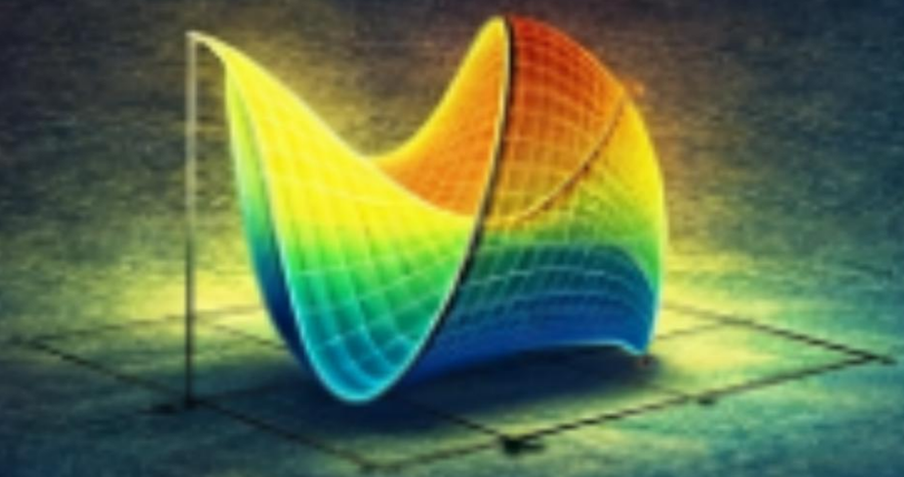
πMA∫ Academy

ACE

Multivariable Calculus



2026

- 100+ Problems
- All Topics
- Detailed Solutions

Aditya Baisakh

Contents

0.1 About TMAS Academy

The Math and Science (TMAS) Academy, formerly known as **Explore Math**, is a nonprofit organization which was started in 2021 to spread competition math resources to those who may not be able to afford existing ones which are usually overpriced. We believe no student should have to struggle merely due to a lack of resources, that a quality education is a right for everyone and should not be dependent on the size of their wallets. Later, it was expanded in March 2024 after the team started developing resources to assist students in preparation for AP STEM courses and exams. These included books from AP Calculus to AP Physics to AP Chemistry to AP Computer Science and more. Currently, the team is working on developing more specialized resources and expand the audience to include university students as well.

0.2 Opportunities For You To Contribute To TMAS Academy

TMAS Academy is very inclusive and you can help support its cause in several ways.
You can **join the team** by filling out the form below, which can also be found on the website: `https://forms.gle/VXGvj27UvcZPGhiJ8`

Donations: If you want to assist us in our monthly payments to run this organization, which includes website costs, Overleaf costs (the platform used to write these books), and filming/editing costs, then please consider donating! For those who are willing to contribute, we have listed some ways below. **Don't forget to write a message so we know who you are and can send you a thank you note!**

- You can donate through PayPal to the email: weexploremath@gmail.com
- If you want to donate and the above method doesn't work for you, then you can send an email to weexploremath@gmail.com

You can also contribute by **subscribing** to the YouTube channel:
`https://www.youtube.com/@tmasacademy`

Also, don't forget to join the Discord server to connect with other hardworking students preparing for AP exams and math competitions such as AMC 10/12 and AIME. `https://discord.gg/tmas-academy-1019082642794229870`

You can also follow all of our social media such as the LinkedIn page and the Instagram account that is run by the marketing team. Also, please join the mailing list to learn about all updates and our upcoming books and videos.

Finally, you can spread our efforts and initiative to anyone you know who may benefit from or support us, be it your classmates, teachers, or other nonprofit organizations focused on education.

0.3 About the Author: Aditya Baisakh

The summer after high school, I randomly decided to try learning multivariable calculus in two weeks. No big plan, no fancy setup, just me, a notebook, an old clunky laptop, and a lot of long nights working through problem sets with my best friend, all in preparation for a proficiency test before my first semester.

It wasn't easy. I got stuck constantly. My scratch paper was riddled with mistakes and crossed-out ideas. I'd replay the same parts of YouTube videos until something finally clicked. But somewhere in the process, I realized that conceptual understanding doesn't come from being a genius. It comes from just not quitting.

This book isn't about me learning Calc III quickly. It's about giving back. I know what it's like to work hard with limited resources, and I wanted to make something free, something that makes the climb a little less steep. If this book can make someone feel more capable, that they belong here too, then everything was 100% worth it.

