( C(t + h) - C(t) \).
2.4. CALCULUS WITH PARAMETRIC CURVES

$C(t)$ and scalar multiplication (dividing by $h$). Since both of those operations are done component-wise, so is differentiation.

The tangent vector $C'(t)$ is usually thought of as having initial point $C(t)$. Notice that $C'(t)$ gives the tangent vector at $C(t)$ for every value of $t$. We say that $C'(t)$ is a vector field along the curve $C(t)$. In Section 2.5 we will consider more general vector fields. The unit tangent vector $T(t)$ is the one with length one pointing in the same direction as $C''(t)$. Normalizing the tangent vector we see

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

**Example 2.4.1. The tangent vector field of a spiral**

Let $C(t) = (t \cos t, t \sin t)$ for $t \geq 0$. The tangent vector is

$$C'(t) = (\cos t - t \sin t, \sin t + t \cos t).$$

See Figure 2.4.2. Note that when the initial point of $C'(t)$ is placed at the terminal point of the position vector $C'(t)$ it is tangent to the curve in the usual sense.

![Figure 2.4.2: Tangent Vectors to a Spiral.](image)

One can calculate that $\|C'(t)\| = \sqrt{(\cos t - t \sin t)^2 + (\sin t + t \cos t)^2} = \sqrt{1 + t^2}$. Thus we have the unit tangent vector field

$$T(t) = \left\langle \frac{\cos t - t \sin t}{\sqrt{1 + t^2}}, \frac{\sin t + t \cos t}{\sqrt{1 + t^2}} \right\rangle.$$

The vectors in $T(t)$ are all one unit long, and a little hard to see in the right-hand picture above since they lie so close to the curve.
Remark: We pause now to remark that equation 2.4.1 can be interpreted in terms of approximations. Instead of saying

\[ C'(t) = \lim_{h \to 0} \frac{C(t+h) - C(t)}{h}, \]

one could say that \( C'(t) \) is approximately the difference quotient

\[ C'(t) \approx \frac{C(t+h) - C(t)}{h} \]

for small values of \( h \). “Solving” this approximation for \( C(t+h) \) shows that the point on the curve at time \( t + h \) can be approximated by adding the position vector \( C(t) \) and the scaled tangent vector \( hC'(t) \). Analytically we have the approximation

\[ C(t+h) \approx C(t) + hC'(t), \]

which is illustrated in Figure 2.4.3. This is completely analogous to, and follows from, the single variable case. It will be used in Section 3.4 to approximate patches of surfaces using pieces of tangent planes.

Example 2.4.2. Parameterizing Tangent Lines

Parameterize the tangent line to the curve \( C(t) = (\cos t, t, \sin t) \) at \( t = \pi/3 \).

Recall that to parameterize a line you need a point on the line and a vector in the direction of the line. In this instance \( C(\pi/3) \) is the point of tangency, so it is on the line, and the vector \( C'(\pi/3) \) is in the direction of the line. Thus the line is parameterized by

\[
\ell(t) = C(\pi/3) + tC'(\pi/3) = \left\langle \frac{1}{2}, \frac{\pi}{3}, \frac{\sqrt{3}}{2} \right\rangle + t \left\langle -\frac{\sqrt{3}}{2}, 1, \frac{1}{2} \right\rangle
\]

\[
= \left\langle \frac{1}{2} - \frac{\sqrt{3}}{2} t, \frac{\pi}{3} + t, \frac{\sqrt{3}}{2} + \frac{1}{2} t \right\rangle, \quad -\infty < t < \infty.
\]
Given a parameteric curve, then we can easily parameterize tangent lines. We summarize the strategy below.

**Tangent Lines to Parametric Curves**

A parameterization of the tangent line \( \ell(t) \) to the parametric curve \( C(t) \) at \( t = a \) is

\[
\ell(t) = C(a) + t C'(a), \quad -\infty < t < \infty.
\]

**Example 2.4.3. Tangent lines to curves of intersection**

Parameterize the tangent line to the curve of intersection of the generalized cylinder \( y^2 + z^2 = 4 \) and the surface \( x = y^2 - z^2 \) at the point \( (0, \sqrt{2}, \sqrt{2}) \) (see Figure 2.4.4).

![Figure 2.4.4: Tangent to curve of intersection](image)

One easily sees that the point is on both surfaces, hence on the curve of intersection. We know how to find tangent lines to parametric curves, so our strategy will be to parameterize the curve first. Recall from Section 1.4 that to parameterize the intersection of a generalized cylinder and a function, one follows:

**Step 1.** Parameterized the curve \( y^2 + z^2 = 4 \) in the \( yz \)-plane. This is a circle, so the parametric equations are

\[
y = 2 \cos t, \quad z = 2 \sin t, \quad 0 \leq t \leq 2\pi.
\]

**Step 2.** Substitute the parametric equations from step one into the function \( x = y^2 - z^2 \) to get the component function for \( x \). Thus \( x = (2 \cos t)^2 + (2 \sin t)^2 = \)
4 \cos 2t, using the fact that \cos 2t = \cos^2 t - \sin^2 t. The curve is then
\[ C(t) = (4 \cos 2t, 2 \cos t, 2 \sin t), \quad 0 \leq t \leq 2\pi. \]

Now that the curve is parameterized, we find out what value of \( t \) gives
\[ C(t) = (4 \cos 2t, 2 \cos t, 2 \sin t) = (0, \sqrt{2}, \sqrt{2}). \] Inspection verifies that \( t = \pi/4 \) does the trick, so the tangent line is
\[ \ell(t) = C(\pi/4) + t C'(\pi/4) = \langle 0, \sqrt{2}, \sqrt{2} \rangle + t \langle -8, -\sqrt{2}, \sqrt{2} \rangle, \quad -\infty < t < \infty. \]

**Example 2.4.4.** A plane normal to a curve

Find a Cartesian equation for the plane normal to the curve \( C(t) = (t, t^2, t^3) \) at \( t = 1 \).

Recall that to find a Cartesian equation for a plane we need a normal vector and a point on the plane. In this context, \( C(1) \) will be a point on the plane, and \( C'(1) \) will be a normal vector. Calculating, we see the point and normal direction are given by
\[ C(1) = (1, 1, 1), \quad C'(1) = \langle 1, 2, 3 \rangle. \]

Thus an equation for the plane is
\[ \langle 1, 2, 3 \rangle \cdot \langle x, y, z \rangle = \langle 1, 2, 3 \rangle \cdot \langle 1, 1, 1 \rangle \]
\[ x + 2y + 3z = 6. \]

**Speed and Arclength:** Suppose we wanted to calculate how fast the bug was flying at a given instant–its speed. Speed is a scalar quantity—it has no direction, just a magnitude. Average speed is the distance traveled divided by time. The distance traveled over the time interval \((t, t + h)\) is approximately \( \|C(t + h) - C(t)\| \), the length of the vector in Figure 2.4.1. Notice that this is not exactly the distance traveled, since the vector is the shortest distance between the points, and the curve may be longer. It is, however, a close enough estimate for small values of \( h \). Since the instantaneous speed is the limit as \( h \to 0 \) of the average speed, we see
\[
\text{Instantaneous Speed} = \lim_{h \to 0} \frac{\|C(t + h) - C(t)\|}{h} = \lim_{h \to 0} \frac{C(t + h) - C(t)}{h} = \|C'(t)\|.
\]

Because of this calculation, we make the following definition.

**Definition 2.4.1.** The speed of a curve \( C(t) \) is the length, \( \|C'(t)\| \), of its tangent vector.

**Example 2.4.5.** The speed of a spiral
The speed of the spiral $C(t) = \langle t \cos t, t \sin t \rangle$ of the previous example is $\|C'(t)\| = \sqrt{1 + t^2}$.

**Example 2.4.6. An elliptical helix**

At what point is the curve $C(t) = \langle 2 \cos t, \sin t, t \rangle$ traveling the fastest (see Figure 2.4.5)?

![Image of an elliptical helix](image)

Figure 2.4.5: Elliptical Helix

We are asked to maximize the speed. To do so, we first find the speed as a function of time. Differentiating we see $C'(t) = \langle -2 \sin t, \cos t, 1 \rangle$. Thus the speed is $\|C'(t)\| = \sqrt{4 \sin^2 t + \cos^2 t + 1}$, and we want to maximize this. To make our computations simpler, we maximize $\|C'(t)\|^2$. In this case,

$$\|C'(t)\|^2 = \|(-2 \sin t, \cos t, 1)\|^2 = 4 \sin^2 t + \cos^2 t + 1 = 3 \sin^2 t + 2$$

This function is maximized when $|\sin t| = 1$, so when $t$ is an odd multiple of $\pi/2$. This is when $C(t)$ is on the “flattest” portion of the ellipse.

A *unit speed* curve is one whose tangent vector has unit length for all values of $t$. Thus the spiral $C = \langle t \cos t, \sin t \rangle$ is not unit speed since $\|C'(t)\| = \sqrt{1 + t^2}$, which is not 1 for all values of $t$. However, the parameterization $C(t) = \langle 2 \cos(t/2), 2 \sin(t/2) \rangle$, $0 \leq t \leq 4\pi$ is unit speed since $C'(t) = \langle -\sin(t/2), \cos(t/2) \rangle$ which is always a unit vector.

Recall that, in single-variable calculus, integrating speed gives distance traveled. The same is true for arbitrary parametric curves. In this context, the distance traveled is the arclength of the curve. One can think of this as approximating the length of the curve by summing the lengths of inscribed line segments. As one can see in Figure 2.4.6, increasing the number of segments used should improve the approximation. The arclength should be the limit of the approximations, as they continue to improve. The limit process turns the sum into an integral.
Definition 2.4.2. The arclength $L(C)$ of the parametric curve $C(t)$, $a \leq t \leq b$, is given by

$$L(C) = \int_a^b \|C'(t)\| \, dt.$$  

Notation: The notation $\|C'(t)\| \, dt$ is called the arclength differential form, and is common enough that writing it becomes cumbersome. For this reason we shorten it to the differential $dC$. When the curve $C$ is defined parametrically, we have

$$dC = \|C'(t)\| \, dt,$$

and the integral for the length $L(C)$ of $C(t)$ is denoted by

$$L(C) = \int dC.$$  

Example 2.4.7. Arclength of a helix

Determine the length of one revolution of the helix $C(t) = (\cos t, \sin t, t)$. Since one revolution occurs for every interval of length $2\pi$, we choose the limits $0 \leq t \leq 2\pi$. Calculating the speed we get

$$\|C'(t)\| = \|(-\sin t, \cos t, 1)\| = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{2},$$

which implies that the arclength is

$$\int_0^{2\pi} \|(-\sin t, \cos t, 1)\| \, dt = \int_0^{2\pi} \sqrt{2} \, dt = \sqrt{2} \left[ t \right]_0^{2\pi} = 2\sqrt{2}\pi.$$  

Example 2.4.8. Another curve

Determine the arclength of $C(t) = (\sqrt{2} t, \ln t, t^2/2)$ for $1 \leq t \leq 2$.

We calculate the speed to be

$$\|C'(t)\| = \left\| \left< \frac{\sqrt{2}}{t}, \frac{1}{t}, t \right> \right\| = \sqrt{2 + \frac{1}{t^2} + t^2} = \sqrt{\left( \frac{1}{t} + t \right)^2} = \frac{1}{t} + t.$$
Thus the arclength is
\[
\int_{1}^{2} \sqrt{1 + t^2} \, dt = \ln t + \frac{t^2}{2} \bigg|_1^2 = \ln 2 + 2 - \frac{1}{2} = \ln 2 + \frac{3}{2}.
\]

**Work and Parametric Curves:** We saw in Section 2.2 that the physical notion of work is defined to be force times distance, where one considers the component of force in the direction of movement. In the case where the force is constant and the movement linear, work turns out to be the dot product of the force and displacement vectors. Eventually, in Section 5.1, we will be able to calculate the work done by a variable force on a particle moving along a curve. For now, we content ourselves with finding the component of force in the direction of movement, when the particle moves along a curve.

**Example 2.4.9. Component of Force**

A helicopter flies along the spiral \( C(t) = (t \cos t, t \sin t, t) \) for \( t \geq 0 \), and the force due to gravity is the vector \( F = (0, 0, -9.8) \). Find and interpret the component of force in the direction of movement.

![Figure 2.4.7: Gravity acting on a helicopter](image)

The key observation is that the direction of movement in this situation is the tangential direction, or \( C'(t) \). Differentiating coordinatewise we get
\[
C'(t) = (\cos t - t \sin t, \sin t + t \cos t, 1).
\]

We now calculate the component of force in the tangential direction:
\[
\text{comp}_{C'(t)} F = \frac{F \cdot C''(t)}{\| C'(t) \|} = \frac{\langle 0, 0, -9.8 \rangle \cdot \langle \cos t - t \sin t, \sin t + t \cos t, 1 \rangle}{\sqrt{(\cos t - t \sin t)^2 + (\sin t + t \cos t)^2 + 1^2}} = \frac{-9.8}{\sqrt{2 + t^2}}.
\]
The fact that \( \text{comp}_{C(t)} F \) is negative indicates that gravity is working against the motion of the helicopter. Note also that \( t \) increases, the component of \( F \) tends to zero.

**Exercises**

1. The following are all parameterizations of the circle \( x^2 + y^2 = 25 \) in \( \mathbb{R}^2 \). Find the tangent vectors for each parameterization, and determine which are unit speed curves.
   (a) \( C(t) = (5 \cos t, 5 \sin t) \), \( 0 \leq t \leq 2\pi \)
   (b) \( C(t) = (5 \cos(t/5), 5 \sin(t/5)) \), \( 0 \leq t \leq 10\pi \)
   (c) \( C(t) = (5 \cos(t^2), 5 \sin(t^2)) \), \( 0 \leq t \leq \sqrt{2\pi} \)

2. Parameterize the tangent line to \( C(t) = (t^2 + 1, 2t - 3, 3 + t^3) \) at \( t = -1 \).

3. Parameterize the tangent line to \( C(t) = (\sec t, \tan t, t) \) at \( t = \pi/4 \).

4. Parameterize the tangent line to \( C(t) = (t \cos t, t \sin t, t^2) \) at \( t = \pi \).

5. Parameterize the tangent line to the curve of intersection of the surfaces \( x^2 + y^2 = 4 \) and \( x + y + z = 5 \) at the point \((0, 2, 3)\).

6. Parameterize the tangent line to the curve of intersection of the generalized cylinders \( z = y^2 \) and \( x + 2y = 4 \) at the point \((-2, 3, 9)\).

7. When is the tangent line to \( C(t) = (t^2 - t, 3t + 1, 2t^3 - 1) \) parallel to the line \( \ell(t) = (2 - t, 3 + t, 5 + 2t) \)?

8. When is the tangent line to \( C(t) = (\sqrt{t}, 2 + t, t^2 - 1) \) normal to the plane \( x + 4y + 32z = 16 \)?

9. Show that the tangent vector to the curve \( C(t) = \left( \cos t, \frac{\sin t}{\sqrt{2}}, \frac{\sin t}{\sqrt{2}} \right) \) is always perpendicular to the position vector.

10. Show that the tangent vector to the curve \( C(t) = (\sin t \cos t, \sin^2 t, \cos t) \), \( 0 \leq t \leq \pi \) is always perpendicular to the position vector.

11. Where does the tangent vector to the helix \( C(t) = (\cos t, \sin t, t) \) make an angle of \( \pi/3 \) with the position vector?

12. Find the arclength of the curve \( C(t) = (e^t \cos t, e^t \sin t) \), \( 0 \leq t \leq \sqrt{2\pi} \).

13. Find the arclength of the curve \( C(t) = (2 \cos t, 2 \sin t, 3t) \), \( 0 \leq t \leq \sqrt{2\pi} \).

14. Find the arclength of the curve \( C(t) = (10t^3 + 2, 9t^2 + 5, 12t^2 - 1) \), \( 0 \leq t \leq \sqrt{2\pi} \).

15. Suppose \( C(t) \), \( a \leq t \leq b \) is a unit speed curve. Show that the arclength of \( C(t) \) is \( b - a \).
16. Show that the parameterization $C(t) = (3 \cos t, 3 \sin t), 0 \leq t \leq 2\pi$ is a constant speed parameterization. Now find a constant $a$ such that $C(t) = (3\cos(t/a), 3\sin(t/a))$ is a unit speed curve. What limits on the parameter are necessary to get once around?

17. A particle travels along the helical path $C(t) = (\cos t, \sin t, t)$. Show that the component of the force $F = \langle 0, 0, -9.8 \rangle$ in the direction of travel is constant.

18. Find the component of $F = \langle 0, 0, -9.8 \rangle$ in the tangential direction of $C(t) = (\sin t \cos t, \sin^2 t, \cos t), \ 0 \leq t \leq \pi$.

19. Show that the helix $C(t) = (\cos t, \sin t, t)$ has constant speed.

20. Find where the tangent line to the curve $C(t) = \langle t, t^2, t^3 \rangle$ has slope $2$.

21. Both $C_1(t) = \langle 3 \cos t, 3 \sin t \rangle, 0 \leq t \leq 2\pi$, and $C_2(t) = \langle 3 \cos(t/3), 3 \sin(t/3) \rangle, 0 \leq t \leq 6\pi$ parameterize the circle centered at the origin with radius three. Compare the speeds of the two parameterizations.

22. The parametric curves $C_1(t) = \langle 0, 2 \sin t, 2 \cos t \rangle, 0 \leq t \leq 2\pi$ and $C_2(t) = (\sqrt{3} \cos t, 2 \sin t, \cos t), 0 \leq t \leq 2\pi$ intersect at the point $(0, 2, 0)$ when $t = \pi/2$. Find an equation for the plane through $(0, 2, 0)$ that contains both tangent vectors $C_1'(t)$ and $C_2'(t)$ there.
2.5 Vector Fields—a first glance

We get our first glimpse of an important, and natural, topic—vector fields. Vector fields are functions that assign vectors (rather than scalars) to points. They arise naturally, and you may have seen some during the weather forecast on your nightly news. One example of a vector field is given in Figure 2.5.1 which assigns to each point on a map the velocity vector of the wind at that point.

Figure 2.5.1: Wind velocities

You may have encountered vector fields in other contexts as well. Electric fields describe the force resulting from charges, and magnetic fields arise in many contexts. Chapter 5 will investigate vector fields in greater detail, but we introduce them here briefly because the topic is a natural conclusion to our introduction to vectors.

Definition 2.5.1. A vector field is a function that assigns a vector to each point in its domain.

Example 2.5.1. Constant Vector Fields

Constant vector fields are perhaps the easiest to visualize. In Figure 2.1.1 we placed the vector \( \mathbf{P} = \langle 3, -3 \rangle \) in the plane with various initial points to emphasize that the vectors are the same regardless of their initial point. We can now interpret that picture differently. The vector field \( \mathbf{F}(x, y) = \langle 3, -3 \rangle \) is the field that attaches to every point in the plane the vector \( \mathbf{P} = \langle 3, -3 \rangle \). Figure 2.1.1 can then be seen as a picture of some representatives of the vector field \( \mathbf{F} \).

Example 2.5.2. Vector fields on \( \mathbb{R}^2 \)

A vector field in \( \mathbb{R}^2 \) is \( \mathbf{F}_1(x, y) = \left\langle \frac{y}{x^2+y^2+1}, \frac{x}{x^2+y^2+1} \right\rangle \). When visualizing a vector field you typically sketch \( \mathbf{F}(x, y) \) with its initial point at \((x, y)\). For example, the vector \( \mathbf{F}(0,1) = \langle -0.5, 0 \rangle \) should be drawn with initial point at
(0,1). Notice that the vector $F_1(x,y)$ is a scalar multiple of $\langle -y, x \rangle$, which is orthogonal to the position vector $\langle x, y \rangle$. This relationship is clear from Figure 2.5.2(a).

\[
F_1(x,y) = \langle -y, x \rangle
\]

This is orthogonal to the position vector $\langle x, y \rangle$.

Figure 2.5.2: Vector Fields on $\mathbb{R}^2$

A second vector field is $F_2(x,y) = \langle y, x \rangle$. You see that the vectors on the line $y = x$ run parallel to the line, while those on either axis are perpendicular to the axis (why?).

We can also evaluate a vector field on a curve simply by restricting the domain. For example, the domain of $F_2$ is the whole plane, but we can evaluate it on the unit circle. Analytically, this just amounts to substituting the parametric equations $\mathbf{x}(t) = (\cos t, \sin t)$ in for $x$ and $y$ in the formula for $F_2$. We have

\[
F_2(x(t)) = F_2(\cos t, \sin t) = \langle \sin t, \cos t \rangle.
\]

Geometrically, this just means drawing those vectors of $F_2$ only on the unit circle (see Figure 2.5.2(c)).

**Example 2.5.3. Electric Field**

A positive point charge placed at $(-1,0)$ and a negative one at $(1,0)$ creates an electric field which exerts a force on a stationary test charge anywhere in the plane. One obtains a vector field by placing the force experienced by the test charge at each point. Up to a constant multiple, the force experienced by a point charge at $(\pm 1, 0)$ is given by the vector field

\[
E(x,y) = \left\langle \frac{y^2 - x^2 + 1}{((x+1)^2 + y^2)((x-1)^2 + y^2)}, \frac{-2xy}{((x+1)^2 + y^2)((x-1)^2 + y^2)} \right\rangle.
\]

Physicists usually use field lines to represent electric fields, obtained by drawing curves tangent to the vectors in Figure 2.5.3. We use vectors to emphasize that electric fields are physical examples of vector fields.

**Example 2.5.4. Vanishing Vector Field**
Find all points on the plane $2x + y - 3z = 6$ where the vector field $\mathbf{F}(x, y, z) = (x + y, y + z, x - z)$ vanishes.

In other words, are there any points on the plane where $\mathbf{F} = \mathbf{0}$? This amounts to solving a system of equations. Since the desired points are on the plane, they must satisfy the equation of the plane. Since the vector field vanishes there, each coordinate must be zero as well. The desired points, then, satisfy the system of equations

$$2x + y - 3z = 6$$

$$x + y = 0$$

$$y + z = 0$$

$$x - z = 0.$$  

The last three equations imply $x = -y = z$, and substituting this into the first equation gives

$$2x + (-x) - 3x = 6,$$

which implies $x = -3$. Thus the point on the plane where $\mathbf{F} = \mathbf{0}$ is $(-3, 3, -3)$.

**Example 2.5.5. Circulation—an introduction**

In this example we introduce notions that motivate path integrals, and are involved in Green’s and Stokes’ Theorems. Suppose the vector field $\mathbf{F}_2$ is the velocity field of a fluid flowing in the plane, and we want to measure the tendency of the fluid to flow around the unit circle. As seen in Figure 2.5.2(c), the fluid is sometimes flowing straight across the unit circle, and is almost tangent at other times. To measure the amount of fluid flowing along the unit circle, it stands to reason that one wants the component of $\mathbf{F}_2$ in the tangential direction. In this case, the tangent vector is a unit vector, so the component of $\mathbf{F}_2$ in the direction of $\mathbf{x}'(t)$ is just the dot product, and we have:

$$\text{comp}_{\mathbf{x}'(t)} \mathbf{F}_2(\mathbf{x}(t)) = \frac{\mathbf{F}_2(\mathbf{x}(t)) \cdot \mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$$

$$= (\sin t, \cos t) \cdot (-\sin t, \cos t) = -\sin^2 t + \cos^2 t = \cos 2t.$$
The dot product $\mathbf{F}_2(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$ is one we’ll encounter again when defining path integrals, so it’s nice to remember this geometric interpretation of it. One can analyze the formula above, for instance, to find when the vector field is orthogonal to the unit circle. That’s when $\mathbf{F}_2(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = \cos 2t = 0$, which happens when $t$ is an odd multiple of $\pi/4$.

Example 2.5.6. Vector Fields in $\mathbb{R}^3$

As in $\mathbb{R}^2$, vector fields in space can be specified by their component functions. The fields $\mathbf{F}_1(x, y, z) = \langle 2x, 2y, 1 \rangle$ and $\mathbf{F}_2(x, y, z) = \langle -x, -y, -z \rangle$ are pictured below. The vectors in $\mathbf{F}_1$ all change one unit vertically. As one moves horizontally away from the origin, the $x$ and $y$ coordinates increase and the vertical component of $\mathbf{F}_1$ is smaller relative to the horizontal components. On the $z$-axis, $\mathbf{F}_1$ yields the vector $(0, 0, 1)$. The field $\mathbf{F}_2$ attaches the vector $(-x, -y, -z)$ to the point $(x, y, z)$. Thus the tip of the vector should touch the origin. The vectors in Figure 2.5.5(b) are not drawn to scale, to give a better geometric understanding of $\mathbf{F}_2$: all vectors point toward the origin.

Example 2.5.7. Flux—an introduction
As in the case of circulation above, consider \( \mathbf{F}_1 \) to represent the velocity field of a fluid flowing through space. Perhaps hot water in a cup of tea. Now suppose the unit sphere represents a tea ball in which tea is steeping. The water flows through the ball, making a delicious cup of tea. The component of the velocity \( \mathbf{F}_1 \) in the normal direction to the sphere measures how fast the fluid is flowing across it.

To see this, you might think of extreme situations. If the direction of flow \( \mathbf{F}_1 \) is normal to the surface, then fluid is flowing straight across it. Since \( \mathbf{F}_1 \) is orthogonal to the surface, \( \| \mathbf{F}_1 \| \) is the component of \( \mathbf{F}_1 \) in the normal direction, and is how fast the fluid is flowing across the surface. The other extreme is when the fluid is flowing tangentially to the surface. At that point none of it crosses the surface, and it’s velocity across the surface is zero. Since \( \mathbf{F}_1 \) is tangent to the surface, it is orthogonal to the normal vector and the component in the normal direction is zero as well. These observations give some intuition about flow across surfaces. There are some details about which of the two normal directions to choose, which we ignore for the time being. Suffice it to say that there are situations where we want to know the component of a vector field in a normal direction to a surface.

![Diagram of F1 in normal direction](image)

**Figure 2.5.6: Component of \( \mathbf{F}_1 \) in normal direction—to calculate flux**

Let’s return to the case of the sphere, or tea ball, and the velocity field \( \mathbf{F}_1(x, y, z) = \langle 2x, 2y, 1 \rangle \). We want to calculate the component of \( \mathbf{F}_1 \) orthogonal to the sphere. We know from geometry that radii are orthogonal to spheres, and this allows us to find unit normal vectors. Let \( (x, y, z) \) be a point on the sphere, then the vector from the origin to \( (x, y, z) \) is a radius of the sphere; therefore the position vector \( \langle x, y, z \rangle \) is an outward pointing normal to the sphere. Typically, we place the vector \( \langle x, y, z \rangle \) with its initial point at \( (x, y, z) \) to see that it’s normal. Since \( (x, y, z) \) is on the unit sphere, \( \langle x, y, z \rangle \) is a unit vector, and the component of \( \mathbf{F}_1 \) in the normal direction is

\[
\mathbf{F}_1(x, y, z) \cdot \langle x, y, z \rangle = \langle 2x, 2y, 1 \rangle \cdot \langle x, y, z \rangle = 2x^2 + 2y^2 + z.
\]

At this point we haven’t used the fact that \( (x, y, z) \) is on the unit sphere \( S^2 \). Since it is, \( x^2 + y^2 = 1 - z^2 \), and we can simplify our calculation making this substitution:
\[ \mathbf{F}_1(x, y, z) \cdot \langle x, y, z \rangle = 2x^2 + 2y^2 + z = 2 - 2z^2 + z. \]

It turns out that this measures how fast the fluid is flowing across the tea ball at any point. To get the total flux of the fluid across the surface we’ll integrate its “velocity” over the surface. The “velocity” function is the component of \( \mathbf{F} \) in the normal direction to the surface. In our case the function we would integrate over the surface is \( 2x^2 + 2y^2 + z \).

This, of course, raises lots of questions: How do you integrate functions with several variables? How can you integrate only on surface? etc. We’ll get to this! For now, recognize that the vector techniques we’re learning are useful and significant in a variety of contexts.

**Example 2.5.8. Normal to a plane**

Find the component of \( \mathbf{F}(x, y, z) = \langle x + y, z, -2y \rangle \) in the upward normal direction to the plane \( x - 2y + z = 7 \). Implicit in this problem is the assumption that we are evaluating \( \mathbf{F} \) on the plane. We will see how we use this analytically in the last step. Since an upward pointing normal to the plane is \( \mathbf{n} = \langle 1, -2, 1 \rangle \) (its negative is a downward pointing normal since the \( z \)-component would be negative), we get the component of \( \mathbf{F} \) in the \( \mathbf{n} \) direction is

\[
\text{comp}_n \mathbf{F} = \frac{\mathbf{F} \cdot \mathbf{n}}{||\mathbf{n}||} = \frac{1}{\sqrt{6}} (x + y, z, -2y) \cdot (1, -2, 1) = \frac{1}{\sqrt{6}} (x - y - 2z).
\]

At no point in this calculation did we use the fact that we evaluated \( \mathbf{F} \) on the plane. We can use this now to reduce our calculation to a function of two variables. The coordinates of any point on the plane satisfy \( x - 2y + z = 7 \), so we can replace \( z \) with \( 7 - x + 2y \). Making this simplification, we get

\[
\text{comp}_n \mathbf{F} = \frac{1}{\sqrt{6}} (x - y - 2z) = \frac{1}{\sqrt{6}} (3x - 5y - 14).
\]

This simplification will be useful later when performing certain integrals.

**Exercises**

1. Sketch the vectors \( \mathbf{F}(\pm 1, \pm 1), \mathbf{F}(0, \pm 1), \text{ and } \mathbf{F}(\pm 1, 0), \) for the vector field \( \mathbf{F}(x, y) = \langle x - y, x + y \rangle \).

2. Describe the vectors of \( \mathbf{F}(x, y) = \langle y + 1, x - 2 \rangle \) on the line \( x = 2 \) in the plane. Sketch a few to illustrate your description.

3. Is there a point on the unit sphere where the vector field \( \mathbf{F}(x, y, z) = \langle x + y, y - 2z, x - z \rangle \) vanishes (i.e. is the zero vector)?

4. Let \( \mathbf{F}(x, y) = \langle x + y, x \rangle \), and consider the restriction of \( \mathbf{F} \) to the unit circle \( \mathbf{x}(t) = (\cos t, \sin t) \). Find the component of \( \mathbf{F} \) in the direction of the tangent vector to the unit circle. How would this change using the circle radius 2, given by \( \mathbf{x}(t) = (\cos t, 2 \sin t) \)?
5. Evaluate the vector field \( \mathbf{F}(x, y, z) = (x^2 + y^2, x + y, x - y) \) on the helix \( \mathbf{x}(t) = (\cos t, \sin t, t) \), then find the component of \( \mathbf{F} \) in the \( \mathbf{x}'(t) \) direction.

6. Find the component of \( \mathbf{F}(x, y, z) = (z, 2y, x) \) in the direction of the upward pointing normal to the plane \( 3x + 2y - 4z = 12 \) (be sure to eliminate \( z \) from the component).

7. Find the component of \( \mathbf{F}(x, y, z) = (y, -x, z) \) in the direction of the outward pointing normal to the unit sphere. Write your answer in terms of \( x \) and \( y \) (i.e. use the equation of the sphere to eliminate \( z \)).
Chapter 3

Differentiation

3.1 Functions

In single variable calculus you studied functions whose domain and range were subsets of $\mathbb{R}^1$, like $f(x) = x^e$. When thinking of these as maps from the real numbers to itself, we use the notation $f : \mathbb{R}^1 \to \mathbb{R}^1$. When introducing surfaces we considered graphs of functions of two variables, which are maps $f : \mathbb{R}^2 \to \mathbb{R}^1$. A map $f : \mathbb{R}^n \to \mathbb{R}$ will be called a function of several variables because it has a independent variables (the “several variables” part), and one dependent variable (the “function” part). For example $f(x, y, z) = \frac{x\sin y}{z}$ is a function of several variables since it requires several (three) inputs to produce a single number output.

We also considered parametric surfaces $\mathbf{x} : \mathbb{R}^2 \to \mathbb{R}^3$, which are examples of vector valued functions of several variables. The “several variables” part refers to the domain having dimension greater than one, and the “vector valued” part to the range being multidimensional. To study parametric surfaces $\mathbf{x}(s, t) = (x(s, t), y(s, t), z(s, t))$ we focused on their component functions $x(s, t), y(s, t)$ and $z(s, t)$. Notice that the component functions of a vector-valued function are functions of several variables since their output is a single number. For example, the vector-valued function of two variables $\mathbf{x}(r, \theta) = (r \cos \theta, r \sin \theta, r^2)$ is made up of three functions of two variables, $x(r, \theta) = r \cos \theta$, $y(r, \theta) = r \sin \theta$ and $z(r, \theta) = r^2$.

Thus the most general functions we consider are mappings of the form $f : \mathbb{R}^n \to \mathbb{R}^m$, but we will emphasize lower dimensions. We give two examples of maps $f : \mathbb{R}^2 \to \mathbb{R}^2$, and illustrate one way to analyze them.

Example 3.1.1. A linear map of the plane to itself

Understand the function $f : \mathbb{R}^2 \to \mathbb{R}^2$ given by $f(x, y) = (x - 3y, x + 3y)$. Both of its component functions are linear in $x$ and $y$, so $f$ maps lines to lines. To understand $f$, we’ll study what it does to horizontal and vertical lines.

Horizontal lines in $\mathbb{R}^2$ are given parametrically by $\mathbf{x}(t) = (t, c)$, where $c$ is a constant. Applying $f$ we see that $f(\mathbf{x}(t)) = f(t, c) = (t - 3c, t + 3c)$. Thus
$f(x(t))$ is a line through the point $(-3c, 3c)$, and in direction $v = (1, 1)$. In other words, horizontal lines get mapped to lines with slope one by $f$.

Similarly, vertical lines are $x(t) = (c, t)$, giving $f(x(t)) = f(c, t) = (c - 3t, c + 3t)$. These are lines through the point $(c, c)$ in the direction of $v = (-3, 3)$, and $f$ maps vertical lines to lines with slope $-1$. See Figure 3.1.1 for the images of several lines.

Finally, we can see the image of the unit circle $x(t) = (\cos t, \sin t)$ is the ellipse given parametrically by $f(x(t)) = (\cos t - 3\sin t, \cos t + 3\sin t)$ pictured on the right.

![Figure 3.1.1: A vector valued function of two variables](image)

**Example 3.1.2. The polar change-of-coordinates**

A second map $f : \mathbb{R}^2 \to \mathbb{R}^2$ is the traditional change of coordinates from the $r\theta$-plane to the $xy$-plane. In this case $f(r, \theta) = (r\cos \theta, r\sin \theta)$. The mapping takes the rectangle $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$ to the unit disk, mapping horizontal segments like $\theta = \pi/3$ to rays and vertical segments like $r = 0.4$ to circles.

![Figure 3.1.2: Change of Coordinates](image)
with a subset of the range (a single value), and looks at the preimage in the domain. Indeed, the level $c$ is a point in the range of $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, while the level curve $f(x, y) = c$ is all points in the domain with image $c$. In this chapter, we will investigate another way to study functions—by considering their rates of change. In other words, we will use derivatives to help analyze functions.

1. Sketch the images of several horizontal and vertical lines to analyze the map $f(x, y) = (x, 2y)$. What is the image of the unit circle?

2. Sketch the images of several horizontal and vertical lines to analyze the map $f(x, y) = (-y, x)$. What is the image of the unit circle?

3. Label the coordinates of $\mathbb{R}^3$ by $(r, \theta, z)$ rather than the traditional $(x, y, z)$, thinking of $r$, $\theta$ and $z$ as perpendicular directions. Sketch the surface $r = 2$ in this copy of $\mathbb{R}^3$. Now sketch the image of $r = 2$ under the coordinate transformation $f(r, \theta, z) = (r \cos \theta, r \sin \theta, z)$.

4. Label the coordinates of $\mathbb{R}^3$ by $(\rho, \theta, \phi)$ rather than the traditional $(x, y, z)$, thinking of $\rho$, $\theta$ and $\phi$ as perpendicular directions. Sketch the surface $\rho = 2$ in this copy of $\mathbb{R}^3$. Now sketch the image of $\rho = 2$ under the coordinate transformation $f(\rho, \theta, \phi) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$. 

3.2 Limits

We’ve taken a brief look at functions of several variables, and now discuss the notion of limits in the several variable setting. We will see there are many similarities between single and several variable limits, as well as some subtleties we didn’t encounter in the single variable case. In this section we first illustrate some multivariable limit techniques that are reminiscent of single-variable limits. After becoming more comfortable with them, we provide a technical definition of multivariable limits and investigate them more carefully.

The intuitive definition of the limit of \( f(x) \) as \( x \to c \) is that it is the expected function value. The key term there is expected. That means that \( \lim_{x \to c} f(x) \) is the value of \( f(c) \) you’d expect to get by looking at the function for \( x \) near \( c \). Lets say that again, slightly differently, because it is worth having some intuition about limits. To calculate \( \lim_{x \to c} f(x) \) you look see what \( f(x) \)-values are trending to as \( x \) gets closer to \( c \). Thus \( \lim_{x \to c} f(x) \) doesn’t care what you get when you plug \( c \) into \( f(x) \), it just cares what happens for \( x \) near \( c \). These intuitive descriptions are worth pondering over a cup of coffee, together with some examples of limits from single-variable calculus. We begin with a sample limit that should jog your memory.

Example 3.2.1. \( \lim_{x \to 3} \frac{\sqrt{x + 1} - 2}{x - 3} \)

When calculating limits, you first evaluate \( f(c) \) to see if it works. Letting \( x = 3 \) in \( f(x) \) you get \( \frac{0}{0}, \) so you have to do tricks, and we rationalize the numerator. Here goes:

\[
\lim_{x \to 3} \frac{\sqrt{x + 1} - 2}{x - 3} = \lim_{x \to 3} \frac{\sqrt{x + 1} - 2}{x - 3} \cdot \frac{\sqrt{x + 1} + 2}{\sqrt{x + 1} + 2}
\]

\[
= \lim_{x \to 3} \frac{x - 3}{(x - 3)(\sqrt{x + 1} + 2)}
\]

\[
= \lim_{x \to 3} \frac{1}{\sqrt{x + 1} + 2} = \frac{1}{4}.
\] \hspace{1cm} (3.2.1)

This example underscores that limits don’t care what happens at \( x = 3 \), since the function is undefined there. Limits only care what happens near \( x = 3 \). The above calculations show that for \( x \neq 3 \) we know \( f(x) = \frac{1}{\sqrt{x + 1} + 2} \), so we can calculate the limit. ▲

This is an example of one trick from single variable calculus for calculating limits. You should be familiar with many others, like factoring and cancelling, L’Hoptal’s Rule, and logarithmic limits. Here’s another single-variable example.

Example 3.2.2. \( \lim_{x \to \infty} \left( 1 - \frac{2}{x} \right)^x \)

In this example the exponent contains a variable. Recall that the strategy in this case is to take the limit of the natural logarithm of the function, then
3.2. LIMITS

Exponentiate it. Formally we have
\[ \lim_{x \to c} f(x) = \lim_{x \to c} e^{\ln f(x)} = e^{\lim_{x \to c} \ln f(x)}, \]
where the last equality is justified by a theorem on limits of compositions of functions. Thus we begin by taking the limit of the natural logarithm of our function.

\[ \lim_{x \to \infty} \left( 1 - \frac{2}{x} \right)^x = \lim_{x \to \infty} x \ln \left( 1 - \frac{2}{x} \right) \text{ log property} \]
\[ = \lim_{x \to \infty} \frac{\ln \left( 1 - \frac{2}{x} \right)}{1/x} \]
\[ = \lim_{x \to \infty} \frac{-2}{x} \text{ L'Hopital's Rule} \]
\[ = \lim_{x \to \infty} \frac{-2}{1 - \frac{2}{x}} = -2. \]

Exponentiating we have
\[ \lim_{x \to \infty} \left( 1 - \frac{2}{x} \right)^x = e^{\lim_{x \to \infty} \ln \left( 1 - \frac{2}{x} \right)^x} = e^{-2}. \]

Limits of functions of several variables have the same intuitive idea as in single-variable calculus. The limit of \( f(x, y) \) as \((x, y) \to (a, b)\) is the expected function value. Some of the same tricks also work, and we illustrate with some examples. In particular, you can evaluate limits of polynomials, rational functions, trigonometric functions, and their compositions just by evaluating the function at the desired point (as long as it is defined there). This follows from some of the same theorems as single variable calculus, ramped up to the several variable case. We illustrate with some examples.

**Example 3.2.3.** \( \lim_{(x,y)\to(2,-1)} x^2 + y^2 \)

Simply by evaluating we get
\[ \lim_{(x,y)\to(2,-1)} x^2 + y^2 = 2^2 + (-1)^2 = 5. \]

**Example 3.2.4.** \( \lim_{(x,y)\to(\sqrt{\pi},0)} \sin(x^2 - 2xy) \)

Since this is a composition of a trigonometric function and polynomial, we simply substitute to evaluate the limit.
\[ \lim_{(x,y)\to(\sqrt{\pi},0)} \sin \left( \frac{x^2}{2} - 2xy \right) = \sin(\pi/2) = 1. \]

Thus some several variable versions of nice functions still behave nicely. Another trick you can do to relate multivariable limits to familiar single variable ones is substitution. We demonstrate with some examples.
Example 3.2.5. \[ \lim_{(x,y) \to (0,0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2} \]

When substituting \((0,0)\) into the function, you get \(0\) which is undefined. However, substituting \(\theta\) for \(x^2 + y^2\) in the function gives a more familiar \(\sin \theta\). Moreover, as \((x,y) \to (0,0)\) we see that \(\theta = x^2 + y^2 \to 0\). Therefore we have

\[ \lim_{(x,y) \to (0,0)} \frac{\sin(x^2 + y^2)}{x^2 + y^2} = \lim_{\theta \to 0} \frac{\sin \theta}{\theta} = 1. \]

Example 3.2.6. \[ \lim_{(x,y) \to (1,1)} \sqrt{\frac{x^2 + y^2 - 2x - 2y + 3 - 1}{x^2 + y^2 - 2x - 2y + 2}} \]

When substituting \((1,1)\) into the function, you get \(0\) which is undefined. However, letting \(z = x^2 + y^2 - 2x - 2y + 2\) in the function gives \(\frac{\sqrt{z + 1} - 1}{z}\). Moreover, as \((x,y) \to (1,1)\) we see that \(z \to 0\). Therefore we have

\[ \lim_{(x,y) \to (1,1)} \sqrt{\frac{x^2 + y^2 - 2x - 2y + 3 - 1}{x^2 + y^2 - 2x - 2y + 2}} = \lim_{z \to 0} \frac{\sqrt{z + 1} - 1}{z} = \lim_{z \to 0} \frac{\sqrt{z + 1} - 1}{z} \cdot \frac{\sqrt{z + 1} + 1}{\sqrt{z + 1} + 1} = \lim_{z \to 0} \frac{z}{z(\sqrt{z + 1} + 1)} = \frac{1}{2}. \]

Given the above examples, one may be tempted to use algebra at will when calculating several variable limits. There is one technicality that we’ve glossed over in the previous two substitution examples. In Examples 3.2.5 and 3.2.6 the function was not defined at the limiting point for \((x,y)\). In both cases it turns out that is the only point nearby where the function is not defined, and that is important. In order to discuss this further we need the technical definition of a limit, which will take some doing so grab some coffee and settle in...here we go.

First we recall the single-variable \(\epsilon - \delta\) definition of a limit.

**Definition 3.2.1.** The limit of \(f(x)\) as \(x\) approaches \(c\) is \(L\) if for every \(\epsilon > 0\) there is a \(\delta > 0\) such that the inequality \(0 < |x-c| < \delta\) implies that \(|f(x)-L| < \epsilon\) is true as well.

This was one of those daunting definitions you hurried to forget, so let’s spend a minute going back over what it says. This definition tells you when a number \(L\) is the limit of a function \(f(x)\). Intuitively (again) it is the technical way of saying that as \(x\) approaches \(c\) the \(f(x)\)-values approach \(L\). To interpret the definition, recall that the inequality \(|f(x)-L| < \epsilon\) means that the function value \(f(x)\) is within \(\epsilon\) of the number \(L\). Equivalently, \(L - \epsilon < f(x) < L + \epsilon\) or \(f(x)\) is in the interval \((L - \epsilon, L + \epsilon)\). I like to say:

The inequality \(|f(x)-L| < \epsilon\) says \(f(x)\) is in an interval centered at \(L\) with radius \(\epsilon\).
Since Definition 3.2.1 says “for every \( \epsilon > 0 \), we are allowed to specify the radius of the interval to be as small as we like. So we can specify how close we want \( f(x) \) to be to \( L \).

Similarly, the inequality \( |x - c| < \delta \) says that \( x \) is in an interval centered at \( c \) with radius \( \delta \).

Thus Definition 3.2.1 says that no matter how close we want \( f(x) \) to be to \( L \) (“for every \( \epsilon > 0 \)”), we can find a radius \( \delta > 0 \), so that whenever \( x \) is within \( \delta \) of \( c \) we know the function value \( f(x) \) is close enough to \( L \).

To generalize Definition 3.2.1 to functions of several variables, the \( \epsilon \) portion of the definition translates directly but the \( \delta \) part needs some work. To say that \( x \) is within \( \delta \) of \( c \) in the one-variable case means \( x \) is in the interval \( (c - \delta, c + \delta) \) of real numbers. For two variables, to say that the point \( (x, y) \) is within \( \delta \) of the point \( (a, b) \), means that \( (x, y) \) is inside the disk centered at \( (a, b) \) with radius \( \delta \). In the 3-variable case, we would want \( (x, y, z) \) to be in the solid ball radius \( \delta \) centered at \( (a, b, c) \). The easiest way to formalize this in higher dimensions is to use vector notation. The point \( x \in \mathbb{R}^n \) is within \( \delta \) of the point \( c \) if the length of the vector \( x - c \) is less than \( \delta \), or formally \( \|x - c\| < \delta \).

Definition 3.2.2. Let \( c \in \mathbb{R}^n \) and let \( f(x) \) be a function of \( n \) variables. The limit of \( f(x) \) as \( x \to c \) is the number \( L \) if for every \( \epsilon > 0 \) there is a \( \delta > 0 \) such that \( 0 < \|x - c\| < \delta \) implies \( |f(x) - L| < \epsilon \).

Thus Definition 3.2.2 says that no matter how close we want \( f(x) \) to be to \( L \) (“for every \( \epsilon > 0 \)”), we can find a radius \( \delta > 0 \), so that whenever \( x \) is within \( \delta \) of \( c \) we know the function value \( f(x) \) is close enough to \( L \).

Notice that the previous paragraph is identical—literally—to the one we used to describe single-variable limits, except the real numbers \( x \) and \( c \) from one variable are replaced by vectors \( x \) and \( c \). I guess we also replace the absolute value in \( 0 < |x - c| < \delta \) with the norm of a vector \( 0 < \|x - c\| < \delta \).

We have already mentioned that multivariable limits of many functions we encounter can be obtained by evaluating \( f(c) \) when it’s defined (e.g. polynomials, rational functions, exponential functions, logarithmic functions, etc.). These functions can be shown to be continuous on their domains, i.e. that \( \lim_{x \to c} f(x) = f(c) \), using techniques similar to the one-variable case. We omit the details. We do want to finish the section with some examples of when multivariable limits do not exist.

We first point out that for Definition 3.2.2 to hold, the function \( f \) must be defined in an open ball centered at \( c \), except possibly at \( c \) itself. Otherwise there is always some \( x \) within \( \delta \) of \( c \) for which \( f(x) \) is not defined, which implies the inequality \( |f(x) - L| < \epsilon \) cannot be true. We use this to show a limit doesn’t exist.

Example 3.2.7. \( \lim_{(x,y) \to (3,3)} \frac{x^2 - y^2}{x - y} \)

You might be tempted to proceed as follows:
We first substitute $x = 3$ and $y = 3$ into the function, and obtain $0 \over 0$ which is undefined. Therefore we do tricks, and one of them is factoring and cancelling.

$$\lim_{(x,y) \to (3,3)} \frac{x^2 - y^2}{x - y} = \lim_{(x,y) \to (3,3)} \frac{(x - y)(x + y)}{x - y} = \lim_{(x,y) \to (3,3)} x + y = 6.$$  

This is not precise, however, since our original function is not defined on the line $y = x$. Therefore any open disk around the point $c = (3, 3)$ contains points where $f$ is not defined. Thus the limit does not exist. ▲

Recall that in single variable calculus a limit didn’t exist if the right-hand and left-hand limits were different. The two-sided limit didn’t exist because coming at $c$ from different directions gave different expected values. In other words, the limit depended on which path you took to get to $c$. Analogously, a second way to show multivariable limits don’t exist is to show the limit is path dependent. In other words, different paths to $c$ lead to different limits. We illustrate with several examples.

Example 3.2.8. $\lim_{(x,y) \to (0,0)} \frac{x^2}{x^2 + y^2}$

To show this limit doesn’t exist, take the limit as $(x, y) \to (0, 0)$ along two different lines and show you get two different numbers. The $x$-axis has equation $y = 0$, so substituting 0 for $y$ then taking the limit gives the expected function values as $(x, y) \to (0, 0)$ along the $x$-axis. We calculate along the $x$-axis

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

Similarly, letting $x = 0$, then taking the limit gives the limit along the $y$-axis. We have

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

Since the limits as $x$ approaches 0 along different directions are different, the limit doesn’t exist.

We can be more general than we have been so far. If we want $x \to 0$ along the line $y = mx$, we can substitute $mx$ for $y$ and take the limit as follows:

$$\lim_{(x,y) \to (0,0)} \frac{x^2}{x^2 + y^2} = \lim_{(x,mx) \to (0,0)} \frac{x^2}{x^2 + m^2 x^2} = \lim_{(x,mx) \to (0,0)} \frac{1}{1 + m^2} = \frac{1}{1 + m^2}.$$  

This calculation shows that the limit changes depending on the slope of the line you approach 0 on. This phenomena is pictured in Figure 3.2.1 ▲

In the previous example we approached the origin along different lines to get different limits. In the next example we approach along different cubics.

Example 3.2.9. $\lim_{(x,y) \to (0,0)} \frac{x^6}{x^6 + y^2}$
3.2. LIMITS

Figure 3.2.1: The surface $z = x^2/(x^2 + y^2)$

Note that approaching the origin along the cubic $y = mx^3$ yields the following limit:

$$
\lim_{(x,y) \to (0,0)} \frac{x^6}{x^6 + y^2} = \lim_{(x,mx) \to (0,0)} \frac{x^6}{x^6 + m^2x^6} = \lim_{(x,mx) \to (0,0)} \frac{1}{1 + m^2} = \frac{1}{1 + m^2}.
$$

Thus different paths yield different limits, and the limit does not exist. ▲

Exercises

1. Calculate the following limits

   (a) $\lim_{x \to (1,2)} \frac{x^2 - 3x}{y^4 + 2y}$

   (b) $\lim_{x \to (3,-1)} \sin(\pi(x^2 + y^2)/4)$

   (c) $\lim_{x \to (e,e^3)} \ln(xy)$

   (d) $\lim_{x \to (\ln 3, \ln 2)} e^{x - 3y}$

2. Calculate the following limits

   (a) $\lim_{x \to (0,0)} \arctan \left( \frac{5}{x^2 + y^2} \right)$

   (b) $\lim_{x \to (0,0)} \left( 1 + \frac{1}{x^2 + y^2} \right)^{x^2 + y^2}$

   (c) $\lim_{x \to (e,e^3)} \ln(xy)$

   (d) $\lim_{x \to (0,0)} \frac{1 - \cos \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}}$
3. Show that the following limits don’t exist

(a) \( \lim_{x \to (0,0)} \frac{x^3}{x^3 + y^3} \)

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

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

(d) \( \lim_{x \to (0,0)} \frac{x^4}{x^4 + y^2} \)
3.3 Partial Derivatives

In single variable calculus we saw that derivatives were instantaneous rates of change, with respect to the independent variable. Analytically, recall that the average rate of change of \( f(x) \) on the interval \( (x, x + h) \) is given by the difference quotient \( \left( f(x + h) - f(x) \right) / h \). This approximates the instantaneous rate of change of \( f \) at \( x \) for small \( h \), and the approximation becomes better as \( h \) gets smaller. Thus we define the instantaneous rate of change \( f'(x) \) to be

\[
f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}, \tag{3.3.1}
\]

if the limit exists.

It doesn’t hurt to recall the geometric interpretation of the pieces of the definition of the derivative, and we refer to Figure 3.3.1 for the following notation. Let \( P \) and \( Q \) be the points \((x, f(x))\) and \((x + h, f(x + h))\), respectively, on the graph of \( y = f(x) \). The difference quotient \( \left( f(x + h) - f(x) \right) / h \) gives the slope \( m_{\text{sec}} \) of the secant line through \( P \) and \( Q \), while the derivative \( f'(x) \) is the slope \( m_{\text{tan}} \) of the tangent line at \( P \). Taking the limit as \( h \) tends to 0 in Equation 3.3.1 is equivalent to letting the point \( Q \) get closer to \( P \). Equation 3.3.1 simply states that the slopes \( m_{\text{sec}} \) approach \( m_{\text{tan}} \) as \( Q \) gets closer to \( P \). Symbolically we have

\[
m_{\text{tan}} = \lim_{Q \to P} m_{\text{sec}},
\]

and in English this just means the slopes of the secant lines approach the slope of the tangent line to \( f \) at \( P \) as \( Q \) gets closer to \( P \).

To begin our study of differentiation we focus on functions \( f : \mathbb{R}^n \to \mathbb{R} \) of several variables. We now have several independent variables, and could ask for the instantaneous rate of change of \( f \) with respect to any of them. This will be called the partial derivative of \( f \), because it focuses on the rate of change with respect to just one of many variables. We treat partial derivatives analytically first, then discuss their geometric interpretation.