0.4 What if there is an error in the book?

There are possibilities for errors such as typos or incorrect solutions to problems. If that is the case, please click on this link and fill out the form to report the error:
Error Form

0.5 Any other questions or concerns?

If you have any questions about TMAS Acaemy and its programs, please contact Aditya Baisakh, the author and current CEO of TMAS Academy at the address
`adityabaisakh123@gmail.com`.

0.6 Acknowledgements

- As the CEO of such a wonderful student organization, I want to thank the entire TMAS Academy community for their unwavering support of the team's educational initiative and for being my inspiration to always put in my best effort.
- I would also like to thank the **Art of Problem Solving (AoPS)** for their supportive community of math and science enthusiasts. In addition, their LaTeX programming forums and tutorials were extremely helpful in writing this book.
- I also would like to thank everyone who supports the work I have done and encourages me to continue.
- Finally, I want to thank my parents for always motivating me to achieve my goals and for everything else they have done.

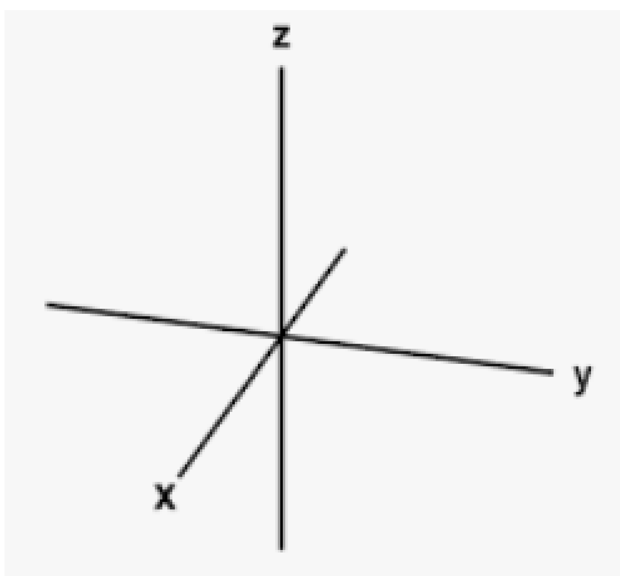
1 Three Dimensional Space

In this chapter, we will introduce vectors and coordinate systems in three-dimensional space. This sets the stage for studying curves in space and functions of two variables (whose graphs are surfaces) in the following chapters. We will also see how vectors make it easy to describe lines and planes in space.

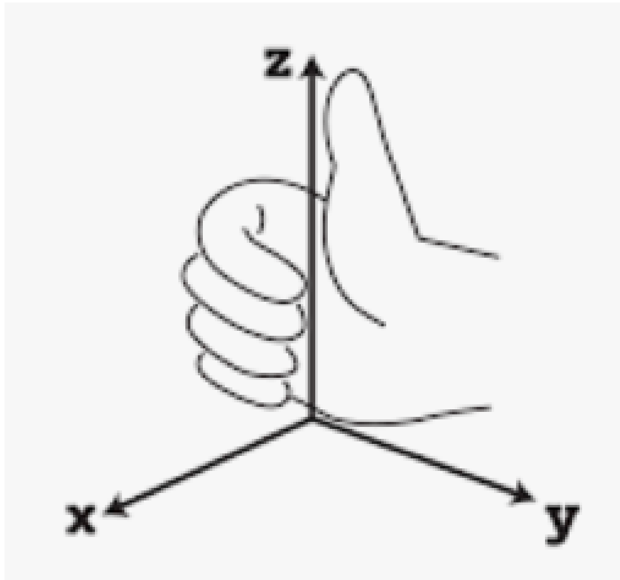
1.1 Three-Dimensional Coordinate System

To locate a point in a plane, we need two numbers. Any point in that plane can be represented by an ordered pair (a, b), where $a, b \in \mathbb{R}$. The x- and y-coordinates are given by a and b, respectively. For this reason, planes are considered two-dimensional. Meanwhile, in order to locate a point in space, three numbers are required. We represent points in space by an ordered triple (a, b, c), again with $a, b, c \in \mathbb{R}$.

3D Space Before representing points in space, we first choose a fixed point O (the origin) and three directed lines through O that are mutually orthogonal. These are labeled as the x-axis, y-axis, and z-axis. To keep things simple, we consider the x- and y-axes as horizontal, and the z-axis as vertical, so we typically orient them as shown below.