For simplicity, we restrict our attention to functions \( f(x, y, z) \) of three variables, and consider the instantaneous rate of change of \( f \) with respect to the
independent variable $x$—the partial derivative of $f$ with respect to $x$. The first observation is that if we want to know how fast $f$ is changing in the $x$-direction, we think of $y$ and $z$ as fixed. In other words: to define the partial derivative of $f$ with respect to $x$, don’t change $y$ and $z$, just change $x$. We now take our cue from single variable calculus. To find the instantaneous rate of change, we approximate it with the average rate $(f(x + h, y, z) - f(x, y, z))/h$ over the interval $(x, x + h)$. Note that $y$ and $z$ stay fixed in this difference quotient because we want the change of $f$ in the $x$-direction. Now take a limit as $h \to 0$ to find the instantaneous rate. Formally, we define

**Definition 3.3.1.** The partial derivative $\frac{\partial f}{\partial x}$ of $f(x, y, z)$ with respect to $x$ is

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

when the limit exists.

Thus partial derivatives can be interpreted as instantaneous rates of change. The difference quotient in the definition is the average rate of change of $f$ on the $x$-interval $(x, x + h)$ (fixing $y$ and $z$). In taking a limit, the average rate of change approaches the instantaneous rate of change, so $\frac{\partial f}{\partial x}(x, y)$ is the instantaneous rate of change of $f$ in the $x$-direction. Of course, $\frac{\partial f}{\partial x}(x, y)$ is the instantaneous rate of change in the $y$-direction, etc. This highlights one difference between single- and multi-variable differentiation. In single-variable calculus there’s essentially one direction to differentiate with respect to. In multivariable calculus, there’s a derivative in every direction of the domain! We’ll make this more precise when we investigate directional derivatives.

To calculate the partial derivative of $f(x, y, z)$ with respect to $x$, think of $y$ and $z$ as constants and differentiate $x$ normally. **WARNING:** This is not implicit differentiation from single-variable calculus! In implicit differentiation we thought of $y$ as being a function of $x$ defined implicitly by an equation. In partial differentiation with respect to $x$ we think of $y$ as a constant.

**Example 3.3.1.** Calculating partial derivatives.

Find all first order partial derivatives of $f(x, y, z) = x^2 \sin(xy^2z^3)$.

$$\frac{\partial f}{\partial x}(x, y, z) = 2x \sin(xy^2z^3) + x^2y^2z^3 \cos(xy^2z^3),$$

$$\frac{\partial f}{\partial y}(x, y, z) = x^2(2xyz^3 \cos(xy^2z^3)) = 2x^3yz^3 \cos(xy^2z^3),$$

$$\frac{\partial f}{\partial z}(x, y, z) = x^2(3xy^2z^2 \cos(xy^2z^3)) = 3x^3y^2z^2 \cos(xy^2z^3).$$

Just as we have second derivatives in single variable calculus, there are second order partial derivatives. Since there are several variables, there are a fair number of second order partial derivatives. The notation for first taking the partial derivative with respect to $x$, then $z$ is $\frac{\partial^2 f}{\partial x \partial z}$. So, as with composition of
functions, the order you differentiate in is read from right to left in this notation. There is another common notation for partial derivatives that uses subscripts. We let \( f_x = \frac{\partial f}{\partial x} \). Higher order partial derivatives in subscript notation are read left-to-right, so \( f_{xyz} \) means differentiate with respect to \( x \), then \( y \), and finally \( z \).

**Example 3.3.2. Higher order partial derivatives**

Find all second order partial derivatives of \( f(x,y) = x^2y - 3xy^3 \). We start by finding the first-order partial derivatives:

\[
\begin{align*}
    f_x(x,y) &= 2xy - 3y^3 \\
    f_y(x,y) &= x^2 - 9xy^2
\end{align*}
\]

Now we calculate the second-order partial derivatives

\[
\begin{align*}
    f_{xx}(x,y) &= 2y \\
    f_{xy}(x,y) &= 2x - 9y^2 \\
    f_{yx}(x,y) &= 2x - 9y^2 \\
    f_{yy}(x,y) &= -18xy
\end{align*}
\]

Notice that the mixed partials \( f_{xy} \) and \( f_{yx} \) are equal. This is no coincidence! In fact, as long as \( f \) is reasonably nice, mixed partials will always be equal.

**Theorem 3.3.1.** If \( f(x,y) \) has continuous second-order partial derivatives, then mixed partial derivatives are equal. Symbolically

\[
\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}, \text{ or } f_{xy} = f_{yx}.
\]

This theorem generalizes to functions of three or more variables, with the analogous conclusion. We haven’t defined continuity and limits for functions of several variables, but hope to do so in some future edition of these notes.

**Example 3.3.3. Verifying the equality of mixed partials**

Verify that the mixed partial derivatives of \( f(x,y) = x^3 - 3x^2y + xy^3 - 5x + 7 \) are equal.

Calculating first partial derivatives gives \( f_x(x,y) = 3x^2 - 6xy + y^3 - 5 \) and \( f_y(x,y) = -3x^2 + 3xy^2 \). From these we compute

\[
\begin{align*}
    f_{xy}(x,y) &= -6x + 3y^2 = f_{yx}(x,y).
\end{align*}
\]

**Example 3.3.4. Partial differentiation backwards**

Find a function \( f(x,y) \) whose partial derivative with respect to \( x \) is \( f_x(x,y) = 2xe^{x^2+y^2} - 3y \).

To find \( f(x,y) \) we do “partial integration”. In other words, we integrate \( f_x \) with respect to \( x \), thinking of any function of \( y \) as a constant. We see

\[
\int 2xe^{x^2+y^2} - 3y \, dx = e^{x^2+y^2} - 3xy + g(y).
\]

This illustrates one interesting feature of “partial” integration—the arbitrary constant is really an arbitrary function of \( y \)!

This makes sense since \( \frac{\partial}{\partial x} g(y) = 0 \), so adding the function of \( y \) does not change the partial derivative with respect to \( x \).
Example 3.3.5. Finding a function given both partial derivatives

Find a function \( f(x,y) \) given

\[
f_x(x,y) = x^2y - 2x + 5, \quad f_y(x,y) = \frac{1}{3}x^3 + 2y.
\]

We outline the steps of the process:

**Step 1:** Do partial integration \( \int f_y(x,y)\,dy \), remembering to add an arbitrary function of \( x \). This determines the portion of \( f(x,y) \) that depends on \( y \).

\[
f(x,y) = \int \frac{1}{3}x^3 + 2y \, dy = \frac{1}{3}x^3y + y^2 + g(x)
\]

**Step 2:** Set the partial derivative of your answer from step 1 equal to the given \( f_x(x,y) = x^2y - 2x + 5 \), and solve for \( g'(x) \). In our case we get

\[
f_x(x,y) = x^2y + g'(x) = x^2y - 2x + 5,
\]

which gives \( g'(x) = 2x - 5 \). The result must be void of \( y \)'s. If there are any \( y \)'s in the equation, something went wrong in Step 1, or the problem is impossible to solve.

**Step 3:** Integrate to find \( g(x) \), and write the concluding sentence. In our case,

\[
g(x) = \int 2x - 5 \, dx = x^2 - 5x + C.
\]

where \( C \) is the friendly arbitrary constant from single-variable calculus. We conclude that

\[
f(x,y) = \frac{1}{3}x^3y + y^2 + x^2 - 5x + C. \quad \blacksquare
\]

The Geometry of partial derivatives: Now that we know how to calculate partial derivatives, let’s take a look at what they tell us geometrically. We’ll focus on the case of a function of two variables \( f(x,y) \). We know that the graph of \( z = f(x,y) \) is a surface in \( \mathbb{R}^3 \), and that fixing \( y \) to be a constant is equivalent to slicing the surface with a plane parallel to the \( xz \)-plane. For example, the surface \( z = 4 - x^2 - y^2 \) is pictured in Figure 3.3.2(a), and the intersection with the plane \( y = 1 \) in Figure 3.3.2(b). Note that the intersection of \( z = f(x,y) \) with the plane is a curve in the plane. It turns out that \( \frac{\partial f}{\partial x}(x,y) \) is the slope of the tangent line to this curve of intersection.

One can see this analytically as well. An equation for the curve of intersection of \( z = 4 - x^2 - y^2 \) and the plane \( y = 1 \) can be obtained by substituting 1 for \( y \). Thus the curve is given by the equation \( z(x) = f(x, 1) = 3 - x^2 \) in the plane \( y = 1 \). The derivative of \( z \), therefore, is the same limit as in Definition 3.3.1 where 1 is substituted for \( y \) (and there is no third variable in \( f \), of course). Symbolically we have

\[
\frac{dz}{dx}(x) = \lim_{h \to 0} \frac{z(x + h) - z(x)}{h} = \lim_{h \to 0} \frac{f(x + h, 1) - f(x, 1)}{h} = \frac{\partial f}{\partial x}(x, 1).
\]
### 3.3. PARTIAL DERIVATIVES

**Extending Partial Differentiation to Vector-Valued Functions:** If \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m \) is a vector valued function of several variables, we can think of it as being comprised of its component functions. For example, if \( f(x, y) = (xy^2, xe^y, x^2 \ln y) \) its component functions are \( f_1(x, y) = xy^2 \), \( f_2(x, y) = xe^y \), and \( f_3(x, y) = x^2 \ln y \). To differentiate \( f \) with respect to \( x \), just differentiate component-wise. This is analogous to how we differentiated parametric curves, the only difference is now we’re taking partial derivatives of component functions rather than single-variable derivatives. Sometimes we combine like partial derivatives of component functions into a single vector, as with parametric curves and as in the following examples.

**Example 3.3.6. Partial differentiation with vector-valued functions**

Find the partial derivatives of \( f(x, y) = (xy^2, xe^y, x^2 \ln y) \). Differentiating component-wise gives

\[
\begin{align*}
    f_x(x, y) &= \left\langle y^2, e^y, 2x \ln y \right\rangle \\
    f_y(x, y) &= \left\langle 2xy, xe^y, x^2/y \right\rangle
\end{align*}
\]

Recall that in function terminology a parametric surface is a vector-valued function of two variables whose range is in three dimensions. We can now differentiate parametric surfaces! We close this section with an example illustrating the calculation, but in the next section we determine the geometric significance of these derivatives.

**Example 3.3.7. Partial Derivatives of parametric surfaces**

Parameterize the sphere \( \rho = 3 \) and find its partial derivatives.

Using the change of coordinates, we recall that the sphere is parameterized by

\[
S(\theta, \phi) = (3 \sin \phi \cos \theta, 3 \sin \phi \sin \theta, 3 \cos \phi), \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq \phi \leq \pi.
\]

One immediately computes the vectors of partial derivatives

\[
\begin{align*}
    S_\phi(\theta, \phi) &= (3 \cos \phi \cos \theta, 3 \cos \phi \sin \theta, -3 \sin \phi) , \\
    S_\theta(\theta, \phi) &= (-3 \sin \phi \sin \theta, 3 \sin \phi \cos \theta, 0) .
\end{align*}
\]
CHAPTER 3. DIFFERENTIATION

(a) The Domain in $\theta\phi$-plane
(b) Image sphere

Figure 3.3.3: Parametric Sphere and singular point

Observe that when $\phi = 0$ we have $S_\theta(\theta, 0)$ is the zero vector. This is because the image of the line segment from $(0, 0)$ to $(2\pi, 0)$ in the $\theta\phi$-plane on our parametric surface is the single point $(0, 0, 3)$. Since a whole line gets mapped to a single point, the parameterization is singular there as indicated by a zero derivative.

Exercises

1. Find all first order partial derivatives for the following functions:
   
   (a) $f(x, y) = \frac{x^2 y}{1 + x^2 + y^2}$
   
   (b) $f(x, y, z) = x \tan^{-1}(yz)$
   
   (c) $f(x, y) = \sec(x^2 + 2xy + y^2)$
   
   (d) $f(x, y, z) = \ln(xy) - e^{yz}$

2. Find all first order partial derivatives for the following functions:
   
   (a) $f(x, y) = \sec(xy) \tan(x^2)$
   
   (b) $f(x, y, z) = xyz e^{-x^2 - y^2}$
   
   (c) $f(x, y) = \sin^{-1}\left(\frac{x}{\sqrt{x^2 + y^2}}\right)$
   
   (d) $f(x, y, z) = 4x^2 + 3y^2 - 2z^2 - 3xy + 2yz - 6xz + 5x - 2y + 4z + 24$

3. A critical point of $f(x, y)$ is any point where both partial derivatives vanish. Find all critical points of
   
   (a) $f(x, y) = x^2 - 3xy + y^2 + 2x + 5$
   
   (b) $f(x, y) = x^2 - 2xy + y^2$
   
   (c) $f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$

4. Find the second-order partial derivatives of the general quadratic in two variables $f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F$. 

5. Find all second order partial derivatives of \( f(x, y) = \cos x \sin y - xy^3 \).

6. Find a function \( f(x, y) \) whose partial derivatives are \( f_x(x, y) = 3x^2 - 2x + 1 \) and \( f_y(x, y) = 2y \).

7. Find a function \( f(x, y) \) whose partial derivatives are \( f_x(x, y) = ye^{x+y} + 2xy + 3 \) and \( f_y(x, y) = xe^{x+y} + x^2 - 3y \).

8. The Cobb-Douglas production function is sometimes used in economics to model production based on labor and capital. One form is \( Y = \frac{L^{2/3}}{K^{1/3}} \), where \( Y \) is the value of all goods produced in a year, \( L \) is the number of hours worked, and \( K \) is the capital input. Find the first order partial derivatives of \( Y \).

9. Find the vectors of partial derivatives of the parametric surface

\[
S(r, \theta) = (r \cos \theta, r \sin \theta, r), \quad r \geq 0, 0 \leq \theta \leq 2\pi.
\]

10. Obtain a parameterization \( S(x, y) \) for the graph of \( f(x, y) = x^2 + y^2 - 2y \), then find the partial derivative vectors \( S_x \) and \( S_y \).

11. Parameterize the surface \( r = 7 \), and find the vectors of partial derivatives.

12. Parameterize the surface \( \phi = \pi/4 \) and find the vectors of partial derivatives.

13. Find the vectors of partial derivatives \( f_x, f_y \) and \( f_z \) of the vector-valued function

\[
f(x, y, z) = \left( \sin(2x - y + 3z), \frac{z}{x^2 + y^2 + 1} \right).
\]

14. Find the vectors of partial derivatives of \( f(x, y, z) = (x^2 z, e^{x^2 + y^2 + z^2}, xy^3) \).
3.4 Tangent Planes to Surfaces

In single variable calculus, differentiation can be introduced as a method for finding the equation of a tangent line to a curve. In multivariable calculus, we can use partial differentiation to find equations for tangent planes to surfaces. Recall that surfaces can be described parametrically, as a graph, or as a level surface. In each case we can use derivatives to find equations for tangent planes, although the level surface techniques will have to wait until we introduce the multivariable chain rule. We begin by defining the tangent plane to a surface.

Definition 3.4.1. Let \( p \) be a point on the surface \( S \). The tangent plane to \( S \) at \( p \) is the plane, if it exists, containing the tangent vectors to all curves on \( S \) that pass through \( p \).

Let \( S(s, t) = (x(s, t), y(s, t), z(s, t)) \) be a parametric surface. We have already seen that fixing \( s \) and letting \( t \) run yields a curve on the surface (they arise as grid lines in the diagrams of section 1.5). The derivative of a parametric curve is a tangent vector to it. Since partial differentiation treats other variables as constant, and fixing a parameter to be constant yields a curve, the partial derivative of a parametric surface is a vector tangent to the surface. Symbolically, we see that

\[
S_s(s, t) = \langle x_s(s, t), y_s(s, t), z_s(s, t) \rangle
\]

and

\[
S_t(s, t) = \langle x_t(s, t), y_t(s, t), z_t(s, t) \rangle
\]

are tangent vectors to the surface \( S(s, t) \). Since both are tangent to the surface, their cross product is a normal vector to the surface. This observation leads to

**Tangent Planes to Parametric Surfaces**

The cross product \( S_s(s_0, t_0) \times S_t(s_0, t_0) \) is normal to the parametric surface \( S(s, t) \) at the point \( S(s_0, t_0) \).

**Example 3.4.1. A paraboloid**

Find an equation for the tangent plane to the paraboloid

\[
S(r, \theta) = (r \cos \theta, r \sin \theta, 4 - r^2)
\]

at the point \( S(1, \pi/4) = (\sqrt{2}/2, \sqrt{2}/2, 3) \).

Rather than merely doing the calculation, we paralleled the justification given in the preceding paragraph. The curve \( S(1, \theta) = (\cos \theta, \sin \theta, 3) \) has tangent vector \( S_\theta(1, \theta) = (-\sin \theta, \cos \theta, 0) \). Evaluating this at \( \theta = \pi/4 \) gives a tangent vector to the curve \( S_\theta(1, \pi/4) = \langle -\sqrt{2}/2, \sqrt{2}/2, 0 \rangle \) as in Figure 3.4.1(a). Since it is tangent to a curve on the surface, it is a vector in the tangent plane by Definition 3.4.1.

Similarly, the curve on the surface obtained by fixing \( \theta = \pi/4 \) is given parametrically by \( S(r, \pi/4) = (\sqrt{2}/2r, \sqrt{2}/2r, 4 - r^2) \). A tangent to the curve is \( S_r(r, \pi/4) = \langle \sqrt{2}/2, \sqrt{2}/2, -2r \rangle \). Thus when \( r = 1 \), we see that \( S_r(1, \pi/4) = \langle \sqrt{2}/2, \sqrt{2}/2, -2 \rangle \).
### 3.4. TANGENT PLANES TO SURFACES

Figure 3.4.1: Parametric tangent plane

\( \left\langle \sqrt{2}/2, \sqrt{2}/2, -2 \right\rangle \) is tangent to a curve on the surface, hence in the tangent plane (see Figure 3.4.1(b)).

Thus we know two vectors in the tangent plane, and their cross product is normal to the plane. We calculate

\[
\mathbf{n} = \mathbf{S}_s \times \mathbf{S}_t = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sqrt{2}/2 & \sqrt{2}/2 & 0 \\ \sqrt{2}/2 & \sqrt{2}/2 & -2 \end{vmatrix} = -\sqrt{2}\mathbf{i} - \sqrt{2}\mathbf{j} - \mathbf{k}.
\]

Any scalar multiple of \( \mathbf{n} \) is normal to the tangent plane, so we choose \( \left\langle \sqrt{2}, \sqrt{2}, 1 \right\rangle \) and get the equation

\[
\sqrt{2}x + \sqrt{2}y + z = \left\langle \sqrt{2}, \sqrt{2}, 1 \right\rangle \cdot \left\langle \sqrt{2}/2, \sqrt{2}/2, 3 \right\rangle = 5.
\]

**Example 3.4.2.** The tangent plane to a torus.

Recall that \( \mathbf{S}(s,t) = ((\cos t + 2) \cos s, (\cos t + 2) \sin s, \sin t) \), for \( 0 \leq s, t \leq 2\pi \), parameterizes a torus. Calculating partial derivatives gives us tangent vectors to the surface, and we calculate them coordinate-wise:

\[
\mathbf{S}_s(s,t) = \left\langle -(\cos t + 2) \sin s, (\cos t + 2) \cos s, 0 \right\rangle,
\]

\[
\mathbf{S}_t(s,t) = \left\langle -\sin t \cos s, -\sin t \sin s, \cos t \right\rangle.
\]

To get the tangent plane to \( \mathbf{S}(s,t) \) we need a point and a normal direction. The parametric equations give the point, while the cross product of the partial derivatives gives the normal direction. Thus a normal direction at \( \mathbf{S}(s,t) \) is

\[
\mathbf{S}_s \times \mathbf{S}_t = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -(\cos t + 2) \sin s & (\cos t + 2) \cos s & 0 \\ -\sin t \cos s & -\sin t \sin s & \cos t \end{vmatrix} = ((\cos t + 2) \cos s \cos t)\mathbf{i} + ((\cos t + 2) \sin s \cos t)\mathbf{j} + ((\cos t + 2) \sin t)\mathbf{k}
\]
We use these calculations to find an equation for the tangent plane at \((s, t) = (\pi/4, \pi/4)\). The point on the tangent plane is \(S(\pi/4, \pi/4) = (\sqrt{2} + 1/2, \sqrt{2} + 1/2, \sqrt{2}/2)\) and a normal vector is
\[
S_s(\pi/4, \pi/4) \times S_t(\pi/4, \pi/4) = (\sqrt{2}/2 + 2) \langle 1/2, 1/2, \sqrt{2}/2 \rangle.
\]
The scalar multiple \(n = \langle 1, 1, \sqrt{2} \rangle\) is also normal, so we get the equation
\[
x + y + \sqrt{2}z = \left(1, 1, \sqrt{2}\right) \cdot \left(\sqrt{2} + 1/2, \sqrt{2} + 1/2, \sqrt{2}/2\right) = 2\sqrt{2} + 2.
\]

**Tangent planes to graphs:** We now use the above remarks to find tangent planes to graphs of functions of two variables. Recall that \(S(x, y) = (x, y, f(x, y))\) parameterizes the graph of \(f(x, y)\). One calculates the normal vector
\[
S_x \times S_y = \begin{vmatrix} i & j & k \\ 1 & 0 & f_x \\ 0 & 1 & f_y \end{vmatrix} = -f_x i - f_y j + k
\]
Using this we get:

**Example 3.4.3. Horizontal Tangent Planes**

Find points on the graph of \(f(x, y) = x^3 - x + y^2\) where the tangent planes are horizontal. This is equivalent to finding where the normal vector is parallel.
to the standard basis vector $k$, so we solve the vector equation $n = \lambda k$. By the previous observations we have

$$n = \langle -f_x(x_0, y_0), -f_y(x_0, y_0), 1 \rangle = \langle 3x^2 - 1, 2y, 1 \rangle = \langle 0, 0, \lambda \rangle.$$

Solving the one vector equation is equivalent to solving the system of three coordinate equations:

$$3x^2 - 1 = 0$$
$$2y = 0$$
$$1 = \lambda$$

Whose solution yields $(x, y) = (\pm 1/\sqrt{3}, 0)$. Substituting into $f(x, y)$ to find the $z$-coordinate, we find the points on the surface that have horizontal tangent planes are $(\pm 1/\sqrt{3}, 0, \mp \frac{2}{3\sqrt{3}})$. ▲

Figure 3.4.3: Horizontal tangents to a graph

Note that one of the points is a saddle similar to Figure 1.5.1, while the other is a local minimum. We’ll see how to determine the nature of these critical points using tests similar to the second derivative test in single variable calculus.

**Example 3.4.4. Approximating Surface Area**

Use a parallelogram in the tangent plane to approximate the area of the patch of the unit sphere parameterized by

$$S(\theta, \phi) = (\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \phi), \ \pi/2 \leq \phi \leq 2\pi/3, \ 0 \leq \theta \leq \pi/6.$$

The portion of the sphere is highlighted in Figure 3.4.4(a), and the question is what parallelogram should we use. The vectors of partial derivatives $S_\theta(0, \pi/2)$ and $S_\phi(0, \pi/2)$ lie in the tangent plane to the patch at the corner $S(0, \pi/2) = \ldots$
(1, 0, 0). It stands to reason that some scalar multiple of these should be used to approximate the area, and the limits on $\theta$ and $\phi$ determine which. In general, you multiply each vector by the change in the parameter over the limits of the parameterization: $(\Delta \theta) S_\theta$ and $(\Delta \phi) S_\phi$. The change $\Delta \theta$ and $\Delta \phi$ over the region we are parameterizing is $\pi/6$. Thus the area of the spherical patch is approximately the area of the parallelogram determined by the vectors $(\pi/6)S_\theta$ and $(\pi/6)S_\phi$. This area, you recall, is the length of the cross product of the two vectors. We proceed with the calculations.

$$S_\theta(\theta, \phi) = (-\sin \phi \sin \theta, \sin \phi \cos \theta, 0)$$

$$S_\phi(\theta, \phi) = (\cos \phi \cos \theta, \cos \phi \sin \theta, -\sin \phi)$$

We are interested in these vectors at the point $(\theta, \phi) = (0, \pi/2)$, which are $S_\theta(0, \pi/2) = (0, 1, 0)$ and $S_\phi(0, \pi/2) = (0, 0, -1)$. Further, to approximate the given patch, we multiply each by the scalar $\pi/6$ and find the length of the cross product. The cross product is:

$$(\pi/6) S_\theta \times (\pi/6) S_\phi = \begin{vmatrix} i & j & k \\ 0 & \pi/6 & 0 \\ 0 & 0 & -\pi/6 \end{vmatrix} = -\frac{\pi^2}{36} i.$$  

The length is evidently $\frac{\pi^2}{36} \approx .2746$. One can compute the actual area of the patch is $\frac{\pi}{12} \approx .2618$. ▲

We wish to generalize the problem slightly. If we wish to approximate the patch of the sphere given by the limits $\pi/2 \leq \phi \leq \pi/2 + \Delta \phi$, $0 \leq \theta \leq \Delta \theta$, we would use the parallelogram determined by $(\Delta \theta) S_\theta$ and $(\Delta \phi) S_\phi$. The cross product calculation becomes

$$(\Delta \theta) S_\theta \times (\Delta \phi) S_\phi = \begin{vmatrix} i & j & k \\ 0 & \Delta \theta & 0 \\ 0 & 0 & -\Delta \phi \end{vmatrix} = -\Delta \theta \Delta \phi i.$$
3.4. TANGENT PLANES TO SURFACES

Thus the approximate area of the patch is $\Delta \theta \Delta \phi$. One can show the actual area of the patch is $\Delta \theta \sin(\Delta \phi)$. Recall that for small $\Delta \phi$ we know $\sin(\Delta \phi) \approx \Delta \phi$, so this approximation gets better as our patch gets smaller. This idea is summarized below.

**Approximating parametric surface area with parallelograms**

The approximate area of the surface

$$S(s, t), \ a \leq s \leq a + \Delta s, \ b \leq t \leq b + \Delta t,$$

is

$$Area \approx \Delta s \Delta t \|S_s(a, b) \times S_t(a, b)\|.$$  

This is the area of the parallelogram determined by $\Delta s S_s(a, b)$ and $\Delta t S_t(a, b)$ (we are assuming both $\Delta s$ and $\Delta t$ are positive and neither $S_s(a, b)$ nor $S_t(a, b)$ are the zero vector).

**Example 3.4.5. Cylindrical surface area**

Approximate the surface area of the cylindrical patch

$$S(s, t) = (3 \cos t, 3 \sin t, s), \ 0 \leq s \leq 2, \ 0 \leq t \leq \pi/6.$$  

Using the above strategy, we need to calculate the length

$$\|2S_s(0, 0) \times (\pi/6)S_t(0, 0)\|.$$  

Since $S_s(s, t) = (0, 0, 1)$ and $S_t(s, t) = (-3 \sin t, 3 \cos t, 0)$, we have

$$2S_s(0, 0) \times (\pi/6)S_t(0, 0) = \begin{vmatrix}
i & j & k \\
0 & 0 & 2 \\
0 & \pi/2 & 0 \\
\end{vmatrix} = -\pi i.$$  

Thus the area is approximately $\pi$ square units. In fact, for cylinders we know the area is the height times the circumference. In this case the height of the patch is 2, and the circumference is $3 \cdot (\pi/6) = \pi/2$. Therefore the actual area is $2(\pi/2) = \pi$. So our approximation gives the actual area! This is not common, but follows from the fact that cylinders can be flattened without stretching. **Concept Connection:** The idea of approximating surface area with parallelogram tiles will play an important role in deriving an integral formula for area of a parametric surface. We will approximate the surface with parallelogram shingles, and take a limit as the approximation gets finer. In this example, we saw that as the patch got smaller, the approximation got better. This is evidence that the integral we derive will give the desired area of the surface.

We have begun our analysis of differentiation by defining partial derivatives. Then we saw how partial differentiation can be used to find tangent planes to surfaces given parametrically and as the graph of $f(x, y)$. These are two ways we
defined surfaces in Chapter 1. We also encountered level surfaces of a function of three variables. After introducing gradients and the multivariable chain rule we will be able to find tangent planes to level surfaces as well.

Exercises

1. Find an equation for the tangent plane to the surface \( \mathbf{S}(s, t) = (\sin s \cos t, \sin s \sin t, \cos s) \) at the point \((s, t) = (\pi/4, \pi/2)\). Sketch the surface and its tangent plane.

2. Parameterize the surface \( \phi = \pi/3 \) and find an equation for the tangent plane to it at the point \((\sqrt{3}/2, 0, 1/2)\).

3. Find where the tangent plane to \( \mathbf{S}(s, t) = (s + t, s - t, 4st + 4t - 2s) \), \(-\infty < s, t < \infty\), is horizontal.

4. Find an equation for the tangent plane to the paraboloid \( f(x, y) = 6 - x^2 - y^2 + 2x - 4y \) at the point \((0, 0, 6)\).

5. Find an equation for the tangent plane to the graph of \( f(x, y) = 2x^2 - 3xy^3 \) at \((x, y) = (2, 1)\).

6. Find where the tangent plane to \( z = x^2 - 2xy - 3y^2 + 4x \) is horizontal.

7. Find where the tangent plane to \( f(x, y) = (x^2 + y^2)^2 - (x^2 + y^2) \) is horizontal. Sketch the intersection of the surface with the \( yz\)-plane, then rotate it about the \( z\)-axis to see what it looks like.

8. Approximate the area of the paraboloid patch
\[
\mathbf{S}(s, t) = (s \cos t, s \sin t, s^2), \quad 1 \leq s \leq 2, \, \pi/6 \leq t \leq \pi/3.
\]

9. Approximate the area of the conical patch
\[
\mathbf{S}(s, t) = (s \cos t, s \sin t, s), \quad 1 \leq s \leq 2, \, 0 \leq t \leq \pi/6.
\]

10. Approximate the area of the conical patch
\[
\mathbf{S}(s, t) = (s \cos t, s \sin t, s), \quad 1 \leq s \leq 1 + \Delta s, \, 0 \leq t \leq \Delta t.
\]

Use geometry to determine the actual area, and compare the two.
3.5 The Chain Rule

In this section we discuss the multivariable chain rule, which generalizes the single variable one. Recall that the single-variable chain rule says that the derivative of a composition of functions is the product of the derivatives. Let’s say that again:

To differentiate a composition of functions, multiply the derivatives.

It turns out that this idea generalizes beautifully to the multivariable setting, with the right interpretation of the terms “derivative” and “multiply”.

We begin this section by making sense of the derivative $Df$

**Definition 3.5.1.** The gradient $\nabla f(x, y, z)$ of a function $f(x, y, z)$ of several variables is the vector of partial derivatives. In symbols,

$$\nabla f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z) \rangle.$$  

The gradient of a function encodes significant information about its rate of change, as well as about the geometry of it. It is also another example of a vector field, since it assigns a vector to every point in the domain of $f$.

**Example 3.5.1. Gradients of functions**

Find the gradients of $g(x, y) = 2x^3y - xy^2$ and $f(x, y, z) = x \tan^{-1} y - y \sin z$.

Taking partial derivatives we see

$$\nabla g(x, y) = \langle 6x^2y - y^2, 2x^3 - 2xy \rangle$$

$$\nabla f(x, y, z) = \langle \tan^{-1} y, \frac{x}{1 + y^2} - \sin z, -y \cos z \rangle.$$  

Notice that the gradient of a function of several variables is a vector field on the domain of the function. In the above example, the domain of $g(x, y) = 2x^3y - xy^2$ is $\mathbb{R}^2$. The gradient $\nabla g(x, y) = \langle 6x^2y - y^2, 2x^3 - 2xy \rangle$ of $g(x, y)$ is a vector field on the domain $\mathbb{R}^2$ (see Figure 3.5.1). More generally, if $f : \mathbb{R}^n \to \mathbb{R}$, then there are $n$ independent variables, so $n$ partial derivatives. Therefore $\nabla f$ is an $n$-vector.

**Example 3.5.2. Electric Potential**

Any electric field turns out to be the (negative of the) gradient of a potential function. In symbols, we have $E = -\nabla V$, where $E$ is the force per unit charge (a vector) and $V$ is the potential (a scalar). We restrict our attention to the plane for this example. Although that is not physically realistic, it allows for pictures that are easier to understand.

For a single point charge $q$ placed at the origin in the plane, its potential function can be shown to be

$$V(x, y) = \frac{q}{4\pi \epsilon_0 \sqrt{x^2 + y^2}}.$$
### DIFFERENTIATION

(a) Gradient of $g(x, y) = 2x^3y - xy^2$

(b) Gradient of the potential $V$

Figure 3.5.1: Gradient vector fields

Taking partial derivatives gives the gradient

$$\nabla V(x, y) = \left\langle -\frac{qx}{4\pi \epsilon_0 (x^2 + y^2)^{3/2}}, -\frac{qy}{4\pi \epsilon_0 (x^2 + y^2)^{3/2}} \right\rangle = -\frac{q}{4\pi \epsilon_0 (x^2 + y^2)^{3/2}} (x, y).$$

The electric field generated by the point charge is then

$$E = -\nabla V = \frac{q}{4\pi \epsilon_0 (x^2 + y^2)^{3/2}} (x, y).$$

For the appropriate choice of $q$, this gradient field is illustrated in Figure 3.5.1(b).

Thus for every function of several variables we can associate a “derivative vector”, its gradient. Parametric surfaces $S(s, t)$ are examples of vector-valued functions of several variables. The coordinate functions $x(s, t)$, $y(s, t)$, and $z(s, t)$ are functions of several variables, and so have gradients. We define the derivative matrix $D_S(s, t)$ of $S(s, t)$ to be

$$D_S(s, t) = \begin{pmatrix} x_s & x_t \\ y_s & y_t \\ z_s & z_t \end{pmatrix}.$$

Notice that there is a row for each component function, and a column for each parameter. The columns are what we called $x_s$ and $x_t$ in the previous section.

We can define the derivative matrix for any vector-valued function of several variables. Suppose

$$f(x_1, x_2, \ldots, x_n) = \langle f_1(x_1, x_2, \ldots, x_n), f_2(x_1, x_2, \ldots, x_n), \ldots, f_m(x_1, x_2, \ldots, x_n) \rangle$$
then the derivative matrix $Df$ is the one with $\frac{\partial f_i}{\partial x_j}$ in the $i^{th}$ row and $j^{th}$ column. So $Df$ is the matrix whose rows are the gradients of the coordinate functions $f_i$.

**Example 3.5.3. Derivative matrices**

Find the derivative matrix of $f(x,y) = (2x + 3y, x - 2y)$. Taking partial derivatives we have

$$Df(x,y) = \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}.$$ 

The derivative matrix for the parameteric torus

$$S(s,t) = ((\cos t + 2) \cos s, (\cos t + 2) \sin s, \sin t)$$

is the $3 \times 2$ matrix

$$DS(s,t) = \begin{pmatrix} -\sin t \cos s & -\sin t \cos s \\ -(\cos t + 2) \sin s & -(\cos t + 2) \cos s \\ 0 & \cos t \end{pmatrix}. \quad \blacksquare$$

It turns out that matrices provide the perfect format for describing the multivariable chain rule. The derivative matrix of a composition of functions is the product of their corresponding derivative matrices! We now set about deriving this. We motivate the chain rule with an example.

**Example 3.5.4. Related Rates...AGAIN**

You thought you were done thinking about related rates problems after your first calculus course...unfortunately, they are back! In this context they provide a nice motivation for the multivariable chain rule. We consider the following example:

A conveyor belt pours sand into a conical pile so that its volume increases at a rate of $10 \text{ ft}^3/\text{min}$ and its radius increases at a rate of $1 \text{ ft/min}$. Determine how fast the height of the pile is increasing when $r = h = 50 \text{ ft}$.

To solve this problem, one differentiates the relationship $V = \frac{\pi}{3} r^2 h$ with respect to $t$, and substitutes the appropriate values. Differentiating gives

$$\frac{dV}{dt} = \frac{\pi}{3} 2rh \frac{dr}{dt} + \frac{\pi}{3} r^2 \frac{dh}{dt}, \quad (3.5.1)$$

and substituting the given constants allows one to solve the problem. Our purpose here, however, is not to solve the problem, but to put it in a multivariable context and motivate the chain rule.

First note that volume is a function of two variables $r$ and $h$. These, in turn, are functions of time $t$. Composing them with the volume formula $V = \frac{\pi}{3} r^2 h$ makes $V$ a function of $t$. Since $V$ is a composition and since it involves several variables, differentiating it requires a multivariable chain rule! Let's look more closely at Equation 3.5.1 to see what that chain rule is.
The function \( V = \frac{\pi}{3} r^2 h \), as a function of two variables, has the gradient
\[
\nabla V(r, h) = \left\langle \frac{\pi}{3} 2rh, \frac{\pi}{3} r^2 \right\rangle.
\]
Further, since \( r \) and \( h \) are functions of \( t \), we can consider the vector-valued function \( x(t) = (r(t), h(t)) \). Differentiating as with parametric curves, we have
\[
x'(t) = \left\langle \frac{dr}{dt}, \frac{dh}{dt} \right\rangle.
\]
The components of both \( \nabla V(r, h) \) and \( x'(t) \) appear in certain parts of Equation 3.5.1. In fact, careful inspection shows that to get \( dV/dt \) you just take the dot product of \( \nabla V(r, h) \) and \( x'(t) \)!
\[
\nabla V(r, h) \cdot x'(t) = \left\langle \frac{\pi}{3} 2rh, \frac{\pi}{3} r^2 \right\rangle \cdot \left\langle \frac{dr}{dt}, \frac{dh}{dt} \right\rangle = \frac{\pi}{3} 2rh \frac{dr}{dt} + \frac{\pi}{3} r^2 \frac{dh}{dt} = \frac{dV}{dt}. \tag{3.5.2}
\]
Thus the derivative of the composition of \( V(r, h) \) and \( x(t) \) is the dot product of their derivatives!

That’s the big deal, and that’s what you want to remember. Now I’ll ramble on about notation a bit. It is important to be precise, and the following describes Equation 3.5.2 while taking a little more care with notation. When we calculate the volume it is at a given time, using the values of \( r \) and \( h \) at that time—in other words, using \( x(t) \). Thus the volume we are interested in is
\[
V(x(t)) = V \circ x(t),
\]
and our function indeed looks like a composition now! Using composition notation, every occurrence of the ordered pair \((r, h)\) should be replaced by the function \( x(t) \). Thus the gradient \( \nabla V(r, h) \) of Equation 3.5.2 is the same as \( \nabla V(x(t)) \). Eliminating the middle terms in Equation 3.5.2 and using composition notation, we have
\[
\frac{d}{dt} V \circ x(t) = \frac{d}{dt} V(x(t)) = \nabla V(x(t)) \cdot x'(t).
\]
Before stating the chain rule more generally, we illustrate with a second example from a physical context.

**Example 3.5.5. Temperature and the chain rule**

Suppose the temperature in the plane at the point \((x, y)\) is given by the formula \( f(x, y) = x^2 - xy + y^3 \), and an ant is crawling along the unit circle so that its position at time \( t \) is given by \( C(t) = (\cos t, \sin t) \). The composition \( f \circ C(t) \) then tells the temperature the ant is experiencing at time \( t \). The derivative \( \frac{d}{dt} f \circ C(t) \) is how fast temperature is changing for the ant.
One method of calculating $\frac{d}{dt} f \circ C(t)$ is to first evaluate $f(x,y)$ on the curve $C(t)$, which gives a one-variable function $f(C(t))$. Then one can use single-variable techniques to differentiate it. The first step, evaluating $f$ on $C(t)$, gives us:

$$f \circ C(t) = f(C(t)) = f(\cos t, \sin t) = (\cos t)^2 - (\cos t)(\sin t) + (\sin t)^3.$$ 

This is a function of one variable, and we can use single variable calculus to differentiate it. In order to make the connection with $\nabla f(x,y)$ and $C'(t)$ we won't simplify completely, and will regroup the terms in a somewhat strange way. The algebra between the following steps is described below.

$$\frac{d}{dt} f \circ C(t) = 2(\cos t)(-\sin t) - ((-\sin t)(\sin t) + (\cos t)(\cos t)) + 3(\sin t)^2(\cos t)$$

$$= (2(\cos t) - \sin t)(-\sin t) + \left(-\cos t + 3(\sin t)^2\right)(\cos t)$$

$$= \langle 2 \cos t - \sin t, -\cos t + 3 \sin^2 t \rangle \cdot \langle -\sin t, \cos t \rangle$$

Between the first and second line we factored out $-\sin t$ from the first two terms and $\cos t$ from the second two. The reason for this is that they are the derivatives of the coordinate functions of $C(t)$, and they appear because of the chain rule in one variable.

Recall that we wish to describe the derivative $\frac{d}{dt} f \circ C(t)$ of Equation 3.5.3 in terms of the gradient $\nabla f(x,y)$ of the temperature function, and the tangent vector $C'(t)$ to the curve. In order to notice the connection we calculate the gradient, evaluate it on the circle (i.e. substitute the parametric equations of $C(t)$ in for $x$ and $y$ in $\nabla f(x,y)$), and find $C'(t)$.

$$\nabla f(x,y) = \langle 2x - y, -x + 3y^2 \rangle$$

$$\nabla f(C(t)) = \nabla f(\cos t, \sin t) = \langle 2 \cos t - \sin t, -\cos t + 3 \sin^2 t \rangle$$

$$C'(t) = \langle -\sin t, \cos t \rangle$$

Comparing these calculations with Equation 3.5.3, we see that in this case

$$\frac{d}{dt} f \circ C(t) = \frac{d}{dt} f(C(t)) = \nabla f(C(t)) \cdot C'(t). \ \blacksquare$$

We can interpret this calculation geometrically. In our case $C'(t)$ is a unit vector. This means that the dot product $\nabla f(C(t)) \cdot C'(t)$ is the component of $\nabla f(C(t))$ in the direction of $C'(t)$. More generally, if $C'(t)$ is not a unit vector, then $\nabla f(C(t)) \cdot C'(t)$ is the speed of $C(t)$ times the component of $\nabla f(C(t))$ in the direction of $C'(t)$.

**Chain Rule on a Curve:** The above examples generalize, giving a chain rule for functions evaluated on parametric curves. We take a moment to remind
you that a parametric curve is equivalently called a vector-valued function of a single variable. The two descriptions are interchangeable. The difference is that the parametric curve description emphasizes the geometry of \( C(t) \) while the vector-valued function description focuses on the analytic nature of it. The analytic approach will be useful when we discuss the general multivariable chain rule.

**Theorem 3.5.1.** Let \( f(x, y, z) \) be a function of several variables with continuous partial derivatives and \( C(t) = (x(t), y(t), z(t)) \) a differentiable parametric curve. The derivative of \( f \circ C(t) \) is given by

\[
\frac{d}{dt} f \circ C(t) = \nabla f(C) \cdot C'(t).
\]

**Proof.** To simplify notation we prove the case when \( f \) is a function of two variables. Let \( C(t) = (x(t), y(t)) \). Then \( f \circ C(t) = f(x(t), y(t)) \), and by the definition of the derivative

\[
\frac{d}{dt} f \circ C(t) = \lim_{h \to 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t))}{h}
= \lim_{h \to 0} \frac{f(x(t+h), y(t+h)) - f(x(t), y(t+h))}{h} + \lim_{h \to 0} \frac{f(x(t), y(t+h)) - f(x(t), y(t))}{h}.
\]

(3.5.4)

The second equality follows from adding \(-f(x(t), y(t+h)) + f(x(t), y(t+h))\) to the numerator and separating the fraction to a sum of two limits. We will show that the second term in this sum is \( f_y(C(t)) \cdot y'(t) \). The analysis is similar to show the first term is \( f_x(C(t)) \cdot x'(t) \). The second term of Equation 3.5.4 can be multiplied by \((y(t+h) - y(t)) / (y(t+h) - y(t))\) and rearranged to give

\[
\lim_{h \to 0} \frac{f(x(t), y(t+h)) - f(x(t), y(t))}{y(t+h) - y(t)} = \frac{\partial f}{\partial y}(x(t), c).
\]

(3.5.5)

The first factor of Expression 3.5.5 holds the \( x \)-coordinate constant, so by the mean value theorem in single variable calculus, there is a \( c \) between \( y(t) \) and \( y(t+h) \) such that

\[
\frac{f(x(t), y(t+h)) - f(x(t), y(t))}{y(t+h) - y(t)} = \frac{\partial f}{\partial y}(x(t), c).
\]
As \( h \to 0 \) we see that \( y(t+h) \to y(t) \), which implies that \( c \to y(t) \) as well. Since we assume \( f_y \) is continuous, the limit of the first factor in Expression 3.5.5 is \( \frac{df}{dy}(x(t), y(t)) = f_y(C(t)). \) The second factor is the definition of the derivative \( y'(t) \) of the coordinate function \( y(t) \). Therefore the second term in the sum of Equation 3.5.4 is \( f_y(C(t)) \cdot y'(t) \). A similar analysis on the first term shows that it is \( f_x(C(t)) \cdot x'(t) \), showing

\[
\frac{d}{dt} f \circ C(t) = f_x(C(t)) \cdot x'(t) + f_y(C(t)) \cdot y'(t) = \nabla f(C(t)) \cdot C'(t). \quad \Box
\]

The formula suggested by certain examples has now been shown to be true in general. Moreover we can interpret the derivative in terms of projections and lengths of the vectors involved. Precisely, we can manipulate the result of Theorem 3.5.1 algebraically as follows:

\[
\frac{d}{dt} f \circ C(t) = \nabla f(C(t)) \cdot C'(t) = \|C'(t)\| \frac{\nabla f(C(t)) \cdot C'(t)}{\|C'(t)\|} = \|C'(t)\| \text{comp}_{C(t)} \nabla f(C(t)).
\]

Thus \( \frac{d}{dt} f(C(t)) \) does represent the speed of the curve times the component of \( \nabla f \) in the tangential direction.

One more remark is in order before moving on to examples. In single-variable calculus important things happened when the derivative was zero (e.g. local extreme values or inflection points). In this course we have found that when the dot product of non-zero vectors is zero, the vectors are orthogonal. Taking these observations together, Theorem 3.5.1 says that important things happen when \( \nabla f(C(t)) \) is perpendicular to \( C'(t) \).

**Example 3.5.6. Verifying the chain rule**

Verify the chain rule for the function \( f(x, y) = xe^{2y-3x} \) and the curve \( C(t) = (t, 3 - 2t) \).

Of course we have a proof that the chain rule is true, so in some sense there is no need to verify it! However, for someone seeing the notation for the first time the verification process can help with understanding. To verify

\[
\frac{d}{dt} f \circ C(t) = \nabla f(C(t)) \cdot C'(t),
\]

we simply calculate both sides separately, step back, and say “Hey, they’re equal.”

To calculate \( \frac{d}{dt} f \circ C(t) \) we first form the composite function \( f \circ C(t) \) then take its derivative as in Calc I. In our case we see

\[
f \circ C(t) = f(t, 3 - 2t) = te^{2(3-2t)-3t} = te^{6-7t},
\]

so the derivative is

\[
\frac{d}{dt} f \circ C(t) = e^{6-7t} - 7te^{6-7t}.
\]
On the other hand we compute $\nabla f(C) \cdot C'(t)$ by finding $\nabla f(x, y)$ and evaluating it on $C(t)$, then dotting the result with the tangent vector $C'(t)$. In our case

$$\nabla f(x, y) = \langle e^{2y-3x} - 3xe^{2y-3x}, 2xe^{2y-3x} \rangle,$$

so the needed dot product is

$$\nabla f(C) \cdot C'(t) = \langle e^{6-7t} - 3te^{6-7t}, 2te^{6-7t} \rangle \cdot \langle 1, -2 \rangle = e^{6-7t} - 7te^{6-7t}. \quad (3.5.7)$$

Now we step back and say “Hey, Equations 3.5.6 and 3.5.7 are equal!” So we have verified the chain rule in this case.

**Example 3.5.7. A Minimization Problem**

Find the point on the parametric curve $C(t) = (3 - 2t, 4 + t, -3t)$, $-\infty < t < \infty$ closest to the origin.

We could use Calc I techniques, but want to cast the problem in multivariable language. To minimize distance one frequently minimizes the distance squared, avoiding the square root in calculations. The square of the distance from the point $(x, y, z)$ to the origin is given by $f(x, y, z) = x^2 + y^2 + z^2$. The squared distance from the point $C(t)$ on the curve to the origin is given by $f(C(t))$, so this is the function we want to minimize. A minimum will occur when the derivative is zero, which is when the gradient $\nabla f(C(t))$ is perpendicular to the tangent vector $C'(t)$. Thus to solve the problem, we need to solve $\nabla f(C(t)) \cdot C'(t) = 0$.

Now that we’ve cast the problem in chain rule language, we commence with the calculations. We see that $\nabla f(x, y, z) = (2x, 2y, 2z)$ and $C'(t) = (1, -2)$, so we need to solve

$$\nabla f(C(t)) \cdot C'(t) = (6 - 4t, 8 + 2t, -6t) \cdot (1, -2, -3) = -4 + 28t = 0.$$

The solution is $t = 1/7$, so the point on the line $C(t)$ that is closest to the origin is $C(1/7) = (19/7, 29/7, -3/7)$. In general, one has to verify that the result is a minimum, but the context of this problem makes that clear.

**Example 3.5.8. Extreme Temperatures**

The temperature at the point $(x, y)$ is given by

$$T(x, y) = 10 - \frac{5}{1 + x^2 - 2xy + y^2}.$$

Find the maximum and minimum temperatures experienced by an ant crawling on the circle $C(t) = (2 \cos t, 2 \sin t)$, $0 \leq t \leq 2\pi$.

In this problem we want to optimize the function $T \circ C(t)$, so we’ll take its
derivative using the chain rule. The necessary ingredients are

\[
\nabla T(x, y) = \begin{pmatrix}
\frac{10(x - y)}{(1 + x^2 - 2xy + y^2)^2} \\
\frac{-10(x - y)}{(1 + x^2 - 2xy + y^2)^2}
\end{pmatrix}
\]

\[
\nabla T(C(t)) = \frac{20}{(5 - 8 \sin t \cos t)^2} \langle \cos t - \sin t, \sin t - \cos t \rangle
\]

\[
C'(t) = \langle -2 \sin t, 2 \cos t \rangle.
\]

We now solve \( \nabla T(C(t)) \cdot C'(t) = 0: \)

\[
\nabla T(C(t)) \cdot C'(t) = \frac{40}{(5 - 8 \sin t \cos t)^2} \left( \sin^2 t - \cos^2 t \right)
\]

\[
= -\frac{40 \cos 2t}{5 - 4 \sin 2t} = 0.
\]

The denominator is never zero, so the derivative is always defined. It equals zero when \( \cos 2t = 0 \), which is when \( t = \pi/4, 3\pi/4, 5\pi/4, 7\pi/4 \). Thus the possible extreme temperatures occur at the points \((\sqrt{2}, \sqrt{2}), (-\sqrt{2}, \sqrt{2}), (\sqrt{2}, -\sqrt{2}), \) and \((-\sqrt{2}, -\sqrt{2})\). Evaluating \( T \) at these points gives

\[
T(\sqrt{2}, \sqrt{2}) = T(-\sqrt{2}, -\sqrt{2}) = 5
\]

as a minimum temperature and

\[
T(-\sqrt{2}, \sqrt{2}) = T(\sqrt{2}, -\sqrt{2}) = 85/9
\]

as a maximum. ▲

**The General Chain Rule:** The above example generalizes to vector valued functions of several variables. The one variable chain rule is sometimes remembered as “the derivative of the outside function, leaving the inside the same, times the derivative of the inside function”. Interpreting this, if we let \( f \circ g(x) = f(g(x)) \) then \( g \) is the inside and \( f \) the outside function. The single variable chain rule states \( (f \circ g)'(x) = f'(g(x)) \cdot g'(x) \). Another way to think of it is that the derivative of a composition is the product of the derivatives, you just have to evaluate them where they make sense.

The multivariable chain rule is the same, just replace “derivative” with “derivative matrix”. So if \( f : \mathbb{R}^m \to \mathbb{R}^p \) and \( g : \mathbb{R}^n \to \mathbb{R}^p \), then \( f \circ g : \mathbb{R}^n \to \mathbb{R}^p \) and the derivative matrix \( D(f \circ g) \) is the product of the derivative matrices \( Df \) and \( Dg \). As in the single variable case, you have to evaluate \( Df \) at \( g \). We have the following theorem:

**Theorem 3.5.2.** Let \( f : \mathbb{R}^m \to \mathbb{R}^p \) and \( g : \mathbb{R}^n \to \mathbb{R}^p \) be differentiable functions and let \( x \in \mathbb{R}^n \). Then

\[
D(f \circ g)(x) = Df(g(x)) \cdot Dg(x).
\]
Proof. The composition \( f \circ g \) has \( p \) component functions, each of which is a function of \( n \) variables. Let \( (f \circ g)_i \) denote the \( i^{th} \) component function, and \( x_j \) the \( j^{th} \) variable. By definition, the entry in the \( i^{th} \) row and \( j^{th} \) column of \( D(f \circ g)(x) \) is \( \frac{\partial (f \circ g)_i}{\partial x_j} \). Since this is partial differentiation, we fix all coordinates other than \( x_j \), and can think of \( g \) as a function of one variable. By Theorem 3.5.1 we have

\[
\frac{\partial (f \circ g)_i}{\partial x_j} = \nabla f_i(g(x)) \cdot \left( \frac{\partial g_1}{\partial x_j}, \frac{\partial g_2}{\partial x_j}, \ldots, \frac{\partial g_m}{\partial x_j} \right).
\]

The vector \( \nabla f_i(g(x)) \) is the \( i^{th} \) row of \( Df(g(x)) \), while the \( j^{th} \) column of \( Dg(x) \) is the vector \( \left( \frac{\partial g_1}{\partial x_j}, \frac{\partial g_2}{\partial x_j}, \ldots, \frac{\partial g_m}{\partial x_j} \right) \). Thus the \( ij \) entry of \( D(f \circ g) \) is the dot product of the \( i^{th} \) row of \( Df \) with the \( j^{th} \) column of \( Dg \). By definition of matrix multiplication, we get

\[
D(f \circ g)(x) = Df(g(x)) \cdot Dg(x). \quad \square
\]

Example 3.5.9. General Chain Rule

Let \( f : \mathbb{R}^2 \to \mathbb{R}^3 \) be given by \( f(x, y) = (x^2, 5x^2 + y^3, y) \), and \( g : \mathbb{R}^2 \to \mathbb{R}^2 \) by \( g(s, t) = (st, s^2 + 2t) \). We can compute the derivative matrices of both \( f \) and \( g \), to find

\[
Df = \begin{pmatrix} 2x & 0 \\ 10x & 3y^2 \\ 0 & 1 \end{pmatrix} \quad Dg = \begin{pmatrix} t & s \\ 2s & 2 \end{pmatrix}
\]

Evaluating \( Df \) at \( g(s, t) \), and multiplying matrices we have:

\[
Df(g(s, t)) \cdot Dg(s, t) = \begin{pmatrix} 2(st) & 0 \\ 10(st) & 3(s^2 + 2t)^2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} t & s \\ 2s & 2 \end{pmatrix} = \begin{pmatrix} 2st^2 & 2s^2t \\ 10st^2 + 6(s^2 + 2t)^2s & 10s^2t + 6(s^2 + 2t)^2 \end{pmatrix}.
\]

Alternatively, we could find explicit formulae for the composition \( f \circ g \). We have \( f \circ g : \mathbb{R}^2 \to \mathbb{R}^3 \) is given analytically by \( f \circ g(s, t) = f(st, s + 2t) = ((st)^2, 5(st)^2 + (s^2 + 2t)^3, s^2 + 2t) \). The \( ij^{th} \) entry is the partial derivative of the \( i^{th} \) coordinate function with respect to the \( j^{th} \) variable. For example, the entry \( (Df \cdot Dg)_{22} \) is the partial derivative \( \frac{\partial (f \circ g)_2}{\partial x_2} \), which is

\[
\frac{\partial (f \circ g)_2}{\partial t}(s, t) = 10(st) \cdot s + 6(s^2 + 2t)^2.
\]

Given enough time and energy, we can verify that the derivative matrix \( D(f \circ g)(s, t) \) is the same as \( Df(g(s, t)) \cdot Dg(s, t) \) in this case. The upshot is:
3.5. THE CHAIN RULE

The derivative of a composition is the product of the derivatives!

Example 3.5.10. Geometric Interpretation of the Derivative Matrix

We defined the derivative matrix as a convenient method for encoding the derivative of a multivariate function. The chain rule underscores this convenience! We now use it to give a geometric interpretation of the derivative matrix. Namely:

The derivative matrix maps tangent vectors of curves to tangent vectors of the image.

We illustrate with a specific example. Let \( x(t) = (\cos t, \sin t) \) be the unit circle, and let \( f(x, y) = (x - 3y, x + 3y) \) be the map of the plane to itself studied in the first section of this chapter. Then the curve \( f(x(t)) \) is an ellipse in the plane, given by \( f(x(t)) = (\cos t - 3\sin t, \cos t + 3\sin t) \) as in Figure 3.1.1. We know that \( \frac{d}{dt} f(x(t)) \) is tangent to the ellipse, and by the chain rule:

\[
\frac{d}{dt} f(x(t)) = Df(x(t)) \cdot x'(t) = \begin{pmatrix} 1 & -3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} = \begin{pmatrix} -\sin t - 3\cos t \\ -\sin t + 3\cos t \end{pmatrix}
\]

For example, if \( t = \pi/4 \), then \( x'(t) = \left\langle -\sqrt{2}/2, \sqrt{2}/2 \right\rangle \) is tangent to the unit circle at \( (\sqrt{2}/2, \sqrt{2}/2) \). To find the tangent to its image under \( f \), multiply it by \( Df \), giving

\[
\frac{d}{dt} f(x(\pi/4)) = Df(x(\pi/4)) \cdot x'(\pi/4) = \begin{pmatrix} 1 & -3 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} -2\sqrt{2} \\ \sqrt{2} \end{pmatrix}
\]

Thus the vector \( \left\langle -2\sqrt{2}, \sqrt{2} \right\rangle \) is tangent to the ellipse at the point \( f(x(\pi/4)) = (-\sqrt{2}, 2\sqrt{2}) \). This is illustrated in Figure 3.5.3

Exercises

1. Find the gradients of the following functions of several variables.
   (a) \( f(x, y) = x^3y - 2\sin(xy) \)
   (b) \( g(x, y) = 2x - \sin^{-1}(xy) \)
   (c) \( f(x, y, z) = \frac{x^2 + 2yz}{y \cos x} \)
   (d) \( h(x, y, z) = e^{y^2z} \sec xy \)
2. Find the derivative matrices of the following vector-valued functions of several variables.
CHAPTER 3. DIFFERENTIATION

\[
\begin{pmatrix}
1 & -3 \\
1 & 3
\end{pmatrix}
\rightarrow
\]

Figure 3.5.3: The geometry of the derivative matrix

(a) \( \mathbf{x}(t) = (t \cos t, t \sin t) \) (is there one row or one column?)
(b) \( f(x, y) = (2x + 4, x - 2y + 7, y - 3) \)
(c) \( f(x, y, z) = (xz^2, x + zy^2) \)

3. Find the derivative matrices for the change-of-coordinate functions, then find their determinants!

(a) \( f(r, \theta) = (r \cos \theta, r \sin \theta) \)
(b) \( f(r, \theta, z) = (r \cos \theta, r \sin \theta, z) \)
(c) \( f(\rho, \theta, \phi) = (\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi) \)

4. Verify the chain rule as in Example 3.5.6 for the function \( f(x, y) = x^2 + y^2 \) and the curve \( C(t) = (3 - 2t, 5t), -\infty < t < \infty \).

5. Verify the chain rule as in Example 3.5.6 for the function \( f(x, y) = xe^y \) and the curve \( C(t) = (2 \sin t, 3 \cos t), 0 \leq t \leq 2\pi \).

6. Verify the chain rule as in Example 3.5.6 for the function \( f(x, y, z) = xy^2 - z^3y \) and the curve \( C(t) = (3t, t^2, 2t) \).

7. Let \( f(x, y) = x^2 + y^2 \) denote the temperature at the point \( (x, y) \), and let \( C(t) = (t, t^2) \) be the path of an ant crawling on the plane. Find how fast the temperature is changing at time \( t = 2 \).

8. An ant crawls on the unit circle \( C(t) = (\cos t, \sin t) \), and \( f(x, y) = \sqrt{3x - y} \) is the temperature at points in the plane. What is the highest temperature the ant encounters? What can you say about the vectors \( \nabla f \) and \( C'(t) \) at this point?
9. Find all points on the helix \( C(t) = (\cos t, \sin t, t) \) where the gradient of \( f(x, y, z) = xy + z \) is normal to the tangent vector \( C'(t) \). Does \( f \) have a global extreme value on \( C(t) \)?

10. Let \( f(x, y, z) \) be a function on \( \mathbb{R}^3 \), and \( C(t) \) a parametric curve. Describe why \( \nabla f \) is orthogonal to \( C(t) \) at extreme values of \( f \) on \( C \).

11. Sketch the ellipse \( C(t) = (2 \cos t, \sin t) \), for \( 0 \leq t \leq 2\pi \). Include the tangent vector \( C'(\pi/4) \) in your sketch. Now consider the map \( f(x, y) = (x, 2y) \). Verify the chain rule for \( (f \circ C)(t) \) by calculating:

   (a) The matrix \( Df \), and multiplying it by the tangent vector \( C'(\pi/4) \).
   
   (b) Parametric equations for \( (f \circ C)(t) \), and finding the tangent vector \( (f \circ C)(\pi/4) \) directly.

   (c) Sketch the curve in part (b), together with its tangent vector.
3.6 Applications of the Gradient

In this section we introduce two applications of Theorem 3.5.1. The first is the notion of directional derivatives, or the rate of change of a function in a given direction. Partial derivatives are rates of change in the positive \(x\), \(y\), and \(z\) directions, and directional derivatives allow us to calculate the rate of change in an arbitrary direction. Secondly, we show how gradients are normal to level sets. This allows us to calculate equations for tangent planes, and will be the basis for the method of Lagrange multipliers.

**Directional Derivatives:** Before defining the directional derivative, we mention some terminology. The direction of the vector \(v\) is a unit vector pointing in the same direction as \(v\). Thus the direction of \(v\) is just the normalized vector \(\frac{v}{\|v\|}\).

**Definition 3.6.1.** Let \(u\) be a unit vector, \(P\) a point, and \(f\) a function of several variables. The directional derivative \(D_u f\) of \(f\) at the point \(P\) in the direction of \(u\) is

\[
D_u f(P) = \nabla f(P) \cdot u.
\]

We remark that the directional derivative of \(f(x, y, z)\) in the \(i\) direction is

\[
D_i f(x, y, z) = \nabla f(x, y, z) \cdot i = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z) \rangle \cdot (1, 0, 0)
\]

\[
= f_x(x, y, z).
\]

Similarly, we have \(D_j f(x, y, z) = f_y(x, y, z)\) and \(D_k f(x, y, z) = f_z(x, y, z)\). In other words, the directional derivatives of \(f\) in the coordinate directions are just the partial derivatives of \(f\). Thus directional derivatives generalize partial differentiation, allowing us to calculate rates of change of \(f\) in any direction, not just the coordinate directions.

**Example 3.6.1. Calculating Directional Derivatives**

Calculate the directional derivative of \(f(x, y) = x^3y - 2x^2y^3\) at the point \((1, 2)\) in the direction of the vector \((3, 4)\). First we get the ingredients necessary to calculate the directional derivative: the gradient and direction. Normalizing the vector, we see the direction is \(u = \langle \frac{3}{5}, \frac{4}{5} \rangle\). The gradient of \(f\) is \(\nabla f(x, y) = \langle 3x^2y - 4xy^3, x^3 - 6x^2y^2 \rangle\), and evaluating at \((1, 2)\) we have \(\nabla f(1, 2) = \langle -26, -23 \rangle\). By the definition of the directional derivative we have

\[
D_u f(1, 2) = \nabla f(1, 2) \cdot u = \langle -26, -23 \rangle \cdot \langle \frac{3}{5}, \frac{4}{5} \rangle = -\frac{170}{5} = -34.
\]

Find the directional derivative of \(f(x, y, z) = xy^3z^2\) at \((3, -1, 2)\) in the direction of \(u = \frac{1}{\sqrt{3}} (1, 1, -1)\). Since the gradient is \(\nabla f(x, y, z) = \langle y^3z^2, 3xy^2z^2, 2xy^3z \rangle\), we have
3.6. APPLICATIONS OF THE GRADIENT

\[ D_u f(3, -1, 2) = (-4, 36, -12) \cdot \left\langle \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right\rangle = \frac{44}{\sqrt{3}}. \]

We now give a geometric interpretation of the directional derivative that is a consequence of the chain rule and justifies the definition. Let \( u \) be any unit vector and \( P \) any point. Then the line through \( P \) in the direction of \( u \) is parameterized by \( x(t) = P + tu \), and \( x'(t) = u \). If we evaluate \( f \) on the curve \( x(t) \), we get a function of one variable, and can find its rate of change at \( P \) by taking the derivative and setting \( t = 0 \). The chain rule states that

\[ \frac{d}{dt} f(x(0)) = \nabla f(x(0)) \cdot x'(t) = \nabla f(P) \cdot u. \]

Thus the directional derivative is the honest-to-goodness calc I derivative of \( f \) restricted to a curve with tangent vector \( u \) at \( P \). In this way the mysterious dot product in the definition of \( D_u f(P) \) is seen to be a rate of change we know and love.

The definition of the directional derivative also provides important information about the gradient of a function. We summarize this in the following:

**Theorem 3.6.1.** Let \( f \) be a differentiable function of several variables with gradient vector \( \nabla f \). Then

1. The vector \( \nabla f \) points in the direction of maximal increase of \( f \), and the maximal rate of increase is \( \|\nabla f\| \).
2. The vector \( -\nabla f \) points in the direction of maximal decrease of \( f \). The maximal rate of decrease is \( \|\nabla f\| \).

**Proof.** For any direction (i.e. unit vector) \( u \) and point \( P \), we have the directional derivative of \( f \) is

\[ D_u f(P) = \nabla f(P) \cdot u = \|\nabla f(P)\| \|u\| \cos \theta = \|\nabla f(P)\| \cos \theta. \]

This function is maximized when \( \theta = 0 \), or when \( \nabla f(P) \) points in the same direction as \( u \). Moreover, in that direction \( D_u f(P) = \|\nabla f(P)\| \). This proves the first statement of the theorem. The second follows similarly, letting \( \theta = \pi \).

\( \square \)

**Example 3.6.2. Rates of Maximal Increase**

In a certain coordinate system, the altitude on a mountain above the point \((x, y)\) is given by \( f(x, y) = xy - 2x + y - 2 \). Which direction should a person standing above \((0, 0)\) walk to climb the fastest?

First of all, please don’t say “up”. Secondly, the direction the person should walk is really a direction in the plane. Typically one might say “North-East”, or use compass directions. These really are directions in the plane, or the domain of \( f \). Thus it makes sense that the answer will be a direction (i.e. unit vector) in the plane.
By the previous theorem $\nabla f(x, y) = (y - 2, x + 1)$ evaluated at $(0, 0)$ points in the direction of maximal increase. Normalizing, we have that the direction of maximal increase is

$$\frac{\nabla f(0, 0)}{\|\nabla f(0, 0)\|} = \left\langle \frac{1}{\sqrt{5}}(-2, 1) \right\rangle,$$

and the maximal rate of increase is $\|\nabla f(0, 0)\| = \sqrt{5}$. ▲

Gradients and Level Sets: We come now to our second application of the chain rule, that the gradient vector is always orthogonal to a level set. Recall that the level surface of $f(x, y, z)$ at level $c$ is the solution set to $f(x, y, z) = c$. Thus the level surface of $f(x, y, z) = x^2 + y^2 - z^2$ at levels $c = 1, 0, -1$ are the one-sheeted hyperboloid, the cone, and the two-sheeted hyperboloid, respectively (see Figure 3.6.1).

![Level Surfaces](image)

Figure 3.6.1: Level Surfaces

**Theorem 3.6.2.** Let $f$ be a differentiable function of several variables. Then $\nabla f$ is orthogonal to the level set $f = c$.

**Proof.** We focus on the case where $f(x, y, z)$ is a function of three variables, although the proof works in general. We have to show that $\nabla f(x, y, z)$ is perpendicular to tangent vectors to the level surface $f(x, y, z) = c$. These tangent vectors are of the form $C'(t)$ for curves $C(t)$ on the surface.

Let $C(t)$ be a curve on the surface $f(x, y, z) = c$, so that $f$ evaluated on $C(t)$ is $f(C(t)) = c$, a constant function. Since $f(C(t))$ is constant, its derivative is 0. By the chain rule we have

$$\frac{d}{dt}f(C(t)) = \nabla f(C(t)) \cdot C'(t) = 0.$$

Thus the vectors $\nabla f(C(t))$ and $C'(t)$ are orthogonal, and $\nabla f$ is perpendicular to tangent vectors to the surface $f(x, y, z) = c$. □

**Theorem 3.6.2** allows us to find equations for tangent planes to surfaces given as solution sets to equations, which we will illustrate in Example 3.6.4 below. Coupled with the techniques of Section 3.4 we now know how to find tangent
planes to surfaces given parametrically, by functions of two variables, and as solution sets to equations in three variables.

**Example 3.6.3. Tangent Line to Level Curve**

Find an equation for the tangent line to the ellipse $5x^2 - 6xy + 5y^2 = 4$ at the point $(-0.5, -1.1)$.

![Figure 3.6.2: Gradient is orthogonal to ellipse](image)

The trick to solving this problem is thinking of the ellipse as a level curve of the function $f(x, y) = 5x^2 - 6xy + 5y^2$ at level $c = 4$. By Theorem 3.6.2 we know $\nabla f(-0.5, -1.1)$ is perpendicular to the level curve $5x^2 - 6xy + 5y^2 = 4$, and therefore perpendicular to the tangent line there. In our case

$\nabla f(x, y) = \langle 10x - 6y, -6x + 10y \rangle,$

so the normal vector is $\nabla f(-0.5, -1.1) = \langle 1.6, -8 \rangle$.

We can use the normal vector $\langle 1.6, -8 \rangle$ to find to find a Cartesian equation for the tangent line in the same way we found Cartesian equations for planes in $\mathbb{R}^3$. Let $(x, y)$ be any point on the line, then the vector $v = \langle x, y \rangle - \langle -0.5, -1.1 \rangle$ is perpendicular to $\nabla f(-0.5, -1.1)$ (see Figure 3.6.2). Thus the dot product

$\nabla f(-0.5, -1.1) \cdot v = \langle 1.6, -8 \rangle \cdot \langle x, y \rangle - \langle -0.5, -1.1 \rangle$

$= 1.6x - 8y - 8,$

must be zero! Therefore an equation for the tangent line is $1.6x - 8y = 8$. If you don’t like decimals, multiplying by $5/8$ gives $x - 5y = 5$. ▲

**Example 3.6.4. Tangent Planes to Level Surfaces**

Find a Cartesian equation for the tangent plane to $x^2 + y^2 - z^2 = 1$ at the point $(1, 1, 1)$.

We need a point, which is given, and a normal direction. Thinking of the surface as a level surface of the function $f(x, y, z) = x^2 + y^2 - z^2$, the previous theorem tells us that $\nabla f(1, 1, 1)$ is normal to the surface. Thus $\nabla f(1, 1, 1) = \langle 2x, 2y, -2z \rangle = \langle 2, 2, -2 \rangle$ at $(1, 1, 1)$. The tangent plane is therefore $2(x - 1) + 2(y - 1) - 2(z - 1) = 0$, or $x + y - z = 1$. ▲
\(\langle 2, 2, -2 \rangle\), is normal to the tangent plane. To find an equation, we use the scalar multiple \(\langle 1, 1, -1 \rangle\), and get
\[
x + y - z = \langle 1, 1, -1 \rangle \cdot \langle 1, 1, 1 \rangle = 1. \quad \blacksquare
\]

**Tangent Planes to Level Surfaces**

A Cartesian equation for the tangent plane to the level surface \(f(x, y, z) = c\) at the point \(P\) on it can be found as follows:

**Step 1:** Find \(\nabla f(P)\).

**Step 2:** \(P\) is on the plane and \(\nabla f(P)\) is normal to it, so an equation is
\[
\nabla f(P) \cdot \langle x, y, z \rangle = \nabla f(P) \cdot P.
\]

**Example 3.6.5. Tangent to an ellipsoid**

Find an equation for the tangent plane to the ellipsoid
\[
9x^2 + y^2 + 4z^2 = 36
\]
at the point \((1, -\sqrt{11}, -2)\).

![Figure 3.6.3: Tangent to an ellipsoid](image)

Thinking of the ellipsoid as a level surface of \(f(x, y, z) = 9x^2 + y^2 + 4z^2\), we find the gradient to be \(\nabla f(x, y, z) = \langle 18x, 2y, 8z \rangle\). Thus a normal to the plane is \(\nabla f(1, -\sqrt{11}, -2) = \langle 18, -2\sqrt{11}, -16 \rangle\), and an equation for it is
\[
18x - 2\sqrt{11}y - 16z = \langle 18, -2\sqrt{11}, -16 \rangle \cdot \langle 1, -\sqrt{11}, -2 \rangle = 72. \quad \blacksquare
\]

**Example 3.6.6. Curves of Intersection**

Find an equation for the tangent line to the curve of intersection of the surface \(z = xy\) and the unit sphere at the point \((1, 0, 0)\).

For an equation of a line we need a point, which is given, and a vector in the line. The key is to notice that the tangent line to the curve of intersection is in
the tangent planes to each surface, and hence orthogonal to both normal vectors. Think of the sphere as a level surface of the function $f(x, y, z) = x^2 + y^2 + z^2$, and the other surface as a level surface of $g(x, y, z) = xy - z$. A vector in the line, then, is the cross product $\nabla f(1, 0, 0) \times \nabla g(1, 0, 0)$. Calculating we have

$$\nabla f(1, 0, 0) \times \nabla g(1, 0, 0) = \langle 2, 0, 0 \rangle \times \langle 0, 1, -1 \rangle = \langle 0, 2, 2 \rangle.$$  

Parametric equations for the line are $C(t) = (1, 0, 0) + t \langle 0, 2, 2 \rangle = (1, 2t, 2t)$. ▲

**Exercises**

1. Find the directional derivatives of the following functions at the given point in the directions of the given vectors
   
   (a) $f(x, y) = x^3 - 4xy^2 + 3y$, $P(1, 2)$, and $v = \langle 2, 3 \rangle$
   
   (b) $f(x, y) = ye^x$, $P(\ln 4, 3)$, and $v = \langle 2, -1 \rangle$
   
   (c) $f(x, y, z) = x^2y + yz^2 - 2x$, $P(-1, 2, 1)$, and $v = \langle 2, -1, 1 \rangle$

2. A hiker is standing above $(0, 0)$ on the mountain whose altitude is given by $f(x, y) = 8 - x^2 - y^2 - 4x + 2y$ (in thousands of feet). Which direction should she walk to get down the fastest? How fast will she be descending?

3. The temperature near a heat source is $T(x, y, z) = \frac{5}{1 + 9x^2 + 4y^2 + z^2}$. What direction should a particle at $(2,1,1)$ travel to heat up the fastest? To cool down as quickly as possible?

4. Find the points on $x^2 + 4y^2 + 9z^2 = 1$ where the tangent plane is perpendicular to the line $C(t) = (3 - t, 4 + t, 4 - 6t)$.

5. Find a Cartesian equation for the tangent plane to $2xy + yz + 5xz = 8$ at the point $(1, 1, 1)$.  

---

Figure 3.6.4: Intersecting Surfaces
6. Find a Cartesian equation for the tangent plane to the two-sheeted hyperboloid \( x^2 + y^2 - z^2 = -1 \) at the point \((0, 1, -\sqrt{2})\).

7. Find an equation for the tangent plane to the surface \( x^2 + 9y^2 - z^2 = 1 \) at the point \((1, 2, 6)\).

8. Find where the tangent plane to the surface \( x^2 + 4(y + z)^2 + 9(y - z)^2 = 1 \) is horizontal.

9. Find an equation for the tangent line to the curve of intersection of the surfaces \( x^2 = 2yz \) and the unit sphere at the point \((\sqrt{2}/2, 1/2, 1/2)\).
3.7 The Hessian Test

In this section we describe a method for optimizing functions of two variables. It is analogous to the second derivative test for extreme values in single-variable calculus. Before considering the two-variable setting, we remind the reader about the single-variable rule.

![Figure 3.7.1: The second derivative test](image)

Recall that \(c\) is a critical point of the function \(f(x)\) if \(f'(c)\) is zero or undefined. This meant the tangent line was horizontal \((f'(c) = 0)\), or weird things were happening \((f'(c)\) undefined). If \(c\) is a critical point of \(f\) and \(f''(c)\) exists, we could tell the concavity of \(f\) near \(x = c\). If \(f''(c) > 0\), then \(f\) was concave up near \(c\) and has a local minimum value at \(x = c\) (see Figure 3.7.1). Similarly, if \(f''(c) < 0\) then \(f\) is concave down and has a relative maximum at \(x = c\). If \(f''(c) = 0\), or is undefined, the second derivative test is inconclusive. We now consider functions of two variables. As in the single variable case, we start by defining critical points.

**Definition 3.7.1.** The point \((x_0, y_0)\) is a critical point of the function \(f(x, y)\) if both partial derivatives of \(f\) are zero at \((x_0, y_0)\).

Thus to find critical points you must solve the system of equations

\[
\begin{align*}
   f_x(x, y) &= 0 \\
   f_y(x, y) &= 0,
\end{align*}
\]

and we illustrate with an example.

**Example 3.7.1. Find Critical Points**

Find the critical points of \(f(x, y) = x^3 + 6xy^2 - 6x\).

This amounts to solving the system of equations

\[
\begin{align*}
   f_x(x, y) &= 3x^2 + 6y^2 - 6 = 0 \\
   f_y(x, y) &= 12xy = 0
\end{align*}
\]
CHAPTER 3. DIFFERENTIATION

The second equation implies that either \( x = 0 \) or \( y = 0 \), and we treat each case separately. If \( x = 0 \), the first equation becomes \( 6y^2 - 6 = 0 \), so \( y = \pm 1 \). Thus \((0, 1)\) and \((0, -1)\) are critical points. If \( y = 0 \), then the first equation becomes \( 3x^2 - 6 = 0 \), so \( x = \pm \sqrt{2} \). Therefore the critical points of \( f \) are \((0, \pm 1)\) and \((\pm \sqrt{2}, 0)\). ▲

Since we are considering functions of two variables, there are three second order partial derivatives: \( \frac{\partial^2 f}{\partial x^2} \), \( \frac{\partial^2 f}{\partial y^2} \), and \( \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \) (in all the functions we consider, the mixed partials will be equal). We define the Hessian matrix for \( f(x, y) \), which will keep track of all possible second partials at once.

**Definition 3.7.2.** The Hessian matrix of \( f(x, y) \) is

\[
H(f) = \begin{pmatrix}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\
\frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2}
\end{pmatrix}
\]

The determinant of the Hessian matrix is called the discriminant \( D \):

\[
D = \left| \begin{array}{cc}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\
\frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2}
\end{array} \right| = \frac{\partial^2 f}{\partial x^2} \cdot \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial y \partial x} \right)^2
\]

**Example 3.7.2. Calculating Hessian Matrices**

Find the Hessian matrix \( f(x, y) = x^3 + 6xy^2 - 6x \), and it’s discriminant.

The first partial derivatives were taken in the previous example. Differentiating those we find:

\[
H(f)(x, y) = \begin{pmatrix}
6x & 12y \\
12y & 12x
\end{pmatrix}, \text{ and } D = \left| \begin{array}{cc}
6x & 12y \\
12y & 12x
\end{array} \right| = 72x^2 - 144y^2. \ ▲
\]

We now have introduced everything we need to state the Hessian test for finding extreme values of \( f(x, y) \).

**The Hessian Test**

Let \((x_0, y_0)\) be a critical point of \( f(x, y) \), and \( D \) be the discriminant of \( f \) evaluated at \((x_0, y_0)\). There are four cases to consider:

1. If \( D > 0 \) and \( f_{xx}(x_0, y_0) > 0 \), then \((x_0, y_0)\) is a local minimum for \( f \).
2. If \( D > 0 \) and \( f_{xx}(x_0, y_0) < 0 \), then \((x_0, y_0)\) is a local maximum for \( f \).
3. If \( D < 0 \), then \( f \) has a saddle point at \((x_0, y_0)\).
4. Otherwise, the test is inconclusive.

**Example 3.7.3. Classifying Critical Points**
3.7. THE HESSIAN TEST

Classify the critical points of \( f(x, y) = x^3 + 6xy^2 - 6x \). We’ve determined the critical points are \((0, \pm 1)\) and \((\pm \sqrt{2}, 0)\). We substitute the critical point \((\sqrt{2}, 0)\) into the discriminant we find \( D(\sqrt{2}, 0) = 72(\sqrt{2})^2 - 144(0)^2 = 144 > 0 \). Since the discriminant is positive, we check \( f_{xx}(\sqrt{2}, 0) = 6\sqrt{2} > 0 \) and conclude that \( f \) has a local minimum at \((\sqrt{2}, 0)\). Similar analyses gives the table below. See if you can find the respective critical points in Figure 3.7.2.

<table>
<thead>
<tr>
<th>Critical Point</th>
<th>Discriminant</th>
<th>( f_{xx} )</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>((\sqrt{2}, 0))</td>
<td>144</td>
<td>(6\sqrt{2})</td>
<td>Minimum</td>
</tr>
<tr>
<td>((-\sqrt{2}, 0))</td>
<td>144</td>
<td>(-6\sqrt{2})</td>
<td>Maximum</td>
</tr>
<tr>
<td>((0, 1))</td>
<td>-144</td>
<td>NA</td>
<td>Saddle</td>
</tr>
<tr>
<td>((0, -1))</td>
<td>-144</td>
<td>NA</td>
<td>Saddle</td>
</tr>
</tbody>
</table>

\[\]^ Example 3.7.4. An inconclusive test

Classify the critical points of \( f(x, y) = x^2 - 2xy + y^2 \). To do so, we note that the solution set of the system

\[
\begin{align*}
    f_x(x, y) &= 2x - 2y = 0 \\
    f_y(x, y) &= -2x + 2y = 0
\end{align*}
\]

is the line \( y = x \). Moreover, the Hessian and discriminant are

\[
H(f)(x, y) = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix}, \quad \text{and} \quad D = \begin{vmatrix} 2 & -2 \\ -2 & 2 \end{vmatrix} = 0.
\]

Since the discriminant is always zero, the Hessian test is inconclusive in this situation. In this case, it’s because the critical points of \( f \) are not isolated:
there is an entire line of critical points! The reason for this is easier to see if one factors \( f(x, y) = x^2 - 2xy + y^2 = (x - y)^2 \). Thus the graph is just a parabola translated back and forth on the line \( y = x \) (see Figure 3.7.3). The line of critical points is the line containing the vertices of the parabola. ▲

Figure 3.7.3: The Hessian Test fails on this parabolic cylinder

**Example 3.7.5. Minimizing Distance**

Find the point on the parametric surface \( S(s, t) = (t + s, t - s, t^2) \) closest to \((12, 4, 1)\).

To do this, we will minimize the square of the distance from \((12, 4, 1)\) to any point on the surface. We square the distance to make our calculations easier. The square of the distance is given by

\[
f(s, t) = (12 - t - s)^2 + (4 - t + s)^2 + (1 - t^2)^2.
\]

Thus our problem of finding the closest point on the surface to \((12, 4, 1)\) is reduced to minimizing \( f(s, t) \), a function of two variables. Using the chain rule to take partial derivatives, we have the system of equations

\[
\begin{align*}
f_s(s, t) &= -2(12 - t - s) + 2(4 - t + s) = -16 + 4s = 0 \\
f_t(s, t) &= -2(12 - t - s) - 2(4 - t + s) - 4t(1 - t^2) = -32 + 4t^3 = 0.
\end{align*}
\]

From these we deduce that the only critical point is \((s_0, t_0) = (4, 2)\). We calculate the Hessian to verify that this is a minimum of the function. We find

\[
H(f)(s, t) = \begin{pmatrix}
4 & 0 \\
0 & 12t^2
\end{pmatrix}, \quad \text{and } D = \begin{vmatrix}
4 & 0 \\
0 & 48
\end{vmatrix} = 192 > 0.
\]

Since \( f_s(s, t) = 4 > 0 \), we know that \((4, 2)\) is a minimum of \( f(s, t) \), and the closest point on the surface to \((12, 4, 1)\) is \( x(4, 2) = (6, -2, 4) \). ▲
3.7. THE HESSIAN TEST

To get an intuitive reason why the Hessian test works, remember that taking partial derivatives of $f$ is equivalent to intersecting the surface with planes, and considering the curves you get. If $D = f_{xx}f_{yy} - f_{xy}^2 > 0$, then $f_{xx}$ and $f_{yy}$ must have the same sign. This means the curves of intersection are either both concave up or both concave down, yielding a minimum or a maximum respectively (see Figure 3.7.4(a)).

(a) $D > 0$, $f_{xx} < 0$, and $f_{yy} < 0$  (b) $f_{xx}$ and $f_{yy}$ have opposite sign

Figure 3.7.4: The Geometry of the Hessian Test

Another situation that is easy to see intuitively is the case where $f_{xx}(x_0, y_0)$ and $f_{yy}(x_0, y_0)$ have opposite sign. In this setting it is clear that the discriminant will be negative, so $(x_0, y_0)$ is a saddle point. This makes sense because if the second partials have opposite sign, then the surface is concave up in one direction and concave down in the other, yielding a saddle point.

Example 3.7.6. Least Squares Regression

One interesting application of the Hessian test is that it finds the formula for the least squares regression line of a collection of paired data. We illustrate this with an over-simplified example.

Find the least squares regression line for the data set (3, 7), (5, 4), and (7, 4).

Before getting started, let’s review least squares regression. The goal is to find a line that “best” fits the data. The definition of “best” might vary, but we describe the sense in which the least squares regression is the best fit. When using a line $y = mx + b$ to approximate data, one could measure how accurate it is by wanting the vertical distance between the observed data $(x_i, y_i)$ and the point on the line $(x_i, mx_i + b)$ to be minimized. The vertical distance is called the residual, and in our notation is $|y_i - (mx_i + b)|$ (see Figure 3.7.5(a)). The least squares regression line is the choice of $m$ and $b$ that minimizes the sum of the squared residuals.
CHAPTER 3. DIFFERENTIATION

Let's make this very explicit in our example. Given any choice of slope and intercept, the sum of the squared residuals for our points is

\[ f(m, b) = (7 - (3m + b))^2 + (4 - (5m + b))^2 + (4 - (7m + b))^2. \]

Our goal is to find the slope \( m \) and \( y \)-intercept \( b \) that minimizes \( f(m, b) \). Thus finding the least squares regression line amounts to optimizing a function of two variables! We use the Hessian test to do this. Taking partial derivatives gives

\[
\begin{align*}
f_m(m, b) &= -6(7 - 3m - b) - 10(4 - 5m - b) - 14(4 - 7m - b) = -138 + 166m + 30b \\
f_b(m, b) &= -2(7 - 3m - b) - 2(4 - 5m - b) - 2(4 - 7m - b) = -30 + 30m + 6b \\
\end{align*}
\]

After some algebra, one finds that \((m_0, b_0) = \left( -\frac{3}{4}, \frac{35}{4} \right) \) is a critical point. In this case, the discriminant is

\[
D = \begin{vmatrix} 166 & 30 \\ 30 & 6 \end{vmatrix} = 996 - 900 > 0,
\]

and \( f_{mm} > 0 \), so the critical point is a minimum. Thus the least squares regression line for \{ (3, 7), (5, 4), (7, 4) \} is \( y = -\frac{3}{4}x + \frac{35}{4} \). ▲

\[\text{\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{least_squares_regression.png}
\caption{Least Squares Regression}
\end{figure}}\]

Exercises

1. Classify the critical points of the following functions
   \[ \text{(a) } f(x, y) = x^2 - y^2 + 6x + 2y + 8 \]
   \[ \text{(b) } f(x, y) = x^3 + 4xy^2 - 12x + 12y \]
   \[ \text{(c) } f(x, y) = (x^3 - x)e^{-y^2} \]
3.7. THE HESSIAN TEST

(d) \( f(x, y) = \frac{x}{1+x^2+y^2} \)

2. Apply the Hessian test to the function \( f(x, y) = x^2 + 4xy + 4y^2 \). What do you find?

3. What can you say about the constants \( A \) through \( F \) in \( f(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F \) when the Hessian test is inconclusive?

4. Find the point on the parametric surface \( x(s, t) = (s, t^2, t - s) \) closest to \( \left( \frac{1}{2}, \frac{1}{4}, -1 \right) \).

5. Find the least squares regression line for \((5, 2), (10, 12), \) and \((12, 10)\).

6. Derive the formula for the slope and intercept of the least squares regression line for the general data set \( \{(x_1, y_1), \ldots, (x_n, y_n)\} \). (Hint: use summation notation and \( n\bar{x} = \sum_{i=1}^{n} x_i \), where \( \bar{x} \) denotes the mean of the \( x_i \).)
3.8 Constrained Optimization

There are many situations where you want to optimize a function subject to some constraint. Formally, these types of questions look like: maximize the function \( f(x, y, z) \) subject to the constraint \( g(x, y, z) = c \). We actually met this type of question in Section ?? in the two-variable case. We maximized the function \( f(x, y) = x^2 - 2x + y^2 + 4y + 1 \) subject to the constraint \( x^2 + y^2 = 1 \). In that instance we argued that \( f(x, y) \) would achieve extreme values on the unit circle where the level curves were tangent to it. The reasoning went as follows: if a level curve crosses the unit circle, then nearby level curves will too. The nearby levels will be different: greater in one direction, smaller in the other (see Figure 3.8.1). This means \( f \) can’t attain an extreme value where level curves cross the constraint curve. Thus any extreme values occur where level curves are tangent to the constraint curve.

\[ \nabla f \text{ is perpendicular to the level curves of } f. \]
\[ \nabla g \text{ will be perpendicular to the constraint curve.} \]

Combining this observation with what we know about gradients and level curves will lead to the method of Lagrange multipliers. Recall that \( \nabla f \) is perpendicular to the level curves of \( f \). Moreover, we can think of the constraint curve \( x^2 + y^2 = 1 \) as the level curve of \( g(x, y) = x^2 + y^2 \) at level 1. Thought of in this way, we see that \( \nabla g \) will be perpendicular to the constraint curve. Of course, when we say a vector is orthogonal to a curve, we really mean it’s orthogonal to the tangent line to the curve.

Now consider what happens when the level curve \( f(x, y) = c \) is tangent to the constraint curve \( g(x, y) = 1 \). At this point, the tangent lines to both curves are the same. Therefore \( \nabla f \) and \( \nabla g \) are perpendicular to the same line, and must be parallel. This means they are scalar multiples, and there must be a scalar \( \lambda \) such that \( \nabla f = \lambda \nabla g \).

These observations lead to the method of Lagrange multipliers, which is used
to optimize functions of several variables subject to a constraint. We outline
the method below and follow with examples.

**The Method of Lagrange Multipliers**

To optimize the function \( f(x, y, z) \) subject to the constraint \( g(x, y, z) = c \),
solve the system of equations

\[
\nabla f(x, y, z) = \lambda \nabla g(x, y, z)
\]

\[g(x, y, z) = c.\]  \hspace{1cm} (3.8.1)

Evaluate \( f \) at the solutions. The largest value is the maximum, the least
is the minimum.

Notice that the first equation in the Lagrange system is a vector equation.
Since two vectors are equal if and only if all their components are, the one
vector equation \( \nabla f(x, y, z) = \lambda \nabla g(x, y, z) \) represents three scalar equations, one
for each coordinate. This method works since, as described above, at extreme
values of \( f \) on the constraint \( g(x, y, z) = c \) the gradients are parallel.

Before illustrating this procedure in several examples, we note that solving
the system of equations (3.8.1) can involve a fair amount of trickery and ingenuity.
For example, we might factor a given equation, divide consecutive equations, or
simply use substitution. We mention that \( \lambda \) could be zero when \( \nabla f \) is the zero
vector, which happens at critical points of \( f \).

**Example 3.8.1. A Two-Variable Example**

Find the extreme values that the function \( f(x, y) = x^2 + y^2 - 2x + 2y \) takes
on the unit circle.

This can be stated in Langrange format: Optimize \( f(x, y) = x^2 + y^2 - 2x + 2y \)
subject to the constraint \( x^2 + y^2 = 1 \). The Lagrange system of equations, in
vector format, is

\[
\langle 2x - 2, 2y + 2 \rangle = \lambda \langle 2x, 2y \rangle
\]

\[x^2 + y^2 = 1\]  \hspace{1cm} (3.8.1)

Separating the vector equation into a system of component equations gives:

\[
2x - 2 = \lambda 2x
\]

\[
2y + 2 = \lambda 2y
\]

\[x^2 + y^2 = 1\]

Rearranging the first equation gives \((1 - \lambda)x = 1\), which implies that \(1 - \lambda \neq 0\). Similarly, the second component equation becomes \((1-\lambda)y = -1\). Combining
these, we get \((1 - \lambda)x = -(1 - \lambda>y\) and dividing by \(1 - \lambda \) gives \(x = -y\) (we
can do this division since we know \(1 - \lambda \neq 0\). Substituting \(x = -y\) into the
constraint equation $x^2 + y^2 = 1$ and solving gives $x = \pm \frac{\sqrt{2}}{2}$. Thus the points where $f$ takes on extreme values are $(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$ and $(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. To find the extreme values, we evaluate $f$ at those points:

$$f \left( \frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2} \right) = \frac{1}{2} + \frac{1}{2} - \sqrt{2} - \sqrt{2} = 1 - 2\sqrt{2}$$

$$f \left( -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right) = \frac{1}{2} + \frac{1}{2} + \sqrt{2} + \sqrt{2} = 1 + 2\sqrt{2}$$

Thus the maximum value $f$ attains on the unit circle is $1 + 2\sqrt{2}$, and the minimum value is $1 - 2\sqrt{2}$. We can interpret this geometrically as follows. The graph of $f(x,y)$ is a paraboloid. Restricting the domain to the unit circle is equivalent to looking at the curve on the paraboloid above the unit circle. We found the highest and lowest point on that curve (see Figure 3.8.2). ▲

![Figure 3.8.2: Geometric interpretation of constrained optimization](image)

**Example 3.8.2.** Consider the possibility $\lambda = 0$

Find the extreme values that the function $f(x,y) = x^2 + y^2 - 2xy$ takes on the unit circle.

The Lagrange system of equations for this problem is

$$\langle 2x - 2y, 2y - 2x \rangle = \lambda \langle 2x, 2y \rangle$$

$$x^2 + y^2 = 1$$

Separating the vector equation into a system of component equations gives:

$$2x - 2y = \lambda 2x$$

$$2y - 2x = \lambda 2y$$

$$x^2 + y^2 = 1$$
The first two equations are satisfied when \( y = x \) and \( \lambda = 0 \). On the unit circle, this corresponds to the points \( \pm(\sqrt{2}/2, \sqrt{2}/2) \), and direct calculation shows \( f(\sqrt{2}/2, -\sqrt{2}/2) = 0 \).

On the other hand, if \( \lambda \neq 0 \), then solving the first equation for \( y \) gives and the second for \( x \) gives the system

\[
\begin{align*}
y &= (1 - \lambda)x \\
x &= (1 - \lambda)y \\
x^2 + y^2 &= 1
\end{align*}
\]

On the unit circle we can’t have both \( x \) and \( y \) be zero simultaneously, so the above system implies \( x \neq 0 \), \( y \neq 0 \) and \( \lambda \neq 1 \) (convince yourself that any of these cases implies both \( x \) and \( y \) are zero—a contradiction). Dividing the first two equations gives

\[
\frac{y}{x} = \frac{(1 - \lambda)x}{(1 - \lambda)y},
\]

which reduces to \( y^2 = x^2 \). Combining this with the constraint \( x^2 + y^2 = 1 \) yields the points \( \pm(\sqrt{2}/2, -\sqrt{2}/2) \). Direct calculation gives \( f(\sqrt{2}/2, -\sqrt{2}/2) = f(-\sqrt{2}/2, \sqrt{2}/2) = 2 \).

Thus \( f \) has a minimum value of 0 and maximum value of 2 on the unit circle.

\[\text{Figure 3.8.3: Extreme values when } \lambda = 0\]

Once again, lets interpret this geometrically. Factoring the function gives \( f(x, y) = (x - y)^2 \), which means that \( f \) is always at least 0 and above the unit disk is the taco shell pictured in Figure 3.8.3. The fold of the taco shell is along the line \( y = x \), and is strange because it corresponds to an entire line of minima for \( f \). As in Section 3.7 these are points where \( \nabla f = 0 \). The Lagrange vector equation \( \nabla f = \lambda \nabla g \) is then satisfied when \( \lambda = 0 \). Thus critical points of \( f \) correspond to solutions to the Lagrange system corresponding to \( \lambda = 0 \).

**Example 3.8.3. Maximizing area for fixed perimeter**
Find the triangle of perimeter two with maximal area.

To solve this problem, let's denote the lengths of the sides of the triangle by \(a, b,\) and \(c\). Since we know information about the lengths of sides and want to maximize area, we need to find a formula that gives the area of a triangle as a function of its side lengths. Toward this end, we recall Heron’s formula. Let \(s = \frac{a+b+c}{2}\) be the semiperimeter of the triangle, then Heron’s formula says the area \(A\) of the triangle is \(A = \sqrt{s(s-a)(s-b)(s-c)}\). We will maximize the square of the area, since it gives the same triangle and simplifies our calculations considerably. Since the perimeter is two, in our case \(s = 1\) and we have

\[
f(a, b, c) = A^2 = (1-a)(1-b)(1-c) = 1 - a - b - c + ab + ac + bc - abc
\]

where the last equality follows from the fact that \(a + b + c = 2\). So we want to maximize \(f\) subject to the constraint that the perimeter is two. The Lagrange system of equations, component-wise rather than in vector form, is

\[
\begin{align*}
b + c - bc &= \lambda \\
a + c - ac &= \lambda \\
a + b - ab &= \lambda \\
a + b + c &= 2
\end{align*}
\]

Equating the first two equations gives \(b + c - bc = a + c - ac\), which simplifies to \((a-b)c = (a-b)\). This implies either \(c = 1\) or \(a-b = 0\). We consider each case separately. If \(c = 1\), the system of equations becomes

\[
\begin{align*}
b + 1 - b &= \lambda \\
a + 1 - a &= \lambda \\
a + b - ab &= \lambda \\
a + b + 1 &= 2.
\end{align*}
\]

The first implies that \(\lambda = 1\) and the fourth that \(a + b = 1\). Substituting these into the third equation gives \(1 - ab = 1\), or \(-ab = 0\). This, of course, implies that at least one of \(a\) or \(b\) is zero, which can’t happen because they represent lengths of a side of a triangle. We conclude that \(c \neq 1\), which leaves that \(a = b\) as the only possible case.

Identical reasoning, using the second and third equations yields \(b = c\). Thus the triangle of perimeter two with maximum area is an equilateral triangle, with edge length \(\frac{2}{3}\). ▲

**Example 3.8.4. Lagrange and three variables**

Find the maximum and minimum values \(f(x, y, z) = 8x - 6y + 10z\) attain on the sphere \(x^2 + y^2 + z^2 = 2\).
3.8. CONSTRAINED OPTIMIZATION

Setting up the Lagrange system of equations gives:

\[(8, -6, 10) = \lambda \langle 2x, 2y, 2z \rangle \]
\[x^2 + y^2 + z^2 = 2\]

Solving the component equations for \(x\), \(y\), and \(z\) gives \(x = \frac{4}{\lambda}\), \(y = -\frac{3}{\lambda}\), and \(z = \frac{5}{\lambda}\). Substituting these into the constraint equations and solving yields

\[\left(\frac{4}{\lambda}\right)^2 + \left(-\frac{3}{\lambda}\right)^2 + \left(\frac{5}{\lambda}\right)^2 = 2\]
\[\frac{50}{\lambda^2} = 2\]
\[25 = \lambda^2\]

Therefore \(\lambda = \pm 5\). We’ve already solved for \(x\), \(y\) and \(z\) in terms of \(\lambda\). If \(\lambda = 5\) we get \(\left(\frac{4}{5}, -\frac{3}{5}, 1\right)\) and if \(\lambda = -5\) we get \(\left(-\frac{4}{5}, \frac{3}{5}, -1\right)\). Substituting into \(f\), the extreme values of \(f\) on the sphere are:

\[f\left(\frac{4}{5}, -\frac{3}{5}, 1\right) = \frac{32}{5} - \frac{18}{5} + 10 = 12.8\]
\[f\left(-\frac{4}{5}, \frac{3}{5}, -1\right) = -12.8. \blacktriangle\]

Example 3.8.5. One Last example

Maximize \(f(x, y, z) = xyz\) on the ellipsoid \(16x^2 + 9y^2 + 25z^2 = 96\). Computing gradients we get \(\nabla f(x, y, z) = \langle yz, xz, xy \rangle\) and \(\nabla g(x, y, z) = \langle 32x, 18y, 50z \rangle\). We set up the gradient and constraint equations:

\[yz = 32\lambda x\]
\[xz = 18\lambda y\]
\[xy = 50\lambda z\]
\[16x^2 + 9y^2 + 25z^2 = 96\]

This system is most easily solved by taking quotients of consecutive equations. Taking the quotient of the first two equations and simplifying gives

\[\frac{yz}{xz} = \frac{32\lambda x}{18\lambda y}\]
\[\frac{y}{x} = \frac{16x}{9y}\]
\[9y^2 = 16x^2\]
Similar calculations show that \( 25z^2 = 16x^2 \). Substituting these into the constraint equation gives

\[
16x^2 + 9y^2 + 25z^2 = 16x^2 + 16x^2 + 16x^2 = 48x^2 = 96.
\]

Therefore we find \( x = \pm \sqrt{2} \). Solving for \( y^2 \) instead will give \( 27y^2 = 96 \), or \( y = \pm \frac{4\sqrt{2}}{3} \). Analogously, one calculates \( z = \pm \frac{4\sqrt{2}}{5} \). The product is maximized when all are positive, so it is \( f(\sqrt{2}, \frac{4\sqrt{2}}{3}, \frac{4\sqrt{2}}{5}) = \frac{32\sqrt{2}}{15} \). ▲

**Exercises**

1. Use the method of Lagrange multipliers to optimize the following functions \( f \) subject to the given constraints.
   (a) \( f(x, y) = x^2 + y^2 + 4x - 4y, \ x^2 + y^2 = 1 \).
   (b) \( f(x, y) = xy, \ x^2 + 4y^2 = 1 \).
   (c) \( f(x, y, z) = x + y + z, \ x^2 + y^2 + z^2 = 1 \).
   (d) \( f(x, y, z) = xyz^2, \ x + y + z = 3 \), with all \( x, y, z \) positive.

2. Find the rectangle with maximum area and perimeter 8.

3. Find the rectangle of maximal area that can be inscribed in the ellipse \( \frac{x^2}{2} + y^2 = 1 \).

4. Recall the Cobb-Douglas production function is \( Q = K^{\alpha}L^{1-\alpha} \), where \( K \) is capital and \( L \) is labor. In economics, one can model the total cost \( C \) of production by \( C = aK + bL \), where \( a \) is the cost per unit capital (the interest rate) and \( b \) the cost per unit labor. Given a fixed cost \( C_0 \), it is reasonable to ask how to allocate resources to Maximize production. In other words, maximize \( Q = K^{\alpha}L^{1-\alpha} \) subject to the constraint \( aK + bL = C_0 \). This is the setting of Lagrange multipliers.
   (a) Maximize the Cobb-Douglas production function \( Q = K^{1/4}L^{3/4} \) subject to the constraint \( 3K + 2L = 7 \).
   (b) Maximize \( Q = K^{\alpha}L^{1-\alpha} \) subject to \( aK + bL = C_0 \).

5. Find three positive numbers whose sum is \( S \) and whose product is maximized. What is the maximum value of the product?

6. Find \( n \) positive numbers whose sum is \( S \) and whose product is maximized.

7. Find the dimensions of the box of greatest volume that can be made from 24 square feet of material. Assume no waste in construction.

8. Find the dimensions of the largest open-top box that can be made from 24 square feet of material.
3.8. CONSTRAINED OPTIMIZATION

Problem Solving

We have learned a powerful technique in Lagrange multipliers. It turns out that Problem 6 has far reaching consequences for problem solving, and we borrow a piece of Polya’s book *Mathematics and Plausible Reasoning, Volume I: Induction and Analogy in Mathematics*. One notices that the solution to Problem 6 implies that if $S$ is fixed and $x_1 + \cdots + x_n = S$, then

$$x_1 x_2 \cdots x_n \leq \left(\frac{S}{n}\right)^n = \left(\frac{x_1 + \cdots + x_n}{n}\right)^n. \quad (3.8.2)$$

Notice that one has equality if and only if each $x_i$ is $S/n$. Taking the $n^{th}$ root of both sides gives the famous relationship between geometric and arithmetic means:

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + \cdots + x_n}{n}.$$

Moreover, this result can be used to easily solve Problem 7 when looked at the right way. If we let $x$, $y$, and $z$ denote the dimensions of the box, and $S$ the surface area (which is 24 in the problem), then the constraint is $2xy + 2yz + 2xz = S$. Thus $xy$, $yz$, and $xz$ are three positive numbers whose sum is $S/2$. By Inequality (3.8.2) we see

$$(xyz)^2 = (xy)(yz)(xz) \leq \left(\frac{S/2}{3}\right)^3 = \left(\frac{S}{6}\right)^3,$$

with equality holding exactly when $xy = yz = xz$ (implying $x = y = z$). Now substitute $V = xyz$ to see that for any choice of dimensions

$$V^2 \leq \left(\frac{S}{6}\right)^3,$$

with equality holding exactly when the box is a cube. So one clever application of the relationship between arithmetic and geometric means shows that a cube maximizes the volume of a box with fixed surface area, and that the relationship is $V^2 = \left(\frac{S}{6}\right)^3$.

Let’s take a minute to discuss the application a little more closely. To apply Inequality (3.8.2) we thought of the constraint $xy + yz + xz = S/2$ as a statement about the arithmetic mean of the three numbers $xy$, $yz$, and $xz$. Even though we want to know $x$, $y$, and $z$ individually, we pushed on, applying the inequality for the numbers $xy$, $yz$, and $xz$. The next key observation is this. We want to maximize the volume. While the product of $xy$, $yz$, and $xz$ is not the volume, it is the square of the volume! With appropriate substitutions, the inequality ends up relating the volume and surface area of a box. Knowing that equality is reached only when the original numbers $xy$, $yz$, and $xz$ are equal solves the optimization problem.