To determine the direction of the z-axis, we apply the **right hand rule**. If you curl the fingers of your right hand around the z-axis in the direction of a 90° counterclockwise rotation from the positive x-axis to the positive y-axis, then your thumb points in the direction of the positive z-axis.



The three coordinate axes determine the three **coordinate planes** shown below. The xy-plane contains the x- and y-axes, the yz-plane contains the y- and z-axes, and the xz-plane contains the x- and z-axes. These three coordinate planes divide space into $2^3 = 8$ components, called **octants**. Most importantly, the **first octant** is determined by the positive axes (x, y, and z).

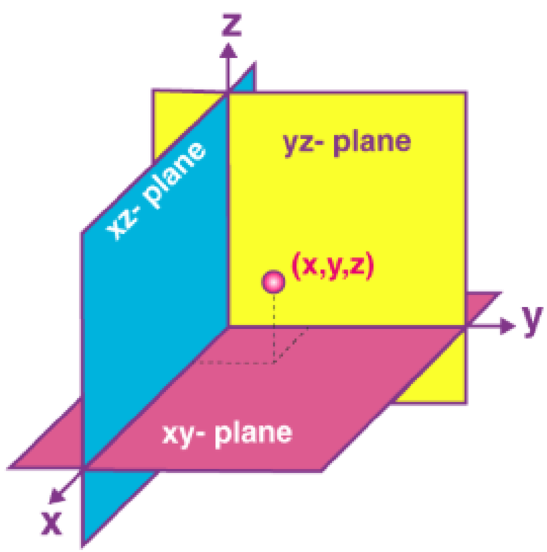


Image Credit: BYJU'S

Many people have some difficulty visualizing diagrams of three-dimensional figures, so it might be helpful to do the following. Consider the bottom corner of any room you're standing in. The wall on your left is the xz-plane, the wall on your right is the yz-plane, and the floor (lower wall) is the xy-plane. You can also observe that the x-axis runs along the intersection of the floor and the left wall, the y-axis runs along the intersection of the floor and the right wall, and the z-axis runs up from the floor toward the ceiling along the the intersection of the two walls. You are situated in the first octant, and you can now imagine seven other rooms situated in the other seven octants (three on the same floor and four on the floor below), all connected by the common corner point, which is the origin, O.

Let P be any point in space: a is the directed distance from the yz-plane to P, b is the directed distance from the xz-plane to P, and c is the directed distance from the xy-plane to P. Thus, P is represented by the ordered triple (a, b, c) of real numbers and the elements of the set $\{a, b, c\}$ are called the **coordinates** of P, i.e. a, b, and c are the x-, y-, and z-coordinates. Thus, to locate a point (a, b, c), we start at the origin O, move a units along the x-axis, then b units parallel to the y-axis, and c parallel to the z-axis, as shown below.

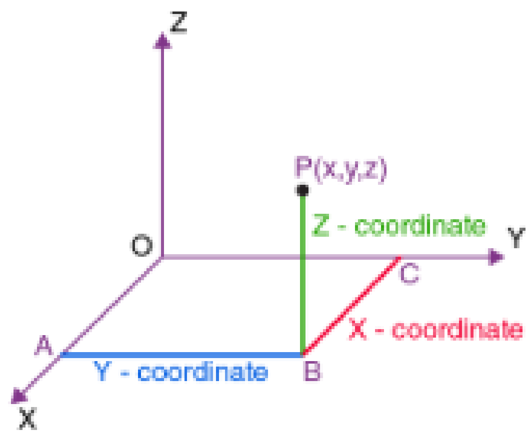


Image Credit: BYJU'S

The point $P(a, b, c)$ determines a rectangular box with dimensions $a \times b \ \times c$. If we were to drop an altitude from P to the xy-plane, we would get a point Q with coordinates of $(a, b, 0)$. Q is called the **projection** of P onto the xy-plane. Similarly, the points $R(0, b, c)$ and $(a, 0, c)$ are the projections of P onto the yz- and xz-plane, respectively.