Notice that the same reasoning also proves the box with least surface area for a fixed volume is the cube, without any more calculation! Can you solve the open-box problem in a similar fashion?
Polya describes more examples of using a problem-solving approach to solve optimization problems. He outlines what he calls the pattern of partial variation to prove the geometric mean is at most the arithmetic mean. He poses many other geometric optimization problems that can be solved in a manner similar to the above reasoning: Think of the constraint as a statement about a mean, apply the geometric-arithmetic mean inequality \( 3.8.2 \) and note that the other side of the inequality is (a function of) the quantity you want to maximize or minimize. The interested reader is referred to Chapter VIII of Polya’s *Mathematics and Plausible Reasoning, Volume 1: Induction and Analogy in Mathematics* for more examples and problems.
Chapter 4

Integration

In this chapter we introduce several extensions of the single-variable notion of integration. In the single-variable setting, the notation \( \int_{a}^{b} f(x) \, dx \) involves three ingredients: the interval \([a, b]\) of integration, the integrand \(f(x)\), and the differential \(dx\) (which is an infinitesimal change in length). We first extend the integrand to functions \(f(x, y)\) of two variables, and the interval \([a, b]\) of the \(x\)-axis to a region \(R\) of the \(xy\)-plane. The infinitesimal length \(dx\) extends to an infinitesimal area \(dA\), resulting in the definition of a double integral \(\iint_{R} f(x, y) \, dA\) in Section 4.1. The definition of \(\iint_{R} f(x, y) \, dA\) is difficult to calculate with, so Section 4.2 introduces techniques for setting up and calculating double integrals.

The skill, introduced in Section 1.6, of describing regions \(R\) using inequalities is reviewed, and gives rise to limits of integration. Section 4.3 extends our abilities one dimension, to triple integrals \(\iiint_{W} f(x, y, z) \, dV\). The region of integration is a solid \(W\) in \(\mathbb{R}^3\), the integrand a function of three variables, and \(dV\) an infinitesimal volume. Section 4.4 shows how to handle integration using coordinate systems other than Cartesian Coordinates. Finally, in Section 4.6 we show how to integrate functions along curves and surfaces. By the end of the chapter, then, you will be able to integrate functions of several variables over regions as diverse as surfaces in space, solids in space, regions in the plane, and curves in either \(\mathbb{R}^2\) or \(\mathbb{R}^3\)–quite an increase in sophistication!

4.1 Double Integrals–The Definition

To motivate the definition of the double integral we recall the definition of the definite integral \(\int_{a}^{b} f(x) \, dx\) from single variable calculus. Recall that if \(f(x) \geq 0\) on the interval \([a, b]\), the integral \(\int_{a}^{b} f(x) \, dx\) is the area \(A\) under the curve, and we focus on this situation to add geometric intuition. The big idea of the definite integral was this:

\textit{Approximate} \(A\) \textit{with areas you can calculate, then take a limit as your approximation gets better.}

Actually working out this idea is an algebraic nightmare, but it’s a recurring
nightmare that gets less scary each time you see it. We remind you of it now, without resorting to complete generality for simplicity of exposition.

To calculate the area $A$ under the graph of $y = f(x)$, first partition the interval $[a, b]$ into $n$ equal subintervals of width $\Delta x = \frac{b-a}{n}$. Label the partition points $a = x_0 < x_1 < \cdots < x_n = b$. With this notation, the $i^{th}$ subinterval is $[x_{i-1}, x_i]$ and the width of each subinterval is $x_i - x_{i-1} = \Delta x$. Now build a rectangle $R_i$ above each subinterval $[x_{i-1}, x_i]$ by choosing the height of the rectangle to be a point on the curve. Analytically this means: arbitrarily pick values $h_i$ in each subinterval $[x_{i-1}, x_i]$ and let $f(h_i)$ be the height of the rectangle $R_i$ (see Figure 4.1.1(a)).

![Diagram](image)

(a) The Rectangle $R_i$

(b) Approximating Rectangles

Figure 4.1.1: The Definite Integral

We now have lots of rectangles, and the sum of their areas is approximately the area under the curve. More importantly, we can calculate the areas of the rectangles! The height of $R_i$ is $f(h_i)$, and its width is $\Delta x$, so its area is $f(h_i)\Delta x$. Adding up the areas of the rectangles gives the familiar Riemann sum approximation for the area under the curve,

$$A \approx \sum_{i=1}^{n} f(h_i)\Delta x.$$

We have accomplished the first step in the process of finding the area $A$ under the curve $y = f(x)$. We have approximated $A$ with areas we can calculate.

Of course there is usually error in this approximation, which is seen geometrically as portions of the rectangles that are either above or below the curve. One asks how they might get a better approximation, and one obvious-ish answer is to use more (and hence smaller) rectangles.

This is the clue to completing the second step in the process—taking a limit as your approximation gets better. Recall that the first thing we did was partition $[a, b]$ into $n$ subintervals of equal length. Since the number of rectangles
constructed is \( n \), and we want to use more rectangles, we simply let \( n \) tend to infinity. This motivated the definition of the definite integral:

\[
\int_a^b f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^n f(x_i) \Delta x.
\]

Let us now turn our attention to a higher dimensional problem—that of finding the volume under the surface \( z = f(x, y) \). Again we make some initial simplifying assumptions, which will be removed later. Let’s assume we want the volume under \( z = f(x, y) \) and over the rectangle \([a, b] \times [c, d]\) in the \( xy \)-plane. Further, we assume that \( f(x, y) \geq 0 \) above this rectangle. To calculate the volume \( V \), we use the same basic idea as in the area problem considered above:

*Approximate \( V \) with volumes you can calculate, then take a limit as your approximation gets better.*

\[
\text{(a) The Box } B_{ij} \quad \text{(b) Approximating Boxes}
\]

Figure 4.1.3: Approximating Volume with Boxes
The idea is completely analogous to the single variable case, but there is a little more to keep track of with two variables. As before, partition the $x$-interval $[a, b]$ into $n$ equal subintervals with partition points $a = x_0 < x_1 < \cdots < x_n = b$. Now do the same with the $y$-interval $[c, d]$ obtaining partition points $c = y_0 < y_1 < \cdots < y_n$. This results in partitioning the rectangle $[a, b] \times [c, d]$ into $n^2$ subrectangles $R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j]$, each with area $\Delta x \Delta y = \frac{b-a}{n} \cdot \frac{d-c}{n}$. Now build a rectangular box above each rectangle $R_{ij}$ by picking a point $(h_i, k_j) \in R_{ij}$ and letting $f(h_i, k_j)$ be the height of the box. As before, we can calculate the volumes of the boxes, and the sum of them approximates the volume under the surface. The difference is that we have to add all the volumes we must sum on both indices $i$ and $j$, resulting in a double sum.

\[ V \approx \sum_{i=1}^{n} \sum_{j=1}^{n} f(h_i, k_j) \Delta x \Delta y \]

Again we observe that our approximation becomes better if we take finer partitions, so we define the double integral of $f$ over the rectangle $R$ to be

\[ \int \int_R f(x, y) \, dA = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(h_i, k_j) \Delta x \Delta y. \]

**Example 4.1.1. Volume of a prism**

Approximate the volume under $f(x, y) = 2 - x$ above the rectangle $R = [0, 1] \times [0, 3]$ (see Figure 4.1.4(a)).

We will use the above technique, partitioning each axis into two subintervals (i.e. $n = 2$). In this case $\Delta x = \frac{1-0}{2} = \frac{1}{2}$ and $\Delta y = \frac{3-0}{2} = \frac{3}{2}$. Our partition points are

\[
\begin{align*}
  x_0 &= 0 \\
  x_1 &= 0 + \Delta x = \frac{1}{2} \\
  x_2 &= 0 + 2\Delta x = 0 + 2 \cdot \frac{1}{2} = 1 \\
  y_0 &= 0 \\
  y_1 &= 0 + \Delta y = \frac{3}{2} \\
  y_2 &= 0 + 2\Delta y = 0 + 2 \cdot \frac{3}{2} = 3.
\end{align*}
\]

It is evident that the general pattern for partition points in this context is

\[ x_i = x_0 + i\Delta x \quad \quad y_j = y_0 + j\Delta y \quad \quad (4.1.1) \]

These partition points divide $R$ into four subrectangles, and we use the “upper right” corner of each to determine the height of our box. Thus we must evaluate $f$ at points $(1/2, 3/2), (1/2, 3), (1, 3/2)$ and $(1, 3)$ to get the heights of the boxes (see Figure 4.1.4(b)). Summing the volumes of the boxes approximates the volume under the plane:

\[
V \approx \sum_{i=1}^{2} \sum_{j=1}^{2} f(x_i, y_j) \Delta x \Delta y
\]

\[
= f(x_1, y_1) \Delta x \Delta y + f(x_1, y_2) \Delta x \Delta y + f(x_2, y_1) \Delta x \Delta y + f(x_2, y_2) \Delta x \Delta y
\]

\[
= \left( 2 - \frac{1}{2} \right) \frac{3}{4} + \left( 2 - \frac{1}{2} \right) \frac{3}{4} + (2 - 1) \frac{3}{4} + (2 - 1) \frac{3}{4} = \frac{15}{4}. \quad \Box
\]
4.1. DOUBLE INTEGRALS–THE DEFINITION

Example 4.1.2. Volume of a prism–again!

This time we want to approximate the volume under \( f(x,y) = 2 - x \) above the rectangle \( R = [0,1] \times [0,3] \) using \( n \) partition points, then taking a limit as \( n \to \infty \).

Recall that for any constant \( k \) we have \( \sum_{j=1}^{n} k = kn \), and that the sum of the first \( n \) integers is given by \( \sum_{i=1}^{n} i = \frac{n(n+1)}{2} \).

When using \( n \) partition points we have \( \Delta x = \frac{1}{n} \) and \( \Delta y = \frac{3}{n} \). Moreover, from Equation 4.1.1 we get the general formula for the partition points:

\[
x_i = i\Delta x = \frac{i}{n} \quad y_j = j\Delta y = \frac{3j}{n}.
\]

Again using the “upper right” corner of each subrectangle we see the height of the box above the \( ij^{th} \) rectangle is \( f(x_i, y_j) = 2 - \frac{i}{n} \). Thus the sum of the volumes of the boxes can be calculated

\[
V \approx \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i, y_j) \Delta x \Delta y
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{n} \left( 2 - \frac{i}{n} \right) \frac{1}{n} \cdot \frac{3}{n} = \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \frac{6}{n^2} - \frac{3i}{n^3} \right)
\]

\[
= \sum_{i=1}^{n} \left( \frac{6}{n} - \frac{3i}{n^2} \right) \quad \text{(multiply by } n \text{ since terms constant wrt } j)\]

\[
= \sum_{i=1}^{n} \frac{6}{n} - \frac{3}{n^2} \sum_{i=1}^{n} i \quad \text{(algebra with summations)}
\]

\[
= 6 - \frac{3}{n^2} \cdot \frac{n(n+1)}{2} = 6 - \frac{3(n+1)}{2n}.
\]

Of course, the volume is the limit as our approximation gets finer, and it is
also the double integral. Thus we can calculate

\[ V = \int\int_R (2 - x) \, dA = \lim_{n \to \infty} 6 - \frac{3(n + 1)}{2n} = 6 - \frac{3}{2} = \frac{9}{2}, \]

We have calculated our first double integral!! Luckily, we won’t calculate many double integrals this way. In the next section we introduce iterated integrals and Fubini’s Theorem. These allow us to integrate functions one variable at a time, much like partial derivatives allow us to differentiate one variable at a time.

Before leaving this example, notice that we can use geometry to calculate the volume of the solid. In fact, the volume will just be the area of the trapezoidal face times the length of the prism. Since the area of the face is \( \frac{3}{2} \) and the length is 3, we have the volume is \( V = \frac{3}{2} \cdot 3 = \frac{9}{2} \), which agrees with our integral calculation.

▲

Math App 4.1.1. Double Integration

In this math app you will be able to control the number of partitions in a double integral, and see the corresponding improvements in volume approximation. As a warning, however, for a 10 \times 10 partition it takes around 15 seconds to plot the graphs—so be patient!

Technical Remarks About \( R \): At this point we have a definition of the double integral \( \iint_R f(x, y) \, dA \) when \( R \) is a rectangle in the \( xy \)-plane. We’d like to integrate functions over much more general regions than these. You can extend to other regions \( R \) by taking a union of rectangles that approximate \( R \) and summing the integrals of \( f(x, y) \) over those. Once that is done, you take a limit as your rectangle approximations of \( R \) improve. These are details we won’t concern ourselves with here. Suffice it to say that if the boundary of \( R \) is a union of piecewise smooth curves, you can define the double integral \( \iint_R f(x, y) \, dA \).
4.2 Calculating Double Integrals

Now that we know what a double integral is, we need to get comfortable calculating them. In the single variable case this was accomplished by the Fundamental Theorem of Calculus. Remember that this theorem relates the definite integral with taking an antiderivative, so you never have to take limits of Riemann sums again! In the two variable case, Fubini’s Theorem states that the double integral we just defined can be calculated by integrating one variable at a time—partial integration, if you will. The fancy math term for this is an iterated integral, and we introduce the notation. An iterated integral is denoted

\[ \int_a^b \int_c^d f(x, y) \, dy \, dx, \]

and means first evaluate \( \int_c^d f(x, y) \, dy \), thinking of \( x \) as a constant. The result will be a function of \( x \), since evaluating at the endpoints \( c \) and \( d \) gets rid of the \( y \)'s. Then integrate that with respect to \( x \). An example illustrates this.

Example 4.2.1. Iterated Integrals

Evaluate \( \int_{-1}^2 \int_0^2 4y + x^2y^2 \, dy \, dx \). Since this is the first example we’ve seen, we’ll separate the steps. Very quickly you will become familiar with the notation and we will carry the double integral notation along as we go. Thinking of \( x \) as a constant, we evaluate

\[ \int_0^2 4y + x^2y^2 \, dy = 2y^2 + \frac{x^3}{3} \bigg|_0^2 = 8 - \frac{8}{3}x^2, \]

which is a function of \( x \). Note that when evaluating at the endpoints, we only substitute the values for \( y \), leaving \( x \) alone. This is because our integration was with respect to \( y \) at this stage. We finish the iterated integral by integrating the result with respect to \( x \):

\[
\int_{-1}^2 \left( \int_0^2 4y + x^2y^2 \, dy \right) \, dx = \int_{-1}^2 \left( 8 - \frac{8}{3}x^2 \right) \, dx
= 8x - \frac{8}{9}x^3 \bigg|_{-1}^2 = 16 - \frac{64}{9} - \left( -8 + \frac{8}{9} \right) = 16. \quad \text{▲}
\]

We note that there is nothing special about integrating with respect to \( y \) first, then with respect to \( x \). Iterated integrals can be calculated in the other order as well, as long as some care is taken with the limits of integration. When the
Before leaving this example we mention again that the limits on the inside integral sign correspond to the first variable of integration, while the limits on the outer integral sign correspond to the second. Take some time to familiarize yourself with the notation of an iterated integral, it will be worth the effort. ▲

Now that we know what an iterated integral is, we can state Fubini’s Theorem. Before doing so, remember that our definition of the double integral is a limit of Riemann sums. It’s very ugly, and difficult to calculate. For this reason, Fubini’s theorem is quite nice.

**Theorem 4.2.1.** Let \( f(x, y) \) be a nice function defined on the rectangle \( R = [a, b] \times [c, d] \) in the \( xy \)-plane. Then

\[
\iint_R f(x, y) \, dA = \int_a^b \int_c^d f(x, y) \, dydx = \int_c^d \int_a^b f(x, y) \, dxdy
\]

Before moving on we do one more example of an iterated integral over a rectangle.

**Example 4.2.2. Introducing an iterated integral**

Evaluate \( \iint_R \frac{y}{1 + x^2} \, dA \) where \( R \) is the rectangle \([0, 1] \times [1, 2] \). By Fubini’s Theorem we see that the double integral can be evaluated as the iterated integral

\[
\int_0^1 \int_1^2 \frac{y}{1 + x^2} \, dydx = \int_0^1 \frac{1}{1 + x^2} \left( \int_1^2 y \, dy \right) \, dx
\]

\[
= \int_0^1 \frac{1}{1 + x^2} \left( \frac{y^2}{2} \right)_{1}^{2} \, dx
\]

\[
= \int_0^1 \frac{3}{2(1 + x^2)} \, dx
\]

\[
= \frac{3}{2} \tan^{-1} x \bigg|_{0}^{1} = \frac{3}{2} \tan^{-1}(1) - \frac{3}{2} \tan^{-1}(0) = \frac{3\pi}{8}. \quad \text{▲}
\]
4.2. CALCULATING DOUBLE INTEGRALS

Limits of Integration from $R$:

There are three aspects of the double integral $\int\int_R f(x,y) \, dA$ to consider when setting up and evaluating it: the region $R$, the integrand $f(x,y)$, and the differential $dA$. The infinitesimal area $dA$ will depend on the coordinates we use to describe the region $R$. Different interpretations of the function $f(x,y)$ will yield different interpretations of the integral. For example, if $f(x,y)$ is the density of a thin plate $R$, then $\int\int_R f(x,y) \, dA$ is the mass of $R$. The region $R$ dictates the limits of integration in iterated integrals. Thus all aspects of the double integral are vital for setting up, evaluating, and interpreting it. We now wish to investigate how the region $R$ determines the limits of integration.

So far we have focused on integrating functions over rectangles. We want to be able to integrate over more general regions in the plane. There are two particularly nice types of regions for which it is easy (sort of) to find limits of integration. We describe them in terms of which variable you will integrate with respect to first.

A $dy$-region in the plane is one in which the top and bottom are functions of $x$, and whose shadow on the $x$-axis is an interval (see Figure 4.2.1(a)). A $dx$-region is one in which the right and left are functions of $y$ and whose shadow on the $y$-axis is an interval (see Figure 4.2.1(b)). Some regions in the plane are both $dx$- and $dy$-regions, and some are neither. It is relatively easy to describe such regions with a system of inequalities. This is an important skill, needed to determine the limits of integration for iterated integrals.

![Figure 4.2.1: Elementary Regions](image)

To describe a $dy$-region:

1. Solve the equations of boundary curves for $y$, so that the top curve is $y = f_2(x)$ and the bottom curve is $y = f_1(x)$. 

2. Find the shadow on the \( x \)-axis. This should be an interval \([a, b]\).

3. The system of inequalities is:

\[
 f_1(x) \leq y \leq f_2(x) \\
 a \leq x \leq b
\]

To integrate a function \( f(x, y) \) over a \( dy \)-region, the system of inequalities gives the limits of integration. You get

\[
 \iint_R f(x, y) \, dA = \int_a^b \int_{f_1(x)}^{f_2(x)} f(x, y) \, dy \, dx.
\]

The technique for finding limits of integration for a \( dx \)-region is analogous to this. We include the description for completeness.

**To describe a \( dx \)-region:**

1. Solve the equations of boundary curves for \( x \), so that the right curve is \( x = g_2(y) \) and the left curve is \( x = g_1(y) \).

2. Find the shadow on the \( y \)-axis. This should be an interval \([c, d]\).

3. The system of inequalities is:

\[
 g_1(y) \leq x \leq g_2(y) \\
 c \leq y \leq d
\]

To integrate a function \( f(x, y) \) over a \( dx \)-region, the system of inequalities gives the limits of integration. You get

\[
 \iint_R f(x, y) \, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y) \, dx \, dy.
\]

We begin our study with some animated Math Apps intended to strengthen your intuition.

**Math App 4.2.1. Limits in Double Integrals**

In the following Math App the limits of integration are found for a \( dy \)-region. The animation assists you in viewing them.
4.2. CALCULATING DOUBLE INTEGRALS

Math App 4.2.2. A Region that is both dx and dy

In this example we use animations to compare the $dx$- and $dy$-limits of integration for the same region $R$. They will turn out to be surprising different!

Example 4.2.3. Describing Regions in the Plane

Use inequalities to describe the triangle $T$ in the first quadrant bounded by the axes and the line $3x + 2y = 6$. Thinking of the triangle as a $dy$-region, so the line is the top of the region, we solve the boundary equation for $y$ to get $y = -\frac{3}{2}x + 3$. The line $y = 0$ is the bottom of $T$ and the line $y = -\frac{3}{2}x + 3$ is the top. Looking at the shadow of $T$ on the $x$-axis gives $0 \leq x \leq 2$. Thus, as a $dy$-region, we can describe $T$ by the system of inequalities:

\[
0 \leq y \leq -\frac{3}{2}x + 3 \\
0 \leq x \leq 2
\]

Figure 4.2.2: The Triangle $T$

To integrate any function $f(x, y)$ over the triangle, you get the integral

\[
\int_0^2 \int_0^{-\frac{3}{2}x+3} f(x, y) \, dy \, dx.
\]
An intuitive way to finding the above limits of integration is the following. Walk toward $\infty$ along a line parallel to the $y$-axis that goes through your region. The equation of the curve where you enter the region is the lower limit, and the equation for the curve where you leave the region is your upper limit.

If one thinks of the line $3x + 2y = 6$ as the right boundary of the region, rather than the top, $T$ becomes a $dx$-region. The left boundary is the $y$-axis, or the line $x = 0$, and one solves the equation $3x + 2y = 6$ for $x$ to find the right boundary. The shadow on the $y$-axis is the interval $[0, 3]$. So we can describe $T$ via the system of inequalities

$$0 \leq x \leq \frac{-2}{3}y + 2$$
$$0 \leq y \leq 3$$

Thus another way to integrate $f(x, y)$ over the triangle is

$$\int_{0}^{3} \int_{0}^{-\frac{2}{3}y+3} f(x, y) \, dx \, dy.$$ 

Again, an intuitive way to find the limits on $x$ in this example is to walk toward $\infty$ on a line through the region and parallel to the $x$-axis. The lower limit is the equation of the curve where you enter the region and the upper limit is the equation of the curve where you leave it. Both these equations are solved for $x$ if you’re integrating with respect to $x$. ▲

**Example 4.2.4. A $dx$-region**

Describe the region $R$ bounded by the coordinate axes and the lines $y = 2$ and $x + 2y = 3$ both as a $dy$-region and as a $dx$-region. To describe $R$ as a $dy$-region, notice that the top function changes at $x = 2$. This means we have to split our description there, and use two systems of inequalities. Since we’re thinking $dy$ first, we solve our boundary equations for $y$ and use the “bottom-top” approach. This gives the systems:

$$0 \leq y \leq 2 \quad \quad \quad \quad 0 \leq y \leq 3 - \frac{1}{2}x$$
$$0 \leq x \leq 2 \quad \quad \quad \quad 2 \leq x \leq 6$$

It’s actually easier to think of the region as a $dx$-region, because there are single boundary curves on the left and right sides of the region. Walking along a line through $R$ and parallel to the $x$-axis, you enter the region at the curve $x = 0$ and leave $R$ at the curve $x = 6 - 2y$. The shadow on the $y$-axis is the interval $[0, 2]$, yielding the system of inequalities

$$0 \leq x \leq 6 - 2y$$
$$0 \leq y \leq 2$$
4.2. CALCULATING DOUBLE INTEGRALS

These inequalities provide the limits of integration when integrating any function over $R$. ▲

So far we have seen that describing $R$ using a system of inequalities gives limits of integration for the region $R$. Conversely, given limits of integration we can find the region $R$. This will be useful when we want to switch the order of integration in certain examples.

**Example 4.2.5. The region from limits**

Determine the region $R$ of integration for the double integral

$$
\int_{-1}^{0} \int_{-y}^{2-y^2} f(x, y) \, dx \, dy.
$$

To accomplish this, we write down the system of inequalities corresponding to the given limits, then sketch the region. The system is

$$
-y \leq x \leq 2 - y^2, \\
-1 \leq y \leq 0.
$$

Sketching the boundary curves gives the region in Figure 4.2.4. ▲

**Example 4.2.6. Integrals over more general regions**
Find the volume of the tetrahedron bounded by the coordinate planes and the plane $3x + 2y + z = 6$ (see Figure 4.2.5). We can think of the tetrahedron as a solid under the graph $z = f(x, y)$. This solid is the region under the graph of $z = 6 - 3x - 2y$, and above the triangle $T$ in the $xy$-plane from the previous example. To find the volume, then, we just need to integrate using the limits found in the previous example.

$$
\int_T 6 - 3x - 2y \, dA = \int_0^2 \int_0^{-\frac{x}{2}+3} (6 - 3x - 2y) \, dy \, dx
$$

$$
= \int_0^2 \left[ (6 - 3x)y - y^2 \right]_{-\frac{x}{2}+3}^0 \, dx
$$

$$
= \int_0^2 (6 - 3x) \left( (6 - 3x)y - y^2 \right) - \left( (6 - 3x)y - y^2 \right)^2 \, dx
$$

$$
= \int_0^2 \left( \frac{9}{4}y^2 - 9x + 9 \right) \, dx
$$

$$
= \left[ \frac{3}{4}y^3 - \frac{9}{2}x^2 + 9x \right]_0^2 = 6. \quad \triangle
$$

Figure 4.2.5: The Tetrahedron

**Example 4.2.7. Double integrals and area**

Recall that if $f(x, y) \geq 0$ over the region $R$ in the plane, then $\iint_R f(x, y) \, dA$ is the volume of the solid under $z = f(x, y)$ and above $R$. If we use the constant function $f(x, y) = c$, then the volume of the solid is the area of $R$ times the height $c$. More specifically, if $f(x, y) = 1$ then $\iint_R dA$ is the area of $R$. We use this observation now.

Find the area of the region bounded by $x = 4 - y^2$ and $x + 2y = 4$. Sketching the region in Figure 4.2.6 we see that the line is the left side and the parabola the right. Moreover, the points of intersection of the curves can be found by
substituting $4 - y^2$ for $x$ in the equation for the line. One obtains

\[
4 - y^2 + 2y = 4 \\
y^2 - 2y = 0
\]

from which we see that $y = 0$ or $y = 2$. This implies that the region can be described by

\[
-2y + 4 \leq x \leq 4 - y^2 \\
0 \leq y \leq 2.
\]

Thus the area is given by

\[
\int_0^2 \int_{4-2y}^{4-y^2} dx dy = \int_0^2 x|_{4-2y}^{4-y^2} dy \\
= \int_0^2 (4 - y^2) - (4 - 2y)dy = \int_0^2 2y - y^2 dy \\
= y^2 - \frac{y^3}{3} \bigg|_0^2 = \frac{4}{3} \triangle
\]

We also observe that after integrating with respect to $x$ we are left with the integral $\int_0^2 (4 - y^2) - (4 - 2y)dy$. This is precisely the integral you get using single variable techniques to find area! The area of a region between two curves is the integral of the “top” curve minus the “bottom” curve. In this case you use the right curve minus the left curve because everything is sideways.

**Example 4.2.8. Switching Order of Integration**

Evaluate $\int_0^{\pi/2} \int_x^{\pi/2} \sin y^2 \, dy \, dx$.

To do this, notice we don’t really know how to integrate $\int \sin y^2 \, dy$, but $\int \sin y^2 \, dx = x \sin y^2 + C$ since $\sin y^2$ is a constant when integrating with respect
to $x$. With this as motivation, we switch the order of integration and hope for the best. To do so, we use the given limits of integration to find the region $R$ of integration. The limits describe the system of inequalities

$$x \leq y \leq \sqrt{\pi/2}$$
$$0 \leq x \leq \sqrt{\pi/2},$$

which is the triangle pictured in Figure 4.2.7. Thinking of it as a $dx$-region, the left side is $x = 0$, the right is $x = y$, and the shadow on the $y$-axis is the interval $[0, \sqrt{\pi/2}]$. Thus, if we want to integrate with respect to $x$ first, the integral becomes:

$$\iint_R \sin y^2 \, dA = \int_0^{\sqrt{\pi/2}} \int_0^y \sin y^2 \, dx \, dy$$
$$= \int_0^{\sqrt{\pi/2}} x \sin y^2 \bigg|_0^y \, dy$$
$$= \int_0^{\sqrt{\pi/2}} y \sin y^2 \, dy$$
$$= \frac{1}{2} \sin y^2 \bigg|_0^{\sqrt{\pi/2}} = \frac{1}{2}.$$ 

The final integration is a simple substitution. This is an example of an integral that is quite direct when viewed one way, but originally very difficult to do.

**Exercises**

In Problems 1-9 describe the given regions first as a $dx$-region, then as a $dy$-region.

1. The triangle in the first quadrant bounded by the coordinate axes and the line $4x + 3y = 12$.

2. The triangle in the second quadrant bounded by the coordinate axes and the line $7x - 2y = 8$. 
3. The trapezoid in the first quadrant bounded by the coordinate axes together with the lines \( y = 2 \) and \( 3x + y = 6 \).

4. The bounded region between the \( x \)-axis and the parabola \( y = 4 - x^2 \).

5. The quarter circle in the first quadrant and under \( x^2 + y^2 = 1 \).

6. The unit disk \( x^2 + y^2 \leq 1 \).

7. The region in the first quadrant between the unit circle and the circle \( x^2 + y^2 = 4 \).

8. The bounded region between the \( y \)-axis and the parabola \( x = 1 - y^2 \).

9. The bounded region between \( x = 1 - y^2 \) and \( x + 2y = 1 \).

In problems 10-17 sketch the regions of integration for the iterated integral.

10. \( \int_{-1}^{3} \int_{-1}^{4} f(x, y) \, dx \, dy \)

11. \( \int_{-1}^{3} \int_{-1}^{4} f(x, y) \, dy \, dx \)

12. \( \int_{-1}^{1} \int_{0}^{3-x} f(x, y) \, dy \, dx \)

13. \( \int_{0}^{2} \int_{y-2}^{0} f(x, y) \, dx \, dy \)

14. \( \int_{-x}^{2-x^2} f(x, y) \, dy \, dx \)

15. \( \int_{-1/2}^{1} \int_{1-y}^{2-2y^2} f(x, y) \, dx \, dy \)

16. \( \int_{-1}^{3} \int_{x^2}^{2x+3} f(x, y) \, dy \, dx \)

17. \( \int_{-1}^{1} \int_{y^2-1}^{2-2y^2} f(x, y) \, dx \, dy \)

In Problems 18-23 evaluate the iterated integrals.

18. \( \int_{-1}^{0} \int_{0}^{2} 2x - 3y^2 \, dy \, dx \)

19. \( \int_{0}^{1} \int_{0}^{\pi/2} xy \sin y \, dy \, dx \)
20. \[ \int_0^1 \int_0^{\sqrt{1-x^2}} y \, dx \, dy \]

21. \[ \int_1^2 \int_0^x (2x-y) \, dy \, dx \]

22. \[ \int_0^{\pi/2} \int_0^x \sin x \, dy \, dx \]

23. \[ \int_{-2}^{2} \int_{x^2-4}^{1-x^2} dy \, dx \]

24. Find the volume of the tetrahedron bounded by the coordinate planes and \(3x + 2y + z = 6\).

25. Find the volume of the solid under the paraboloid \(z = 4 - x^2 - y^2\) and above the rectangle \([0, 1] \times [-1, 1]\).

26. Find the volume of the solid under the plane \(2x + y + z = 10\) and over the rectangle \([1, 3] \times [-1, 0]\).

27. Find the volume of the solid under the plane \(x + y + z = 20\) in \(\mathbb{R}^3\) and over the region in the \(xy\)-plane bounded by the curves \(x = 2 - y^2\) and \(y = -x\).

28. Find the area between \(y = x^2 - 2\) and \(y = x\) using double integration.

29. Find the area between \(x + y = 1\) and \(x = y^2 - 1\) using double integration.

30. Evaluate \( \int_0^1 \int_x^1 e^{y^2} \, dy \, dx \) by first switching the order of integration.

31. Evaluate \( \int_0^1 \int_0^{\sqrt[3]{1-x^2}} \sqrt{1-y^3} \, dy \, dx \) by first switching the order of integration.

32. Evaluate \( \iint_R x - y \, dA \), where \(R\) is the triangle in the plane with vertices \((0, 0), (1, 1)\) and \((2, 1)\).

33. Evaluate \( \iiint_R \, dA \) where \(R\) is the bounded region between the parabolas \(y = 1 - x^2\) and \(y = x^2 - 1\).
4.3 Triple Integrals

The integral of a function \( f(x, y, z) \) of three variables over a solid \( W \) is a triple integral, and denoted \( \iiint_W f(x, y, z) \, dV \). It is defined completely analogously to the double integral. Partition the solid \( W \) into rectangular boxes with length \( \Delta x_i \), width \( \Delta y_j \) and height \( \Delta z_k \). The volume of the \( ijk^{th} \) box is \( \Delta x_i \Delta y_j \Delta z_k \).

Choose a point \( (a_i, b_j, c_k) \) in each box, evaluate \( f \) there, and sum the values to estimate the triple integral. Taking a limit as the partition gets finer yields

\[
\iiint_W f(x, y, z) \, dV = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} f(a_i, b_j, c_k) \Delta x_i \Delta y_j \Delta z_k.
\]

Motivating the Triple Integral—Mass from Density: This process can seem pretty abstract, so we discuss a physical example in which integration arises naturally: calculating mass from density. Recall that the density of an object is mass per unit volume. If two objects are exactly the same size, the more dense object is heavier.

**Example 4.3.1. Constant density**

If the density of an object is constant, then the mass is simply density times volume. Suppose the tetrahedron of Example 4.2.6 is made from material with constant density 10 grams per cubic centimeter, then the mass \( M \) is

\[
M = \text{Volume} \cdot \text{density} = 6 \cdot 10 = 60 \text{ gm}.
\]

This follows because in Example 4.2.6 we calculated the volume of the tetrahedron to be 6 \( \text{cm}^3 \), assuming the units on each axis to be centimeters.

**Example 4.3.2. Mass from variable density**

Now suppose the tetrahedron is more dense at its base than at its apex, and we still want to calculate its mass. More precisely, suppose the density function is given by \( \rho(x, y, z) = 40 - z^2 \). Our approach could be the standard calculus approach to such problems: Approximate the quantity using things you can calculate, then take a limit as your approximation gets better.

We describe an example and then the more general, yet still finite, setting. The actual calculations are tedious, and not that important. It is significant that:

The process of approximating mass from variable density then taking a limit as the approximations improve leads naturally to integration.

A particularly simple approximation would be the following. We could say that the tetrahedron is approximated by the three boxes

\[
B_1 = [0, \frac{4}{3}] \times [0, 1] \times [0, 2],
B_2 = [0, \frac{4}{3}] \times [1, 2] \times [0, 2], \text{ and}
B_3 = [0, \frac{4}{3}] \times [0, 1] \times [2, 4].
\]
CHAPTER 4. INTEGRATION

Figure 4.3.1: Approximating the mass of a tetrahedron

Note the corresponding volumes of the boxes are \( V_1 = \frac{8}{3}, \ V_2 = \frac{4}{3}, \) and \( V_3 = \frac{4}{3}. \) To further simplify our calculation we could assume that the box has constant density, the density \( \rho(x, y, z) = 40 - z^2 \) of its geometric center. The sum of the masses of the boxes with these assumptions approximates the mass of the tetrahedron. Then our approximate mass of the tetrahedron is

\[
M \approx V_1 \rho(2/3, 1/2, 1) + V_2 \rho(1/3, 3/2, 1) + V_3 \rho(1/3, 1/2, 3) \\
= \frac{8}{3} \cdot 39 + \frac{4}{3} \cdot 39 + \frac{4}{3} \cdot 36 = 204 \text{ gm.} \uparrow
\]

The assumptions we made were that our object was made up of boxes (so we can easily calculate volume) and that the density of each box was constant (which facilitates the mass calculation). If we use many small boxes to approximate the tetrahedron, with constant densities which are close to that given by \( \rho(x, y, z) \) for each box, we could add up the masses of the boxes to approximate the mass of the tetrahedron.

In our case, we could partition the tetrahedron into tiny boxes

\[
B_{ijk} = [x_i, x_{i+1}] \times [y_j, y_{j+1}] \times [z_k, z_{k+1}]
\]

and assume the density is constant on each box. If the boxes are small, this assumption is reasonable. To determine the approximate density on each cube, pick a point \((a_i, b_j, c_k)\) in \(B_{ijk}\) and use its density \( \rho(a_i, b_j, c_k) \). Again, if the boxes are small, then \( \rho(a_i, b_j, c_k) \) should be close to the density of the material at each point in box \(B_{ijk}\). The mass \( M_{ijk} \) of the box \( B_{ijk} \) is approximated by

\[
M_{ijk} \approx (\text{Density at } (a_i, b_j, c_k)) \cdot (\text{Volume of } B_{ijk}) \\
= \rho(a_i, b_j, c_k) \Delta x_i \Delta y_j \Delta z_k,
\]  

where \( \Delta x_i = x_{i+1} - x_i, \) and similarly for \( \Delta y_j \) and \( \Delta z_k. \) The sum of the masses of the boxes \( M_{ijk} \) approximates the total mass \( M \) of the tetrahedron. Each
4.3. TRIPLE INTEGRALS

$M_{ijk}$ can be approximated as in Equation 4.3.1, so we can approximate $M$ by

$$M \approx \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} M_{ijk} \approx \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \rho(a_i, b_j, c_k) \Delta x_i \Delta y_j \Delta z_k. \quad (4.3.2)$$

Since the approximation gets better as we use smaller and smaller boxes, we say

$$M = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \rho(a_i, b_j, c_k) \Delta x_i \Delta y_j \Delta z_k. \quad (4.3.3)$$

This limit is what we defined to be the triple integral, so we find that the integral of the density function is the mass of a solid. We summarize this discussion in the following:

**Calculating Mass from Density**

Let $W$ be a solid, and let the density at each point $(x, y, z)$ of $W$ be given by the function $\rho(x, y, z)$. The mass $M$ of $W$ is given by

$$M = \int \int \int_W \rho(x, y, z) \, dV.$$ 

Thus far we have introduced the notion of a triple integral and motivated the definition using a physical application. We now turn our attention to calculating triple integrals. It turns out that Fubini’s Theorem generalizes, and we can compute triple integrals via iterated integrals. Thus over the box $W = [a, b] \times [c, d] \times [e, f]$ we have

$$\int \int \int_W f(x, y, z) \, dV = \int_a^b \int_c^d \int_e^f f(x, y, z) \, dz \, dy \, dx.$$ 

**Example 4.3.3. Mass of a box**

The density of the box $W = [0, 2] \times [0, 3] \times [0, 1]$ is given by $\rho(x, y, z) = x^2 y + 3z$. Find the mass of $W$.

$$\int_0^2 \int_0^3 \int_0^1 x^2 y + 3z \, dz \, dy \, dx = \int_0^2 \int_0^3 x^2 y z + \frac{3}{2} \, dz \Big|_{z=0}^1 \, dy \, dx$$

$$= \int_0^2 \int_0^3 x^2 y + \frac{3}{2} \, dy \, dx = \int_0^2 \left( \frac{1}{2} y^2 + \frac{3}{2} y \right) \Big|_{y=0}^3 \, dx$$

$$= \int_0^2 \frac{9}{2} x^2 + \frac{9}{2} \, dx = \frac{3}{2} x^3 + \frac{9}{2} x \Big|_{x=0}^2 = 21. \, \text{▲}$$

**Solids from Limits:** Just as in double integrals, there are three features of triple integrals that are important to consider: the solid $W$ of integration, the integrand $f(x, y, z)$, and the differential (or volume element) $dV$. We have already seen that the meaning of the integrand gives meaning to the triple integral.
In particular, the integral of density is mass. We will see many interpretations of the integrand that yield other interesting integrals as we proceed. For now, however, we wish to focus on the relationship between the limits and solid of integration. We begin by sketching the solid of integration given the integral.

The region $R$ of integration in $\int \int_R f(x,y) \, dA$ is a rectangle when the limits of integration are all constants. Analogously, the solid $W$ of integration in $\int \int \int_W f(x,y,z) \, dV$ is a rectangular prism (a box) when all the limits are constant. One obtains the solid, as in the double integral setting, by first translating the limits into a system of inequalities.

**Example 4.3.4. Constant Limits**

Sketch the solid of integration for

$$\int_{-1}^{0} \int_{0}^{3} \int_{3}^{4} x^2y + 2z \, dx\, dy\, dz.$$

Since we only care about the solid of integration, we can ignore the integrand entirely. Recalling that the notation for iterated integrals is nested, so that the innermost limits correspond to the first variable of integration, we get the system of inequalities:

$$3 \leq x \leq 4$$
$$0 \leq y \leq 3$$
$$-1 \leq z \leq 0.$$ 

The box is pictured in Figure 4.3.2.

**Example 4.3.5. A generalized cylinder**

We now look at solids of integration where two sets of limits are constant and one set contains variable expressions. The variable expression in the first example only has the variable $y$, which makes the bounding surface a generalized cylinder. The second example has a variable expression in $x$ and $z$, making one bounding surface the graph of a function (a paraboloid in this case).
4.3. TRIPLE INTEGRALS

Sketch the solid of integration for
\[ \int_{-1}^{0} \int_{0}^{1} \int_{-y^2}^{0} f(x, y, z) \, dz \, dy \, dx. \]

We get the corresponding system of inequalities from the limits:
\[
\begin{align*}
0 & \leq z \leq 1 - y^2 \\
-1 & \leq y \leq 0 \\
0 & \leq x \leq 3.
\end{align*}
\]
The shadow of the solid \( W \) in the \( xy \)-plane, then, is the rectangle \( R = [0, 3] \times [-1, 0] \). The inequality \( 0 \leq z \leq 1 - y^2 \) indicates that the solid itself lies above the plane \( z = 0 \) and below the generalized cylinder \( z = 1 - y^2 \). The solid is pictured in Figure 4.3.2.

Example 4.3.6. The graph of a function

Sketch the solid of integration for
\[ \int_{-1}^{1} \int_{-1}^{1} \int_{0}^{4 - z^2 - x^2} f(x, y, z) \, dy \, dx \, dz. \]

We get the corresponding system of inequalities from the limits:
\[
\begin{align*}
0 & \leq y \leq 4 - z^2 - x^2 \\
-1 & \leq x \leq 1 \\
-1 & \leq z \leq 1.
\end{align*}
\]
The shadow of the solid \( W \) in the \( xz \)-plane, then, is the rectangle \( R = [-1, 1] \times [-1, 1] \). The inequality \( 0 \leq y \leq 4 - z^2 - x^2 \) indicates that the solid itself lies to the right of the plane \( y = 0 \) and to the left of the paraboloid \( y = 4 - z^2 - x^2 \). The solid is pictured in Figure 4.3.3.

![Graph General Solid](image)

Figure 4.3.3: Solid of integration from Limits

We finish the “limits to solids” section with an example in which two sets of limits have variable expressions, and only the last set is constants. This is, of course, the most general case.
Example 4.3.7. The General Case

Sketch the solid of integration for
\[ \int_{-1}^{1} \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \int_{0}^{6+2y-3z} f(x, y, z) \, dx \, dy \, dz. \]

We get the corresponding system of inequalities from the limits:
\[
0 \leq x \leq 6 + 2y - 3z \\
-\sqrt{1-z^2} \leq y \leq \sqrt{1-z^2} \\
-1 \leq z \leq 1.
\]

Since the last two inequalities describe the unit disk in the yz-plane, that is the shadow of the solid \( W \). The inequality \( 0 \leq x \leq 6 + 2y - 3z \) indicates that the solid itself lies in front of the plane \( x = 0 \) and behind the plane \( x = 6 + 2y - 3z \).

The solid is pictured in Figure 4.3.3.

Limits from solids: In the previous subsection we looked at finding solids of integration from given limits. We now reverse the process, and discuss finding limits of integration from given solids. This can be slightly challenging, simply because it’s harder to visualize solids in three dimensions. The goal, then, is to reduce the problem to two dimensions as quickly as possible. We make the convention that the \( x \)-axis is pointing out of the paper at you, the \( y \)-axis to your right, and the \( z \)-axis points straight up.

A \( dz \)-solid is one that has a nice top and bottom surface. When integrating over such a solid, you integrate with respect to \( z \) first. To find limits of integration, walk along a line parallel to the \( z \)-axis which pierces the solid. The equation for the surface where you enter the solid is the lower limit of integration and the equation for the surface where you leave is the upper limit (see Figure 4.3.4). Both equations must be solved for \( z \), since that is the variable you are integrating with respect to. Once you have the limits on \( z \), you immediately take the shadow \( R \) of the solid \( W \) in the \( xy \)-plane. Now follow double-integral techniques to find limits on \( x \) and \( y \).

A \( dy \)-solid has a nice left surface and right surface. To find limits, walk along a line parallel to the \( y \)-axis. The equation for the surface (solved for \( y \) of course) where you enter the solid is the lower limit, and where you leave is the upper limit (see Figure 4.3.8). A \( dx \)-solid is similar, but with nice front and back surfaces (see Figure 4.3.6).

Example 4.3.8. Finding limits for a \( dz \)-solid

Let \( W \) be the solid below the paraboloid \( x^2 + y^2 + z = 4 \) and above the \( xy \)-plane (i.e. the plane \( z = 0 \)). Describe the solid as a \( dz \)-solid.

A vertical line piercing the solid enters at the plane \( z = 0 \) and leaves at the paraboloid \( z = 4 - x^2 - y^2 \). Thus limits on \( z \) are \( 0 \leq z \leq 4 - x^2 - y^2 \). Once these have been determined, immediately reduce to a two-dimensional problem.
4.3. TRIPLE INTEGRALS

by taking the shadow $R$ of $W$ in the $xy$-plane. The boundary of the region $R$
can easily be seen to be the disk of radius 2 and centered at the origin. Using
techniques from the previous section we can describe the solid by the system of
inequalities

\begin{align*}
0 &\leq z \leq 4 - x^2 - y^2 \\
-\sqrt{4 - x^2} &\leq y \leq \sqrt{4 - x^2} \\
-2 &\leq x \leq 2.
\end{align*}

Figure 4.3.4: The Solid $W$

In more complicated examples one can sometimes find the boundary of the
region $R$ from the equations for the top and bottom surfaces. The equations for
the surfaces should already be solved for $z$, so setting them equal to each other
gives a single equation in just $x$ and $y$. This is the equation for the boundary of
$R$. Geometrically, setting the equations equal is like finding the intersection of
the two surfaces. Ignoring the $z$ variable altogether is equivalent to projecting
into the $xy$-plane. Thus the boundary of $R$ is the projection into the $xy$-plane
of the curve of intersection of the two surfaces.

Example 4.3.9. Finding limits for $dy$-solid.

Express the same solid as a $dy$-solid.

Now walk along a line parallel to the $y$-axis that pierces the solid. The
equation for the surface where you enter the solid is your lower limit on $y$, the
surface where you leave is your upper limit. This line enters and leaves at the
paraboloid $x^2 + y^2 + z = 4$. Since we’re integrating $dy$ first we solve for $y$,
yielding $y = \pm\sqrt{4 - z - x^2}$. The shadow in the $xz$-plane is the region $R$ under
the curve $z = 4 - x^2$. A line parallel to the $z$-axis enters $R$ at the curve $z = 0$
and leaves at $z = 4 - x^2$. The shadow on the $x$-axis is the interval $[-2, 2]$. Thus
the system of inequalities describing $W$ as a $dy$-solid are:
Example 4.3.10. A slight variation

Let $W$ be the solid between the same paraboloid $z = 4 - x^2 - y^2$ and the plane $2y + z = 1$. Describe $W$ as a $dz$-solid (see Figure 4.3.5).

A line parallel to the $z$-axis enters $W$ at the plane and leaves at the paraboloid. Thus limits on $z$ are $1 - 2y \leq z \leq 4 - x^2 - y^2$. We reduce to two-dimensions, and the equation for the boundary of the shadow is obtained by setting the two surface equations equal to each other. Thus we get the boundary curve

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

Thus the region $R$ is the disk radius 2 and centered at $(0, 1)$. Solving the circle equation for $x$, then taking the shadow on the $y$-axis gives that $W$ is described by:

$$1 - 2y \leq z \leq 4 - x^2 - y^2$$
$$-\sqrt{4 - (y - 1)^2} \leq x \leq \sqrt{4 - (y - 1)^2}$$
$$-1 \leq y \leq 3.$$
4.3. TRIPLE INTEGRALS

For the limits on \( x \), note that a line parallel to the \( x \)-axis enters the disk at the left semi-circle (the negative square root) and leaves at the right one (the positive square root).

We finish this example by noting that any integration problem with these limits would be rather messy. This example is just intended to illustrate how to set up the limits of integration. ▲

**Math App 4.3.1. Finding limits for a \( dx \)-solid**

In the following Math App, limits for a \( dx \)-solid are illustrated using animations. Click the hyperlink below to investigate the process.

![Image](image)

**Math App 4.3.2. Finding limits for a tetrahedron**

In the following Math App, limits for certain tetrahedra are found. In particular, we consider tetrahedra with vertices \((0, 0, 0)\), \((x, 0, 0)\), \((0, y, 0)\), and \((0, 0, z)\) where you specify values for the \( x \)-, \( y \)- and \( z \)-intercepts. Experiment with different values and see if you can predict the limits.

![Image](image)

**Some Examples**  Recall that the double integral of the constant function \( f(x, y) = 1 \) over the region \( R \) gives the area of \( R \). Symbolically we say: \( \iint_R dA = \text{Area of } R \). Analogously, the triple integral of \( f(x, y, z) = 1 \) over the solid \( W \) is the volume of \( W \), or

\[
\text{Volume of } W = \iiint_W dV.
\]

This can also be seen by assuming unit density, using the mass calculation in the box at the beginning of this section, and noting that for unit density the mass is the same magnitude as the volume (different units, of course).

**Example 4.3.11. Volume**
Find the volume of the solid bounded by the coordinate planes, the plane 
\( y + z = 1 \) and \( x + 2y + z = 6 \). Note that the plane \( y + z = 1 \) is a cylinder, since 
the equation is void of \( x \), and the plane \( x + 2y + z = 6 \) slices through the first 
octant since its intercepts are \((6, 0, 0), (0, 3, 0), \) and \((0, 0, 6)\). The solid bounded 
by the planes, then, is a \( dx \)-solid. A line parallel to the \( x \)-axis enters the solid 
at \( x = 0 \) and leaves at \( x = 6 - z - 2y \). The shadow of the solid in the \( yz \)-plane 
is the triangle bounded by the coordinate axes and the line \( y + z = 1 \). Thus we 
get the volume by calculating the iterated integral

\[
V = \int_0^1 \int_{1-y}^{1-y} \int_0^{6-z-2y} dx \, dz \, dy
\]

\[
= \int_0^1 \int_0^{1-y} \left. x \right|_{z=0}^{6-z-2y} \, dz \, dy = \int_0^1 \int_0^{1-y} 6 - z - 2y \, dz \, dy
\]

\[
= \int_0^1 (6 - 2y)z - \frac{z^2}{2} \bigg|_{z=0}^{1-y} \, dy = \int_0^1 11 - 7y + \frac{3}{2} y^2 \, dy
\]

\[
= \frac{11}{2}y - \frac{7}{2} y^2 + \frac{1}{2} y^3 \bigg|_0^1 = \frac{5}{2}. \quad \Box
\]

**Example 4.3.12. Solid made with a Generalized Cylinder**

Find the volume of the solid bounded by the generalized cylinder \( y = 4 - x^2 \), 
the \( xy \)-plane and the plane \( z = 2y \).

**Step 1.** Determine limits of integration.

The equation \( y = 4 - x^2 \) is missing a \( z \), so the surface is obtained by trans-
lating the parabola \( y = 4 - x^2 \) vertically. The planes then cut a wedge out of 
the generalized cylinder as in Figure 4.3.7. The plane \( z = 2y \) is the top, while 
the \( xy \)-plane \((z = 0)\) is the bottom of the wedge. This implies the limits on \( z \) 
are

\[ 0 \leq z \leq 2y. \]
Since we have limits for $z$, we take the shadow in the $xy$-plane to find limits on $x$ and $y$. The shadow is the region between the parabola $y = 4 - x^2$ and the $x$-axis. Thinking of this as a $dy$-region we get

$$
0 \leq y \leq 4 - x^2 \\
-2 \leq x \leq 2.
$$

Intersecting surfaces

The resulting wedge

Figure 4.3.7: Volume of a Wedge

Step 2. Integrate

The volume is given by

$$
V = \int_{-2}^{2} \int_{0}^{4-x^2} \int_{0}^{2y} dz dy dz = \int_{-2}^{2} \int_{0}^{2y} \left(\int_{0}^{4-x^2} dz\right) dy dx
$$

$$
= \int_{-2}^{2} \int_{0}^{4-x^2} 2y dy dx = \int_{-2}^{2} y^2\bigg|_{y=0}^{4-x^2} dx
$$

$$
= \int_{-2}^{2} (4 - x^2)^2 dx = \int_{-2}^{2} 16 - 8x^2 + x^4 dx
$$

$$
= 16x - \frac{8}{3}x^3 + \frac{1}{5}x^5\bigg|_{-2}^{2} = \frac{512}{15}. \Box
$$

Example 4.3.13. Mass of the tetrahedron

Calculate the mass of the tetrahedron of example 4.2.6 if the density is given by $\rho(x, y, z) = 40 - z^2$.

We need to evaluate $\iiint_W \rho(x, y, z) \, dV$.

Step 1. Determine the limits of integration
Considering the tetrahedron to be a $dz$-solid, we walk through it parallel the the $z$-axis. The equation for the surface where we enter the solid is the lower limit and where we leave is the upper limit (the equations must be solved for $z$). We enter at the $xy$-plane, or $z = 0$, and leave at the plane $3x + 2y + z = 6$. Solving for $z$ gives

$$0 \leq z \leq 6 - 3x - 2y.$$  

After finding limits on $z$, we find limits on the projection in the $xy$-plane. Notice that we reduce the number of dimensions as soon as we take care of limits on $z$. In this case, the projection is the triangle in the first quadrant bounded by the line $3x + 2y = 6$. Thinking of this as a $dy$-region in the plane gives the limits

$$0 \leq y \leq (6 - 3x)/2$$
$$0 \leq x \leq 2.$$  

**Step 2. Integrate**

We see that the mass is given by

$$m = \int_0^2 \int_0^{(6-3x)/2} \int_0^{6-3x-2y} 40 - z^2 \, dz \, dy \, dx$$

$$= \int_0^2 \int_0^{(6-3x)/2} 40z - \frac{z^3}{3} \bigg|^{6-3x-2y}_{z=0} \, dy \, dx$$

$$= \int_0^2 \int_0^{(6-3x)/2} 40(6 - 3x - 2y) - \frac{(6 - 3x - 2y)^3}{3} \, dy \, dx$$

$$= \int_0^2 -10(6 - 3x - 2y)^2 + \frac{(6 - 3x - 2y)^4}{24} \bigg|_{y=0}^{(6-3x)/2} \, dx$$

$$= \int_0^2 10(6 - 3x)^2 - \frac{(6 - 3x)^4}{24} \, dx$$

$$= -\frac{10(6 - 3x)^3}{9} + \frac{(6 - 3x)^5}{360} \bigg|_0^2 = \frac{1092}{5} = 218.4 \ gm. \ ▲$$

When working with double integrals, we encountered problems in which we switched the limits of integration. To do so, we had to sketch the region of integration given the limits. Similar problems can arise with triple integrals.

**Example 4.3.14. The Solid of Integration**

Sketch the solid of integration corresponding to the following iterated integral, then evaluate the integral.

$$\int_1^0 \int_{\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+z^2}^{2-x^2-z^2} dy \, dz \, dx.$$  

The first limits indicate that $x^2 + z^2 \leq y \leq 2 - x^2 - z^2$. Thus $y$ is between the paraboloids $y = x^2 + z^2$ and $y = 2 - x^2 - z^2$. The last two sets of limits are for
the unit circle in the $xz$-plane, so the shadow of the solid lies inside $x^2 + z^2 = 1$. Since the curve of intersection of the paraboloids is the circle $x^2 + z^2 = 1$ lying in the plane $y = 1$, the solid of integration is the intersection of the paraboloids $y = x^2 + z^2$ and $y = 2 - x^2 - z^2$ (see Figure 4.3.8). The triple integral gives the volume, and by symmetry it is

$$
\int_{-1}^{1} \int_{\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+z^2}^{2-x^2-z^2} dydzdx = 4 \int_{0}^{1} \int_{0}^{\sqrt{1-x^2}} \int_{x^2+z^2}^{2-x^2-z^2} dydzdx
$$

$$
= 4 \int_{0}^{1} \int_{0}^{\sqrt{1-x^2}} (2 - 2x^2 - 2z^2) dzdx
$$

$$
= 4 \int_{0}^{1} 2 \left(1 - x^2\right) z - \frac{2}{3} z^3 \right|_{z=0}^{\sqrt{1-x^2}} dx
$$

$$
= 4 \int_{0}^{1} \frac{4}{3} \left(1 - x^2\right)^{3/2} dx = \frac{16}{3} \int_{0}^{1} \left(1 - x^2\right)^{3/2} dx
$$

Now make the trigonometric substitution $x = \sin \theta$, $dx = \cos \theta d\theta$, and change the limits of integration to get
\[
\int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+z^2}^{2-x^2-z^2} dydzdx = \frac{16}{3} \int_{0}^{1} (1-x^2)^{3/2} \, dx \\
= \frac{16}{3} \int_{0}^{\pi/2} (\cos^2 \theta)^{3/2} \, \cos \theta d\theta \\
= \frac{16}{3} \int_{0}^{\pi/2} \cos^4 \theta d\theta \\
= \frac{16}{3} \int_{0}^{\pi/2} \left( \frac{1}{2} (1 + \cos 2\theta) \right)^2 \, d\theta 
\]

where the last equality follows from the half-angle formula \( \cos^2 \theta = \frac{1}{2} (1 + \cos 2\theta) \).

Expanding the integrand, and another application of the half-angle formula gives

\[
\int_{-1}^{1} \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{x^2+z^2}^{2-x^2-z^2} dydzdx = \frac{16}{3} \int_{0}^{\pi/2} \left( \frac{1}{2} (1 + \cos 2\theta) \right)^2 \, d\theta \\
= \frac{4}{3} \int_{0}^{\pi/2} 1 + \cos 2\theta + \cos^2 2\theta \, d\theta \\
= \frac{4}{3} \int_{0}^{\pi/2} 1 + \cos 2\theta + \frac{1}{2} (1 + \cos 4\theta) \, d\theta \\
= \frac{4}{3} \int_{0}^{\pi/2} \frac{3}{2} + \cos 2\theta + \frac{1}{2} \cos 4\theta \, d\theta \\
= \frac{4}{3} \left( \frac{3}{2} \theta + \sin 2\theta + \frac{1}{8} \sin 4\theta \right)_{0}^{\pi/2} = \pi. \quad \blacksquare
\]

This was a fairly complicated integral. In the next section we’ll see a much easier way to evaluate it. Stay tuned!

**Exercises**

1. The vertices of tetrahedra are given below. Express each as a \( dx-\), \( dy-\), and \( dz-\) solid.
   - (a) \((0,0,0), (-2,0,0), (0,1,0), (0,0,-3)\).
   - (b) \((0,0,0), (1,0,0), (0,1,0), (0,0,1)\).
   - (c) \((-1,0,0), (1,0,0), (0,1,0), (0,0,1)\).

2. Express the tetrahedron in the first octant bounded by \( x + 5y + 2z = 10 \) as a \( dy-\) solid.

3. Express the tetrahedron in the first octant bounded by \( 7x + y + 3z = 21 \) as a \( dx-\) solid.
4. Express the tetrahedron in the first octant bounded by \( 2x + 3y + z = 6 \) as a \( dz \)-solid.

In exercises 5-17, express each solid as a \( dx \)-, \( dy \)-, or \( dz \)-solid. Choose the most convenient direction.

5. The solid inside the cone \( z = \sqrt{x^2 + y^2} \) and under the plane \( z = 3 \).

6. The solid between the paraboloid \( z = 4 - x^2 - y^2 \) and the plane \( 2x + z = 4 \).

7. The solid inside \( x^2 + z^2 = 4 \) and between \( y = 0 \) and \( 2y + z = 2 \).

8. The solid inside \( x^2 + y^2 = 9 \) and between the \( xy \)-plane and the plane \( 9x + 4y + 6z = 36 \).

9. The solid bounded by the \( xy \)-plane, the cylinder \( x^2 + y^2 = 4 \), and the plane \( 5x + 10y + z = 50 \).

10. The solid bounded by the \( yz \)-plane, the cylinder \( z^2 + y^2 = 1 \), and the plane \( x + y + z = 4 \).

11. The solid bounded by the \( xz \)-plane, the cylinder \( x^2 + z^2 = 4 \), and the plane \( x - 2y + z = 6 \).

12. The solid bounded by the \( xz \)-plane and the paraboloid \( y = x^2 + z^2 - 4 \).

13. The solid bounded by the \( xz \)-plane and the paraboloid \( y = 4 - x^2 - z^2 \).

14. The solid bounded by the \( x^2 + y^2 + z = 0 \) and the plane \( 2x + 2y - z = 0 \).

15. The solid inside the sphere \( x^2 + y^2 + z^2 = 1 \) and above the cone \( x^2 + y^2 - z^2 = 0 \).

16. The solid inside the sphere \( x^2 + y^2 + z^2 = 1 \) and to the right of the plane \( y = 1/2 \).

17. The solid inside the sphere \( x^2 + y^2 + z^2 = 1 \) and below the plane \( z = -1/2 \).

18. Let \( T \) be the tetrahedron bounded by the coordinate planes and \( 3x + y + 5z = 15 \). Evaluate \( \iiint_T \, dV \).

19. Sketch the solid of integration, then evaluate the following triple integrals:

   (a) \( \int_0^2 \int_0^{2-y} \int_0^{2-z-y} \, dx \, dz \, dy \)

   (b) \( \int_0^1 \int_{4x}^1 \int_{\sqrt{4x^2 - 4x}} \, dy \, dx \)

   (c) \( \int_0^1 \int_x^{\sqrt{4z}} \int_{\sqrt{4z}} \, y \, dy \, dz \)

20. Find the volume of the solid bounded by the coordinate planes, the plane \( x + 2y = 4 \) and the plane \( x + y + z = 5 \).
21. Find the volume of the solid inside the parabolic cylinder $z = x^2$ and the planes $y = 0$ and $z + y = 1$.

22. Let $W$ be the solid in the first octant which is inside the paraboloid $z = x^2 + y^2$ and under the plane $z = 1$. Find the mass of $W$ if the density is given by $\delta(x, y, z) = y$.

23. Compute the volume of the solid inside the cylinders $y^2 + z^2 = 1$ and $x^2 + z^2 = 1$. 
4.4 Integration in Different Coordinate Systems

Thus far we have discussed double and triple integrals in Cartesian coordinates, and found that there are three features of multiple integrals to consider when setting up and evaluating them: the domain of integration, the integrand, and the differential. In this section we’ll consider multiple integrals in other coordinate systems, and how each of these features are handled. We’ll begin by studying double integrals in polar coordinates, then move on to treat triple integrals in cylindrical and spherical coordinates in turn. You might want to review the other coordinate systems in Section 1.1 as familiarity with them is essential for this section. We will also find that the area form \( dA \) and volume form \( dV \) are more subtle and significant than you might think! Without further ado, let’s begin.

**Polar Double Integrals:** We begin by comparing Cartesian and polar coordinates with regard to limits of integration. We’ve seen that finding limits of integration in multiple integrals can be a challenging part of the problem, and the limits themselves can make integration fairly complicated. Consider example 4.3.14 from the previous section. The shadow in the \( xz \)-plane was the unit disk \( x^2 + z^2 \leq 1 \), so limits in Cartesian coordinates were

\[
-\sqrt{1-x^2} \leq z \leq \sqrt{1-x^2} \\
-1 \leq x \leq 1.
\]

These limits, in turn, made the integration difficult, requiring a trigonometric substitution. The region of integration, however, has a very simple description using polar coordinates in the \( xz \)-plane, where \( r \) is the distance to the origin in the \( xz \)-plane and \( \theta \) the angle with the positive \( x \)-axis. The unit disk is all points within one unit of the origin (regardless of the angle \( \theta \)), so the following system of inequalities describes the region in polar coordinates:

\[
0 \leq r \leq 1 \\
0 \leq \theta \leq 2\pi.
\]

In this coordinate system the limits of integration are constants, as opposed to unpleasant square roots. Rather than integrating with respect to the Cartesian coordinates, then, we’d rather integrate with respect to polar. In this section we describe how to change coordinates in multiple integration.

There are three things that need to change when translating a Cartesian integral to one in another coordinate system. You must change the limits of integration, make appropriate substitutions in the integrand, and change the area form \( dA \) for double integrals or the volume form \( dV \) for triple integrals. We introduce the mechanics of the process with some examples, then discuss why it works.

**Example 4.4.1. Double Integral in polar coordinates**
Integrate the function \( f(x, y) = x^2 \) over the portion \( R \) of the unit disk in the first quadrant. In Cartesian coordinates we have

\[
\iint_R f(x, y) \, dA = \int_0^1 \int_0^{\sqrt{1-x^2}} x^2 \, dy \, dx = \int_0^1 \int_0^{\sqrt{1-y^2}} x^2 \, dx \, dy.
\]

Either order of integration gets messy because of the square root.

**Step 1.** Limits in Polar Coordinates.

Using polar coordinates, we see the quarter circle \( R \) is described by the system of inequalities

\[
0 \leq r \leq 1 \\
0 \leq \theta \leq \frac{\pi}{2}.
\]

**Step 2.** Integrand in Polar Coordinates.

Since we're working in polar coordinates, we have to substitute \( x = r \cos \theta \) and \( y = r \sin \theta \) in the integrand \( f(x, y) = x^2 \). This gives the integrand

\[
f(r \cos \theta, r \sin \theta) = r^2 \cos^2 \theta.
\]

**Step 3.** \( dA \) in Polar Coordinates.

One final change is required, and it may be surprising at first glance. Historically it represented an infinitesimal area, coming from the limit in the partitioning process of the Riemann sum. We’ll go deeper into the description later, for now it suffices to say that in polar coordinates

\[
dA = r \, dr \, d\theta.
\]

Do not forget the “extra” \( r \) that pops up out of nowhere! Using
• The limits we found on polar coordinates,
• Our integrand translated into polar coordinates, and
• The polar version of $dA$,

we have

$$
\iint_{R} f(x, y) \, dA = \int_{0}^{\pi/2} \int_{0}^{1} r^2 \cos^2 \theta \, r \, dr \, d\theta = \int_{0}^{1} \int_{0}^{1} r^3 \cos^2 \theta \, dr \, d\theta \\
= \int_{0}^{\pi/2} \cos^2 \theta \left[ \frac{r^4}{4} \right]_{r=0}^{1} \, d\theta = \frac{1}{4} \int_{0}^{\pi/2} \cos^2 \theta \, d\theta \\
= \frac{1}{8} \int_{0}^{\pi/2} 1 + \cos 2\theta \, d\theta = \frac{1}{8} \left( \theta + \frac{\sin 2\theta}{2} \right) \bigg|_{0}^{\pi/2} = \frac{\pi}{16}.
$$

This example illustrates the essential features for evaluating integrals in polar coordinates. We summarize the process below. There are analogous steps for integrating in cylindrical and spherical as well.

**Double Integrals in Polar Coordinates**

To evaluate $\iint_{R} f(x, y) \, dA$ using polar coordinates:

1. **Limits**: Find limits for $R$ in terms of $r$ and $\theta$.
2. **Integrand**: Evaluate and simplify $f(r \cos \theta, r \sin \theta)$.
3. **Integrate**: Let $dA = r \, dr \, d\theta$ and integrate.

**Math App 4.4.1. Constant Limits of Integration**

When using Cartesian coordinates we found the region of integration for $\int_{a}^{b} \int_{c}^{d} f(x, y) \, dx \, dy$ is the rectangle $R = [a, b] \times [c, d]$. One wonders what types of regions arise in polar coordinates when the limits are constants, i.e. when evaluating $\int_{a}^{b} \int_{c}^{d} f(r \cos \theta, r \sin \theta) \, r \, dr \, d\theta$. Click the following hyperlink to investigate the constant limit setting in polar coordinates.
Another situation in which polar integrals are useful is when the region of integration is given by polar equations. Here is such an example.

**Example 4.4.2. Regions given by polar equations**

Find the area inside the cardioid \( r = 1 + \cos \theta \).

\[ \begin{align*}
\text{Step 1. Polar Limits.} \\
\text{To find limits of integration on } r, \text{ sketch a ray eminating from the origin that intersects the region. The equation for the curve where it enters your region is the lower limit, and the equation for the curve where your ray leaves the region is your upper limit (both equations should be solved for } r \). Limits on } \theta \text{ are those angles for which the ray intersects the region. In this case, the region is described by} \\
0 \leq r \leq 1 + \cos \theta \\
0 \leq \theta \leq 2\pi.
\end{align*} \]

\[ \begin{align*}
\text{Step 2. Polar Integrand.} \\
The integrand is } f(x, y) = 1 \text{ since we’re trying to find area, which is the same in polar coordinates.}
\end{align*} \]

\[ \begin{align*}
\text{Step 3. Polar } dA. \\
\text{Since } dA = r \, dr \, d\theta, \text{ the area enclosed by the cardioid is given by} \\
\int \int_R dA = \int_0^{2\pi} \int_0^{1+\cos \theta} r \, dr \, d\theta = \int_0^{2\pi} \frac{r^2}{2} \bigg|_{r=0}^{1+\cos \theta} \, d\theta \\
= \frac{1}{2} \int_0^{2\pi} (1 + 2 \cos \theta + \cos^2 \theta) \, d\theta \\
= \frac{1}{2} \int_0^{2\pi} \left( \frac{3}{2} + 2 \cos \theta + \frac{\cos 2\theta}{2} \right) \, d\theta \\
= \frac{1}{2} \left[ \frac{3}{2} \theta + 2 \sin \theta + \frac{\sin 2\theta}{4} \right]_0^{2\pi} = \frac{3\pi}{2}. \quad \blacktriangle
\end{align*} \]
This is the first example where we used functions to find limits of integration in polar coordinates. For your reference we summarize the process in the following box.

**Finding Limits in Polar Coordinates**

- Limits on \( r \): Draw a ray emanating from the origin that intersects the region \( R \). The equation for the curve where it enters \( R \) is your lower limit and the equation for the curve where it leaves \( R \) is your upper limit (both equations solved for \( r \) of course).
- Limits on \( \theta \): Determine for which angles \( \theta \) do the rays go through \( R \).

**Example 4.4.3. Limits in polar coordinates**

Let \( R \) be the region inside the unit circle and outside the cardioid \( r = 1 - \sin \theta \).

Describe \( R \) using a system of inequalities in polar coordinates.

\[
1 - \sin \theta \leq r \leq 1 \quad 0 \leq \theta \leq \pi. \quad \blacktriangleleft
\]

**Draw a ray from the origin that passes through \( R \), as in Figure 4.4.3.** It enters the region at the cardioid \( r = 1 - \sin \theta \) and leaves at the unit circle \( r = 1 \). Moreover, rays pass through \( R \) for angles \( \theta \) between 0 and \( \pi \). Thus \( R \) is described by the system

\[
1 - \sin \theta \leq r \leq 1 \\
0 \leq \theta \leq \pi. \quad \blacktriangleleft
\]

**Justification that \( dA = rdrd\theta \):** We now know the mechanics of integration in polar coordinates, so let’s look at where the extra \( r \) comes from in the area form. Recall that to define the double integral over a rectangle \( R \), we started by partitioning \( R \) into small sub-rectangles. These came from using partitions in the \( x \) and \( y \) directions. Alternatively, one could use the \( r \) and \( \theta \) directions to partition the region as in Figure 4.4.4. The result is a partitioning of the region into portions of circular sectors.
To integrate a function $f(x, y)$ in polar coordinates, then we approximate the integral $\iint_R f(x, y) \, dA$ by:

1. Using the above partition, find the area of each sub-sector. To do this, recall that the area that an angle $\Delta \theta$ cuts out of a circle radius $r$ is given by $\frac{\Delta \theta}{2} \pi r^2 = \frac{\Delta \theta}{2} r^2$. This follows from the fact that the ratio of the areas of the sector to the circle is the same as the ratio of the angle to $2\pi$. Now let the inner radius be $r$ and $\Delta r$ be the change in radius, so the outer radius is $r + \Delta r$. Then the area of the sub-sector is the outer area minus the inner:

$$\frac{\Delta \theta}{2} (r + \Delta r)^2 - \frac{\Delta \theta}{2} r^2 = \frac{\Delta \theta}{2} (2r + \Delta r) \Delta r = \left( r + \frac{\Delta r}{2} \right) r \Delta \theta = r_a \Delta r \Delta \theta$$

where $r_a$ is the average of the two boundary radii.

2. Pick a point $(r_i \cos \theta_j, r_i \sin \theta_j)$ in each subsector, and evaluate $f$ there. It turns out the limit is independent of choice here, so we make the choice that $r_i = r_a$. For each sub-sector we get the contribution

$$f(r_i \cos \theta_j, r_i \sin \theta_j) r_i \Delta r_i \Delta \theta_j$$

to the Riemann sum.

3. Approximate the integral with the double-sum of the contributions from each sector

$$\iint_R f(x, y) \, dA \approx \sum_{i=1}^{n} \sum_{j=1}^{n} f(r_i \cos \theta_j, r_i \sin \theta_j) r_i \Delta r_i \Delta \theta_j.$$
4. Take a limit as the partition gets finer to find the integral. Notice that the Riemann sum in the previous step is formally the same as our original double integral, but where the function is \( f(r \cos \theta, r \sin \theta) r \) followed by the change in variables. Therefore in the limit we get the extra \( r \):

\[
\iint_R f(x, y) \, dA = \int_a^b \int_c^d f(r \cos \theta, r \sin \theta) \, rdrd\theta.
\]

Thus we’ve seen the “extra” \( r \) in the polar differential \( dA \) arises naturally when calculating areas of circular sectors, and in fact are a necessary component of polar integration.

**Cylindrical triple integrals:** We now introduce triple integrals in cylindrical and spherical coordinates. The techniques are similar to double integrals in polar coordinates, with minor modifications. We summarize them in the following boxes.

**Triple Integrals in Cylindrical Coordinates**

To evaluate \( \iiint_W f(x, y, z) \, dV \) using cylindrical coordinates:

1. Find limits for \( W \) in terms of \( z, r \) and \( \theta \).
2. Evaluate and simplify \( f(r \cos \theta, r \sin \theta, z) \).
3. Let \( dV = r \, dz \, dr \, d\theta \) and integrate.

**Math App 4.4.2. Constant Limits in Cylindrical Coordinates**

You may have noticed that the previous Math App on constant limits in polar coordinates had cylindrical and spherical as well. Feel free to review the solids you obtain in those coordinate systems now.

**Example 4.4.4. Mass of a cone**

Find the mass of the cone below \( z = 2 - \sqrt{x^2 + y^2} \) and above the \( xy \)-plane if the density is given by \( \delta(x, y, z) = 3 - z \) (see Figure 4.4.5).
Before we begin the process, note that the equation of the cone in cylindrical coordinates is \( z = 2 - r \), which is considerably nicer than its Cartesian equation. This is one hint that using cylindrical coordinates might calculations easier.

**Step 1.** Cylindrical Limits.

We find limits of integration on \( z \) first, then project into the \( xy \)-plane and use polar coordinates there. The cone is above the plane \( z = 0 \) and below the surface \( z = 2 - r \), so limits on \( z \) are

\[
0 \leq z \leq 2 - r.
\]

Once we find limits on \( z \), project into the \( xy \)-plane and use polar coordinates. The projection into the \( xy \)-plane is the disk radius 2, centered at the origin. Using polar coordinates, any point \((r, \theta)\) in the region has \( r \) at most 2 and no restriction on \( \theta \), so the disk is described by the system of inequalities

\[
0 \leq r \leq 2,
0 \leq \theta \leq 2\pi.
\]

**Step 2.** Cylindrical Integrand.

We must translate the integrand into cylindrical coordinates. Since the integrand is \( \delta(x, y, a) = 3 - z \), and \( z \) is a cylindrical coordinate, there is no change required.

**Step 3.** Cylindrical \( dV \).

Now that we have the limits of integration, and the integrand, we use the appropriate volume form and integrate.

\[
\iiint_W \delta \, dV = \int_0^{2\pi} \int_0^2 \int_0^{2-r} (3 - z)r \, dz \, dr \, d\theta
= \int_0^{2\pi} \int_0^2 3rz - \frac{r}{2}z^2 \bigg|_{z=0}^{2-r} \, dr \, d\theta
= \int_0^{2\pi} \int_0^2 4r - r^2 - \frac{1}{2}r^3 \, dr \, d\theta
= \int_0^{2\pi} \left( \frac{10}{3}r^2 - \frac{1}{8}r^4 \right)_{r=0}^{2} \, d\theta
= \int_0^{2\pi} \frac{20\pi}{3} \, d\theta
= \frac{20\pi}{3}. \quad \blacksquare
\]
Remark: In cylindrical coordinates we usually use polar coordinates in the $xy$-plane, together with the Cartesian coordinate $z$. There is nothing special, however, about which coordinate is Cartesian and which two are polar. We could just as easily have used polar coordinates in the $yz$-plane and $x$ for the Cartesian. The appropriate adjustment in the volume form would be $dV = r dx dr d\theta$. The following example uses $y$ as the Cartesian coordinate.

**Example 4.4.5. Paraboloids revisited**

Recall Example 4.3.14 where we calculated the volume between the paraboloids $x^2 + y + z^2 = 2$ and $y = x^2 + z^2$. The integration was fairly complex, and will be greatly simplified by using cylindrical coordinates $r$, $\theta$, $y$. Notice that we’re not using $z$ as the third axis, but we’re using polar coordinates in the $xz$-plane. Thus we’ll let $x = r \cos \theta$, $y = y$ and $z = r \sin \theta$.

**Step 0. Cylindrical Surface Equations.**

Before starting the problem, let’s get all the equations in the appropriate coordinates. Straightforward substitutions give the equations $y = 2 - r^2$ and $y = r^2$.

**Step 1. Cylindrical Limits.**

To find limits of integration, first find limits on $y$, then project into the $xz$-plane as before. In cylindrical coordinates we see $r^2 \leq y \leq 2 - r^2$. The shadow in the $xz$-plane is still the unit disk so limits are $0 \leq r \leq 1$ and $0 \leq \theta \leq 2\pi$.

**Step 2. Cylindrical Integrand.**

The integrand is the constant one, in any coordinate system, since we are finding volume.

**Step 3. Cylindrical $dV$.**

The volume form is $dV = r \, dy \, dr \, d\theta$, so the volume is given by

\[
\int_{W} \int_{r_0}^{r_1} \int_{\theta_0}^{\theta_1} \, r \, dy \, dr \, d\theta
\]

\[
= \int_{r_0}^{r_1} \int_{\theta_0}^{\theta_1} \left. r y \right|_{y=r^2}^{2-r^2} \, dr \, d\theta
= \int_{r_0}^{r_1} \int_{\theta_0}^{\theta_1} 2r - 2r^3 \, dr \, d\theta
= \int_{r_0}^{r_1} \int_{\theta_0}^{\theta_1} 2 \frac{r^2}{2} \, d\theta = \pi.
\]

Notice how much easier cylindrical coordinates were than Cartesian in this problem. ▲

**Example 4.4.6. A cylinder in the $x$-direction**

Find the volume of the solid inside the cylinder $y^2 + z^2 = 4$ and between the $yz$-plane and $x + y + z = 10$.

**Step 0. Cylindrical Surface Equations.**

The cylinder is the circle radius 2 in the $yz$-plane translated along the $x$-axis. The plane $x + y + z = 10$ cuts the cylinder in front of the $yz$-plane as in Figure 4.16.
Using cylindrical coordinates in the $yz$-plane, letting $y = r \cos \theta$ and $z = r \sin \theta$ and $x$ be the coordinate that remains Cartesian, we can translate the equations for the bounding surfaces into cylindrical coordinates. The cylinder $y^2 + z^2 = 4$ becomes $r = 2$, the $yz$-plane remains $x = 0$ while the slanted plane becomes $x = 10 - r \cos \theta - r \sin \theta$.

**Step 1. Cylindrical Limits.**

The solid has a nice front and back surface, so it is a $dx$-solid. The back surface is $x = 0$, while the front is the slanted plane, yielding the limits on $x$:

$$0 \leq x \leq 10 - r \cos \theta - r \sin \theta.$$  

The projection of the solid in the $yz$-plane is the disk centered at the origin with radius 2. Using polar coordinates in the $yz$-plane, the disk can be described by $0 \leq r \leq 2$, $0 \leq \theta \leq 2\pi$. Now we have the system

$$
0 \leq x \leq 10 - r \cos \theta - r \sin \theta \\
0 \leq r \leq 2 \\
0 \leq \theta \leq 2\pi
$$

**Step 2. Cylindrical Integrand.**

We want volume, so the integrand is the constant function 1, which is the same in any coordinate system.

**Step 3. Cylindrical $dV$.**
Using the volume form $dV = r dx dr d\theta$, we calculate

$$
\iiint_W dV = \int_0^{2\pi} \int_0^2 \int_0^{10-\rho \cos \theta - \rho \sin \theta} \rho \ dx \ dr \ d\theta
$$

$$
= \int_0^{2\pi} \int_0^2 \rho x \bigg|_{x=0}^{10-\rho \cos \theta - \rho \sin \theta} \ dr \ d\theta
$$

$$
= \int_0^{2\pi} \int_0^2 10r - \rho^2 \cos \theta - \rho^2 \sin \theta \ dr \ d\theta
$$

$$
= \int_0^{2\pi} \int_0^2 5\rho^2 - \rho^3 \cos \theta - \rho^3 \sin \theta \bigg|_{r=0}^2 \ dr \ d\theta
$$

$$
= \int_0^{2\pi} 20 - \frac{8}{3} \cos \theta - \frac{8}{3} \sin \theta \ d\theta
$$

$$
= 20\theta - \frac{8}{3} \sin \theta + \frac{8}{3} \cos \theta \bigg|_0^{2\pi} = 40\pi. \ 
\text{▲}
$$

**Spherical Triple Integrals**

We’ve investigated several examples of integration in polar and cylindrical coordinates, and it is time to turn our attention to spherical coordinates. As in the previous cases, it turns out that certain solids of integration, and certain integrands, are best described using spherical coordinates. Moreover, the steps to integrating in spherical coordinates are very similar to those we’ve already explored.

**Triple Integrals in Spherical Coordinates**

To evaluate $\iiint_W f(x, y, z) \ dV$ using spherical coordinates:

1. Find limits for $W$ in terms of $\rho$, $\phi$ and $\theta$.
2. Evaluate and simplify $f(\rho \sin \phi \cos \theta, \rho \sin \phi \sin \theta, \rho \cos \phi)$.
3. Let $dV = \rho^2 \sin \phi \ d\rho \ d\phi \ d\theta$ and integrate.

**Math App 4.4.3. Spherical limits of integration**

In this Math App we explore finding limits of integration in spherical coordinates. We have to split the limits into two sets of inequalities because the surfaces where rays from the origin leave the solid change for different values of $\phi$. 

Example 4.4.7. Volume of a snowcone

Find the volume of the region above the cone \( z^2 = x^2 + y^2 \) and inside the unit sphere (see Figure 4.4.7).

**Step 0.** Spherical Surface Equations.

Recall that the cone has spherical equation \( \phi = \frac{\pi}{4} \) and the sphere \( \rho = 1 \). The cone equation can be obtained making substitutions \( x = \rho \sin \phi \cos \theta, \ y = \rho \sin \phi \sin \theta, \ z = \rho \cos \phi \), and simplifying, but that is rather laborious. Hopefully you remember that the cone is a constant coordinate surface, as described in Section 1.1.

**Step 1.** Spherical Limits.

To find limits on \( \rho \), draw a ray emanating from the origin that intersects your solid. The equation for the surface where it enters the solid is your lower limit, and the equation for the surface where it leaves is your upper limit. In our case, the ray starts in the solid and leaves at \( \rho = 1 \), so limits are \( 0 \leq \rho \leq 1 \).

To find limits on \( \phi \), start with the positive z-axis \( \phi = 0 \), and let the ray drop until it reaches the boundary of your solid. In our case, the ray starts in our solid, and leaves at \( \phi = \frac{\pi}{4} \). Thus limits are \( 0 \leq \phi \leq \frac{\pi}{4} \). Finally, ask for which angles \( \theta \) do the rays intersect the solid, leaving \( 0 \leq \theta \leq 2\pi \).

**Step 2.** Spherical Integrand.

Since volume is obtained by integrating the constant function \( f(x, y, z) = 1 \), the integrand is unchanged by switching to spherical coordinates.

**Step 3.** Spherical \( dV \).
The volume is given by

\[
\iiint_W dV = \int_0^{2\pi} \int_0^{\pi/4} \int_0^1 \rho^2 \sin \phi \ d\rho d\phi d\theta
\]

\[
= \int_0^{2\pi} \int_0^{\pi/4} \left[ \frac{\rho^3}{3} \sin \phi \right]_0^1 \ d\phi d\theta = \int_0^{2\pi} \int_0^{\pi/4} \frac{1}{3} \sin \phi \ d\phi d\theta
\]

\[
= \int_0^{2\pi} \left[ -\frac{1}{3} \cos \phi \right]_0^{\pi/4} \ d\theta = \int_0^{2\pi} \frac{2 - \sqrt{2}}{6} \ d\theta = \left(2 - \sqrt{2}\right)\pi. \ □
\]

**Example 4.4.8.** Another volume example

Find the volume of the solid inside \( \rho = 1 - \cos \phi \), for \( 0 \leq \phi \leq \pi \) (see Figure 4.4.8).

![Figure 4.4.8: The surface \( \rho = 1 - \cos \phi \)](image)

**Step 1.** Spherical Limits.

A ray emanating from the origin enters the solid at \( \rho = 0 \) and leaves at the surface \( \rho = 1 - \cos \phi \). The limits for \( \phi \) are given, and there is no restriction on \( \theta \). Thus the solid is described by

\[
0 \leq \rho \leq 1 - \cos \phi
\]

\[
0 \leq \phi \leq \pi
\]

\[
0 \leq \theta \leq 2\pi.
\]

**Step 2.** Spherical Integrand.

We want volume, so the integrand is 1.

**Step 3.** Spherical \( dV \).
Letting \( dV = \rho^2 \sin \phi \, d\rho \, d\phi \, d\theta \) we get
\[
\iiint_W dV = \int_0^{2\pi} \int_0^\pi \int_0^{1-\cos \phi} \rho^2 \sin \phi \, d\rho \, d\phi \, d\theta
= \int_0^{2\pi} \int_0^\pi \rho^3 \frac{1-\cos \phi}{3} \sin \phi \, d\phi \, d\theta
= \int_0^{2\pi} \int_0^\pi \frac{(1-\cos \phi)^3}{3} \sin \phi \, d\phi \, d\theta.
\]

Using the substitution \( u = 1 - \cos \phi, \, du = \sin \phi \, d\phi \), and changing the limits of integration gives
\[
\iiint_W dV = \int_0^{2\pi} \int_0^\pi \frac{u^3}{3} \, dud\theta = \int_0^{2\pi} \frac{u^4}{12} \, d\theta = \int_0^{2\pi} \frac{4}{3} \, d\theta = \frac{8\pi}{3}. \quad \blacksquare
\]

**Example 4.4.9. Mass with Variable Density**

Determine the mass of the unit ball \( B^3 \) if the density (in \( gm/cm^3 \)) is
\[
\delta(x, y, z) = \frac{1}{1 + x^2 + y^2 + z^2}.
\]

**Step 1. Spherical Limits.**

Find limits of integration in spherical coordinates.

In this case, the solid of integration is the unit ball in \( \mathbb{R}^3 \), which is described using spherical coordinates by the system of inequalities
\[
0 \leq \rho \leq 1
0 \leq \phi \leq \pi
0 \leq \theta \leq 2\pi.
\]

**Step 2. Spherical Integrand.**

To find mass, we integrate density, so we must translate the density function into spherical coordinates. Since \( \rho^2 = x^2 + y^2 + z^2 \), the density function is
\[
\delta = \frac{1}{1 + \rho^2}.
\]

**Step 3. Spherical \( dV \).**

We integrate in spherical coordinates:
\[
M = \iiint_{B^3} \delta(x, y, z) \, dV = \int_0^{2\pi} \int_0^\pi \int_0^1 \frac{1}{1 + \rho^2} \rho^2 \sin \phi \, d\rho \, d\phi \, d\theta
\]
Now \( \frac{\rho^2}{1+\rho^2} = 1 - \frac{1}{1+\rho^2} \), so we have
\[
M = \int_0^{2\pi} \int_0^\pi \int_0^1 \left( 1 - \frac{1}{1+\rho^2} \right) \sin \phi \, d\rho \, d\phi \, d\theta
= \int_0^{2\pi} \int_0^\pi \rho - \rho \tan^{-1} \left( \frac{1}{\rho} \right) \bigg|_{\rho=0}^{1} \sin \phi \, d\phi \, d\theta
= \left( 1 - \frac{\pi}{4} \right) \int_0^{2\pi} \int_0^\pi \sin \phi \, d\phi \, d\theta = \left( 1 - \frac{\pi}{4} \right) 4\pi = 4\pi - \pi^2. \quad \blacksquare
4.4. INTEGRATION IN DIFFERENT COORDINATE SYSTEMS

Exercises

1. Sketch the region of integration for the following polar integrals:

   a) \( \int_{0}^{\pi/4} \int_{1}^{2} r \, dr \, d\theta \)
   b) \( \int_{-\pi/2}^{\pi/3} \int_{0}^{3} r \, dr \, d\theta \)
   c) \( \int_{\pi/6}^{\pi} \int_{1/2}^{|\sin \theta|} r \, dr \, d\theta \)
   d) \( \int_{-\pi/2}^{\pi/2} \int_{1}^{1+\cos \theta} r \, dr \, d\theta \)

2. Evaluate \( \int \int_{R} \sqrt{1 + x^2 + y^2} \, dA \) where \( R \) is the quarter unit disk in the first quadrant.

3. Evaluate \( \int \int_{R} xy \, dA \) where \( R \) is the region in the first quadrant between \( x^2 + y^2 = 1 \) and \( x^2 + y^2 = 2 \).

4. Evaluate \( \int \int_{R} e^{-(x^2+y^2)} \, dA \) where \( R \) is the disk centered at the origin and radius 10. How does your answer change if \( R \) has radius \( a \)? What happens to the integral as \( a \to \infty \)?

5. Find the area between the unit circle and the spiral \( r = e^\theta \) for \( 0 \leq \theta \leq 2\pi \).

6. Find the area of the region inside the cardioid \( r = 1 + \cos \theta \) and outside the unit circle.

7. Find the area enclosed by one petal of the 4-leafed rose \( r = \cos 2\theta \).

8. Find the area enclosed by the Lemniscate \( r^2 = \cos 2\theta \). (Hint: find the area for one side and double it)

9. Find the volume of the solid between the \( xy \)-plane and the cone \( z = 10 - 2\sqrt{x^2 + y^2} \).

10. Find the volume between the paraboloid \( x = y^2 + z^2 - 1 \) and the \( yz \)-plane.

11. Find the volume inside the cylinder \( x^2 + z^2 = 1 \), and between the plane \( y = 0 \) and the paraboloid \( y = x^2 + z^2 \).

12. Find the volume between the paraboloids \( x = 2 - 2y^2 - 2z^2 \) and \( x = y^2 + z^2 - 1 \).

13. Find the volume of the solid under \( z = \sqrt{x^2 + y^2} \) and over the circle \( x^2 + (y-1)^2 = 1 \) in the \( xy \)-plane.

14. Find the volume of the solid between the paraboloid \( z = x^2 + y^2 \) and the plane \( z = 4 \).

15. Find the volume of the solid inside the cylinder \( y^2 + z^2 = 1 \) and between the planes \( x = 0 \) and \( 2x + y + 3z = 6 \).
16. Find the volume of the solid inside the cylinder $x^2 + z^2 = 1$ and between the planes $y + z = 1$ and $y - z = -2$.

17. Find the mass of the cylinder inside $x^2 + y^2 = 1$ and between $z = \pm 1$ if the density is $\delta(x, y, z) = 4 - y - z$.

18. Sketch the solids of integration for the following spherical integrals:

   a) $\int_0^{2\pi} \int_{\pi/4}^{\pi/2} \int_0^1 \rho^2 \sin \phi d\rho d\phi d\theta$

   b) $\int_0^{2\pi} \int_{\pi/2}^{\pi} \int_0^1 \rho^2 \sin \phi d\rho d\phi d\theta$

   c) $\int_0^{\pi/2} \int_0^{\pi/2} \int_0^1 \rho^2 \sin \phi d\rho d\phi d\theta$

   d) $\int_0^{2\pi} \int_0^{\pi} \int_0^1 \rho^2 \sin \phi d\rho d\phi d\theta$

19. Find the volume of the solid outside the cone $z^2 = x^2 + y^2$ and inside the unit sphere.

20. Find the volume of the solid inside the sphere $\rho = 1$ and outside the cylinder $r = 1/2$.

21. Find the mass of the unit ball if the density is given by $\delta(x, y, z) = z^2$.

22. Find the volume inside $\rho = 1 + \cos \phi$ and outside $\rho = 1$.

23. Use integration to show that the volume of a sphere of radius $R$ is $\frac{4\pi}{3} R^3$.

   Set up the triple integrals that give the solids in exercises 24 to 27 in a) Cartesian, b) Cylindrical, and c) Spherical coordinates.

24. The solid above $\phi = \pi/6$ and below the unit sphere.

25. The solid under $z = 4 - x^2 - y^2$ and above the $xy$-plane.

26. The solid inside $\rho = 1$ and above $z = 1/2$.

27. The solid outside the cylinder $r = 1/2$ and inside the unit sphere.
4.5 Applications of Integration

In this section we introduce a couple applications of integration. We begin with calculating the surface area of parametric surfaces, then use this to derive the formula for the surface area of the graph of a function of two variables. We finish with finding centers of mass of planar regions as well as solids.

Recall that a parametric surface \( S(s,t) \) takes a region in the \( st \)-plane and bends, twists and stretches it into a surface in \( \mathbb{R}^3 \). For example, the surface \( S(s,t) = (s,t,1-s^2-t^2) \) takes the rectangle \( R = [0,1] \times [0,1] \) in the \( st \)-plane and maps it to a portion of a paraboloid (see Figure 4.5.1).

The area of the image \( S(R) \) is not necessarily the same as that of the domain \( R \). It is our goal to develop a method for calculating the area of \( S(R) \). We use the typical calculus approach:

*Approximate the surface area with things you can calculate, then take a limit as your approximation gets better.*

To approximate the surface area, first partition \( R \) into small rectangles \( R_{ij} \) of width \( \Delta s_i \) and height \( \Delta t_j \). As in Example 3.4.4 the area of \( S(R_{ij}) \) (the image of \( R_{ij} \) under the map \( S \)) is approximately the area of the parallelogram spanned by the vectors \( S_s(s_i,t_j) \Delta s_i \) and \( S_t(s_i,t_j) \Delta t_j \) (see Figure 4.5.2). The area of that parallelogram is the length of the cross product

\[
\left\| (S_s(s_i,t_j) \Delta s_i) \times (S_t(s_i,t_j) \Delta t_j) \right\|
\]

Since scalars can factor out of cross-products and taking lengths, we have
Area of $S(R_{ij}) \approx \|S_s(s_i, t_j) \Delta s_i \times S_t(s_i, t_j) \Delta t_j\|
= \|S_s(s_i, t_j) \times S_t(s_i, t_j)\| \Delta s_i \Delta t_j.$

Figure 4.5.2: Approximating Surface Area

The area of $S(R)$ is approximated by summing the approximations of areas of $S(R_{ij})$, so

$$\text{Area of } S(R) \approx \sum_{i=1}^{n} \sum_{j=1}^{m} \|S_s(s_i, t_j) \times S_t(s_i, t_j)\| \Delta s_i \Delta t_j.$$  \hspace{1cm} (4.5.1)

Taking finer partitions improves the approximation. This is illustrated for the parametric surface $S(s, t) = (s, t, 1 - s^2 - t^2), \ 0 \leq s, t \leq 1$ in Figure 4.5.3.

The original surface is pictured first, then the approximating parallelograms for different partitions of the domain $[0,1] \times [0,1]$ into subrectangles. For example, Figure 4.5.3(d) represents the approximating parallelograms when $[0,1] \times [0,1]$ is divided into 9 subsquares (the interval $[0,1]$ on each axis is divided into three subintervals of equal length). It is evident from the picture that sum of the areas of the parallelograms gets closer to the area of the surface as the partition gets finer. Taking the limit of Approximation 4.5.1 gives

**Area of a Parametric Surface**

Let $S(s, t)$ be a parametric surface over the region $R$ in the $st$-plane. The image $S(R)$ has area

$$\iint_R \|S_s(s, t) \times S_t(s, t)\| \, ds \, dt = \iint_R dS.$$  

Thus we introduce the notation $dS = \|S_s(s, t) \times S_t(s, t)\| \, ds \, dt$ for a para-
4.5. APPLICATIONS OF INTEGRATION

Notation: The notation for integrals is important, and can be daunting. For example, $dA$ represents an infinitesimal area in a plane. Depending on which coordinate system you use, it could mean $dxdy$ or $rdrd\theta$. We are now introduced to the area element $dS$ which represents an infinitesimal area on the surface $S$. If $S$ is given parametrically by $S(s,t)$, then $dS = \|S_s \times S_t\| \, dsdt$. In Chapter 5 we will want to integrate vector fields over surfaces, and the notation $dS$ will be introduced. While the symbols $dS$ and $dS$ are similar, they mean quite different things. We have seen that $dS$ is a scalar, and it will turn out that $dS$ is a vector normal to the surface $S$. Try your best to pay attention to notation, it is subtle.

After such a lengthy discussion, it is high time to illustrate these ideas with some examples.

Example 4.5.1. Paraboloid patch

Find the area of the patch of the paraboloid

$$S(s,t) = (s,t,1-s^2-t^2), \ 0 \leq s, t \leq 1.$$ 

Step 1. Find the integrand $\|S_s \times S_t\|$.  

We first compute the cross product:

$$S_s \times S_t = \begin{vmatrix} i & j & k \\ 1 & 0 & -2s \\ 0 & 1 & -2t \end{vmatrix} = (2s, 2t, 1),$$
whose length is $\|S_s \times S_t\| = \sqrt{4s^2 + 4t^2 + 1}$.

**Step 2.** Evaluate the integral

$$\iint_R ds = \int_0^1 \int_0^1 \sqrt{4s^2 + 4t^2 + 1} \, ds \, dt = \frac{7}{12} \ln(5) + 1 - \frac{1}{12} \arctan \left( \frac{4}{3} \right),$$

where the integral is done by inspection. Ok, Ok, we’re kidding. The integral is really hard and we used software to calculate it! However, if we change the domain $R$ in the $st$-plane, we end up with an integral we can calculate.

Let $S$ be the same parametric surface, $S(s, t) = (s, t, 1 - s^2 - t^2)$ over the unit disk in the $st$-plane. Find the area of $S$.

**Step 1.** The parameteric equations are the same, so the integrand is $\|S_s \times S_t\| = \sqrt{4s^2 + 4t^2 + 1}$.

**Step 2.** Evaluate the integral over the unit disk in the $st$-plane. Since $R$ is the unit disk, we switch to polar coordinates to evaluate below.

$$\iint_R dS = \iint_{\text{unit disk}} \sqrt{4s^2 + 4t^2 + 1} \, dA$$

$$= \int_0^{2\pi} \int_0^1 \sqrt{1 + 4r^2} \, r \, dr \, d\theta = \frac{1}{8} \int_0^{2\pi} \int_1^{\sqrt{5}} \sqrt{u} \, du \, d\theta$$

$$= \frac{1}{12} \int_0^{2\pi} u^{3/2} \bigg|_{u=1}^{\sqrt{5}} \, d\theta = \frac{5\sqrt{5} - 1}{12} \int_0^{2\pi} \, d\theta = \frac{(5\sqrt{5} - 1)\pi}{6}. \quad \blacksquare$$

**Example 4.5.2. Area of a torus**

Let $S(s, t) = ((\cos t + 2) \cos s, (\cos t + 2) \sin s, \sin t)$ for $0 \leq s, t \leq 2\pi$. In Example 3.4.2 we analyzed this torus, and found that

$$S_s \times S_t = \langle (\cos t + 2) \cos s \cos t, (\cos t + 2) \sin s \cos t, (\cos t + 2) \sin t \rangle.$$

![Figure 4.5.4: A torus](image)
To find the area of the surface of the torus, we integrate \( \| \mathbf{S}_s \times \mathbf{S}_t \| \) over the domain and get

\[
\iint_R dS = \int_0^{2\pi} \int_0^{2\pi} \left\| (\cos t + 2) \cos s \cos t, (\cos t + 2) \sin s \cos t, (\cos t + 2) \sin t \right\| \, dsdt
\]

\[
= \int_0^{2\pi} \int_0^{2\pi} |\cos t + 2| \left\| (\cos s \cos t, \sin s \sin t, \sin t) \right\| \, dsdt
\]

\[
= \int_0^{2\pi} \int_0^{2\pi} \cos t + 2 \, ds \, dt = 2\pi \int_0^{2\pi} (\sin t + 2t) \, dt = 2\pi \left( \sin t + 2t \right) \bigg|_0^{2\pi} = 4\pi^2. \quad \blacksquare
\]

Recall that the graph of \( z = f(x,y) \) can be parameterized by \( \mathbf{S}(x,y) = (x,y,f(x,y)) \). One readily computes

\[
\| \mathbf{S}_x \times \mathbf{S}_y \| = \left\| (1,0,f_x) \times (0,1,f_y) \right\| = \sqrt{1 + f_x^2 + f_y^2}.
\]

Applying the above formula for surface area we get the following criteria:

**Surface Area of a Graph**

The area of the graph of \( z = f(x,y) \) over the region \( R \) in the \( xy \)-plane is

\[
\iint_R dS = \iint_R \sqrt{1 + f_x^2 + f_y^2} \, dx \, dy.
\]

Thus if \( S \) is the graph of \( z = f(x,y) \), then the area element \( dS \) is given by

\[
dS = \sqrt{1 + f_x^2 + f_y^2} \, dx \, dy,
\]

and represents the area of an infinitesimally small patch on \( S \).

**Example 4.5.3. Area of a Plane**

Find the area of the portion of \( 2x + 3y + z = 6 \) in the first octant. This is the triangle with vertices \((3,0,0)\), \((0,2,0)\) and \((0,0,6)\) so we could use geometry to find the area. However, we can also think of the triangle as the portion of the graph of \( z = 6 - 2x - 3y \) in the first octant. We need to find the region \( R \) in the \( xy \)-plane that it lies above. Of course, \( R \) is just the right triangle in the first quadrant with legs on the coordinate axes and hypotenuse \( 2x + 3y = 6 \).
\[
\iint_R \sqrt{1 + f_x^2 + f_y^2} \, dxdy = \int_0^2 \int_0^{(6-3y)/2} \sqrt{1 + (-2)^2 + (-3)^2} \, dxdy \\
= \int_0^2 \sqrt{14} x \left| \frac{(6-3y)^{6-3y}/2}{2} \right|_0^2 \, dy = \frac{\sqrt{14}}{2} \int_0^2 6 - 3y \, dy \\
= \frac{\sqrt{14}}{2} \left( 6y - \frac{3}{2}y^2 \right) \biggr|_0^2 = 3\sqrt{14}. \quad \Box
\]

**Centers of Mass:** We now look at another application of integration. We’ve seen that integrating the density function over a solid gives the mass of the solid. Since the center of mass is the weighted average position, we have

Let \( W \) be a solid with density \( \rho(x, y, z) \), mass \( M \) and center of mass \((\bar{x}, \bar{y}, \bar{z})\). They are related by the following formulas:

\[
M = \iiint_W \rho(x, y, z) \, dV \\
\bar{x} = \frac{\iiint_W x\rho(x, y, z) \, dV}{M} \quad \bar{y} = \frac{\iiint_W y\rho(x, y, z) \, dV}{M} \quad \bar{z} = \frac{\iiint_W z\rho(x, y, z) \, dV}{M}
\]

There are analogous formulas in two dimensions. We finish this section with a three- and a two-dimensional example.

**Example 4.5.4. Center of Mass of Tetrahedron**

Find the center of mass of the tetrahedron with vertices \((0, 0, 0), (2, 0, 0), (0, 3, 0), \) and \((0, 0, 6)\) if the density function is given by \( \delta(x, y, z) = 40 - z^2 \).

This is the tetrahedron of Examples 4.2.6 and 4.3.13 which is bounded by the coordinate planes and \( 3x + 2y + z = 6 \). We found our limits of integration to be

\[
0 \leq z \leq 6 - 3x - 2y \\
0 \leq y \leq (6 - 3x)/2 \\
0 \leq x \leq 2.
\]

And we found the mass to be \( M = 1092/5 \). To find the center of mass, we have
to calculate the numerators of each fraction above. We have

\[
\iiint_W x\delta(x, y, z)dV = \int_0^2 \int_0^{(6-3x)/2} \int_0^{6-3x-2y} x(40 - z^2)dzdydx \\
= \int_0^2 \int_0^{(6-3x)/2} x \left(40(6-3x-2y) - \frac{(6-3x-2y)^3}{3}\right)dydx \\
= \int_0^2 x \left(-10(6-3x)^2 + \frac{(6-3x)^4}{24}\right)dx = \frac{564}{5}
\]

So the $x$-coordinate of the center of mass is

\[\overline{x} = \frac{\frac{564}{5}}{1092} = \frac{47}{91}.
\]

Similar brute force calculations yield

\[
\iiint_W y\delta(x, y, z)dV = \frac{846}{5} \\
\iiint_W z\delta(x, y, z)dV = \frac{1476}{5},
\]

which gives $\overline{y} = \frac{141}{182}$ and $\overline{z} = \frac{123}{91}$. ▲

**Example 4.5.5. A thin plate**

A thin plate covering the quarter of the unit disk in the first quadrant has density $\rho(x, y) = \sqrt{x^2 + y^2}$. Find its center of mass.

First, since the region and the density function are symmetric in $x$ and $y$, we know $\overline{x} = \overline{y}$. The mass is most easily calculated using polar coordinates, so

\[
M = \iint_R \sqrt{x^2 + y^2}dA = \int_0^{\pi/2} \int_0^1 r^2 drd\theta = \int_0^{\pi/2} \frac{r^3}{3} \bigg|_0^1 d\theta = \int_0^{\pi/2} \frac{1}{3}d\theta = \frac{\pi}{6}.
\]

Again using polar coordinates, we compute the numerator of $\overline{x}$.

\[
\iint_R x\sqrt{x^2 + y^2}dA = \int_0^{\pi/2} \int_0^1 r^4 \cos \theta \sqrt{r^2}drd\theta = \int_0^{\pi/2} \int_0^1 r^3 \cos \theta drd\theta \\
= \int_0^{\pi/2} \left[\frac{r^4}{4} \cos \theta\right]_0^1 d\theta = \frac{1}{4} \int_0^{\pi/2} \cos \theta d\theta \\
= \frac{1}{4} \sin \theta|_0^{\pi/2} = \frac{1}{4}.
\]

Combining the calculations gives $\overline{x} = \overline{y} = \frac{1/4}{\pi/6} = \frac{3}{2\pi}$. ▲

**Exercises**
1. Find the center of mass of the rectangle \( R = [0, 1] \times [0, 1] \) if the density is \( \rho(x, y) = xy \).