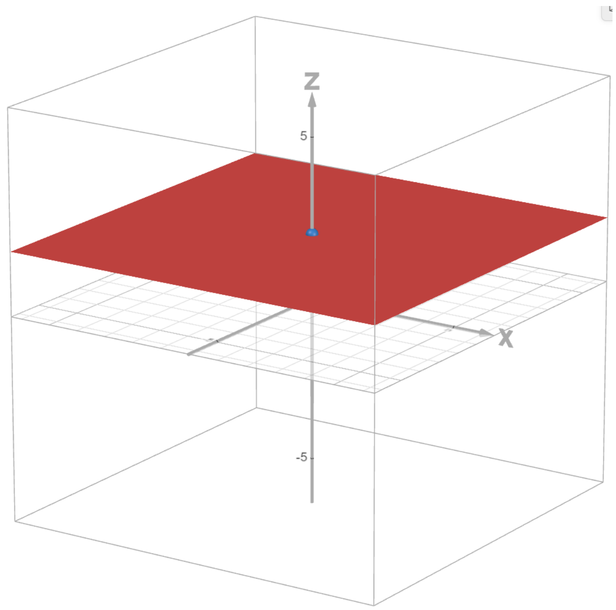
The Cartesian product $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \{(x, y, z) \,|\, x, y, z \in \mathbb{R}\}$ is the set of all ordered triples of real numbers and is denoted simply by $\mathbb{R}^3$. There exists a one-to-one correspondence between points P in space and ordered triples (a, b, c) in $\mathbb{R}^3$. This is called a **three-dimensional rectangular coordinate system**. In terms of coordinates, the first octant can be described by the subset of $\mathbb{R}^3$, where $x, y, z > 0$.

Surfaces and Solids In two-dimensional analytic geometry, we know that any equation involving the variables x and y defines a curve in $\mathbb{R}^2$. Extending this to three dimensions, an equation in x, y, and z represents a *surface* in $\mathbb{R}^3$.

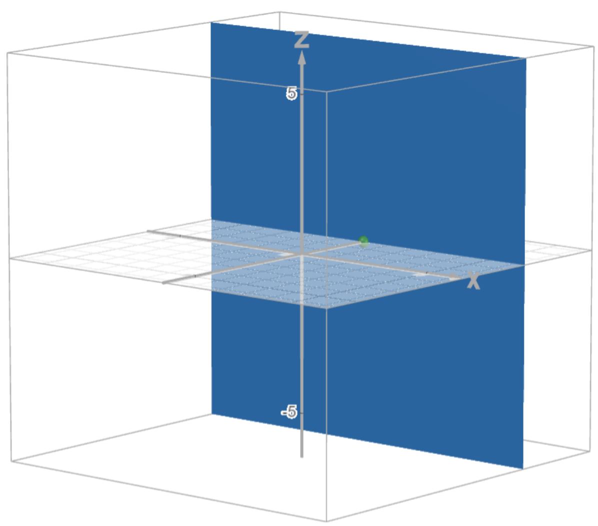
Problem 1.1.1. *What surface in $\mathbb{R}^3$ is represented by each equation?*

$$(a)\ \ z = 2 \qquad\qquad (b)\ \ y = 3$$

Solution to part a: The equation $z = 2$ represents the set $\{(x, y, z) \,|\, z = 2\}$, which is the set of all points with z-coordinate of 2, while x and y are free to vary. The surface is the horizontal plane parallel to the xy-plane and two units above it.



Solution to part b: The equation $y = 3$ represents the set of all points in $\mathbb{R}^3$ with a y-coordinate with 3, where x and z are free to vary. The surface is the vertical plane parallel to the xz-plane and three units to the right of it.

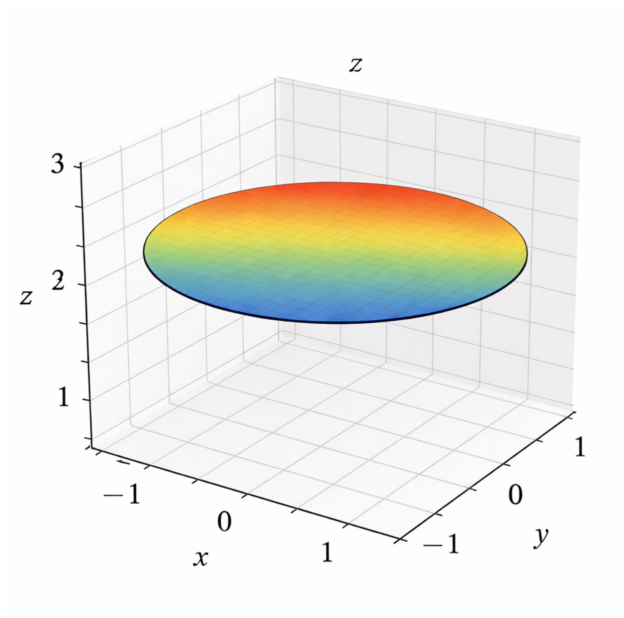


***Remark*.** When an equation is given, its interpretation depends on the context: it may describe a curve in $\mathbb{R}^2$ or a surface in $\mathbb{R}^3$. For instance, the equation $x = 4$ represents a plane in $\mathbb{R}^3$, but in two-dimensional analytic geometry it represents a line in $\mathbb{R}^2$.