2. Find the center of mass of the rectangle \( R = [0, \pi/4] \times [0, \pi/4] \) if the density is \( \rho(x, y) = \sin(x + y) \).

3. Find the center of mass of the region between \( y = x^2 \) and \( y = 1 \) if the density function is \( \rho(x, y) = y \).

4. Find the center of mass of the constant density \( \rho(x, y, z) = 1 \) tetrahedron with vertices at \((0, 0, 0), (3, 0, 0), (0, 2, 0), \) and \((0, 0, 6)\).

5. Find the area of the portion of the plane \( 3x - 2y + z = 5 \) above the rectangle \([0, 1] \times [0, 1]\).

6. Find the area of the portion of the plane \( 2x - 3y + z = 40 \) above the rectangle \([0, 3] \times [1, 2]\).

7. Find the area of the portion of the plane \( x + 2y - z = 2 \) above the region \( R \) between the \( y \)-axis and \( x = 1 - y^2 \).

8. Find the area of the cone \( z = 10 - \sqrt{x^2 + y^2} \) above the unit disk.

9. Find the area of the portion of the cone \( z = 2 + 4\sqrt{x^2 + y^2} \) over the rectangle \([-1, 1] \times [-1, 1]\).

10. Find the area of the portion of \( z = \sqrt{1 - x^2} \) lying above the rectangle \([0, 1] \times [-2, 2]\).

11. Find the area of the portion of the surface \( z = x^2 + y^2 \) below the plane \( z = 9 \).

12. Find the area of the parametric surface
    \[
    S(s, t) = (2s + t - 1, 3s + 5t + 2, -s + 2t + 4),
    \]
    for \( 0 \leq s, t \leq 1 \).

13. Find the area of the parametric surface
    \[
    S(s, t) = (s + t - 2, 2s - t - 3, s + 3t + 5),
    \]
    for \( 0 \leq s, t \leq 1 \).

14. Find the area of the parametric surface \( S(s, t) = (s \cos t, 4 - s^2, s \sin t) \), for \( 0 \leq s \leq 2 \) and \( 0 \leq t \leq 2\pi \).

15. Find the area of \( S(s, t) = (s \cos t, s \sin t, s^2 \cos 2t) \), for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq \pi \).

16. Find the area of \( S(s, t) = (s \cos t, 3 - 2s, s \sin t) \), for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq \pi \).
17. Find the area of the parametric surface $\mathbf{S}(s,t) = (3s, s\cos t, s\sin t)$, for $0 \leq s \leq 2$ and $0 \leq t \leq 2\pi$.

18. Find the area of the portion of the sphere $\mathbf{S}(s,t) = (\sin s \cos t, \sin s \sin t, \cos s)$, for $0 \leq s,t \leq \pi/6$.

19. Find the area of a sphere radius $R$. Recall that one parameterization of the sphere radius $R$ is $\mathbf{S}(s,t) = (R \sin s \cos t, R \sin s \sin t, R \cos s)$, for $0 \leq s \leq \pi$ and $0 \leq t \leq 2\pi$.

20. Find the surface area of the torus $\mathbf{S}(s,t) = ((2 \cos t + 3) \cos s, (2 \cos t + 3) \sin s, 2 \sin t)$ for $0 \leq s,t \leq 2\pi$. Note that this torus is obtained by revolving the circle $(y - 3)^2 + z^2 = 4$ around the $z$-axis.
4.6 Curve and Surface Integrals

In first year calculus you learned how to integrate functions of a single variable \( \int_a^b f(x) \, dx \) defined on intervals \([a, b]\) of real numbers. In this chapter we defined double integrals \( \iint_R f(x, y) \, dA \) of functions of two variables whose domain was the region \( R \) in the \( xy \)-plane. Since the domain of \( f(x, y) \) is two-dimensional, we need two integrals to evaluate it. Similarly, a single integral suffices to integrate \( f(x) \) since its domain is one-dimensional. In each case, the dimension of the domain of the function determined the number of integrals. In this section we investigate what it means to integrate functions defined on curves and surfaces.

If \( f \) is a function defined on a curve \( C \), we will define a single integral \( \int_C f \, dC \); while if \( f \) is defined on a surface \( S \) we will define the double integral \( \iint_S f \, dS \) of \( f \) over \( S \).

Thus we will extend the notion of single integrals of functions defined on intervals \([a, b]\) in \( \mathbb{R} \) to single integrals of functions defined on curves \( C \) in \( \mathbb{R}^3 \). Analogously, double integrals of functions defined on regions \( R \) in \( \mathbb{R}^2 \) will be generalized to double integrals of functions defined on surfaces \( S \) in \( \mathbb{R}^3 \). The strategy in each case will be to define the generalized integral via ordinary single and double integrals using a choice of parameterization for \( C \) or \( S \). Then we show that the value of the integral is independent of our choice of parameterization. Since the integral is independent of how we parameterize our curves or surfaces, we think of it as integrating the function on the curve or surface itself!

Our notation, \( \int_C f \, dC \) and \( \iint_S f \, dS \), is chosen to emphasize this point of view.

**Integrals along a curve**  Let \( C \) be a curve in \( \mathbb{R}^3 \) and \( f \) a function defined on \( C \). We wish to define the integral of \( f \) over \( C \), denoted \( \int_C f \, dC \). To do so, we recall how the definite integral \( \int_a^b f(x) \, dx \) was defined in Calc I. First partition the interval \([a, b]\) into subintervals of length \( \Delta x_i \), and choose a point \( x_i^* \) in each subinterval. Then evaluate \( f \) at the chosen points \( x_i^* \) and form the approximation

\[
\int_a^b f(x) \, dx \approx \sum_{i=1}^n f(x_i^*) \Delta x_i.
\]

Finally, take the limit as the partition becomes finer and define

\[
\int_a^b f(x) \, dx = \lim_{n \to \infty} \sum_{i=1}^n f(x_i^*) \Delta x_i.
\]

The idea for defining \( \int_C f \, dC \) is completely analogous. Partition the curve \( C \) into sub-curves \( C_i \) of length \( \Delta C_i \), and choose representative points \( P_i \) from each sub-curve (see Figure UUU). For each sub-curve evaluate \( f(P_i) \Delta C_i \), and take a limit as the partition of \( C \) becomes finer to define

\[
\int_C f \, dC = \lim_{n \to \infty} \sum_{i=1}^n f(P_i) \Delta C_i. \tag{4.6.1}
\]
This is a natural conceptual generalization of single-variable integration, but how does one calculate it? To answer this question, recall that curves are usually given to us parametrically. Our first task is to calculate the ingredients of Equation 4.6.2 given a parameterization \( C(t) \), \( a \leq t \leq b \) of the curve \( C \). We will make several approximations along the way.

A partition \( t_0 = a < t_1 < \cdots < t_n = b \) of the interval \([a, b]\) yields a partition of the curve \( C \) into sub-curves. Indeed, just let \( C_i \) be the portion of \( C \) from \( C(t_{i-1}) \) to \( C(t_i) \). Since our choice of \( P_i \) was arbitrary above, we might as well choose \( P_i = C(t_{i-1}) \). Now that we’ve made the choices necessary, we need to calculate lengths and function values. As long as \( \Delta t_i \) is small, the length \( \Delta C_i \) of the sub-curve \( C_i \) is approximated by \(|C'(t_{i-1})|\Delta t_{i-1}\) (recall that \(|C'(t_{i-1})|\) is the speed of the curve, so distance traveled, \( \Delta C_i \), is speed times time, or approximately \(|C'(t_{i-1})|\Delta t_{i-1}\)). We then get the approximation

\[
\int_C f\,dC \approx \sum_{i=1}^{n} f(C(t_{i-1}))|C'(t_{i-1})|\Delta t_{i-1}.
\]

Taking a limit, the \( \Delta t_{i-1} \) becomes \( dt \), and the rest \( f(C(t))|C'(t)| \). With this motivation, we make the following definition.

**Definition 4.6.1.** Let \( C(t) \), \( a \leq t \leq b \), be a parametric curve, and \( f \) a function defined on it. The integral of \( f \) along \( C \) is

\[
\int_a^b f(C(t))|C'(t)| \, dt.
\]

At this point we have defined the integral of a function along a given parameterization. We will show that any parameterization of the same geometric curve \( C \) leads to the same value of the integral, and hence we think of it as integrating \( f \) along \( C \). Before we do, let’s calculate some examples.

**Example 4.6.1. Integrating along a line.**

Integrate the function \( f(x, y, z) = x^2 + yz \) along the curve \( C(t) = (1 - t, 2 + t, 3t) \), \( 0 \leq t \leq 1 \).

We first find the ingredients for the integrand, then put them together and integrate. We calculate

\[
f(C(t)) = f(1 - t, 2 + t, 3t) = (1 - t)^2 + (2 + t)3t = 4t^2 + 4t + 1
\]

\[
|C'(t)| = \|(−1, 1, 3)\| = \sqrt{11}.
\]

Putting these together, we get the integral of \( f \) over the line segment is

\[
\int_0^1 (4t^2 + 4t + 1)\sqrt{11} \, dt = \sqrt{11} \left( \frac{4}{3}t^3 + 2t^2 + t \right) \bigg|_0^1 = \frac{13\sqrt{11}}{3}. \quad \▲
\]

**Example 4.6.2. Mass from density**
Let \( C(t) = (3 \cos t, 3 \sin t, t), \ 0 \leq t \leq \pi/2 \), represent a thin wire in space with density function \( \delta(x, y, z) = x^2 + y^2 + z^2 \). Find the mass of the wire.

We know that integrating density gives mass, so we compute the ingredients for the integrand

\[
\delta(C(t)) = \delta(3 \cos t, 3 \sin t, t) = (3 \cos t)^2 + (3 \sin t)^2 + t^2 = 9 + t^2,
\]

\[
\| C'(t) \| = \|(-3 \sin t, 3 \cos t, 1)\| = \sqrt{10}.
\]

The mass of the wire \( M \) is then

\[
M = \int_0^{\pi/2} (9 + t^2) \sqrt{10} \, dt = \sqrt{10} \left. \left( 9t + \frac{t^3}{3} \right) \right|_0^{\pi/2} = \sqrt{10} \left( \frac{9\pi}{2} + \frac{\pi^3}{24} \right). \quad \blacksquare
\]

Math App 4.6.1. The area of a vase

In this Math App we provide another context for a line integral. In particular, the curve \( C \) is planar, and the function \( f \) represents the height of a vase lying above the curve. Then the integral \( \int_C fdC \) represents the surface area of the vase. Click on the link below to investigate.

Recall that the same curve \( C \) in \( \mathbb{R}^3 \) can have many different parameterizations. For example both \( C_1(t) = (\cos t, \sin t), \ 0 \leq t \leq \pi \), and \( C_2(t) = (\cos 2t, \sin 2t), \ 0 \leq t \leq \pi/2 \) parameterize the top half of the unit circle oriented counterclockwise. Definition 5.1.1 does not guarantee that using two different parameterizations for the same curve will yield the same integral, however we prove that now.

**Theorem 4.6.1.** The line integral of \( f \) along a curve \( C \) is independent of the parameterization chosen for \( C \).

**Proof.** The proof relies on a fact relating two parameterizations of the same curve, together with the chain rule for differentiation. Let \( C_1(t), \ a \leq t \leq b \) be a parameterization of \( C \). It turns out that any other (regular) parameterization of \( C \) can be obtained by composing with a function \( g : [c, d] \to [a, b] \) such that \( g'(s) > 0 \). Let \( C_2(s) = C_1(g(s)), \ c \leq s \leq d, \) be any other parameterization of \( C \). Using the chain rule we have

\[
\frac{d}{ds} C_2(s) = \frac{d}{ds} C_1(g(s)) = \frac{d}{ds} (x(g(s)), y(g(s)), z(g(s)))
= (x'(g(s))g'(s), y'(g(s))g'(s), z'(g(s))g'(s))
= (x'(g(s)), y'(g(s)), z'(g(s))) g'(s) = C'_1(g(s))g'(s).
\]
This calculation shows that the line integral is

\[
\int_c^d f(C_2(s)) \cdot C_2'(s) \, ds = \int_c^d f(C_1(g(s))) \cdot C_1'(g(s))g'(s) \, ds = \int_a^b f(C_1(t)) \cdot C_1'(t) \, dt,
\]

where the last equality follows by substituting \( t = g(s) \) and \( dt = g'(s) \, ds \) and using the fact that \( a = g(c) \) and \( b = g(d) \). □

**Notation**  Because the line integrals are independent of parameterization, we use the notation \( \int_C f \, dC \) for the line integral. Notice that, given a parameterization \( C(t) \) for \( C \), the differentials are related by \( dC = \|C'(t)\| \, dt \). It will be important to keep track of the differential notations as we proceed. This also outlines a method for integrating functions along curves:

---

**Integrating a function along a curve**

To evaluate \( \int_C f \, dC \):

1. Parameterize the curve \( C(t) \), for \( a \leq t \leq b \), and find \( C'(t) \) (be sure to find the limits \( a \) and \( b \) on your parameterization since they become the limits of integration).
2. Evaluate \( f \) on the curve \( C(t) \) and find \( f(C(t))\|C'(t)\| \).
3. Integrate: \( \int_a^b f(C(t))\|C'(t)\| \, dt \).

---

**Example 4.6.3. Integrating along a spiral**

Let \( f(x,y,z) = \sqrt{x^2 + y^2 + z^2} \) and let \( C \) be the parametric curve \( C(t) = (t \cos t, t \sin t, t) \), \( 0 \leq t \leq 2\pi \). Evaluate \( \int_C f \, dC \).

Since the parameterization is given, we can begin by finding the integrand. One computes that

\[
f(C(t))\|C'(t)\| = f(t \cos t, t \sin t, t)\|(\cos t \sin t, \sin t \cos t, 1)\|
\]

\[= \sqrt{t^2 \cos^2 t + t^2 \sin^2 t + t^2}.
\]

\[= \sqrt{\cos^2 t - 2t \sin t \cos t + t^2 \sin^2 t + \sin^2 t + 2t \sin t \cos t + t^2 \cos^2 t + 1}
\]

\[= \sqrt{2} \cdot t \sqrt{2 + t^2}.
\]
Integrating we get
\[
\int_C f \, dC = \sqrt{2} \int_0^{2\pi} t \sqrt{2 + t^2} \, dt \\
= \frac{\sqrt{2}}{2} \int_2^{2+4\pi^2} u^{1/2} \, du \quad \left( \text{letting } u = 2 + t^2 \right) \\
= \frac{\sqrt{2}}{3} \left. u^{3/2} \right|_2^{2+4\pi^2} = \frac{\sqrt{2}}{3} \left( 2 + 4\pi^2 \right)^{3/2} - \frac{4}{3}.
\]

Thus far we have seen how to integrate functions along curves. In Chapter 5 we will see how to integrate vector fields along curves! Before that, we discuss integrating functions on surfaces.

**Surface Integrals** To define the integral \( \iint_S f \, dS \) of a function \( f \) over a surface \( S \), we generalize double integrals. This is analogous to the way the integral along a curve generalizes that along an interval. Recall that \( \iint_R f(x,y) \, dA \) was defined by first partitioning the region \( R \) into subrectangles of area \( \Delta A_{ij} \), and choosing a point \((x_i^*,y_j^*)\) in each subrectangle. Then evaluate \( f(x_i^*,y_j^*) \) and form the approximation
\[
\iint_R f(x,y) \, dA \approx \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i^*,y_j^*) \Delta A_{ij}.
\]

Finally, take the limit as the partition becomes finer and define
\[
\iint_R f(x,y) \, dA = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i^*,y_j^*) \Delta A_{ij}.
\]

To define \( \iint_S f \, dS \), partition the surface into sub-surfaces \( S_{ij} \) of area \( \Delta S_{ij} \), and choose representative points \( P_{ij} \) from each sub-surface (see Figure UUU). Evaluate \( f(P_{ij}) \Delta S_{ij} \), and take a limit as the partition of \( S \) becomes finer to define
\[
\iint_S f \, dS = \lim_{n \to \infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(P_{ij}) \Delta S_{ij}. \tag{4.6.2}
\]

We perform the calculations given a parametric surface \( S(s,t) \) over some region \( R \) in the \( st \)-plane. Partitioning \( R \) into subrectangles \( R_{ij} \) leads to a partition of the surface, letting \( S_{ij} \) be the image of \( R_{ij} \) under \( S \). As in Section 4.5, the area of \( S_{ij} = S(R_{ij}) \) is approximately \( \| S_s \times S_t \| \) times the area of \( R_{ij} \). Thus \( \Delta S_{ij} \approx \| S_s \times S_t \| \Delta s_i \Delta t_j \), and we get the approximation
\[
\iint_S f \, dS \approx \sum_{i=1}^{n} \sum_{j=1}^{n} f(P_{ij}) \| S_s \times S_t \| \Delta s_i \Delta t_j.
\]

Taking a limit, the \( \Delta s_i \Delta t_j \) becomes \( dsdt \), and the rest \( f(S(s,t)) \| S_s \times S_t \| \). With this motivation, we make the following definition.
**Definition 4.6.2.** Let $S$ be a parametric surface $S(s, t)$ defined over the region $R$ in the $st$-plane, and $f$ a function defined on it. The integral of $f$ on $S$ is

$$\int_S f dS = \int_R f(S(s, t)) \| S_s \times S_t \| ds dt.$$ 

**Remarks:**

1. Notice from the formula for surface area, if $f$ is the constant function 1, the surface integral gives the area of the surface $S$. Symbolically,

$$\text{Area of } S = \int_S dS.$$ 

2. The differential $dS = \| S_s \times S_t \| ds dt$, and $\| S_s \times S_t \|$ is the length of the normal vector to $S(s, t)$. Thus one can think that integrating a function on a parametric surface is like double integration, with the added factor of the length of the normal vector in the integrand.

**Example 4.6.4. Mass of a hemisphere**

The upper hemisphere of the unit sphere has density $\delta(x, y, z) = 1 - z^2$, find its mass.

We know that mass is the integral of density, so we have to integrate the density function on the hemisphere.

**Step 1.** Parameterize the surface.

Recall that the unit sphere has spherical equation $\rho = 1$, so the coordinate transformations give the parameterization

$$S(s, t) = (\sin s \cos t, \sin s \sin t, \cos s), \quad 0 \leq s \leq \pi/2, \quad 0 \leq t \leq 2\pi.$$ 

**Step 2.** Find the integrand. The cross product of the partial derivatives of $S(s, t)$ is

$$S_s \times S_t = \begin{vmatrix} i & j & k \\ \cos s \cos t & \cos s \sin t & -\sin s \\ -\sin s \sin t & \sin s \cos t & 0 \end{vmatrix} = \langle \sin^2 s \cos t, \sin^2 s \sin t, \sin s \cos s \rangle,$$

so the length of the normal vector is

$$\| S_s \times S_t \| = \| \langle \sin^2 s \cos t, \sin^2 s \sin t, \sin s \cos s \rangle \| = \sin s \| \langle \sin s \cos t, \sin s \sin t, \cos s \rangle \| = \sin s.$$ 

Observing that the density function evaluated on the surface is

$$\delta(S(s, t)) = \delta(\sin s \cos t, \sin s \sin t, \cos s) = 1 - \cos^2 s,$$
the integrand is
\[ \delta(S(s,t))\parallel S_s \times S_t \parallel = \left(1 - \cos^2 s\right) \sin s. \]

**Step 3. Integrate.**
\[
\int_0^{2\pi} \int_0^{\pi/2} \left(1 - \cos^2 s\right) \sin s ds dt = \int_0^{2\pi} \left( - \int_1^0 1 - u^2 du \right) dt \\
= \int_0^{2\pi} \frac{u^3}{3} - u \bigg|_1^0 dt = \frac{2}{3} \int_0^{2\pi} dt = \frac{4\pi}{3}. \square
\]

**Example 4.6.5. Mass of a generalized cylinder**

Let \( S \) be the parametric surface \( S(s,t) = (\sin t, s, t) \) for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq \pi/2 \) with density function \( \delta(x, y, z) = x \). Find the mass of \( S \).

Evaluating the function on the surface we have
\[ \delta(S(s,t)) = \delta(\sin t, s, t) = \sin t. \]

The normal vector to the surface is
\[ S_s \times S_t = \begin{vmatrix}
  i & j & k \\
  0 & 1 & 0 \\
  \cos t & 0 & 1 
\end{vmatrix} = (1, 0, -\cos t), \]
so the length of the normal vector is \( \sqrt{1 + \cos^2 t} \). The mass of the surface is then given by the integral
\[
\iint_S \delta \ dS = \int_0^{\pi/2} \int_0^1 \sin t \sqrt{1 + \cos^2 t} \ ds dt \\
= \int_0^{\pi/2} \sin t \sqrt{1 + \cos^2 t} \ dt = - \int_1^0 \sqrt{1 + u^2} du \\
= - \left( \frac{u}{2} \sqrt{1 + u^2} + \frac{1}{2} \ln \left| u + \sqrt{1 + u^2} \right| \right)|_1^0 \\
= \frac{\sqrt{2}}{2} + \frac{1}{2} \ln \left( 1 + \sqrt{2} \right). \]

We made the substitution \( u = \cos t \), then used an integral table for the rest! ▲

**Integrating a function on a parametric surface**

The integral \( \iint_S f dS \) of the function \( f \) over the surface \( S \) parameterized by \( S(s,t) \) over the region \( R \) in the \( st \)-plane is found as follows:

1. Find the integrand \( f(S(s,t))\parallel S_s \times S_t \parallel \). This amounts to
   (a) Evaluating \( f \) on the surface (i.e find \( f(S(s,t)) \)).
(b) Find the length of the normal vector to the surface (i.e. find $\|S_s \times S_t\|$).

2. Integrate the function from Step 1 over the region $R$ in the $st$-plane.

$$\iint_S f \, ds = \iint_R f(S(s,t))\|S_s \times S_t\| \, dA.$$ 

**Integration over $z = g(x, y)$**

We now know how to integrate functions over parametric surfaces, and can apply that knowledge to integrating functions over graphs of $z = g(x, y)$. In fact, the graph of $z = g(x, y)$ can be parameterized by $S(x, y) = (x, y, g(x, y))$ and the techniques for parametric surfaces apply. As in our surface area calculations in Section 4.5, one calculates that the length of the normal vector is $\|n\| = \sqrt{g_x^2 + g_y^2 + 1}$, and evaluating $f$ on the surface gives

$$\iint_S f \, ds = \iint_R f(x, y, g(x, y))\sqrt{g_x^2 + g_y^2 + 1} \, dA.$$ 

We illustrate how to apply this formula with an example.

**Example 4.6.6. Integrating a function over a graph**

Integrate $f(x, y, z) = x^2 + y^2 + z$ over the portion of the paraboloid $z = x^2 + y^2$ lying over the unit disk.

**Step 1.** Find the integrand.

The function evaluated on the surface is

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

while the length of the normal vector (infinitesimal area) is

$$\|n\| = \|(-2x, -2y, 1)\| = \sqrt{4x^2 + 4y^2 + 1}.$$ 

**Step 2.** Integrate. The surface integral is given by

$$\iint_S f \, ds = \iint_{\text{unit disk}} 2(x^2 + y^2)\sqrt{4x^2 + 4y^2 + 1} \, dA,$$

which is best calculated in polar coordinates. The limits for the unit disk are $0 \leq \theta \leq 2\pi$, $0 \leq r \leq 1$, and $dA = r \, dr \, d\theta$. Using $r^2 = x^2 + y^2$, the surface integral becomes

$$\iint_S f \, ds = \int_0^{2\pi} \int_0^1 2r^2\sqrt{1 + 4r^2} \, r \, dr \, d\theta.$$

Letting $u = 1 + 4r^2$, one can solve for $r^2$ to get $r^2 = (u - 1)/4$. Substituting...
using the differential $du = 8rdr$ yields

\[
\int_S f \, dS = \int_0^{2\pi} \frac{1}{16} \int_1^5 (u-1)\sqrt{u} \, du \, d\theta = \int_0^{2\pi} \frac{1}{16} \int_1^5 u^{3/2} - u^{1/2} \, du \, d\theta
\]

\[
= \int_0^{2\pi} \frac{1}{16} \left( \frac{2}{5} \frac{5^{5/2}}{2} - \frac{2}{3} \frac{3^{3/2}}{2} - \left( \frac{2}{5} - \frac{2}{3} \right) \right) \, d\theta
\]

\[
= \int_0^{2\pi} \frac{1}{16} \left( \frac{20\sqrt{5}}{3} + \frac{4}{15} \right) \, d\theta = \frac{\pi}{8} \left( \frac{20\sqrt{5}}{3} + \frac{4}{15} \right).
\]

\[
\text{▲}
\]

Exercises

1. A fence lies above the curve $C(t) = (t, t^2)$ for $1 \leq t \leq 4$. Its height above any point $(x, y)$ on the curve is given by $f(x, y) = x + \sqrt{y}$. Find the area of the fence.

2. A fence lies above the curve $C(t) = (2t + 1, 3t - 5)$ for $0 \leq t \leq 2$. Its height above any point $(x, y)$ on the curve is given by $f(x, y) = 3x^2 - 4y$. Find the area of the fence.

3. Integrate the function $f(x, y) = 5x^3 + 4y$ along the curve $C(t) = (t, t^3)$, $0 \leq t \leq 1$.

4. Evaluate the integral of $f(x, y, z) = xy + z$ along the line segment from the origin to $(1, 2, 3)$.

5. Evaluate the integral of $f(x, y, z) = x^2 + y^2 + z^2$ along the curve $C(t) = (\cos t, \sin t, t)$, $0 \leq t \leq 2\pi$.

6. Evaluate the integral of $f(x, y, z) = 2x - y + z$ along the line segment $C(t) = (t, 1 + 2t, 1 + t)$, $0 \leq t \leq 3$.

7. Find the mass of the wire $C(t) = (t, t^2)$, $0 \leq t \leq 1$ if the density is given by $\delta(x, y) = xy$.

8. Find the mass of the unit circle if the density is given by $\delta(x, y) = x^2 - 2xy + y^2$.

9. Find the mass of the wire $C(t) = (2 - t, 3 + 2t, -1 + t)$, $-1 \leq t \leq 1$ if the density is given by $\delta(x, y) = x^2 + z^2$.

10. Find the mass of the wire $C(t) = (t \cos t, t \sin t, t)$, $0 \leq t \leq 2\pi$ if the density is given by $\delta(x, y) = \sqrt{x^2 + y^2 + z^2}$.

11. Evaluate the integral of $f(x, y, z) = xyz$ over the cylinder $S(s, t) = (\cos t, s, \sin t)$, for $0 \leq s \leq 1$ and $0 \leq t \leq 2\pi$.

12. Evaluate the integral of $f(x, y, z) = x + y + z$ on the surface $S(s, t) = (2s - t + 1, 3 + 2t - 1, 3s + 2t - 3)$, $0 \leq s, t \leq 1$. 

\[
\text{▲}
\]
13. The density of $S(s,t) = (s,t,s^2)$, $0 \leq s, t \leq 1$ is given by $\delta(x,y,z) = xy$. Find the mass of the surface.

14. Evaluate the integral of $f(x,y,z) = yz$ on the surface $S(s,t) = (2t^3,3t^2,s)$, $0 \leq s, t \leq 1$.

15. Evaluate the integral of $f(x,y,z) = x^2 + y + z$ on the surface $S(s,t) = (2t^3,3t^2,s)$, $0 \leq s, t \leq 1$.

16. Find the mass of the surface $S(s,t) = (2t^3,3t^2,s)$, $0 \leq s, t \leq 1$, if the density is $\delta(x,y,z) = z^2 + 2z$.

17. Evaluate the integral of $f(x,y,z) = x^2yz$ on the surface $S(s,t) = (\cos t, \sin t,s)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$.

18. Evaluate the integral of $f(x,y,z) = e^x$ on the surface $S(s,t) = (s \cos t, s \sin t,s)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$.

19. The density of the surface $S(s,t) = (\cos t, \sin t,s)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$ is $\delta(x,y,z) = 2x^2 + 3y^2 + z + 5$. Find its mass.

20. Evaluate the integral of $f(x,y,z) = x^2 + y^2 + z^2$ on the surface $S(s,t) = (s \cos t, s \sin t,s)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$.

21. Evaluate the integral of $f(x,y,z) = xy$ on the surface $S(s,t) = (s \cos t, s \sin t,s^2)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$.

22. Evaluate the integral of $f(x,y,z) = \frac{x^2}{z}$ on the surface $S(s,t) = (s \cos t, s \sin t,s^2)$, $1 \leq s \leq 2, 0 \leq t \leq \pi/2$.

23. The density of the surface $S(s,t) = (s \cos t, s \sin t,s^2)$, $0 \leq s \leq 1, 0 \leq t \leq \pi/2$ is $\delta(x,y,z) = x^2 + y^2$. Find its mass.

24. Evaluate the integral of $f(x,y,z) = z^2$ on the surface $S(s,t) = (\sin s \cos t, \sin s \sin t, \cos s)$, $0 \leq s \leq \pi/2, 0 \leq t \leq \pi/2$.

25. Evaluate the integral of $f(x,y,z) = xyz$ on the surface $S(s,t) = (\sin s \cos t, \sin s \sin t, \cos s)$, $0 \leq s \leq \pi/2, 0 \leq t \leq \pi/2$.

26. Find the mass of the portion of the cone $z = \sqrt{x^2 + y^2}$ that lies over the unit disk if the density function is $\delta(x,y,z) = 1 - z$.

27. Evaluate the integral of $f(x,y,z) = yz$ over the portion of $2x + y + z = 6$ that lies above the rectangle $R = [0,2] \times [0,1]$.

28. The density of the portion of $z = 4 - x^2 - y^2$ that lies above the $xy$-plane is $\delta(x,y,z) = 4 - z$. Find the mass of the surface.
Chapter 5

Vector Analysis

5.1 Line Integrals

In this section we describe how to integrate a vector field along a curve. Such an integral is called a line integral, even though it’s really an integral along a curve. Line integrals have many applications, and can be motivated by physical considerations. We do that now.

Example 5.1.1. Work by Constant Force

Recall that, in physics, the work done by a force $\mathbf{F}$ in moving an object from one point to another is defined to be the component of force in the direction of movement times the distance traveled. In short, work is force times distance. We saw in Section 2.2 that the work $w$ done by a constant force $\mathbf{F}$ on an object moving through the displacement vector $\mathbf{d}$ is given by

$$w = \mathbf{F} \cdot \mathbf{d} = \mathbf{F} \cdot \frac{\mathbf{d}}{\|\mathbf{d}\|} \|\mathbf{d}\|.$$

In English, the work done by a constant force on an object moving in a straight line is the dot product of the force and displacement vectors.

We now want to consider what happens if we try to calculate work for a variable force along a path that is not straight. The answer is just what you would expect in a calculus class:

Approximate the solution with work you can calculate, then take a limit as your approximation becomes better.

In order to approximate the work done we will approximate the path with small line segments, and assume the force is constant on each. To approximate the curve, recall that the direction of travel at any point on the curve $C$ is the tangent vector to $C$. If $C$ is parameterized by $\mathbf{x}(t)$, the tangent vector is $\mathbf{x}'(t)$. At any point $\mathbf{x}(t)$, we make the simplifying assumption that the force is the constant vector $\mathbf{F}(\mathbf{x}(t))$ (see Figure 5.1.1). Locally the component of $\mathbf{F}$ in the direction of movement, then, should be $\frac{\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)}{\|\mathbf{x}'(t)\|}$. The distance traveled is
speed times time, so for small time intervals $\Delta t$ it’s $\|x'(t)\| \Delta t$. Locally, then, the work done is

$$\frac{\mathbf{F}(x(t)) \cdot x'(t)}{\|x'(t)\|} \|x'(t)\| \Delta t = \mathbf{F}(x(t)) \cdot x'(t) \Delta t.$$  

Adding the local contributions approximates the total work, and taking a limit turns the sum into the integral of the following definition.

**Definition 5.1.1.** The line integral of the vector field $\mathbf{F}(x, y, z)$ along the parametric curve $x(t) = (x(t), y(t), z(t))$, for $a \leq t \leq b$, is

$$\int_{a}^{b} \mathbf{F}(x(t)) \cdot x'(t) \, dt.$$  

The formula of Definition 5.1.1 will be called vector notation for the line integral because it emphasizes the vector aspect of it (as opposed to differential notation defined later in the section). Note that the integrand is the dot product of two vectors, so it is a scalar function and the integral is a single-variable integral. Implicit in the notation $\mathbf{F}(x(t))$ is the assumption that you have a parameterization, and the notation means that you substitute the parametric equations into the component functions of the vector field. This is outlined in the following procedure for evaluating line integrals.

### Parametric Method for Evaluating Line Integrals

To evaluate $\int_{a}^{b} \mathbf{F}(x(t)) \cdot x'(t) \, dt$:

1. Parameterize the curve $x(t)$, for $a \leq t \leq b$, and find $x'(t)$ (be sure to find the limits $a$ and $b$ on your parameterization since they become the limits of integration).

2. Evaluate $\mathbf{F}$ on the curve $x(t)$ and find $\mathbf{F}(x(t)) \cdot x'(t)$.

3. Integrate: $\int_{a}^{b} \mathbf{F}(x(t)) \cdot x'(t) \, dt$. 

We carry this procedure out in a few examples. It will be helpful to recall that
\[ x(t) = (1 - t)P + tQ \text{ for } 0 \leq t \leq 1. \]
parameterizes the line segment from \( P \) to \( Q \).

**Example 5.1.2. Work by variable force**

Let \( F(x, y, z) = \left( \frac{-x}{\sqrt{x^2 + y^2 + z^2}}, \frac{-y}{\sqrt{x^2 + y^2 + z^2}}, \frac{-z}{\sqrt{x^2 + y^2 + z^2}} \right) \), and find the work done by \( F \) on a particle that travels along the line segment from the origin to the point \( (1, 2, 3) \).

**Step 1.** Parameterize the curve and differentiate it.

In our case, the line segment from \( (0, 0, 0) \) to \( (1, 2, 3) \) can be parameterized by
\[ x(t) = (1 - t)(0, 0, 0) + t(1, 2, 3) = (t, 2t, 3t) \text{ for } 0 \leq t \leq 1. \]
So \( x'(t) = (1, 2, 3) \).

**Step 2.** Find \( F(x(t)) \cdot x'(t) \)

Evaluating \( F \) on the curve gives
\[
F(x(t)) = F(t, 2t, 3t)
= \left( \frac{-t}{\sqrt{t^2 + (2t)^2 + (3t)^2}}, \frac{-2t}{\sqrt{t^2 + (2t)^2 + (3t)^2}}, \frac{-3t}{\sqrt{t^2 + (2t)^2 + (3t)^2}} \right)
= \left( \frac{-1}{\sqrt{10}}, \frac{-2}{\sqrt{10}}, \frac{-3}{\sqrt{10}} \right).
\]
We remark that it is not common for \( F(x(t)) \) to be a constant vector, that just happens in this example. The integrand of the line integral is
\[
F(x(t)) \cdot x'(t) = \left( \frac{-1}{\sqrt{10}}, \frac{-2}{\sqrt{10}}, \frac{-3}{\sqrt{10}} \right) \cdot (1, 2, 3) = -\frac{10}{\sqrt{10}} = -\sqrt{10}.
\]

**Step 3.** Integrate
\[
\int_a^b F(x(t)) \cdot x'(t) \, dt = \int_0^1 -\sqrt{10} \, dt = -\sqrt{10}t \bigg|_0^1 = -\sqrt{10}.
\]
The fact that the work done by \( F \) is negative means that the force was hindering movement rather than helping it. Geometrically, it means that the component of \( F \) in the \( x' \) direction is negative.

**Independence of Parameterization** The first step in integrating a vector field along a curve is to parameterize the curve. Recall that there are many parameterizations for the same curve. For example both
\[
x_1(t) = (\cos t, \sin t), \ 0 \leq t \leq \pi, \text{ and } x_2(t) = (\cos 2t, \sin 2t), \ 0 \leq t \leq \pi/2,
\]
parameterize the top half of the unit circle oriented counterclockwise. A priori, using different parameterizations of the same curve \( C \) could result in different
values of the line integral. It turns out they don’t (see Theorem 5.1.1). This allows us to talk about the integral of a vector field $\mathbf{F}$ along a curve $C$, rather than having to specify the parameterization we use. Before stating the theorem, we illustrate it with an example to get more familiar with the parameteric method of integrating vector fields along curves.

**Example 5.1.3. Independence of Parameterization**

Evaluate the line integral of $\mathbf{F}(x, y) = \langle xy, 2x \rangle$ along the top half of the unit circle using the parameterizations:

(a) $\mathbf{x}(t) = (\cos t, \sin t)$, $0 \leq t \leq \pi$, and

(b) $\mathbf{x}(t) = (\cos 2t, \sin 2t)$, $0 \leq t \leq \pi/2$

We follow the above steps in each case:

(a) Using the standard parameterization of the unit circle.

1. $\mathbf{x}(t) = (\cos t, \sin t)$, for $0 \leq t \leq \pi$, and $\mathbf{x}'(t) = (\sin t, \cos t)$.

2. $\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = (\cos t \sin t, 2 \cos t) \cdot (\sin t, \cos t)$

\[ = -\cos t \sin^2 t + 2 \cos^2 t. \]

3. $\int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \, dt = \int_0^\pi -\cos t \sin^2 t + 2 \cos^2 t \, dt$. The first sum can be integrated using the substitution $u = \sin t$, and the second using the half-angle formula $\cos^2 t = \frac{1}{2}(1 + \cos 2t)$.

Explicitly, we have:

\[ \int_0^\pi -\cos t \sin^2 t + 2 \cos^2 t \, dt = \int_0^\pi -\cos t \sin^2 t \, dt + \int_0^\pi 2 \cos^2 t \, dt \]

\[ = -\frac{\sin^3 t}{3} + t + \frac{\sin 2t}{2} \bigg|_0^\pi = \pi. \]

(b) Using the parameterization that is twice as fast.

1. $\mathbf{x}(t) = (\cos 2t, \sin 2t)$, for $0 \leq t \leq \pi/2$, and $\mathbf{x}'(t) = (\sin 2t, 2 \cos 2t)$.

2. $\mathbf{F}(\mathbf{x}(t)) = (\cos 2t \sin 2t, 2 \cos 2t)$, so

\[ \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = (\cos 2t \sin 2t, 2 \cos 2t) \cdot (\sin 2t, 2 \cos 2t) \]

\[ = -2 \cos 2t \sin^2 2t + 4 \cos^2 2t. \]

3. $\int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \, dt = \int_0^{\pi/2} -2 \cos 2t \sin^2 2t + 4 \cos^2 2t \, dt$. Making the substitution $u = 2t$, $du = 2dt$, and changing the limits gives $\int_0^\pi -\cos u \sin^2 u + 2 \cos^2 u \, du$. Since this is the same integral as part (a), the value of the line integral of $\mathbf{F}$ for either parameterization is $\pi$. 
As mentioned above, this is not an isolated phenomena, but a general fact. Let \( C \) be any geometric curve from point \( A \) to point \( B \) (for example, the semi-circle from \((1,0)\) to \((-1,0)\) in the plane). As long as we parameterize \( C \) with the correct orientation (from \( A \) to \( B \)) and don’t backtrack while traversing \( C \), the line integral is independent of the parameterization we choose. We state it as a theorem:

**Theorem 5.1.1.** The line integral of \( F \) along a curve \( C \) is independent of the parameterization chosen for \( C \).

**Proof.** The proof relies on a fact relating two parameterizations of the same curve, together with the chain rule for differentiation. Let \( x(t), a \leq t \leq b \) be a parameterization of \( C \). It turns out that any other (regular) parameterization \( g \) of \( C \) can be obtained by composing with a function \( g: [c,d] \rightarrow [a,b] \) such that \( g'(s) \geq 0 \). Let \( y(s) = x(g(s)), c \leq s \leq d \), be any other parameterization of \( C \). Using the chain rule we have

\[
\frac{d}{ds} y(s) = \frac{d}{ds} x(g(s)) = \frac{d}{ds} (x(g(s)), y(g(s)), z(g(s)))
\]

\[
= (x'(g(s))g'(s), y'(g(s))g'(s), z'(g(s))g'(s))
\]

\[
= (x'(g(s)), y'(g(s)), z'(g(s))) g'(s) = x'(g(s))g'(s).
\]

This calculation shows that the line integral is

\[
\int_{C}^d F(y(s)) \cdot y'(s) \, ds = \int_{C}^d F(x(g(s))) \cdot x'(g(s))g'(s) \, ds = \int_{a}^b F(x(t)) \cdot x'(t) \, dt,
\]

where the last equality follows by substituting \( t = g(s) \) and \( dt = g'(s)ds \) and using the fact that \( a = g(c) \) and \( b = g(d) \). \( \square \)

Because the line integral depends only on the curve \( C \) from point \( A \) to point \( B \), and not on a choice of parameterization \( x(t) \), we often use the following notation for line integrals:

\[
\int_{a}^{b} F(x(t)) \cdot x'(t) \, dt = \int_{C} F \cdot dx.
\]

In this shorthand notation \( \int_{C} F \cdot dx \), the parameterization \( x(t) \) is suppressed, and we've also introduced the differential notation \( dx \) to represent \( x'(t)dt \). This is consistent with the use of differential notation in single-variable calculus, where if \( x = f(t) \) we know \( dx = f'(t)dt \). One difference in notation, of course, is that \( dx = x'(t)dt \) represents a vector times the differential \( dt \).

**Curve Notation:** Sometimes we encounter curves that are made by concatenating pieces of smooth curves. For example, the boundary of the triangle with vertices \((0,0)\), \((1,0)\) and \((0,1)\) consists of three line segments. Let \( C \) denote the boundary of the triangle, oriented counterclockwise. Then \( C \) is made up of the line segments \( C_1 \) from \((0,0)\) to \((1,0)\), \( C_2 \) from \((1,0)\) to \((0,1)\), and \( C_3 \) from \((0,1)\) to \((0,0)\).
CHAPTER 5. VECTOR ANALYSIS

back to \((0, 0)\). In situations like this, we say that the line integral of a vector field \(F\) along the boundary of the triangle \(C\) is the sum of the line integrals along the curves that combine to make \(C\). Symbolically we have

\[
\int_C F \cdot dx = \int_{C_1+C_2+C_3} F \cdot dx = \int_{C_1} F \cdot dx + \int_{C_2} F \cdot dx + \int_{C_3} F \cdot dx.
\]

Similarly we let \(-C\) denote the curve \(C\) backwards. More precisely any parameterization of the curve \(C\) induces an orientation of \(C\), in other words, a direction of travel along \(C\). The orientation is the direction of increasing \(t\)-values. For example \(x(t) = (\cos t, \sin t)\) for \(0 \leq t \leq \pi\) starts at \((1, 0)\) and induces a counterclockwise orientation of the unit circle. On the other hand, \(x(t) = (\sin t, \cos t)\) for \(0 \leq t \leq 2\pi\) starts at \((0, 1)\) and traverses the unit circle in a clockwise fashion. If \(C\) is a curve oriented from \(A\) to \(B\), then we denote the same curve from \(B\) to \(A\) as \(-C\). It follows from the chain rule (or rather, from substitution in integration) that

\[
\int_{-C} F \cdot dx = -\int_C F \cdot dx,
\]

and we will use this fact regularly. This is intuitively obvious as well. The line integral of a vector field \(F\) along \(C\) is just the integral of the component of \(F\) in the tangential direction. The curve \(-C\) has the opposite tangential direction to \(C\). Since the tangents to \(C\) and \(-C\) are negatives of each other, the component of \(F\) in the \(-C\) direction will be the negative of that in the \(C\) direction. Hence the line integrals will have the same magnitude, but opposite signs.

We illustrate these notions with some examples.

**Example 5.1.4. Integrals along piecewise smooth curves**

Find the line integral \(\int_C F \cdot dx\) of \(F(x, y) = \langle xy, -x^2 \rangle\) along the boundary \(C\) of the triangle with vertices \((0, 0)\), \((1, 0)\), and \((0, 1)\) oriented counterclockwise (see Figure 5.1.2). We still use the approach outlined above of parameterizing, evaluating and taking dot products to find the integrand.

\(C_1\): This segment is parameterized by \(x_1(t) = (t, 0)\) for \(0 \leq t \leq 1\), so \(x'_1(t) = \langle 1, 0 \rangle\). Evaluating the vector field on the curve we have \(F(t, 0) = \langle 0, -t^2 \rangle\), so the line integral is

\[
\int_{C_1} F \cdot dx = \int_0^1 \langle 0, -t^2 \rangle \cdot \langle 1, 0 \rangle \, dt = 0.
\]

\(C_2\): This curve is parameterized by \(x_2 = (1 - t, t)\), for \(0 \leq t \leq 1\), so that \(x'_2(t) = \langle -1, 1 \rangle\). Evaluating the vector field on the curve gives \(F(1 - t, t) = \langle (1 - t)t, -(1 - t)^2 \rangle\), and the line integral becomes

\[
\int_{C_2} F \cdot dx = \int_0^1 \langle (1 - t)t, -(1 - t)^2 \rangle \cdot \langle -1, 1 \rangle \, dt
\]

\[
= \int_0^1 -t + 2t^2 dt = t - \frac{3t^3}{2} \bigg|_0^1 = \frac{1}{6}.
\]
5.1. LINE INTEGRALS

Figure 5.1.2: Line integrals along piecewise defined curves

$C_3$: This curve is parameterized by $x_3(t) = (0, 1-t)$, for $0 \leq t \leq 1$ so that $x_3'(t) = (0, -1)$. Since each component function of $F$ has an $x$ in it, we see that $F(0, 1-t) = (0, 0)$ and the integrand in the line integral will be zero. Thus $\int_{C_3} F \cdot dx = 0$. Putting these together we have

$$\int_{C} F \cdot dx = \int_{C_1} F \cdot dx + \int_{C_2} F \cdot dx + \int_{C_3} F \cdot dx = 0 + \frac{1}{6} + 0 = \frac{1}{6}.$$

Example 5.1.5. Integrating along $-C_2$

In the previous example we calculated the line integral of $F(x, y) = (xy, -x^3)$ along the line segment $C_2$ from $(1, 0)$ to $(0, 1)$. We found $\int_{C_2} F \cdot dx = \frac{1}{6}$.

The curve $-C_2$ is the line segment from $(0, 1)$ to $(1, 0)$, which is parameterized by $x(t) = (t, 1-t)$ for $0 \leq t \leq 1$ (so $x'(t) = (1, -1)$). As $F(t, (1-t)) = (t(1-t), -t^2)$, we have

$$\int_{-C_2} F \cdot dx = \int_{0}^{1} \left\langle t(1-t) - t^2 \right\rangle \cdot (1, -1) \, dt = \int_{0}^{1} t - 2t^2 \, dt = \left. \frac{t^2}{2} - \frac{2t^3}{3} \right|_{0}^{1} = -\frac{1}{6}.$$

This illustrates that when you integrate backwards along a curve you get the negative of the forward integral.

Differential Notation: There is yet another notation for line integrals that is common, and useful. So far we have denoted the line integral of $F$ along a curve $C$ by $\int_{C} F \cdot dx$, which is relatively simple notation but can hide some of the details of what we mean. A more descriptive notation was given initially, and is $\int_{a}^{b} F(x(t)) \cdot x'(t) \, dt$. This notation indicates that to calculate line integrals we use a parameterization of the curve, and makes it clearer that we’re just integrating the component of $F$ in the direction tangent to $x(t)$. We now
introduce differential notation, which will be convenient when discussing some of the main theorems in vector analysis.

Let \( \mathbf{F}(x, y, z) = P(x, y, z) \mathbf{i} + Q(x, y, z) \mathbf{j} + R(x, y, z) \mathbf{k} \) be a vector field, and \( C \) any curve. We introduce the notation

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \int_C P(x, y, z)\, dx + Q(x, y, z)\, dy + R(x, y, z)\, dz,
\]

which will sometimes be shortened to \( \int_C P\, dx + Q\, dy + R\, dz \), suppressing the independent variables.

We now describe why this notation makes sense, given our experience with differentials \( dx, dy \) and \( dz \). Let \( x(t) \), for \( a \leq t \leq b \), be a parameterization of \( C \).

We then have

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \int_a^b F(x(t)) \cdot \langle x'(t) \rangle \, dt
\]

\[
= \int_a^b \langle P(x(t)), Q(x(t)), R(x(t)) \rangle \cdot \langle x'(t), y'(t), z'(t) \rangle \, dt
\]

\[
= \int_a^b P(x(t))x'(t) + Q(x(t))y'(t) + R(x(t))z'(t) \, dt
\]

\[
= \int_a^b P(x(t))\, dx' + \int_a^b Q(x(t))\, dy' + \int_a^b R(x(t))\, dz'.
\]

By now you’re getting tired of writing \( t \)’s all the time, so you decide to substitute \( x \) for \( x(t) \), \( y \) for \( y(t) \) and \( z \) for \( z(t) \). Of course, if you substitute \( x = x(t) \), then taking the differential you get \( dx = x'(t)\, dt \). Moreover, you’re suppressing the parameterization, so limits of integration should change to reflect that. Instead of integrating for \( a \leq t \leq b \), we say we integrate along the curve \( C \) again. After substitution, then, we see

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \int_a^b P(x(t))x'(t)\, dt + \int_a^b Q(x(t))y'(t)\, dt + \int_a^b R(x(t))z'(t)\, dt.
\]

We describe how to evaluate line integrals given in differential notation.

---

**Evaluating Line Integrals in Differential Notation**

To evaluate \( \int_C P\, dx + Q\, dy + R\, dz \),

1. Parameterize \( C \) by \( x(t) = (x(t), y(t), z(t)) \) for \( a \leq t \leq b \)
2. Evaluate the component functions $P$, $Q$, and $R$ on $\mathbf{x}(t)$, and let $dx = x'(t)dt$, $dy = y'(t)dt$ and $dz = z'(t)dt$.

3. Integrate $\int_a^b P(\mathbf{x}(t))x'(t) + Q(\mathbf{x}(t))y'(t) + R(\mathbf{x}(t))z'(t)dt$.

**Example 5.1.6. Evaluating a line integral in differential notation**

Let $C$ be the helical arc given by $\mathbf{x}(t) = (\cos t, \sin t, t)$ for $0 \leq t \leq \pi$. Evaluate $\int_C zdx + xydy + xzdz$.

Notice that we already have a parameterization, so the first step is done. The differentials are given by $dx = -\sin t dt$, $dy = \cos t dt$ and $dz = dt$. Substituting we get

$$
\int_C zdx + xydy + xzdz = \int_0^\pi t(-\sin t dt) + \cos t \sin t(\cos t dt) + t \cos t dt
$$

$$
= -\int_0^\pi t \sin t dt + \int_0^\pi \cos^2 t \sin t dt + \int_0^\pi t \cos t dt
$$

$$
= -(-t \cos t + \sin t) + \frac{\cos^3 t}{3} + (t \sin t + \cos t) \bigg|_0^\pi
$$

$$
= -\pi - \frac{8}{3}.
$$

Integration by parts and substitution were used to evaluate the integrals.

**Things to Know**

- The line integral $\int_C \mathbf{F} \cdot d\mathbf{x}$ of a vector field $\mathbf{F}$ along a curve $C$ is the integral of the component of $\mathbf{F}$ in the tangential direction.

- Vector notation

$$
\int_a^b \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \ dt
$$

emphasizes the dot product, and the need for a parameterization.

- Differential notation

$$
\int_C Pdx + Qdy + Rdz
$$

emphasizes the components of $\mathbf{F}$ and is used in other contexts.

- Know how to calculate $\int_C \mathbf{F} \cdot d\mathbf{x}$ in either notation using parameterizations.
\[ \int_C \mathbf{F} \cdot d\mathbf{x} \] is independent of the parameterization of \( C \).

**Exercises**

1. Let \( \mathbf{F}(x, y, z) = \langle -x, -y, -z \rangle \), and let \( C \) be the line segment from the origin to the point \((3, -1, 2)\). Do you expect \( \int_C \mathbf{F} \cdot d\mathbf{x} \) to be positive or negative, and why?

2. Assume both curves \( A \) and \( B \) in the Figure 5.1.3 are oriented counterclockwise.

\[ \text{Figure 5.1.3: Relative size of line integrals} \]

(a) Are \( \int_A \mathbf{F} \cdot d\mathbf{x} \) and \( \int_B \mathbf{F} \cdot d\mathbf{x} \) positive or negative? Why?

(b) Is \( \int_A \mathbf{F} \cdot d\mathbf{x} < \int_B \mathbf{F} \cdot d\mathbf{x} \)? Be careful of the signs, and justify your conclusion.

3. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{C} \) where \( \mathbf{C}(t) = (2 - t, 2t), \, 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y) = \langle x + 2y, y^2 \rangle \).

4. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{C} \) where \( \mathbf{C}(t) = (t, t^2), \, 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y) = \langle 3x^2 - y, 2y + x \rangle \).

5. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{C} \) where \( \mathbf{C}(t) = (1 - t, 1 - t), \, 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y) = \langle 3x^2 - y, 2y + x \rangle \).

6. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{C} \) where \( \mathbf{C}(t) = (1 - t, 1 - t), \, 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y) = \langle 3x^2 - y, 2y + x \rangle \).

7. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \) where \( \mathbf{C}(t) = (2 + t, 3 - 2t, 4t), \, 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y, z) = \langle 2x - z, y + 2z, x - 3y \rangle \).
8. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \) where \( \mathbf{C}(t) = (\pi + t, 2t, t^2) \), \( 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y, z) = (\cos x, \sin y, z) \).

9. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \) where \( \mathbf{C}(t) = (1 + t, t, 1 - t) \), \( 0 \leq t \leq 1 \) and \( \mathbf{F}(x, y, z) = (xy, xz, yz) \).

10. Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{C} \) where \( C \) is the elliptical arc \( \mathbf{x}(t) = (2 \cos t, 3 \sin t) \), \( 0 \leq t \leq \pi/2 \) and \( \mathbf{F}(x, y) = (y, -x + 2) \).

11. Let \( \mathbf{F}(x, y) = (x, y) \) and \( C \) be the line segment from \((2, 0)\) to \((0, 1)\).

   (a) Find \( \int_C \mathbf{F} \cdot d\mathbf{x} \) using the parameterization \( \mathbf{x}(t) = (2 - 2t, t) \) for \( 0 \leq t \leq 1 \).

   (b) Find \( \int_C \mathbf{F} \cdot d\mathbf{x} \) using the parameterization \( \mathbf{x}(t) = (2 - 2t^2, t^2) \) for \( 0 \leq t \leq 1 \). (you should convince yourself that both parameterize the segment of \( x + 2y = 2 \) from \((2, 0)\) to \((0, 1)\).)

12. Let \( \mathbf{F}(x, y) = (y, x^2) \). Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \) where \( C \) is

   (a) The line segment from \((1, 0)\) to \((-1, 0)\).

   (b) The upper semicircle from \((1, 0)\) to \((-1, 0)\).

13. Find the work done by the constant vector field \( \mathbf{F}(x, y, z) = (0, 0, -9.8) \) in moving a particle from \((1, 0, 0)\) to \((1, 0, 2\pi)\) along

   (a) The vertical line segment.

   (b) The helical arc \( \mathbf{x}(t) = (\cos t, \sin t, t) \), \( 0 \leq t \leq 2\pi \). Are your answers the same or different?

14. Find the work done by the force \( \mathbf{F}(x, y) = (2y - x, y) \) in moving a particle from the origin to \((1, 1)\) along the parabola \( y = x^2 \).

15. Find the work done by \( \mathbf{F}(x, y) = (xy, 1 + x) \) along the curve \( \mathbf{x}(t) = (2 - 3t^{10}, 1 + t^{10}) \) for \( 0 \leq t \leq 1 \) (Hint: find a different parameterization of the same curve that makes your calculation easier).

16. Let \( \mathbf{F}(x, y) = (y, -x) \).

   (a) Find \( \int_C \mathbf{F} \cdot d\mathbf{x} \) where \( C \) is the unit circle oriented counterclockwise.

   (b) Now parameterize the unit circle in a clockwise fashion to calculate \( \int_{-C} \mathbf{F} \cdot d\mathbf{x} \). This should verify that \( \int_C \mathbf{F} \cdot d\mathbf{x} = -\int_{-C} \mathbf{F} \cdot d\mathbf{x} \).

17. Evaluate both \( \int_C \mathbf{F} \cdot d\mathbf{x} \) and \( \int_{-C} \mathbf{F} \cdot d\mathbf{x} \) where \( \mathbf{F}(x, y) = (x^2, x + y) \) and \( C \) is the line segment from \((0, 3)\) to \((4, 3)\).

18. Let \( C \) start at \((0, 0)\) and travel counterclockwise around the triangle with vertices \((0, 0)\), \((1, 0)\) and \((0, 2)\). Find the line integral of \( \mathbf{F}(x, y) = (x + 2y, 3x - y) \) around \( C \).
19. Let $C$ be any curve on the cone $x^2 + y^2 - z^2 = 0$, and let $\mathbf{F}(x, y, z) = \langle x, y, -z \rangle$. Describe why $\int_C \mathbf{F} \cdot d\mathbf{x} = 0$.

20. Evaluate $\int_C xy dx - 2y dy$ where $C$ is the line segment from $(2, 1)$ to $(4, 0)$.

21. Evaluate $\int_C -y dx + x dy$ where $C$ is the unit circle.

22. Evaluate $\int_C -y dx + x dy$ where $C(t) = (3 \cos t, 2 \sin t), \ 0 \leq t \leq 2\pi$.

23. Evaluate $\int_C \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$ where $C$ is the unit circle.

24. Evaluate $\int_C \frac{-y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$ where $C$ is the circle centered at the origin with radius 2.

25. Evaluate $\int_C 2xy dx + x^2 dy$ where $C$ is the triangle with vertices $(0, 0), (1, 0)$ and $(0, 2)$ oriented counterclockwise.

26. Evaluate $\int_C z dx + x dy - y dz$ where $C$ is the helical arc $\mathbf{x}(t) = (\cos t, \sin t, t), \ 0 \leq t \leq \pi$.

27. Let $C$ be a curve parameterized by $\mathbf{x}(t) = (x(t), y(t)), \ a \leq t \leq b$, where the product $x(t)y(t)$ is a constant. Show that $\int_C y dx + x dy = 0$. 


5.2 Conservative Vector Fields

Path Dependence: We have seen that the integral of a vector field along a curve does not depend on the choice of parameterization. One might also focus on the endpoints of the curve and ask if line integrals depend on the path between two points. In general two paths $C_1$ and $C_2$ with the same endpoints yield different line integrals, as we see in the next example. There is an important class of vector fields, called conservative vector fields, for which line integrals are independent of the path. We will study conservative vector fields in this section, and prove the Fundamental Theorem of Calculus for line integrals.

Example 5.2.1. Two examples in the plane

Let $\mathbf{F}(x, y) = (-y, x)$. Find the work done by $\mathbf{F}$ on a particle moving from $(1, 0)$ to $(-1, 0)$ along

(a) The straight line segment connecting them.

(b) The top semicircle of the unit circle.

We treat each case separately, using the three step strategy of parameterizing, finding the integrand and integrating.

(a) The line segment

Step 1. Parameterize and differentiate

A parameterization of the line segment is $\mathbf{x}(t) = (1-t)(1,0) + t(-1,0) = (1-2t,0)$ for $0 \leq t \leq 1$. Thus $\mathbf{x}'(t) = (-2,0)$.

Step 2. Find $\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t)$

Evaluating $\mathbf{F}$ on $\mathbf{x}(t)$ gives $\mathbf{F}(1-2t,0) = (0,1-2t)$. We obtain the integrand:

$$\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = (0,1-2t) \cdot (-2,0) = 0.$$

Step 3. Integrate

$$\int_0^1 \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \, dt = \int_0^1 0 \, dt = 0.$$

Notice that in this example the integrand is always zero. This means that the vector field $\mathbf{F}$ is always perpendicular to the tangent vector of the line segment (see Figure 5.2.1).

(b) The integral of $\mathbf{F}$ over the semicircle.

Step 1. $\mathbf{x}(t) = (\cos t, \sin t)$ for $0 \leq t \leq \pi$ and $\mathbf{x}'(t) = (-\sin t, \cos t)$.

Step 2. $\mathbf{F}(\cos t, \sin t) = (-\sin t, \cos t)$, and

$$\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = (-\sin t, \cos t) \cdot (-\sin t, \cos t) = \sin^2 t + \cos^2 t = 1.$$

Step 3. Integrate

$$\int_0^\pi \mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) \, dt = \int_0^\pi 1 \, dt = \pi.$$
Figure 5.2.1: Work depends on the path

Notice that the line integral of $F$ along the semicircle is different from that along the line, even though the endpoints are the same. We say the line integral is path-dependent. More precisely, if $F$ is a vector field and $C_1$ and $C_2$ are both paths from point $A$ to point $B$, then it may be that $\int_{C_1} F \cdot dx \neq \int_{C_2} F \cdot dx$.

**Conservative Fields:** It will turn out that the vector field $F$ is conservative if and only if line integrals of $F$ are independent of the path. Even though this characterization of conservative fields is a statement about paths, we use closed loops to define a conservative vector fields. A *closed loop* is any curve that starts and stops at the same point. A *simple closed loop* is a closed loop be a curve that doesn’t intersect itself. A lemniscate is a closed loop since it starts and stops in the same place, but it is not simple because it intersects itself. An ellipse is an example of a simple closed loop (see Figure 5.2.2). For now we consider closed loops in general, while simple closed loops will be used in our discussion of Green’s Theorem.

![Figure 5.2.2: Simple vs. non-simple closed loops](image)

(a) The lemniscate $r^2 = \cos \theta$  \hspace{1cm} (b) A simple closed loop

We can now define what we mean by a conservative vector field. We then give an equivalent definition using paths rather than loops, and characterize conservative vector fields in terms of antiderivatives.
Definition 5.2.1. A vector field is conservative if \( \int_C \mathbf{F} \cdot d\mathbf{x} = 0 \) for every closed loop \( C \).

An equivalent definition is: The vector field \( \mathbf{F} \) is conservative if all line integrals of \( \mathbf{F} \) are independent of the path taken. More precisely, let \( C_1 \) and \( C_2 \) be any two curves from point \( A \) to point \( B \). Then \( \mathbf{F} \) is conservative if for every \( C_1 \), \( C_2 \) we have \( \int_{C_1} \mathbf{F} \cdot d\mathbf{x} = \int_{C_2} \mathbf{F} \cdot d\mathbf{x} \). To see that this is equivalent, notice that if \( C_1 \) and \( C_2 \) both start and stop at the same points, then the piecewise defined curve \( C = C_1 - C_2 \) is a closed loop. Path independence now implies

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \int_{C_1} \mathbf{F} \cdot d\mathbf{x} - \int_{C_2} \mathbf{F} \cdot d\mathbf{x} = 0.
\]

Since any closed loop can be expressed as the difference of two paths with the same endpoints, path independence of line integrals implies integrals around closed loops are zero. The same equation implies the converse: if the line integral of \( \mathbf{F} \) around any closed loop is zero, then line integrals of \( \mathbf{F} \) along depend only on their endpoints.

Thus there is a closed loop definition of a conservative vector field, and an equivalent definition in terms of paths. While this is a nice intuitive definition of a conservative field, it isn’t very helpful in determining if a given vector field is conservative or not. It turns out that gradient fields \( \nabla f \) are exactly conservative vector fields, and we state this as a theorem (without proof).

Theorem 5.2.1. The vector field \( \mathbf{F} \) is conservative if and only if there is a function \( f \) such that \( \mathbf{F} = \nabla f \).

This fact gives us a method for testing whether or not a vector field is conservative. If \( \mathbf{F} = \langle P, Q \rangle \) is a gradient field, then there is a function \( f \) such that \( \mathbf{F} = \nabla f \). We call \( f \) the potential function, and note that \( P = \frac{\partial f}{\partial x} \) and \( Q = \frac{\partial f}{\partial y} \). Since mixed partials are equal we notice that \( \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \), which becomes our test for determining if a vector field is conservative or not.

Example 5.2.2. Determining Conservative Vector Fields

Determine if the following vector fields are conservative:

1. \( \mathbf{F} = \langle -y, x \rangle \) is not conservative since \( \frac{\partial P}{\partial y} = -1 \), but \( \frac{\partial Q}{\partial x} = 1 \).

2. \( \mathbf{F}(x, y) = \langle x^2 - 3 \cos y, \tan y + 3x \sin y \rangle \) is conservative since

\[
\frac{\partial P}{\partial y} = 3 \sin y = \frac{\partial Q}{\partial x}.
\]

Once we know a vector field is conservative, it will be important to find its potential function. The key to doing this is partial integration. Suppose \( \mathbf{F} = \langle P, Q \rangle \) is conservative, then \( P = f_x \) for some function \( f \). Thus integrating \( P \) with respect to \( x \) will give \( f \) back, up to a constant. One key fact to remember is that any function of \( y \) is a constant when differentiating with respect to \( x \). Thus when integrating \( \int P \, dx \) you have to add not just a constant \( C \), but an arbitrary function of \( y \). We illustrate with an example.
Example 5.2.3. Finding a two-dimensional potential function

Find a potential function for \( \mathbf{F}(x, y) = \langle x^2 - 3 \cos y, \tan y + 3x \sin y \rangle \). First we integrate \( P \) with respect to \( x \), adding an arbitrary constant function.

\[
 f(x, y) = \int P \, dx = \int (x^2 - 3 \cos y) \, dx = \frac{x^3}{3} - 3x \cos y + g(y).
\]

Now we also know \( Q \) = \( f_y \), so we set the partial derivative equal to \( Q \) in order to find \( g(y) \).

\[
 \frac{\partial f}{\partial y} = 3x \sin y + g'(y) = \tan y + 3x \sin y.
\]

Solving for \( g'(y) \) gives \( g'(y) = \tan y \), and integrating gives \( g(y) = -\ln |\cos y| + C \). Putting it all together gives the potential function

\[
 f(x, y) = \frac{x^3}{3} - 3x \cos y - \ln |\cos y|.
\]

We ignore the arbitrary constant \( C \) since we won’t use it for our purposes.

In the three dimensional setting there are more mixed partials to consider. To show the vector field \( \mathbf{F} = \langle P, Q, R \rangle \) is conservative, verify that the following three equations are satisfied:

\[
 \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}; \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}; \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}.
\]

Once you know a vector field is conservative, just apply the two-dimensional process twice to find a potential function.

Example 5.2.4. A three-dimensional example.

Show that

\[
 \mathbf{F}(x, y, z) = \langle 2xyz - 3y^2 + 1, x^2z - 6xy + 5z^2, x^2y + 10yz \rangle
\]
5.2. CONSERVATIVE VECTOR FIELDS

is a conservative vector field, and find a potential function for it.

Taking partials we find
\[
\frac{\partial P}{\partial y} = 2xz - 6y = \frac{\partial Q}{\partial x}
\]
\[
\frac{\partial P}{\partial z} = 2xy = \frac{\partial R}{\partial x}
\]
\[
\frac{\partial Q}{\partial z} = x^2 + 10z = \frac{\partial R}{\partial y},
\]
so \(F\) is conservative.

Integrating \(P\) with respect to \(x\), and noting that the arbitrary constant is now a function of both \(y\) and \(z\), we have
\[
f(x, y, z) = \int 2xyz - 3y^2 + 1\, dx = x^2yz - 3xy^2 + x + g(y, z).
\]
Knowing that \(f_y(x, y, z) = Q\), we solve
\[
x^2z - 6xy + g_y(y, z) = x^2z - 6xy + 5z^2
\]
\[
g_y(y, z) = 5z^2.
\]
Integrating with respect to \(y\), and adding the arbitrary constant \(h(z)\), gives
\[
g(y, z) = \int 5z^2\, dy = 5z^2y + h(z).
\]
Thus far we have \(f(x, y, z) = x^2yz - 3xy^2 + x + 5z^2y + h(z)\). Solving \(f_z(x, y, z) = R\) shows that \(h(z) = 0\), and we have found a potential function
\[
f(x, y, z) = x^2yz - 3xy^2 + x + 5z^2y.
\]

The Fundamental Theorem of Line Integrals: The significance of conservative vector fields comes from physics, in the context of conservation of kinetic and potential energy. For us, the significance of conservative vector fields is that line integrals become relatively straightforward when the vector field is conservative. We now state the analogue of the fundamental theorem of calculus for line integrals.

**Theorem 5.2.2.** Let \(F\) be a conservative vector field with potential function \(f\), and let \(C\) be a path from point \(A\) to point \(B\). Then
\[
\int_C F \cdot dx = f(B) - f(A).
\]

**Proof.** Let \(x(t)\) for \(a \leq t \leq b\) be a parameterization of the curve \(C\). Thus \(x(a) = A\) and \(x(b) = B\). Since \(F = \nabla f\) we have
\[
\int_C F \cdot dx = \int_a^b F(x(t)) \cdot x'(t)\, dt = \int_a^b \nabla f(x(t)) \cdot x'(t)\, dt.
\]
But the chain rule for curves tells us that the integrand \( \nabla f(x(t)) \cdot x'(t) = \frac{d}{dt} f \circ x(t) \). The Fundamental Theorem of Calculus for a single variable then gives
\[
\int_a^b \nabla f(x(t)) \cdot x'(t) dt = \int_a^b \frac{d}{dt} f \circ x(t) dt
\]
\[
= f \circ x(t)|_a^b = f(x(b)) - f(x(a)) = f(B) - f(A). \quad \square
\]

Thus to evaluate the line integral of a conservative vector field, we find the potential function and evaluate at the endpoints. More explicitly, we have:

**Conservative Fields: The Fundamental Theorem of Line Integrals**

Let \( C \) be any path from point \( A \) to point \( B \), and let \( F \) be a conservative vector field. To evaluate \( \int_C F \cdot dx \):

1. Find a potential function \( f \) for \( F \) (this can be thought of as an antiderivative).
2. Use the fact that \( \int_C F \cdot dx = f(B) - f(A) \).

**Example 5.2.5. Using the Fundamental Theorem**

Evaluate the integral of \( F(x,y) = \langle 2xy - 2\cos x, x^2 - \sin y \rangle \) along the quarter circle from \((\pi/2,0)\) to \((0,\pi/2)\). In this instance the vector field is conservative since

\[
\frac{\partial P}{\partial y} = 2x = \frac{\partial Q}{\partial x}.
\]

Figure 5.2.4: Integrating a conservative field
5.2. CONSERVATIVE VECTOR FIELDS

So we find a potential function and apply the theorem. Integrating we have

\[ f(x, y) = \int P \, dx = \int 2xy - 2\cos x \, dx = x^2y - 2\sin x + g(y). \]

Taking the partial with respect to \( y \), and equating to \( Q \) gives

\[ f_y(x, y) = x^2 + g'(y) = x^2 - \sin y. \]

Thus \( g'(y) = -\sin y \), and \( g(y) = \cos y \). The potential function is \( f(x, y) = x^2y - 2\sin x + \cos y \), and we use it to evaluate the integral of \( \mathbf{F} \) over \( C \):

\[ \int_C \mathbf{F} \cdot d\mathbf{x} = f(\pi/2, 0) - f(0, \pi/2) = -1. \]

**Example 5.2.6. Work in three dimensions**

Find the work done by \( \mathbf{F}(x, y, z) = (y - z, x + z, y - x) \) on a particle that travels along the helical arc \( \mathbf{x}(t) = (\cos t, \sin t, t), 0 \leq t \leq \pi \).

Since

\[
\begin{align*}
\frac{\partial P}{\partial y} &= 1 = \frac{\partial Q}{\partial x} \\
\frac{\partial P}{\partial z} &= -1 = \frac{\partial R}{\partial x} \\
\frac{\partial Q}{\partial z} &= 1 = \frac{\partial R}{\partial y}
\end{align*}
\]

the vector field is conservative. By integrating we know \( f(x, y, z) = xy - xz + g(y, z) \), and inspection shows that \( g(y, z) = yz \). Thus a potential function is \( f(x, y, z) = xy - xz + yz \). Since \( \mathbf{x}(0) = (1, 0, 0) \) and \( \mathbf{x}(\pi) = (-1, 0, \pi) \), the fundamental theorem of line integrals implies

\[ w = \int_C \mathbf{F} \cdot d\mathbf{x} = f(-1, 0, \pi) - f(1, 0, 0) = \pi. \]

**Things to Know**

- Different paths \( C_1 \) and \( C_2 \) from point \( A \) to point \( B \) lead to different integrals, in general. In other words:

\[ \int_{C_1} \mathbf{F} \cdot d\mathbf{x} \neq \int_{C_2} \mathbf{F} \cdot d\mathbf{x} \]

in general.

- The following are equivalent definitions of conservative vector fields:

  1. \( \int_C \mathbf{F} \cdot d\mathbf{x} = 0 \) for any closed loop \( C \).
2. \( \int_C \mathbf{F} \cdot d\mathbf{x} \) is independent of the path.

3. There is a potential function \( f \) such that \( \mathbf{F} = \nabla f \).

- To see if \( \mathbf{F} = (P, Q, R) \) is conservative, check equality of partials:
  \[
  \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}; \quad \frac{\partial P}{\partial z} = \frac{\partial R}{\partial x}; \quad \frac{\partial Q}{\partial z} = \frac{\partial R}{\partial y}.
  \]

- To find a potential function for a conservative field \( \mathbf{F} \), use partial integration several times.

- When \( \mathbf{F} \) is conservative, with potential function \( f \), you don’t have to parameterize \( C \) to evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \)! The Fundamental Theorem says
  \[
  \int_C \mathbf{F} \cdot d\mathbf{x} = f(B) - f(A),
  \]
  where \( A \) is the initial and \( B \) the terminal point of the curve \( C \).

**Exercises**

1. Determine which of the following vector fields are conservative. If they are conservative, find a potential function.
   
   (a) \( \mathbf{F}(x, y) = (x^2 - 3xy, y^2 - 3xy) \)
   
   (b) \( \mathbf{F}(x, y) = (3x^2y^2 - 2y\cos x, 2x^3y - 2\sin x - \sin y) \)
   
   (c) \( \mathbf{F}(x, y) = \left( \frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right) \)
   
   (d) \( \mathbf{F}(x, y) = (x, y). \)

2. For each conservative vector field in the previous problem, evaluate the line integral along the curve \( \mathbf{x}(t) = (t, \sin t) \) for \( 0 \leq t \leq 2\pi \).

3. Find the work done by \( \mathbf{F}(x, y) = (\cos y, -x\sin y) \) on a particle moving along the line segment from \((0, 0)\) to \((1, \pi)\).

4. Let the curve \( C \) be parameterized by
   
   \( \mathbf{x}(t) = (e^t \cos t, e^t \sin t), \; 0 \leq t \leq \pi/4, \)
   
   and let \( \mathbf{F} = (3x^2y, x^3) \). Evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \).

5. The electric force that a certain charge placed at the origin exerts on a unit charge at \((x, y, z)\) is given by
   
   \[
   \mathbf{F}(x, y, z) = \frac{1}{(x^2 + y^2 + z^2)^{3/2}} \langle x, y, z \rangle.
   \]

   Find the work done by the electric force on a particle traveling from \((0, 0, 1)\) to \((1, 3, 2)\).
5.3 Green’s Theorem

The physical notion of work is defined to be the force in the direction of movement times the distance traveled. For constant force and direction, we saw that calculating force reduced to taking a dot product. The line integral arose naturally out of an effort to calculate the work done by a variable force along a more general path. So far we have two methods for evaluating line integrals: either parameterize the curve and use the original definition or, in the case of a conservative vector field, find a potential function and evaluate at the endpoints. In this section we will consider another method for calculating line integrals, which can be used when the curve(s) involved form the boundary of a region in the plane.

Another physical situation gives a second interpretation of line integrals. When \( F \) denotes the velocity field of a fluid, then the component of \( F \) in the tangential direction measures how much of the fluid is flowing along the curve. If \( C \) happens to be a closed loop, \( \int_C F \cdot dx \) is called the circulation of \( F \) around \( C \), and it is a measure of how much the fluid is circulating around \( C \) and in what direction (clockwise or counterclockwise).

We begin by considering the special case when \( C \) is a (piecewise) simple closed loop in the plane, oriented counterclockwise, and bounding the rectangle \( R = [a, b] \times [c, d] \). We have the special case of Green’s Theorem:

**Theorem 5.3.1.** Let \( F(x, y) = P(x, y)i + Q(x, y)j \) be a vector field in the plane where \( P \) and \( Q \) have continuous partial derivatives in the rectangle \( R = [s, b] \times [c, d] \). Moreover, let \( C \) be the boundary of \( R \) oriented counterclockwise. Then

\[
\int_C Pdx + Qdy = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA.
\]

**Proof.** The proof amounts to calculating the integrals on both sides of the equation, and noting that they are equal. Since integration “distributes” across sums, we treat the integrals involving \( P \) separately from those involving \( Q \). Let’s start by evaluating those involving \( P \).

To evaluate the line integral \( \int_C Qdy \) we parameterize \( C \), thinking of \( C = C_1 + C_2 + C_3 + C_4 \) where

- \( C_1 \) is the horizontal line segment from \((a, c)\) to \((b, c)\), parameterized by \( x_1(t) = (t, c) \) for \( a \leq t \leq b \).
- \( C_2 \) is the vertical line segment from \((b, c)\) to \((b, d)\), parameterized by \( x_2(t) = (b, t) \) for \( c \leq t \leq d \).
- \( C_3 \) is the horizontal line segment from \((b, d)\) to \((a, d)\), parameterized by \( x_3(t) = (-t, d) \) for \(-b \leq t \leq -a \).
- \( C_4 \) is the vertical line segment from \((a, d)\) to \((a, c)\), parameterized by \( x_4(t) = (a, -t) \) for \(-d \leq t \leq -c \).
CHAPTER 5. VECTOR ANALYSIS

Notice that the parameterizations are a little strange to account for the orientation on the different line segments. When evaluating \( \int_C Q \, dy \) we need to find \( dy \), and notice that \( dy = 0 \) for \( C_1 \) and \( C_3 \) since they are horizontal line segments. Thus

\[
\int_C Q \, dy = \int_{C_2} Q \, dy + \int_{C_4} Q \, dy \\
= \int_c^d Q(b,t) \, dt + \int_a^{−c} Q(a,−t) \, (−dt) \\
= \int_c^d Q(b,t) \, dt + \int_c^d Q(a,u) \, du \\
= \int_c^d Q(b,t) \, dt - \int_c^d Q(a,t) \, dt,
\]

(5.3.1)

where we’ve made the substitution \( u = −t \) to get the third equality. To get the last equality, we changed the label from \( u \) back to \( t \) and changed the order of the limits by taking the negative of the integral.

Now let’s consider the double integral involving \( Q \), and integrate with respect to \( x \) first. We calculate

\[
\iint_R \frac{\partial Q}{\partial x} \, dA = \int_c^d \int_a^b \frac{\partial Q}{\partial x} \, dx \, dy \\
= \int_c^d Q(x,y) \bigg|_{x=a}^b \, dy \\
= \int_c^d Q(b,y) - Q(a,y) \, dy.
\]

(5.3.2)

Comparing Equations 5.3.1 and 5.3.2 we see that \( \int_C Q \, dy = \iint_R \frac{\partial Q}{\partial x} \, dA \), as desired. The analysis that \( \int_C P \, dx = -\iint_R \frac{\partial P}{\partial y} \, dA \) is similar, and left as an exercise for the reader. \( \square \)

Theorem 5.3.1 gives a third method for calculating line integrals: replacing the line integral with a double integral of a related function (when the curve is the boundary of a rectangle oriented counterclockwise).

Example 5.3.1. An initial application

Let \( R \) be the rectangle \([0,1] \times [2,3] \), and \( C \) be its boundary oriented counterclockwise. Moreover, let \( \mathbf{F}(x,y) = (xe^x - y^2, \tan^{-1} y + x) \), and evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \).

Using previous techniques to evaluate \( \int_C P \, dx + Q \, dy \), we’d start by parameterizing the curve, substitute and integrate. Just along the right side of the rectangle \( C_2 \), we’d get \( x(t) = (1,t) \), for \( 2 \leq t \leq 3 \). Substituting gives
\[ \int_{C_2} P \, dx + Q \, dy = \int_2^3 (e - t^2) \cdot 0 + (\tan^{-1} t + 1) \, dt = \int_2^3 \tan^{-1} t + 1 \, dt. \]

This integral alone is mildly unsavory, and we'd have to calculate three others to complete the problem. Luckily Theorem 5.3.1 provides an alternative method of evaluation. To apply it, we first find the integrand, then evaluate the integral. In our case we have

\[ \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial}{\partial x} (\tan^{-1} y + x) - \frac{\partial}{\partial y} (xe^x - y^2) = 1 + 2y. \]

Applying Theorem 5.3.1 gives

\[ \int_C \mathbf{F} \cdot d\mathbf{x} = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \, dA \]
\[ = \int_0^1 \int_2^3 1 - 2ydydx = \int_0^1 y - y^2 \bigg|_2^3 \, dx \]
\[ = \int_0^1 -4dx = -4. \]

Green’s Theorem applies to much more general regions than the rectangles of Theorem 5.3.1. We state a more general form, which still isn’t the most general result.

**Theorem 5.3.2.** Let \( R \) be any region in the plane bounded by the simple closed curve \( C \), oriented counterclockwise. Further, let \( \mathbf{F}(x, y) = P(x, y)\hat{i} + Q(x, y)\hat{j} \) be a vector field where \( P \) and \( Q \) have continuous partial derivatives. Then

\[ \int_C P \, dx + Q \, dy = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \, dA. \]

Rather than a formal proof, we provide a plausible argument. Essentially the same proof generalizes to the case where \( R \) is any region which is either a \( dx \)-region or a \( dy \)-region (see Figure 5.3.1). The only difference is the parameterizations are a little more involved than in the proof of Theorem 5.3.1. For more general regions, subdivide them into sub-regions that are either \( dx \) or \( dy \) and apply the theorem in each of those. Boundaries of regions must be oriented so that the region lies on your left as you traverse the boundary. Thus if a curve bounds a region on both the left and the right, it will be oriented opposite directions in the two regions. Since the integrals along \( C \) and \(-C\) are negatives of each other, the net contribution from one of these curves is zero. For example, the vertical line labeled \( C \) in Figure 5.3.1 contributes positively to the integral around \( R_1 \) and negatively around \( R_3 \), thereby cancelling out. Thus the sum of all the double integrals over the sub-regions is equal to the line integral around the boundary.
Example 5.3.2. Green’s Theorem on a trapezoid

Let \( R \) be the trapezoid with vertices \((0,0), (1,0), (1,1), \) and \((0,2)\), and let \( C \) be its boundary curve oriented counterclockwise. Evaluate \( \int_C xydx - x^2dy \) using Green’s Theorem. We calculate, setting up the limits on \( R \).
5.3. GREEN’S THEOREM

\[ \int_C xy\,dx - x^2\,dy = \int \int_R \frac{\partial}{\partial x}(-x^2) - \frac{\partial}{\partial y}(xy)\,dA = \int \int_R -3x\,dA = \int_0^1 \int_0^{2-x} -3x\,dy\,dx = \int_0^1 -3xy|_{y=0}^{2-x} \,dx = \int_0^1 3x^2 - 6xdx = x^3 - 3x^2|_0^1 = -2. \]

**Scalar Curl:** The integrand in Green’s Theorem is at this point unmotivated, and rather unintuitive. The quantity \( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \) is called the scalar curl of the vector field \( \mathbf{F}(x,y) = P(x,y)i + Q(x,y)j \). It turns out that the scalar curl of a vector field in the plane is one measure of the tendency of the vector field to rotate counterclockwise about a point. To get some intuition for this we look at two vector fields and their scalar curls.

If \( \mathbf{F}(x,y) = \langle x, y \rangle \), then all vectors point away from the origin (see Figure 5.3.3(a)). Intuitively the vector field does not seem to rotate, so a reasonable measure of the rotation of \( \mathbf{F} \) should be zero. The scalar curl is \( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial}{\partial x}y - \frac{\partial}{\partial y}x = 0 \).

As a second example, the vector field \( \mathbf{F}(x,y) = \langle -y, x \rangle \) seems to rotate in a counterclockwise fashion, so we would expect a positive scalar curl (see Figure 5.3.3(b)). Indeed, we calculate \( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial}{\partial x}x - \frac{\partial}{\partial y}(-y) = 2 \).

On the other hand, \( \mathbf{F}(x,y) = \langle y, -x \rangle \) rotates in a clockwise fashion, and has scalar curl \( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial}{\partial x}(-x) - \frac{\partial}{\partial y}y = -2 \) (Figure 5.3.3).

\[ \text{(a) No scalar curl} \quad \text{(b) Positive scalar curl} \quad \text{(c) Negative scalar curl} \]

Figure 5.3.3: Intuitive pictures for the scalar curl

These pictures help our intuition, but of course the situation can be more complicated than this. For example, any conservative vector field is of the form \( \mathbf{F} = \langle P, Q \rangle = \langle f_x, f_y \rangle \) so the scalar curl is \( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} = 0! \) This means that the vector fields in Figures 5.2.3(b) and 5.2.4 have no scalar curl.

There is a more precise interpretation of the scalar curl: it is the (infinitesimal) circulation per unit area. To see this, let \( x_0 \) be a point \( R \) be a disk radius
$\epsilon$ centered at $x_0$. By the mean value theorem for integrals, there is a point $x_\epsilon$ in the disk $R_\epsilon$ such that

$$\int_{R_\epsilon} \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA = \left( \frac{\partial Q}{\partial x}(x_\epsilon) - \frac{\partial P}{\partial y}(x_\epsilon) \right) (\text{Area of } R_\epsilon)$$

Solving for the scalar curl and substituting using Green’s Theorem gives

$$\left( \frac{\partial Q}{\partial x}(x_\epsilon) - \frac{\partial P}{\partial y}(x_\epsilon) \right) = \frac{1}{\text{Area of } R_\epsilon} \int_{R_\epsilon} \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA = \frac{1}{\text{Area of } R_\epsilon} \int_C \mathbf{F} \cdot d\mathbf{x}$$

Now as $\epsilon \to 0$, we see $x_\epsilon \to x_0$, and we have

$$\left( \frac{\partial Q}{\partial x}(x_0) - \frac{\partial P}{\partial y}(x_0) \right) = \lim_{\epsilon \to 0} \frac{1}{\text{Area of } R_\epsilon} \int_C \mathbf{F} \cdot d\mathbf{x},$$

where the integral is the circulation of $\mathbf{F}$ around the boundary of $R_\epsilon$. The expression

$$\frac{1}{\text{Area of } R_\epsilon} \int_C \mathbf{F} \cdot d\mathbf{x}$$

is the circulation per unit area of $\mathbf{F}$, and taking a limit measures the circulation per unit area of $\mathbf{F}$ at the point $x_0$.

We remark at this point that Green’s theorem is consistent with what we know about integrals of conservative vector fields (whose scalar curl is zero). If $\mathbf{F}$ is conservative and $C$ is a closed loop in the plane, then Green’s Theorem gives

$$\int_C \mathbf{F} \cdot d\mathbf{x} = \iint_{R_\epsilon} \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} dA = \iint_R 0 \, dA = 0,$$

as expected.

**Example 5.3.3. Using polar coordinates**

Let $R$ be the quarter of the unit circle in the first quadrant, and $C$ its boundary oriented counterclockwise. Evaluate $\int_C \sin x \, dx + \sqrt{x^2 + y^2} \, dy$ (see Figure 5.3.4). The integrand is

$$\frac{\partial}{\partial x} \sqrt{x^2 + y^2} - \frac{\partial}{\partial y} \sin x = \frac{x}{\sqrt{x^2 + y^2}}.$$

The double integral becomes

$$\int_C \sin x \, dx + \sqrt{x^2 + y^2} \, dy = \iint_R \frac{x}{\sqrt{x^2 + y^2}} \, dA,$$

which is most easily calculated using polar coordinates. Doing so gives:
5.3. GREEN’S THEOREM

Figure 5.3.4: Using polar coordinates in Green’s Theorem

\[ \int_C \sin x \, dx + \sqrt{x^2 + y^2} \, dy = \int \int_R \frac{x}{\sqrt{x^2 + y^2}} \, dA = \int_0^{\pi/2} \int_0^1 r \cos \theta \, r \, dr \, d\theta \]
\[= \int_0^{\pi/2} \int_0^1 r \cos \theta \, dr \, d\theta = \int_0^{\pi/2} \frac{r^2}{2} \cos \theta \bigg|_0^1 \, d\theta \]
\[= \int_0^{\pi/2} \frac{1}{2} \cos \theta \, d\theta = \frac{1}{2} \sin \theta \bigg|_0^{\pi/2} = \frac{1}{2}. \]

We now give an interesting application of Green’s Theorem to calculating area, arising from the fact that the scalar curl of \( \mathbf{F}(x, y) = \langle -y, x \rangle \) is two. Let \( C \) be a simple closed loop bounding the region \( R \) in the plane. Then Green’s Theorem gives

\[ \int_C -y \, dx + x \, dy = \int \int_R 2 \, dA = 2(\text{The Area of } R). \]

**Example 5.3.4. The area of an ellipse**

Find the area enclosed by the ellipse \( \frac{x^2}{9} + \frac{y^2}{4} = 1. \)

Using double integrals, we’d have to calculate

\[ \int_{-3}^{3} \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} dy \, dx = \int_{-3}^{3} \frac{4}{3} \sqrt{9-x^2} \, dx, \]

which can be done using several different approaches, which may not be immediately apparent. We choose to calculate the area using the line integral.

To calculate the line integral, we first parameterize the ellipse by \( \mathbf{x}(t) = (3 \cos t, 2 \sin t) \), for \( 0 \leq t \leq 2\pi \). Then \( dx = -3 \sin t \, dt \) and \( dy = 2 \cos t \, dt \), so that
\[ \int_C -ydx + xdy = \int_0^{2\pi} - (2\sin t)(-3\sin t)dt + 3\cos t 2\cos t dt = \int_0^{2\pi} 6dt = 12\pi. \]

The area of the ellipse is half this, so it’s 6π.

In the previous examples we have used the double integral of Green’s Theorem to make evaluating the line integral easier. This is the first example where the line integral told us something about the double integral.

**Exercises**

1. Verify that \( \int_C Pdx = -\int_R \frac{\partial P}{\partial y} dA \) where \( C \) is the boundary of the rectangle \( R = [a, b] \times [c, d] \) oriented counterclockwise.
2. Evaluate \( \int_C (3x - 2y)dx - xdy \), where \( C \) is the boundary of the region under \( y = 1 - x^2 \) and above the x-axis in two ways:
   
   (a) By parameterizing \( C \) and computing the line integral directly.
   
   (b) By using Green’s Theorem.
3. Evaluate \( \int_C ydx - x^2dy \), where \( C \) is the boundary of the top half of the unit circle in two ways:
   
   (a) By parameterizing \( C \) and computing the line integral directly.
   
   (b) By using Green’s Theorem.
4. Let \( C \) be the square with vertices \( (0, 0) \), \( (1, 0) \), \( (1, 1) \) and \( (0, 1) \) oriented counterclockwise. Find the circulation \( \oint_C F \cdot dx \) of \( F = \langle 3xy^2, -y^3 \rangle \) around \( C \).
5. Evaluate \( \int_C \cos ydx + x^2dy \) where \( C \) is the boundary of the rectangle \( [0, \pi] \times [0, \pi] \) oriented counterclockwise.
6. Compute the work done by the force \( \mathbf{F}(x, y) = \langle x + 2y, xy \rangle \) on a particle traveling once counterclockwise around the triangle with vertices \( (0, 0) \), \( (1, 0) \) and \( (0, 1) \).
7. Evaluate \( \int_C x^2ydx - y^3xdy \) where \( C \) is the unit circle oriented counterclockwise.
8. Let \( R \) be the region in the first quadrant between the circles \( x^2 + y^2 = 1 \) and \( x^2 + y^2 = 4 \), and let \( C \) be its boundary oriented counterclockwise. Evaluate the line integral \( \int_C xy^2dx + x^2ydy \).
9. Let \( R \) be the region in the first quadrant between the circles \( x^2 + y^2 = 1 \) and \( x^2 + y^2 = 4 \), and let \( C \) be its boundary oriented counterclockwise. Evaluate the line integral \( \int_C xy^2dx - x^2ydy \).
10. Use the line integral of Green’s Theorem to find the area enclosed by the ellipse \( \frac{x^2}{16} + \frac{y^2}{25} = 1 \).

11. Use the line integral of Green’s Theorem to find the area enclosed by the ellipse \( \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \).
5.4 Curl and Divergence

In this section we introduce two ways to differentiate vector fields, curl and divergence. The “del” operator,

\[ \nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k} \]

is a standard notational convenience that we will use. It is a vector whose components are differential operators, so \( \nabla \) by itself really has no geometric meaning. Algebraically, however, we will treat it like a vector and use it in cross products and dot products. The convention is that when the vector product involves \( \nabla \) you differentiate rather than multiply. We illustrate with an example.

**Example 5.4.1. A first look at divergence**

Let \( \mathbf{F} = \langle x^2 y, \sin z e^y, 3z^2 \cos x \rangle \), and evaluate \( \nabla \cdot \mathbf{F} \). Remembering to differentiate rather than multiply, we see

\[
\nabla \cdot \mathbf{F} = \left( \frac{\partial}{\partial x} , \frac{\partial}{\partial y} , \frac{\partial}{\partial z} \right) \cdot \left( x^2 y, \sin z e^y, 3z^2 \cos x \right) \\
= \frac{\partial}{\partial x} (x^2 y) + \frac{\partial}{\partial y} (\sin z e^y) + \frac{\partial}{\partial z} (3z^2 \cos x) = 2xy + 2ye^y \sin z + 6z \cos x.
\]

Thus the “del operator” performs specified differentiation operations on vector fields. We make the following definition:

**Definition 5.4.1.** Let \( \mathbf{F} = \langle P, Q, R \rangle \) be a vector field. The divergence \( \text{div} \mathbf{F} \) of \( \mathbf{F} \) is

\[
\text{div} \mathbf{F} = \nabla \cdot \mathbf{F} = \left( \frac{\partial}{\partial x} , \frac{\partial}{\partial y} , \frac{\partial}{\partial z} \right) \cdot \langle P, Q, R \rangle = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.
\]

For vector fields in the plane, just ignore the third coordinate in the above definition. If our vector field is the velocity field of a fluid, one can think of divergence as a measure of the tendency of the fluid to compress or expand at a given point. If \( \nabla \cdot \mathbf{F} = 0 \), the field \( \mathbf{F} \) is sometimes called *incompressible* for this reason. Such a field is also called divergence free.

**Example 5.4.2. Some vector fields in the plane**

When discussing the scalar curl of a vector field we considered the three simple cases of Figure 5.3.3. The divergence of \( \mathbf{F} = \langle x, y \rangle \) is

\[
\text{div} \mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial}{\partial x} (x) + \frac{\partial}{\partial y} (y) = 2,
\]

indicating that the field tends to expand everywhere. This makes intuitive sense by considering Figure 5.3.3(a).
5.4. CURL AND DIVERGENCE

One also immediately calculates that \( \mathbf{F} = \langle -y, x \rangle \) is divergence free since
\[
\text{div}\mathbf{F} = \nabla \cdot \mathbf{F} = \frac{\partial}{\partial x}(-y) + \frac{\partial}{\partial y}(x) = 0,
\]
so that Figure 5.3.3(b) depicts an incompressible field. These examples serve to illustrate that the notions of divergence and scalar curl are different.

We now generalize the notion of scalar curl of a vector field in the plane to the curl of a vector field in \( \mathbb{R}^3 \).

**Definition 5.4.2.** The curl \( \text{curl}\mathbf{F} \) of the vector field \( \mathbf{F} \) is defined to be
\[
\text{curl}\mathbf{F} = \nabla \times \mathbf{F}.
\]

The above definition is the easy way to remember how to calculate curl. If \( \mathbf{F} = \langle P, Q, R \rangle \), then we compute
\[
\text{curl}\mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix}
i & j & k \\
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
P & Q & R
\end{vmatrix}
= \left( \frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right)i + \left( \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right)j + \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)k.
\]

Again, remember that instead of scalar multiplication we differentiate when the del operator is involved.

Thus the curl of a vector field is a vector. One should notice that if \( \mathbf{F} = \langle P, Q, 0 \rangle \) is a vector field in the plane (so \( P \) and \( Q \) are functions of only \( x \) and \( y \)), then
\[
\text{curl}\mathbf{F} = \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right)k.
\]

This follows from the fact that all the partial derivatives in the first two components of curl\( \mathbf{F} \) are zero in this instance. Therefore the scalar curl of a planar vector field \( \mathbf{F} \) is the coefficient of \( k \) in curl\( \mathbf{F} \).

**Example 5.4.3.** Calculating Curl

Let \( \mathbf{F} = \langle xy^2, x^2z, 3yz \rangle \) and calculate curl\( \mathbf{F} \).
\[
\text{curl}\mathbf{F} = \nabla \times \mathbf{F} = \begin{vmatrix}
i & j & k \\
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
xy^2 & x^2z & 3yz
\end{vmatrix}
= \left( \frac{\partial}{\partial y}3yz - \frac{\partial}{\partial z}x^2z \right)i + \left( \frac{\partial}{\partial z}xy^2 - \frac{\partial}{\partial x}3yz \right)j + \left( \frac{\partial}{\partial x}x^2z - \frac{\partial}{\partial y}xy^2 \right)k
= (3z - x^2)i + (2xz - 2xy)k.
\]
The curl of a vector field measures the direction and speed of its tendency to rotate. Intuitively, if $\mathbf{F}$ is the velocity field of a fluid and we place a top in it, the top would tilt to a specific direction and start spinning. The vector $\text{curl}\mathbf{F}$ points in the direction that the top’s axis would point, and the length of $\text{curl}\mathbf{F}$ measures the speed of rotation. For this reason, if $\text{curl}\mathbf{F} = \mathbf{0}$ the vector field $\mathbf{F}$ is called \textit{irrotational}.

There are several facts about curl and divergence that we now mention.

\textbf{Theorem 5.4.1.} The vector field $\mathbf{F}$ is conservative if and only if $\text{curl}\mathbf{F} = \mathbf{0}$.

\textit{Proof.} The fact that an irrotational vector field is conservative follows from Stoke’s Theorem in the next section, so we only prove that conservative vector fields are irrotational here.

Let $\mathbf{F}$ be a conservative vector field. Then it has a potential function $f$, and $\mathbf{F} = \langle f_x, f_y, f_z \rangle$. Using Equation 5.4.1 we see

$$\text{curl}\mathbf{F} = \left( \frac{\partial f}{\partial y} \frac{\partial f}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial f}{\partial y} \right) \mathbf{i} + \left( \frac{\partial f}{\partial z} \frac{\partial f}{\partial x} - \frac{\partial f}{\partial x} \frac{\partial f}{\partial z} \right) \mathbf{j} + \left( \frac{\partial f}{\partial x} \frac{\partial f}{\partial y} - \frac{\partial f}{\partial y} \frac{\partial f}{\partial x} \right) \mathbf{k}.$$ 

Since mixed partials are equal, the right hand side is the zero vector. Thus conservative vector fields are irrotational. □

Another way to say this is $\text{curl}(\nabla f) = \mathbf{0}$. Notice that the $\nabla$ in this notation is not the del operator, rather $\nabla f$ is the gradient of the function $f$. This gives us an easy way to check if a given vector field is conservative: compute the curl. If the curl is the zero vector, then $\mathbf{F}$ is conservative; otherwise, it’s not.

\textbf{Example 5.4.4.} \textit{Conservative Fields}

Determine which fields are conservative. If a field is conservative, find its potential function.

Let $\mathbf{F} = \langle 3x^2y, x + 2y, 2xz \rangle$. Calculating the curl gives

$$\text{curl}\mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial z} & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \\ 3x^2y & x + 2y & 2xz \end{vmatrix}$$

$$= \left( \frac{\partial}{\partial y} 2xz - \frac{\partial}{\partial z} x + 2y \right) \mathbf{i} + \left( \frac{\partial}{\partial z} 3x^2y - \frac{\partial}{\partial x} 2xz \right) \mathbf{j} + \left( \frac{\partial}{\partial x} x + 2y - \frac{\partial}{\partial y} 3x^2y \right) \mathbf{k}$$

$$= -2x \mathbf{j} + (1 - 3x^2) \mathbf{k}.$$ 

Thus our original vector field is not conservative.

Let $\mathbf{F} = \langle 2xy, x^2 + 2yz, y^2 \rangle$. Calculating the curl gives
5.4. CURL AND DIVERGENCE

\[ \text{curl} \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial z} & \frac{\partial}{\partial y} & \frac{\partial}{\partial x} \\ 2xy & y^2 & x^2 + 2yz \end{vmatrix} \]

\[ = \left( \frac{\partial}{\partial y} y^2 - \frac{\partial}{\partial z} x^2 + 2yz \right) \mathbf{i} + \left( \frac{\partial}{\partial z} 2xy - \frac{\partial}{\partial x} y^2 \right) \mathbf{j} + \left( \frac{\partial}{\partial x} x^2 + 2yz - \frac{\partial}{\partial y} 2xy \right) \mathbf{k} \]

\[ = 0. \]

Thus \( \mathbf{F} = \langle 2xy, x^2 + 2yz, y^2 \rangle \) is conservative, and has a potential function \( f(x,y,z) \). To find it we integrate \( P \) with respect to \( x \), but now the arbitrary constant needs to be a function \( g(y,z) \) of both \( y \) and \( z \).

\[ f(x,y,z) = \int 2xy \, dx = x^2y + g(y,z). \]

Now the partial derivative of \( f \) with respect to \( z \) has to be the third coordinate function of \( \mathbf{F} \). Therefore we have

\[ \frac{\partial f}{\partial z} = \frac{\partial}{\partial z} \left( x^2y + g(y,z) \right) = \frac{\partial g}{\partial z} = y^2. \]

Integrating with respect to \( z \) gives that \( g(y,z) = y^2z + h(y) \) for some arbitrary function \( h(y) \) of \( y \). So far we have

\[ f(x,y,z) = x^2y + y^2z + h(y), \]

and taking partials with respect to \( y \) indicates that we can choose \( h(y) = 0 \). Therefore the potential function for \( \mathbf{F} \) is \( f(x,y) = x^2y + y^2z \).

We have one more relationship to highlight, and do so in the following theorem.

**Theorem 5.4.2.** Let \( \mathbf{F} \) be a nice vector field (the component functions have continuous second order partials), then

\[ \text{div curl } \mathbf{F} = 0. \]

Thus the curl of a vector field is always a divergence free vector field. The proof is straightforward, and just relies on equality of mixed partial derivatives.

**Exercises**

1. Determine the curl and divergence of the following vector fields.
   (a) \( \mathbf{F} = \langle x^2 - y, 2xz + y^2, z - 2xy \rangle \)
   (b) \( \mathbf{F} = \langle \sin x, e^y, z^2 \rangle \)
   (c) \( \mathbf{F} = \langle y^3z, xz^2, x + 2y \rangle \)
   (d) \( \mathbf{F} = \langle x - y + 2z, 3x + y - 2z, 3x - 3y + 2z \rangle \)
2. Determine whether the following vector fields are conservative. If so, find their potential functions.

(a) \( \mathbf{F} = \langle x^3, y^2, z \rangle \)
(b) \( \mathbf{F} = \langle x + y, x - y + z, x + y - z \rangle \)

3. Let \( f(x), g(y) \) and \( h(z) \) be single-variable functions. Show that \( \mathbf{F} = \langle f(x), g(y), h(z) \rangle \) is irrotational.

4. Let \( f(y, z), g(x, z) \) and \( h(x, y) \) be two-variable functions. Show that \( \mathbf{F} = \langle f(y, z), g(x, z), h(x, y) \rangle \) is incompressible.

5. Let \( \mathbf{F} = \langle P, Q, R \rangle \), where \( P, Q \) and \( R \) are all linear functions of \( x, y \) and \( z \) (i.e. \( P(x, y, z) = Ax + By + Cz \) for some constants \( A, B \) and \( C \)). Find linear functions for \( P, Q \) and \( R \) so that

(a) \( \mathbf{F} \) is irrotational but not incompressible.
(b) \( \mathbf{F} \) is incompressible but not irrotational.
(c) \( \mathbf{F} \) is both incompressible and irrotational.
(d) \( \mathbf{F} \) is neither incompressible nor irrotational.
5.5 Surface Integrals and Stokes’ Theorem

In this section we introduce integrals of vector fields across surfaces, and establish Stokes’ Theorem, which is a generalization of Green’s Theorem to three dimensions.

We consider the velocity field of fluid flow to motivate the definition of a surface integral. To begin with, suppose a fluid is flowing uniformly with constant velocity vector $\mathbf{v} = 2\mathbf{j}$. In other words, each particle in the fluid is moving 2 meters per second in the positive $y$-direction. Insert a rectangular sieve into the fluid, perpendicular to the flow. Our goal is to compute the volume of water that passes through our sieve in one second. Figure 5.5.1(a) illustrates the sieve together with a few velocity vectors of the fluid flow.

![Figure 5.5.1: Computing Volume crossing a surface orthogonal to flow]

To find the volume of fluid passing through the sieve in one second, notice that particles on the sieve travel to the ends of the velocity vectors. Particles that were 2 units back will end up on the sieve. Thus the rectangular box in Figure 5.5.1(b) represents the fluid that has passed through the sieve in one unit of time. Its volume is the area of the sieve $A$ times the length of the velocity vector $\|\mathbf{v}\|$.

Now consider the case where the sieve is placed at an angle to the flow of the fluid, as in Figure 5.5.2. It is still true that particles starting on the sieve at time zero travel to the tip of the velocity vector, it’s just that this vector is not at a right angle to the sieve. Thus the fluid that flows through the sieve in one second can be represented by the parallelepiped in Figure 5.5.2(b). From three-dimensional geometry, we know this volume can be calculated as the area of the sieve times the distance to the opposite side. This distance is the component of $\mathbf{v}$ in the direction of the normal vector to the surface. Thus the volume of fluid passing through the angled sieve (per unit time) is $V = \mathbf{v} \cdot \|\mathbf{n}\| A$, where $A$ is the area of the sieve.

Let’s now generalize this to a variable velocity field, and calculate the volume of fluid per unit time across an arbitrary surface $S$. The key to this calculation,
like most of calculus, is

\textit{Approximate the desired quantity with ones you can calculate, then take a limit as your approximation gets better.}

Let us suppose our surface $S$ is given to us parametrically by $\mathbf{x}(s,t)$, for $a \leq s \leq b$ and $c \leq t \leq d$, and that our velocity field is $\mathbf{F}$. Since we can calculate the fluid flow of a constant velocity field across a planar region, as in Figure 5.5.2, we will approximate the surface $S$ by small parallelograms. Think of them as shingles on a roof. Then we’ll make the simplifying assumption that our velocity field $\mathbf{F}$ is constant on the parallelogram shingles. With these assumptions, we use the calculation of Figure 5.5.2 to make local calculations, which we add up to approximate the total volume. This is similar to the technique we used to derive the integral formula for the surface area of a parametric surface.

Partition the $s$- and $t$-intervals into subintervals, giving rectangular patches in the domain which get mapped by $\mathbf{x}$ to small patches of the surface $S$ (see Figure 4.5.2). Each small patch of surface is approximated by the parallelogram spanned by the vectors $\mathbf{x}_s \Delta s$ and $\mathbf{x}_t \Delta t$, and the parallelograms have area $A = \|\mathbf{x}_s \times \mathbf{x}_t\| \Delta s \Delta t$. Now assume $\mathbf{F}$ is the constant vector $\mathbf{F}(\mathbf{x}(s_i,t_j))$ over the whole parallelogram, even though it varies over the surface. Finally, recall that $\mathbf{x}_s \times \mathbf{x}_t$ is normal to the surface. We now have the ingredients to calculate (approximately) the volume of fluid flowing across the shingles per unit time (see Figure 5.5.3). Using the formula of Figure 5.5.2 together with our current discussion, we have

$$V_{ij} = \frac{\mathbf{v} \cdot \mathbf{n}}{\|\mathbf{n}\|} A = \frac{\mathbf{F}(\mathbf{x}(s_i,t_j)) \cdot (\mathbf{x}_s \times \mathbf{x}_t)}{\|\mathbf{x}_s \times \mathbf{x}_t\|} \|\mathbf{x}_s \times \mathbf{x}_t\| \Delta s \Delta t.$$  

Cancelling the norms of the cross product, and summing the $V_{ij}$, we get that the total volume of fluid crossing $S$ in unit time is approximately the Riemann sum

$$V \approx \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{F}(\mathbf{x}(s_i,t_j)) \cdot (\mathbf{x}_s \times \mathbf{x}_t) \Delta s \Delta t.$$  

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{vector_analysis.png}
\caption{(a) Angled sieve in the flow (b) Volume per unit time calculation}
\end{figure}
5.5. SURFACE INTEGRALS AND STOKES’ THEOREM

Figure 5.5.3: Local approximation of flux

Taking a limit as our approximation gets better yields the integral

$$\int_a^b \int_c^d \mathbf{F}(\mathbf{x}(s,t)) \cdot (\mathbf{x}_s \times \mathbf{x}_t) \, dt \, ds,$$

which motivates our definition of the integral of a vector field across a surface.

**Definition 5.5.1.** Let $\mathbf{F}$ be a vector field and $S$ a surface given parametrically by $\mathbf{x}(s,t)$ over the domain $R$ in the $st$-plane. The flux of $\mathbf{F}$ across $S$ is the integral

$$\iint_R \mathbf{F}(\mathbf{x}(s,t)) \cdot (\mathbf{x}_s \times \mathbf{x}_t) \, ds \, dt.$$

By the previous discussion, if $\mathbf{F}$ is the velocity field of a fluid flow then the flux of $\mathbf{F}$ across $S$ can be interpreted as the volume of fluid flowing across $S$ in unit time. We now outline the steps for computing flux.

**Computing flux across a parametric surface**

To compute the flux of the vector field $\mathbf{F}$ across the parametric surface $\mathbf{x}(s,t)$, for a region $R$ in the $st$-plane:

1. Calculate the normal vector $\mathbf{x}_s \times \mathbf{x}_t$.
2. Evaluate $\mathbf{F}$ on $\mathbf{x}$ and determine the integrand $\mathbf{F}(\mathbf{x}(s,t)) \cdot (\mathbf{x}_s \times \mathbf{x}_t)$.
3. Evaluate the integral $\iint_R \mathbf{F}(\mathbf{x}(s,t)) \cdot (\mathbf{x}_s \times \mathbf{x}_t) \, ds \, dt$.

One should notice the similarity between the procedure for calculating a flux integral and a line integral. The main difference (other than the dimension) is
that for line integrals you project the vector into the tangential direction while for flux integrals you project into the normal direction to the surface. Because of the similarities with line integrals, and for simplicity, we introduce the notation
\[ \int \int_S F \cdot dS = \int \int_R F(x(s, t)) \cdot (x_s \times x_t) ds dt, \]
for the surface integral of \( F \) across the surface \( S \). For parametric surfaces, then, the area form \( dS \) is shorthand notation for \((x_s \times x_t)dA\).

**Example 5.5.1. Flux across a parametric surface**

Evaluate the flux of \( F(x, y, z) = \langle zx, zy, z \rangle \) across the parametric surface \( x(s, t) = \langle s \cos t, s \sin t, 1 - s^2 \rangle \), for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq \pi/2 \) (see Figure 5.5.4).

Figure 5.5.4: Flux across a parametric surface

First we find the normal vector to the surface:
\[ x_s \times x_t = \begin{vmatrix} i & j & k \\ \cos t & \sin t & -2s \\ -s \sin t & s \cos t & 0 \end{vmatrix} = \langle 2s^2 \cos t, 2s^2 \sin t, s \rangle. \]

Evaluating \( F(s \cos t, s \sin t, 1 - s^2) = \langle (1 - s^2)s \cos t, (1 - s^2)s \sin t, 1 - s^2 \rangle \), so the integrand of the flux integral is
\[ F(x(s, t)) \cdot x_s \times x_t = \langle (1 - s^2)s \cos t, (1 - s^2)s \sin t, 1 - s^2 \rangle \cdot \langle 2s^2 \cos t, 2s^2 \sin t, s \rangle \\
= (1 - s^2) \left( 2s^3 \cos^2 t + 2s^3 \sin^2 t + s - s^3 \right) \\
= (1 - s^2)(s + s^3) = s - s^3. \]

The flux, then is given by the integral
5.5. SURFACE INTEGRALS AND STOKES’ THEOREM

\[
\int_0^{\pi/2} \int_0^1 s - s^5 ds \, dt = \int_0^{\pi/2} \frac{s^2}{2} - \frac{s^6}{6} \, dt = \int_0^{\pi/2} \frac{1}{3} \, dt = \frac{\pi}{6}.
\]

Since the graph of the function \( z = f(x, y) \) can be thought of as the parametric surface \( x(x, y) = (x, y, f(x, y)) \), the formula for flux across a parametric surface translates into one for flux across the graph of a function. Straightforward substitution shows that the flux of \( F(x, y, z) \) across the surface \( S \) given as the graph of \( z = f(x, y) \) above the region \( R \) in the plane is:

\[
\iint_S F \cdot dS = \iint_R F(x, y, f(x, y)) \cdot \langle -f_x(x, y), -f_y(x, y), 1 \rangle \, dA.
\]

This formula leads to the following process for evaluating flux across a graph.

<table>
<thead>
<tr>
<th>Computing flux across the graph of a function</th>
</tr>
</thead>
<tbody>
<tr>
<td>To compute the flux of the vector field ( F ) across the surface ( z = f(x, y) ) over a region ( R ) in the ( xy )-plane:</td>
</tr>
<tr>
<td>1. Calculate the normal vector ( \langle -f_x, -f_y, 1 \rangle ).</td>
</tr>
<tr>
<td>2. Evaluate ( F ) on the surface ( z = f(x, y) ) and determine the integrand ( F(x, y, f(x, y)) \cdot \langle -f_x, -f_y, 1 \rangle ).</td>
</tr>
<tr>
<td>3. Evaluate the integral ( \iint_R F(x, y, f(x, y)) \cdot \langle -f_x(x, y), -f_y(x, y), 1 \rangle , dA ).</td>
</tr>
</tbody>
</table>

**Example 5.5.2. Flux across the graph of a function**

To compute the flux of the field \( F = \langle z, yz, x \rangle \) across the portion of the plane \( 3x + 2y + z = 6 \) above the rectangle \([0, 1] \times [0, 1]\) (see Figure 5.5.5).

Solving for \( z \), we get the plane is the graph of \( f(x, y) = 6 - 3x - 2y \). Thus the normal vector is \( \langle -f_x, -f_y, 1 \rangle = \langle 3, 2, 1 \rangle \). Further, evaluating the vector field on the surface gives \( F(x, y, 6 - 3x - 2y) = \langle 6 - 3x - 2y, 6y - 3xy - 2y^2, x \rangle \), so the flux is given by

\[
\iint_S F \cdot dS = \int_0^1 \int_0^1 \langle 6 - 3x - 2y, 6y - 3xy - 2y^2, x \rangle \cdot \langle 3, 2, 1 \rangle \, dy \, dx
\]

\[=
\int_0^1 \int_0^1 (18 - 8x + 6y - 6xy - 4y^2) \, dy \, dx
\]

\[=
\int_0^1 (18 - 8x)y + 3y^2 - 3xy^2 - \frac{4y^3}{3} \bigg|_{y=0}^{y=1} \, dx
\]

\[=
\int_0^1 59/3 - 11x \, dx = \frac{59}{3} x - \frac{11}{2}x^2 \bigg|_0^1 = \frac{85}{6}.\]

▲
Remark: The flux of a vector field across a surface measures how much of the vector field points across the surface. More precisely, it’s the integral of the component \( \text{comp}_n \mathbf{F} \) of \( \mathbf{F} \) in the direction \( \mathbf{n} \) normal to the surface. If the vector field \( \mathbf{F} \) is always tangent to the surface, then the component \( \text{comp}_n \mathbf{F} \) is zero everywhere and the flux is as well. We formalize these observations in the following results.

**Theorem 5.5.1.** Let \( \mathbf{F} \) be a vector field and \( S \) an oriented surface with normal direction \( \mathbf{n} \) on which \( \mathbf{F} \) is defined. The flux of \( \mathbf{F} \) across \( S \) is the integral on \( S \) of the scalar component of \( \mathbf{F} \) in the \( \mathbf{n} \) direction. In symbols

\[
\iint_S \mathbf{F} \cdot d\mathbf{S} = \int_S \text{comp}_n \mathbf{F} d\mathbf{S}.
\]

Proof. We prove it for the parametric surface \( S(s,t) \) defined over the region \( R \) in the \( st \)-plane with normal vector \( S_s \times S_t \). By definition 5.5.1, some algebraic trickery, and definition ???, we have

\[
\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_R \mathbf{F}(S(s,t)) \cdot (S_s \times S_t) dsdt
\]

\[
= \iint_R \mathbf{F}(S(s,t)) \cdot \frac{(S_s \times S_t)}{\|S_s \times S_t\|} \|S_s \times S_t\| dsdt
\]

\[
= \iint_R \mathbf{F}(S(s,t)) \cdot \frac{(S_s \times S_t)}{\|S_s \times S_t\|} \|S_s \times S_t\| dsdt
\]

\[
= \iint_S \mathbf{F}(S(s,t)) \cdot \frac{(S_s \times S_t)}{\|S_s \times S_t\|} dS
\]

\[
= \iint_S \text{comp}_n \mathbf{F} d\mathbf{S}. \quad \blacksquare
\]

Note that Theorem 5.5.1 interprets the flux of a vector field across a surface as the integral of a scalar function over the surface as defined in Section 4.6.
Thus we see that the integral of a function over a surface can provide interesting information, depending on what the function is. We take the time to point out two corollaries of Theorem 5.5.1.

**Corollary 5.5.1.** If the vector field $\mathbf{F}$ is always tangent to the surface $S$, then

$$\iint_S \mathbf{F} \cdot dS = 0.$$ 

*Proof.* If $\mathbf{F}$ is tangent to $S$, the component of $\mathbf{F}$ in the normal direction is zero. By Theorem 5.5.1 we have

$$\iint_S \mathbf{F} \cdot dS = \iint_S \text{comp}_n \mathbf{F} dS = \iint_S 0 dS = 0.$$ 

**Corollary 5.5.2.** Let $S$ be a surface with normal vector $\mathbf{n}$ and area $A$. Further, suppose $\mathbf{F}$ is a vector field on $S$ whose component in the $\mathbf{n}$ direction is constant, so that $\text{comp}_n \mathbf{F} = c$ for some constant $c$. Then the flux of $\mathbf{F}$ across $S$ is

$$\iint_S \mathbf{F} \cdot dS = cA.$$ 

*Proof.* Recall that $\iint_S dS = A$. Applying Theorem 5.5.1 we get

$$\iint_S \mathbf{F} \cdot dS = \iint_S \text{comp}_n \mathbf{F} dS = \iint_S c dS = c \iint_S dS = cA.$$ 

We now consider several examples of Corollaries 5.5.1 and 5.5.2.

**Example 5.5.3.** Tangential $\mathbf{F}$ means $\iint_S \mathbf{F} \cdot dS = 0$

Show that the flux of the vector field $\mathbf{F}(x,y,z) = \langle x, y, -2z \rangle$ across any portion of the surface $(x^2 + y^2)z = 1$ is zero.

Our approach will be to show that $\text{comp}_n \mathbf{F} = 0$. The surface is the level surface of the function $g(x,y,z) = (x^2 + y^2)z$ at level one. Since gradients are perpendicular to level surfaces, a normal vector to the surface is

$$\mathbf{n} = \nabla g = \langle 2xz, 2yz, x^2 + y^2 \rangle.$$ 

But then

$$\mathbf{F} \cdot \mathbf{n} = (x, y, -2z) \cdot \langle 2xz, 2yz, x^2 + y^2 \rangle = 2x^2z + 2y^2z - 2z(x^2 + y^2) = 0,$$

which implies $\mathbf{F}$ is tangent to the surface everywhere, and the flux is zero (see Figure 5.5.6). In this last example the gradient $\nabla g$ was perpendicular to $\mathbf{F}$ for all points $(x,y,z)$. We never had to use the fact that $(x^2 + y^2)z = 1$ to make the dot product $\mathbf{F} \cdot \mathbf{n}$ equal zero. This means that the flux across every level surface of $g$ will be zero, not just $g(x,y,z) = 1$. We include one more example which requires one to evaluate $\mathbf{F}$ on the given surface to ensure.
Example 5.5.4. Another tangent field with zero flux.

Show that the vector field $F = \langle x - yz, y + xz, z \rangle$ has zero flux across the cone $z = \sqrt{x^2 + y^2}$.

To do this we’ll show that $F$ is tangent to the cone by showing it is perpendicular to the normal vector. Since the cone is the graph of a function $f(x, y) = \sqrt{x^2 + y^2}$, normal vectors are scalar multiples of

$$\langle -fx, -fy, 1 \rangle = \left\langle -\frac{x}{\sqrt{x^2 + y^2}}, -\frac{y}{\sqrt{x^2 + y^2}}, 1 \right\rangle.$$

We multiply by the scalar $\sqrt{x^2 + y^2}$ and use the normal vector

$$n = \left\langle -x, -y, \sqrt{x^2 + y^2} \right\rangle.$$

Calculating the dot product gives

$$F \cdot n = (x - yz, y + xz, z) \cdot \left\langle -x, -y, \sqrt{x^2 + y^2} \right\rangle = -x^2 - y^2 + z \sqrt{x^2 + y^2}.$$

This is not zero everywhere, but we are only concerned with the field on the cone. Therefore we substitute $\sqrt{x^2 + y^2}$ for $z$ in the dot product, obtaining

$$F \cdot n = -x^2 - y^2 + z \sqrt{x^2 + y^2} = -x^2 - y^2 + \left(\sqrt{x^2 + y^2}\right)^2 = 0. \blacktriangle$$

Example 5.5.5. Constant normal component

Show that the vector field

$$F(x, y, z) = \left\langle \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, 0 \right\rangle$$

is constant normal component.
has a constant component in the normal direction to the cone \( z = \sqrt{x^2 + y^2} \). Then use Corollary 5.5.2 to calculate the flux of \( \mathbf{F} \) across the portion \( S \) of the cone with \( 0 \leq z \leq 1 \).

The induced orientation on the surface \( z = \sqrt{x^2 + y^2} \) is the normal vector

\[
\mathbf{n} = \left\langle -\frac{x}{\sqrt{x^2 + y^2}}, -\frac{y}{\sqrt{x^2 + y^2}}, 1 \right\rangle,
\]

so the component of \( \mathbf{F} \) normal to the surface is

\[
\text{comp}_n \mathbf{F} = \left\langle \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \frac{y}{\sqrt{x^2 + y^2 + z^2}}, 0 \right\rangle \cdot \left\langle -\frac{x}{\sqrt{x^2 + y^2}}, -\frac{y}{\sqrt{x^2 + y^2}}, 1 \right\rangle \sqrt{x^2 + y^2 + z^2}.
\]

Now on the cone we have \( z^2 = x^2 + y^2 \), and substituting gives

\[
\text{comp}_n \mathbf{F} = -\frac{1}{\sqrt{2}} \sqrt{\frac{x^2 + y^2}{x^2 + y^2 + z^2}} = -\frac{1}{2} \sqrt{\frac{x^2 + y^2}{2(x^2 + y^2)} = -\frac{1}{2}}.
\]

By Corollary 5.5.2 and the fact that the lateral surface area of a cone radius \( r \) and height \( h \) is \( A_L = 2\pi r \sqrt{r^2 + h^2} \), we calculate the flux to be

\[
\int_S \mathbf{F} \cdot d\mathbf{S} = -\frac{1}{2} 2\pi \sqrt{2} = -\pi \sqrt{2}. \quad \checkmark
\]

We conclude with an important example, the flux of an electric field due to a point charge across a sphere.

**Example 5.5.6. Electric flux across a sphere**

Find the flux of

\[
\mathbf{E}(x, y, z) = \left\langle \frac{x}{(x^2 + y^2 + z^2)^{3/2}}, \frac{y}{(x^2 + y^2 + z^2)^{3/2}}, \frac{z}{(x^2 + y^2 + z^2)^{3/2}} \right\rangle
\]

across the sphere \( S \) radius 2, centered at the origin, with outward pointing normal.

To do so, we calculate the component of \( \mathbf{E} \) in the normal direction of \( S \). On the sphere we have \( x^2 + y^2 + z^2 = 4 \), so evaluating \( \mathbf{E} \) on \( S \) gives

\[
\mathbf{E}(x, y, z) = \left\langle \frac{x}{(4)^{3/2}}, \frac{y}{(4)^{3/2}}, \frac{z}{(4)^{3/2}} \right\rangle = \frac{1}{8} \left\langle x, y, z \right\rangle.
\]
Moreover, an outward pointing normal vector is \( \mathbf{n} = \langle x, y, z \rangle \), with length \( \| \mathbf{n} \| = 2 \) since \((x, y, z)\) is on \( S \). Using these observations we calculate

\[
\text{comp}_n \mathbf{E} = \mathbf{E} \cdot \frac{\mathbf{n}}{\| \mathbf{n} \|} = \frac{1}{8} \langle x, y, z \rangle \cdot \frac{\langle x, y, z \rangle}{2} = \frac{1}{16} \left( x^2 + y^2 + z^2 \right) = \frac{1}{4}.
\]

Since \( \text{comp}_n \mathbf{E} \) is constant we can apply Corollary 5.5.2 to evaluate the flux. Further, the area of \( S \) is \( 16\pi \), so

\[
\iint_S \mathbf{E} \cdot dS = cA = \frac{1}{4} 16\pi = 4\pi. \quad \blacktriangle
\]

We are now in position to discuss Stokes’ Theorem, which is a generalization of Green’s Theorem to surfaces in three dimensions. Recall that Green’s Theorem states that under the right circumstances

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \iint_R \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \, dA.
\]

Green’s Theorem equates the circulation of \( \mathbf{F} \) around a closed loop \( C \) to the integral of the scalar curl over the region in the plane that \( C \) bounds. Thus it can be thought of as equating a line integral along a closed loop to a surface integral over a (planar) surface that the loop bounds. In three dimensions the surface need not be planar, and the scalar curl is replaced by the vector curl of \( \mathbf{F} \).

In Green’s theorem it was assumed that \( C \) was oriented so that the region \( R \) lies to your left. Similarly, there are orientation issues involved with Stokes’ Theorem. Given a surface \( S \) in space, there are two options for choosing a normal direction. For simplicity, we’ll say up and down. Let \( \mathbf{n} \) be a choice of normal direction for \( S \), then \( \mathbf{n} \) induces an orientation on the boundary \( C = \partial S \) as follows: choose a direction on \( C \) so that if you walk that direction along \( C \), standing in the \( \mathbf{n} \) direction, the surface \( S \) lies to your left. Given a surface \( S \) and choice of normal direction \( \mathbf{n} \), we call this the \textit{induced} orientation on \( C \). We now state Stokes’ Theorem.

\textbf{Theorem 5.5.2.} Let \( \mathbf{F} \) be a vector field and \( S \) an oriented surface with boundary curve \( C \) having the orientation induced by \( S \). Then

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \iint_S (\text{curl} \mathbf{F}) \cdot dS.
\]

The proof of this theorem has some clever components to it. One evaluates the surface integral pretty much by brute force, and it gets ugly. The line integral also involves some formal analysis with parameterizations, but a nice application of the multivariable chain rule enables one to use Green’s Theorem in the plane to turn the line integral into a double integral over a region. This is the slick part of the proof. Once that’s accomplished one checks and voila—the two integrals are equal!
The details for parametric surfaces are rather cumbersome. For graphs of functions, they’re not quite so bad, but the algebra is a little lengthy. For now, we content ourselves with looking at some applications.

Example 5.5.7. Verifying Stokes’ Theorem

Let \( \mathbf{F} = \langle x + y, y - z, z - x \rangle \) and \( S \) be the portion of the plane \( x + y - z = 1 \) inside the cylinder \( x^2 + y^2 = 1 \) oriented by the upward pointing normal. The boundary curve \( C \), then, is the intersection of the plane and cylinder, and the induced orientation on \( C \) is counterclockwise when viewed from above. Verify Stokes’ Theorem in this case by computing both the line and surface integrals.

Computing \( \int_C \mathbf{F} \cdot d\mathbf{x} \)

To compute the line integral directly, we first parameterize the boundary curve. The \( x \) and \( y \) coordinates satisfy \( x^2 + y^2 = 1 \), so they can be parameterized by \( x = \cos t \) and \( y = \sin t \), for \( 0 \leq t \leq 2\pi \). Now substitute these into the equation \( z = x + y - 1 \) of the plane to find parametric equations for the ellipse. Thus our curve, and its tangent vector, are

\[
\mathbf{x}(t) = (\cos t, \sin t, \cos t + \sin t - 1) \quad \text{for } 0 \leq t \leq 2\pi
\]

\[
\mathbf{x}'(t) = (-\sin t, \cos t, \cos t - \sin t).
\]

The integrand for the line integral is

\[
\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = (\cos t + \sin t, 1 - \cos t, \sin t - 1) \cdot (-\sin t, \cos t, \cos t - \sin t)
\]

\[
= \sin t - \sin^2 t - 1 = \sin t + \frac{1}{2} \cos 2t - \frac{3}{2},
\]

where the last equality follows from the half-angle formula \( \sin^2 t = \frac{1}{2}(1 - \cos 2t) \). Integrating gives

\[
\int_0^{2\pi} \sin t + \frac{1}{2} \cos 2t - \frac{3}{2} \, dt = -\cos t + \frac{1}{4} \sin 2t - \frac{3}{2} \left|_0^{2\pi} = -3\pi.\right.
\]

Computing \( \iint_S (\text{curl} \, \mathbf{F}) \cdot d\mathbf{S} \)

To compute the surface integral we need to describe \( S \) parametrically or as the graph of a function of two variables. In this case, \( S \) is the portion of the plane \( z = x + y - 1 \) lying over the unit disk with upward pointing normal. We want to evaluate

\[
\iint_{\text{unit disk}} \text{curl} \, \mathbf{F} \cdot \langle -f_x, -f_y, 1 \rangle \, dA.
\]

Since in our case \( z = f(x, y) = x + y - 1 \), we observe that \( \langle -f_x, -f_y, 1 \rangle = \langle -1, -1, 1 \rangle \). Calculating the curl gives
CHAPTER 5. VECTOR ANALYSIS

(a) Circulation of $\mathbf{F}$ around $C$  
(b) Flux of $\text{curl} \mathbf{F}$ across $S$

Figure 5.5.7: Verifying Stokes’ Theorem

\[
\text{curl } \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ x + y & y - z & z - x \end{vmatrix} = \langle 1, 1, -1 \rangle.
\]

Thus the surface integral in Stokes’ Theorem becomes

\[
\int \int_{\text{unit disk}} (1, 1, -1) \cdot (-1, -1, 1) \, dA = \int \int_{\text{unit disk}} -3 \, dA = -3\pi,
\]

since the integral is $-3$ times the area of the unit disk. Thus, in this case, Stokes’ Theorem has been verified.

Example 5.5.8. Using the surface integral of Stokes’ Theorem

Use Stokes’ Theorem to evaluate $\int_C \mathbf{F} \cdot d\mathbf{x}$, where $C$ is the triangle with vertices $(-1, 0, 2)$, $(1, 1, -3)$, and $(0, 3, 5)$ oriented counterclockwise when viewed from above and $\mathbf{F} = \langle yz, xz, xy \rangle$.

In this case

\[
\text{curl } \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz & xz & xy \end{vmatrix} = \langle 0, 0, 0 \rangle,
\]

so the surface integral of Stokes’ Theorem is always zero. Thus $\int_C \mathbf{F} \cdot d\mathbf{x} = 0$. Of course, we’ve seen that $\text{curl } \mathbf{F} = 0$ means that $\mathbf{F}$ is conservative, so the integral around any closed loop will be zero. Let’s do another example:
5.5. SURFACE INTEGRALS AND STOKES' THEOREM

Use Stokes' Theorem to evaluate \( \int_C \mathbf{F} \cdot d\mathbf{x} \), where \( C \) is the triangle with vertices \((2, 0, 0), (0, 1, 0), \) and \((0, 0, 2)\) oriented counterclockwise when viewed from above and \( \mathbf{F} = \langle x^2, z^2, y^2 \rangle \).

First we calculate curl, in case that makes our lives easier. This time we get

\[
\text{curl } \mathbf{F} = \left| \begin{array}{ccc}
\mathbf{i} & \mathbf{j} & \mathbf{k} \\
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
x^2 & z^2 & y^2
\end{array} \right| = \langle 2y - 2z, 0, 0 \rangle.
\]

Now we must describe our surface, given by the vertices of a triangle, parametrically or as the graph of a function of two variables. The triangle is part of a plane with the given intercepts. It isn’t too hard to verify an equation for the plane is \( x + 2y + z = 2 \), so solving for \( z \) we get the surface \( S \) is a portion of the graph of \( z = 2 - x - 2y \). In particular, it’s the portion above the triangle in the first quadrant of the \( xy \)-plane and bounded by \( x + 2y = 2 \). The upward pointing normal vector is \( \langle -f_x, -f_y, 1 \rangle = \langle 1, 2, 1 \rangle \).

We now evaluate curl \( \mathbf{F} \) on the surface, which means analytically that we use the equation of the surface to eliminate \( z \) from the curl, giving

\[
\text{curl } \mathbf{F} = \langle 2y - 2(2 - x - 2y), 0, 0 \rangle = \langle -4 + 6y + 2x, 0, 0 \rangle.
\]

The surface integral is
\[
\iint_S (-4 + 6y + 2x, 0, 0) \cdot (1, 2, 1) \, dA = \int_0^1 \int_0^{2-2y} -4 + 6y + 2x \, dx \, dy \\
= \int_0^1 (-4 + 6y)x + x^2 \bigg|_{x=0}^{2-2y} \, dy \\
= \int_0^1 (-4 + 6y)(2 - 2y) + (2 - 2y)^2 \, dy \\
= \int_0^1 -2y^2 + 3y - 1 \, dy = \frac{-2}{3}.
\]

**Example 5.5.9. Using the line integral of Stokes’ Theorem**

Use Stokes’ Theorem to evaluate \( \iint_S \text{curl} \mathbf{F} \cdot dS \) where \( \mathbf{F} = (xz, xy, yz) \) and \( S \) is the portion of the graph of \( z = 1 - \frac{x^2}{4} - y^2 \) lying above the \( xy \)-plane with upward pointing normal.

To evaluate the line integral of Stokes’ Theorem, we need to parameterize the boundary of \( S \), which is the intersection of the surface \( z = 1 - \frac{x^2}{4} - y^2 \) with the \( xy \)-plane. This is the ellipse \( \frac{x^2}{4} + y^2 = 1 \) which is parameterized by \( \mathbf{x}(t) = (2 \cos t, \sin t, 0) \) (recall that the curve is in \( \mathbb{R}^3 \), so we need the third coordinate even though it’s always zero). The integrand for the line integral is

\[
\mathbf{F}(\mathbf{x}(t)) \cdot \mathbf{x}'(t) = \langle 0, 2 \cos t \sin t, 0 \rangle \cdot \langle -2 \sin t, \cos t, 0 \rangle = 2 \cos^2 t \sin t.
\]

Evaluating we have

\[
\int_C \mathbf{F} \cdot d\mathbf{x} = \int_0^{2\pi} 2 \cos^2 t \sin t \, dt = -\frac{2}{3} \cos^3 t \bigg|_0^{2\pi} = 0.
\]
Notice that the circulation is zero because the positive and negative circulation cancel each other out. Just because the circulation is zero that doesn’t mean \( \mathbf{F} \) is orthogonal to \( C \) everywhere.

We make a final remark regarding Stokes’ Theorem. Notice that the same curve \( C \) is the boundary of many surfaces, and that Stokes’ Theorem applies to all of them. Thus the flux integral of \( \text{curl} \ \mathbf{F} \) along all of them is the same value, just the value of the line integral \( \int_C \mathbf{F} \cdot d\mathbf{x} \). We use this in the following example.

**Example 5.5.10. Switching surfaces**

Let \( S \) be the sides of the pyramid that has base \([0, 2] \times [0, 2]\) in the \( xy\)-plane, and top vertex at the point \((1, 1, 1)\). Thus \( S \) consists of four triangles meeting at \((1, 1, 1)\). Suppose \( S \) is oriented with an upward pointing normal, then \( C \) is the square in the \( xy\)-plane bounding the base of the pyramid. Let \( \mathbf{F} = \langle xz^2, x^2 + y^2z^2, x^3 \rangle \), and evaluate \( \int\int_S (\text{curl} \ \mathbf{F}) \cdot d\mathbf{S} \).

By the remarks preceding the example, we can exchange \( S \) for a more convenient surface \( S' \) with the same boundary to do the calculation. We choose the base of the pyramid, with upward pointing normal \( \langle 0, 0, 1 \rangle \). We compute the curl, evaluate it on the surface \( S' \), and compute the surface integral.

\[
\text{curl} \ \mathbf{F} = \begin{vmatrix}
i & j & k \\
\frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\
xz^2 & x^2 + y^2z^2 & x^3
\end{vmatrix} = \langle -2y^2z, 2xz - 3x^2, 2x \rangle.
\]

Evaluating the curl on the surface \( z = 0 \) gives \( \text{curl} \ \mathbf{F} = \langle 0, -3x^2, 2x \rangle \), and the integrand is \( \text{curl} \ \mathbf{F} \cdot \langle 0, 0, 1 \rangle = \langle 0, -3x^2, 2x \rangle \cdot \langle 0, 0, 1 \rangle = 2x \). We have

\[
\int\int_{S'} (\text{curl} \ \mathbf{F}) \cdot d\mathbf{S} = \int_0^2 \int_0^2 2xdxdy = \int_0^2 4dy = 8.
\]

**Exercises**

1. Evaluate the flux of \( \mathbf{F} = \langle 1, 1, 1 \rangle \) across the portion of the plane \( z = xy \) above the unit disk.

2. Evaluate the flux of \( \mathbf{F} = \langle 2x - z, x + 3z, y - x \rangle \) across the paraboloid \( z = 4 - x^2 - y^2 \) above the rectangle \([-1, 1] \times [-1, 1]\).

3. Evaluate the flux of \( \mathbf{F} = \langle x, z^2, 2y - x \rangle \) across the portion of the cone \( z = 1 - \sqrt{x^2 + y^2} \) above the unit disk.

4. Evaluate the flux of \( \mathbf{F} = \langle xy, z^2 - y, x^2 - 2zy \rangle \) across the portion of the plane \( 2x + 3y + z = 6 \) in the first octant.

5. Evaluate the flux of \( \mathbf{F} = \langle x, y, z \rangle \) across the portion of the paraboloid \( z = 3 + x^2 + y^2 \) above the unit disk.

6. Evaluate the flux of \( \mathbf{F} = \langle 3x + 2y, z, y \rangle \) across the portion of the plane \( 2x - y + z = 6 \) above the rectangle \([0, 1] \times [0, 1]\).
CHAPTER 5. VECTOR ANALYSIS

7. Evaluate the flux of \( \mathbf{F} = \langle x - 2y, x + y, z \rangle \) across the portion of the plane \( y + z = 1 \) inside the cylinder \( x^2 + y^2 = 1 \).

8. Evaluate the flux of \( \mathbf{F} = \langle y + z, x + z, y \rangle \) across the portion of the saddle \( z = x^2 - y^2 \) above the rectangle \([0, 1] \times [0, 1]\).

9. Find the Flux of \( \mathbf{F} = \langle x^2, y + z, x - z \rangle \) across the parallelogram with vertices \((0, 0, 0), (1, 2, -1), (3, 2, 1)\) and \((4, 4, 0)\).

10. Find the Flux of \( \mathbf{F} = \langle 2x - y, 3y + 2z, z - 2x \rangle \) across the parallelogram with vertices \((1, 0, -3), (2, 4, 0), (-2, 1, 3),\) and \((-1, 5, 0)\).

11. Parameterize the portion of the cone \( z = \sqrt{x^2 + y^2} \) above the unit disk, and find the flux of \( \mathbf{F} = \langle y, -z, x \rangle \) across it.

12. Evaluate the flux of \( \mathbf{F} = \langle 0, 0, 3 \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (2s - 3t + 4, s + 2t, 3s - t) \) for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq 1 \).

13. Evaluate the flux of \( \mathbf{F} = \langle 1, 1, 2 \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (t + s, t - s, s) \) for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq 1 \).

14. Evaluate the flux of \( \mathbf{F} = \langle xy, 2z, z + 2x \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (2 - t + 3s, s, t - s) \) for \( 0 \leq s \leq 1 \) and \( 0 \leq t \leq 1 \).

15. Evaluate the flux of \( \mathbf{F} = \langle y^2 - x, z, y \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (t^2 - s, st, 2 - s^2) \) for \( 0 \leq s, t \leq 1 \).

16. Evaluate the flux of \( \mathbf{F} = \frac{1}{\sqrt{x^2 + y^2 + z^2}} \langle -x, -y, -z \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (\cos t, \sin t, s) \) for \(-1 \leq s \leq 1\) and \(0 \leq t \leq \pi/2\).

17. Evaluate the flux of \( \mathbf{F} = \langle y, -x, z^2 \rangle \) across the parametric surface \( \mathbf{x}(s, t) = (s \cos t, s \sin t, t) \) for \(0 \leq s \leq 1\) and \(0 \leq t \leq \pi\).

18. Evaluate the flux of \( \mathbf{F} = \langle xz^2, yz^2, z^3 \rangle \) across the unit sphere \( \mathbf{x}(s, t) = (\sin s \cos t, \sin s \sin t, \cos s) \) for \(0 \leq s \leq \pi\) and \(0 \leq t \leq 2\pi\).

19. Find the flux of \( \mathbf{F} = \langle z, xy, x + y + 2 \rangle \) across the portion of the paraboloid \( z = x^2 + y^2 \) that lies above the unit disk.

20. A current moving positively along the y-axis induces the magnetic field

\[
\mathbf{B}(x, y, z) = \left( \frac{z}{x^2 + y^2}, 0, \frac{-x}{x^2 + y^2} \right).
\]

Find the magnetic flux of \( \mathbf{B} \) across the surface \( \mathbf{S}(s, t) = (t + 1, s, t - 1) \) for \(0 \leq s, t \leq 1\).

21. A point charge placed at the origin induces (a scalar multiple of) the electric field

\[
\mathbf{E}(x, y, z) = \left( \frac{x}{(x^2 + y^2 + z^2)^{3/2}}, \frac{y}{(x^2 + y^2 + z^2)^{3/2}}, \frac{z}{(x^2 + y^2 + z^2)^{3/2}} \right).
\]
5.5. **SURFACE INTEGRALS AND STOKES’ THEOREM**

Find the flux of \( \mathbf{E} \) across the hemisphere \( z = \sqrt{1 - x^2 - y^2} \).

22. Show that the flux of \( \mathbf{F}(x, y, z) = \langle zy, xz, 2xy \rangle \) across any portion of the one-sheeted hyperboloid \( x^2 + y^2 - z^2 = 1 \) is zero.

23. Show that the flux of \( \mathbf{F}(x, y, z) = \langle x, y, z \rangle \) across any portion of the cone \( \mathbf{S}(s, t) = (s \cos t, s \sin t, s) \), \( 0 \leq s, 0 \leq t \leq 2\pi \) is zero.

24. Show that the flux of \( \mathbf{F}(x, y, z) = \langle x, y, 2z \rangle \) across any portion of the paraboloid \( z = x^2 + y^2 \) is zero.

25. Show that the flux of \( \mathbf{F}(x, y, z) = \langle 1 - z - x^2, -xy, x - xz \rangle \) across any portion of the unit sphere is zero.

26. Show that the flux of \( \mathbf{F}(x, y, z) = \langle -xy, 1 - z - y^2, y - yz \rangle \) across any portion of the unit sphere is zero. (The vector field restricted to the sphere is the image of \( \mathbf{j} \) under stereographic projection).

27. Verify Stokes’ Theorem for the constant vector field \( \mathbf{F}(x, y, z) = \langle 2, 1, 3 \rangle \) and the portion of the paraboloid \( z = 4 - x^2 - y^2 \) lying on or above the \( xy \)-plane.

28. Verify Stokes’ Theorem for the constant vector field \( \mathbf{F}(x, y, z) = \langle 1, 4, 7 \rangle \) and the portion of the cone \( z = 1 - \sqrt{x^2 + y^2} \) lying on or above the \( xy \)-plane.

29. Verify Stokes’ Theorem for the vector field \( \mathbf{F}(x, y, z) = \langle y, 2z, 3x \rangle \) and the portion of the saddle \( z = x^2 - y^2 \) lying “above” the unit disk. Note the boundary curve is \( \mathbf{C}(t) = (\cos t, \sin t, \cos^2 t - \sin^2 t) \).

30. Verify Stokes’ Theorem for the vector field \( \mathbf{F}(x, y, z) = \langle 2z, 3x, y \rangle \) and the portion of the saddle \( z = 2x + 3y \) lying “above” the unit disk. Note the boundary curve is \( \mathbf{C}(t) = (\cos t, \sin t, 2\cos t + 3\sin t) \).

Use Stokes’ Theorem to evaluate \( \int_{C} \mathbf{F} \cdot d\mathbf{x} \) where \( C \) is the boundary of the surfaces with the induced orientation in exercises 31-33.

31. \( \mathbf{F} = \langle \cos x, y^3, \frac{1}{1+x^2} \rangle \), \( S \) is the portion of the plane \( x - 2y + z = 4 \) inside the cylinder \( x^2 + y^2 = 1 \) with upward pointing normal.

32. \( \mathbf{F} = \langle x + 2y, 3x - 4y, 2z + 4x \rangle \), \( S \) is the portion of the surface \( z = 1 - x^2 - y^2 \) above the \( xy \)-plane with upward pointing normal.

33. \( \mathbf{F} = \langle xz, yz, x^2 \rangle \), \( S \) is the surface \( \mathbf{x}(s, t) = (s \cos t, s \sin t, t) \), for \( 0 \leq s \leq 1 \), \( 0 \leq t \leq \pi/2 \) with normal vector \( \mathbf{x}_s \times \mathbf{x}_t \).

In exercises 34-36 use Stokes’ Theorem to evaluate \( \iint_{S} \text{curl} \ \mathbf{F} \cdot d\mathbf{S} \).

34. \( \mathbf{F} = \langle x, z + 2x - 1, x^2 + y^2 \rangle \), \( S \) is the portion of the plane \( 2x + y + z = 1 \) inside the cylinder \( x^2 + y^2 = 1 \) with upward pointing normal.
35. \( \mathbf{F} = (-y, x, z) \), and \( S \) is the portion of the sphere \( x^2 + y^2 + z^2 = 4 \) lying above the plane \( z = 1 \) with upward pointing normal.

36. \( \mathbf{F} = \langle xy + z, 2z, y^2 + 2z \rangle \), and \( S \) is the portion of the ellipsoid \( \frac{x^2}{9} + y^2 + \frac{z^2}{16} = 1 \) lying above the \( xy \)-plane with upward pointing normal.

37. Evaluate the integral \( \iint_S \text{curl} \mathbf{F} \cdot d\mathbf{S} \) where \( \mathbf{F} = \langle 3x - 2y, y + z, 2z - x \rangle \), and \( S \) is the sides of the tetrahedron with vertices \((0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\) with upward pointing normal (here \( S \) is the sides, not including the base triangle with vertices \((0, 0, 0), (1, 0, 0) \) and \((0, 1, 0)\)). (Hint: switch surfaces)

38. Evaluate the integral \( \iint_S \text{curl} \mathbf{F} \cdot d\mathbf{S} \) where \( \mathbf{F} = \langle x + y - 2, x^2 + z, y - z \rangle \), and \( S \) is the portion of the paraboloid \( z = 1 - x^2 - y^2 \) lying on or above the \( xy \)-plane. (Hint: switch surfaces)

39. Evaluate the integral \( \iint_S \text{curl} \mathbf{F} \cdot d\mathbf{S} \) where \( \mathbf{F} = \langle yz^2 + xz, xy, xy^2 z^3 \rangle \), and \( S \) is the portion of the cone \( z = 1 - \sqrt{x^2 + y^2} \) lying on or above the \( xy \)-plane. (Hint: switch surfaces)

40. Evaluate the integral \( \iint_S \text{curl} \mathbf{F} \cdot d\mathbf{S} \) where \( \mathbf{F} = \langle 3x - y, x + z^2, x - y \rangle \), and \( S \) is the hemisphere \( z = \sqrt{1 - x^2 - y^2} \). (Hint: switch surfaces)
5.6. THE DIVERGENCE THEOREM

5.6 The Divergence Theorem

Recall that divergence is a measure of how much a vector field is spreading out. In this section we introduce Gauss’ Divergence Theorem, which relates the flux across a closed surface to the integral of the divergence over the solid. More precisely we have:

**Theorem 5.6.1.** Let $S$ be a closed surface bounding the solid $W$, and let $\mathbf{F}$ be a vector field defined on $W$. Then

$$\int_S \mathbf{F} \cdot d\mathbf{S} = \iiint_W \text{div } \mathbf{W} \, dV.$$  

The proof of the Divergence Theorem, or Gauss’ Theorem as it is also called, is pretty much brute force. One just integrates both sides of the equation to a certain point, and observes that the results are the same. This is similar in flavor to the proof of Green’s Theorem. We prove it in the special case where $W$ is a rectangular box.

**Proof.** Let $W$ be the box $[a, b] \times [c, d] \times [e, f]$ and $S$ its boundary with outward pointing normal. Thus $S$ is the union of six rectangles: the back $S_1 = \{a\} \times [c, d] \times [e, f]$, the front $S_2 = \{b\} \times [c, d] \times [e, f]$, the left $S_3$, right $S_4$, bottom $S_5$, and top $S_6$. The outward pointing normal to $S_1$ is $-\mathbf{i}$ and to $S_2$ is $\mathbf{i}$. Similarly, the normals to the other faces are parallel to coordinate axes. This helps in calculating the surface integral in the Divergence Theorem.

Let $\mathbf{F}(x, y, z) = \langle P(x, y, z), Q(x, y, z), R(x, y, z) \rangle$, so that

$$\iiint_W \text{div} \mathbf{W} \, dV = \int_c^f \int_d^b \int_a^e \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} \, dx \, dy \, dz.$$  

To prove the theorem, we calculate the contributions of each component function of $\mathbf{F}$ separately. Calculating the contribution from $\frac{\partial P}{\partial x}$ to the triple integral we see

$$\int_c^f \int_d^b \int_a^e \frac{\partial P}{\partial x} \, dx \, dy \, dz = \int_c^f \int_d^b P(x, y, z) \bigg|_a^b \, dy \, dz = \int_c^f \int_d^b P(b, y, z) - P(a, y, z) \, dy \, dz.$$  

(5.6.1)

To calculate the contribution of $P(x, y, z)$ to the flux, we use the observations regarding normal vectors to the faces of the cube to see

$$\int_S \mathbf{F} \cdot d\mathbf{S} = \int_{S_1} \mathbf{F} \cdot (-\mathbf{i}) \, dA + \int_{S_2} \mathbf{F} \cdot \mathbf{i} \, dA + \int_{S_3} \mathbf{F} \cdot (-\mathbf{j}) \, dA + \int_{S_4} \mathbf{F} \cdot \mathbf{j} \, dA + \int_{S_5} \mathbf{F} \cdot (-\mathbf{k}) \, dA + \int_{S_6} \mathbf{F} \cdot \mathbf{k} \, dA.$$
The component function $P(x, y, z)$ only contributes to the first two integrals, since it vanishes in the remaining dot products. The integrand in each will be $P(x, y, z)$ evaluated on the surface. Since $S_1$ is the surface $x = a$, we integrate $-P(a, y, z)$ for the first integral. Similarly the second integrand is $P(b, y, z)$. Thus the contribution from $P(x, y, z)$ to the surface integral is

$$\int_{S_1} \mathbf{F} \cdot (-\mathbf{i}) \, dA + \int_{S_2} \mathbf{F} \cdot (\mathbf{i}) \, dA = \int_C \int_S -P(a, y, z)dydz + \int_C \int_S P(b, y, z)dydz. \tag{5.6.2}$$

Comparing the contributions of $P$ to the triple integral in Equation $5.6.1$ to the contribution in Equation $5.6.2$, we see that they are the same. Similar calculations show that the component functions $Q$ and $R$ both contribute the same to the triple and surface integrals. This proves the Divergence Theorem for the special case of a rectangular box. The arguments generalize to other special domains using similar techniques, and to even more complicated domains by decomposing them into simpler pieces and adding them. We do not treat the more general settings here.

Before doing some examples, we take a moment to compare the important theorems of vector analysis. Recall that both curl and divergence are ways to differentiate vector fields, and that integrals involve antidifferentiation. Notice that Green’s, Stokes’, and Gauss’ Theorems all relate the integral of a derivative over a region to the integral of its antiderivative over the boundary of the region. For example, to integrate div$\mathbf{F}$ (a derivative of $\mathbf{F}$) over a solid $W$, you integrate its antiderivative $\mathbf{F}$ over the boundary of $W$. This is analogous to the one dimensional case, where the integral $\int_a^b f(x)dx$ of the derivative $F'(x) = f(x)$ over the interval $[a, b]$ is related to the antiderivative evaluated on its boundary. In this case, the boundary of an interval is its endpoints. Thus we get $\int_a^b f(x)dx = F(b) - F(a)$.

All of the important integrals we’ve discussed, then, follow the same general pattern: integrating a derivative over a region equals the integral of the antiderivative of the boundary of the region.

Applications of Gauss’ Theorem usually involve using a triple integral to evaluate a related surface integral. We illustrate with a couple examples.

**Example 5.6.1. Calculating Flux across a sphere**

Let $\mathbf{F}(x, y, z) = \langle x^3 - y^3, y^3 - z^3, z^3 - x^3 \rangle$. Calculate the flux across the unit sphere.

By the divergence theorem, we will calculate the flux over the sphere by integrating the divergence over the entire ball. Moreover, the domain of integration and the divergence both indicate that spherical coordinates will be the easiest way to evaluate the integral. The calculations proceed as follows:
5.6. THE DIVERGENCE THEOREM

∫∫ₐ S \mathbf{F} \cdot d\mathbf{S} = ∫∫∫ₐ \text{div} \mathbf{F} dV

= ∫∫∫ₐ \left( \frac{∂(x^3 - y^3)}{∂x} + \frac{∂(y^3 - z^3)}{∂y} + \frac{∂(z^3 - x^3)}{∂z} \right) dV

= ∫∫∫ₐ 3(x^2 + y^2 + z^2) dV

= ∫₀^2π ∫₀^π ∫₀¹ 3ρ^4 \sin φ dρ dφ dθ

= ∫₀^2π ∫₀^π 3ρ^5 \sin φ \bigg|₀^¹ dφ dθ

= \frac{3}{5} ∫₀^2π ∫₀^π \sin φ dφ dθ = \frac{3}{5} ∫₀^2π - \cos φ \bigg|₀^π dθ

= \frac{6}{5} ∫₀^2π dθ = \frac{12\pi}{5}.

Example 5.6.2. Calculating flux across a cone

Calculate the flux of \( \mathbf{F}(x, y, z) = (xy^2, yz^2, zx^2) \) across the right circular cone of height one with unit disk as base. Thus the sides of the cone are given by the equation \( z = 1 - \sqrt{x^2 + y^2} \), and the inside of the cone is above the \( xy \)-plane and below the cone.

The triple integral of the divergence theorem will be easier to do using cylindrical coordinates. In cylindrical coordinates the solid is described by the system of inequalities

\[ x = \rho \cos \theta, \quad y = \rho \sin \theta, \quad z = 1 - \rho \]

with \( 0 ≤ \rho ≤ 1, 0 ≤ \theta ≤ 2\pi, \) and \( 0 ≤ φ ≤ \pi \).
Figure 5.6.2: Flux across a cone

\[ 0 \leq z \leq 1 - r \]
\[ 0 \leq r \leq 1 \]
\[ 0 \leq \theta \leq 2\pi \]

A quick calculation shows \( \text{div} \mathbf{F} = y^2 + z^2 + x^2 \), so the flux is given by

\[
\iint_S \mathbf{F} \cdot d\mathbf{S} = \int_0^{2\pi} \int_0^1 \int_0^{1-r} (r^2 + z^2) r dz dr d\theta \\
= \int_0^{2\pi} \int_0^1 \int_0^{1-r} r^3 z + rz^3 \bigg|_0^1 dr dz d\theta \\
= \int_0^{2\pi} \int_0^1 r(1 - r) \left( r^2 + \frac{(1 - r)^2}{3} \right) dr d\theta \\
= \frac{1}{3} \int_0^{2\pi} \int_0^1 -4r^4 + 6r^3 - 3r^2 + r dr d\theta \\
= \frac{1}{3} \int_0^{2\pi} \left[ -\frac{4r^5}{5} + \frac{3r^4}{2} - r^3 + \frac{r^2}{2} \right]_0^1 d\theta \\
= \frac{1}{3} \int_0^{2\pi} \frac{1}{5} d\theta = \frac{2\pi}{15}.
\]

**Exercises**

1. Find the flux of \( \mathbf{F}(x, y, z) = \langle 3, 1, -4 \rangle \) across the unit sphere with outward pointing normal direction.
2. Find the flux of \( \mathbf{F}(x, y, z) = \langle 1, 4, -5 \rangle \) across the surface of the cube \( W = [0, 1] \times [0, 1] \times [0, 1] \) with outward pointing normal.

3. Find the flux of \( \mathbf{F}(x, y, z) = \langle x^2, 1, z \rangle \) across the boundary of the solid inside \( y^2 + z^2 = 1 \) and between \( x = 0 \) and \( x + z = 2 \) with outward pointing normal.

4. Find the flux of \( \mathbf{F}(x, y, z) = \langle 3x + 2y, z - y, x + y - z \rangle \) across the boundary of the solid inside \( x^2 + y^2 = 1 \) and between \( 0 \leq z \leq 1 \) with outward pointing normal.

5. Find the flux of \( \mathbf{F}(x, y, z) = \langle xz, x^2yz, \frac{x^2z^2}{2} \rangle \) across the cone \( \sqrt{x^2 + y^2} \leq z \leq 1 \) and between \( x = 0 \) and \( x + z = 2 \) with outward pointing normal.

6. Find the flux of \( \mathbf{F}(x, y, z) = \langle 7x - y + 2z, 3 + 4z, x - 2y + z \rangle \) across the boundary of the solid in the first octant under both \( x + z = 1 \) and \( x + y + z = 2 \) with outward pointing normal.

7. Find the flux of \( \mathbf{F}(x, y, z) = \langle xy^2 - yz^2, yz^2 - zxz^2, zx^2 - xy^2 \rangle \) across the surfaces (with outward pointing normal vectors)
   (a) The unit sphere.
   (b) The cylindrical can with lateral surface \( x^2 + y^2 = 1 \), top at \( z = 1 \) and bottom at \( z = 0 \).

8. Find the flux of \( \mathbf{F}(x, y, z) = \langle xy - yz, x^2z + 2y, 3y - 2z \rangle \) across the cube \( [0, 1] \times [0, 1] \times [0, 1] \) with outward pointing normal.

9. Find the flux of \( \mathbf{F}(x, y, z) = \langle 3xz^2, y^2, 3x - 2y \rangle \) across the surface between \( x = 0 \) and \( x = 1 \), above the \( xy \)-plane and below the cylinder \( y^2 + z^2 = 1 \).

10. Show that the flux of a constant vector field across a closed surface is zero.

11. Let \( W \) be a solid with boundary \( S \). Show that the flux of \( \mathbf{F}(x, y, z) = \langle 3x, x^2 - y, y - z + 3x \rangle \) across \( S \) is the volume of \( W \).

12. Let \( W \) be the cylinder \( x^2 + y^2 \leq R^2 \) and \( 0 \leq z \leq H \), and let \( S \) be the boundary of \( W \) with outward pointing normal. Evaluate the flux of

    \[
    \mathbf{F}(x, y, z) = \langle \sin(xy), x^2, -3y \cos(xy) \rangle
    \]

across \( S \) without integrating.

13. Show that the flux of \( \mathbf{F}(x, y, z) = \langle f(y, z), g(x, z), h(x, y) \rangle \) across any closed surface is zero.

14. Let \( W \) be a solid with boundary \( S \) and volume \( \text{vol}(W) \). Show that if the component functions of \( \mathbf{F} \) are linear, then the flux of \( \mathbf{F} \) across \( S \) can be calculated without integrating, just using \( \text{vol}(W) \).