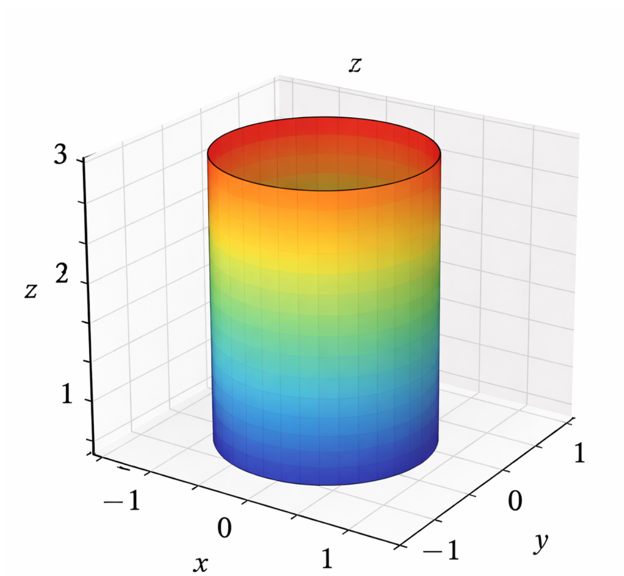
In general, let k be a real number. Then $x = k$ represents a plane parallel to the yz-plane, $y = k$ represents a plane parallel to the xz-plane, and $z = k$ represents a plane parallel to the xy-plane. The faces of the rectangular box we conceptualized earlier are formed by the three coordinate planes $x = 0$ (yz-plane), $y = 0$ (xz-plane), $z = 0$ (xy-plane), and the planes $x = a$, $y = b$, and $z = c$.

Problem 1.1.2. *Answer the following questions.*
(a) What is the set of all points (x, y, z) that satisfy the equations $x^2 + y^2 = 1$ and $z = 3$?
(b) What surface does the equation $x^2 + y^2 = 1$ describe in $\mathbb{R}^3$?
(c) What is the solid in $\mathbb{R}^3$ represented by the inequalities $x^2 + y^2 \leq 1$, $2 \leq z \leq 4$?

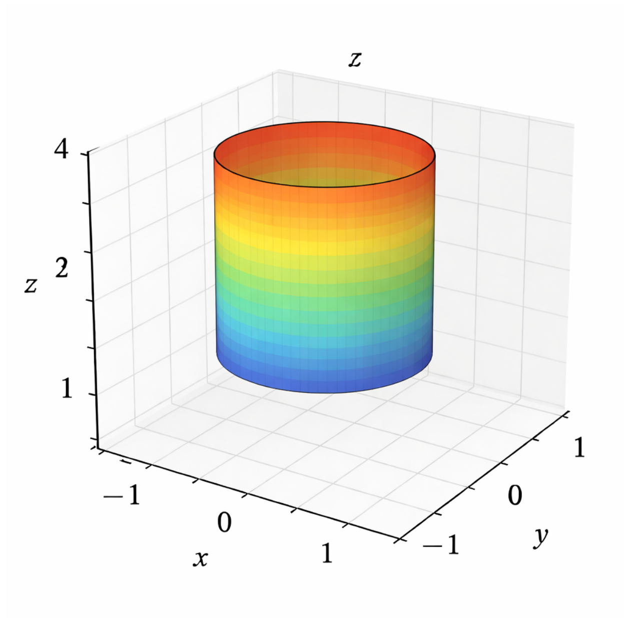
Solution to part a: The equation $z = 3$ indicates that the points lie on the horizontal plane $z = 3$. With $x^2 + y^2 = 1$, the points lie on the circle of radius 1 centered on the z-axis.



Solution to part b: If $x^2 + y^2 = 1$ is given with no restriction on z, the point (x, y, z) could lie on a circle in any horizontal plane of the equation $z = k$. So the surface $x^2 + y^2 = 1$ in $\mathbb{R}^2$ consists of all horizontal circles with radius 1 lying the planes $z = k$, where k is a real number, and is essentially a cylinder of radius 1 centered on the z-axis.

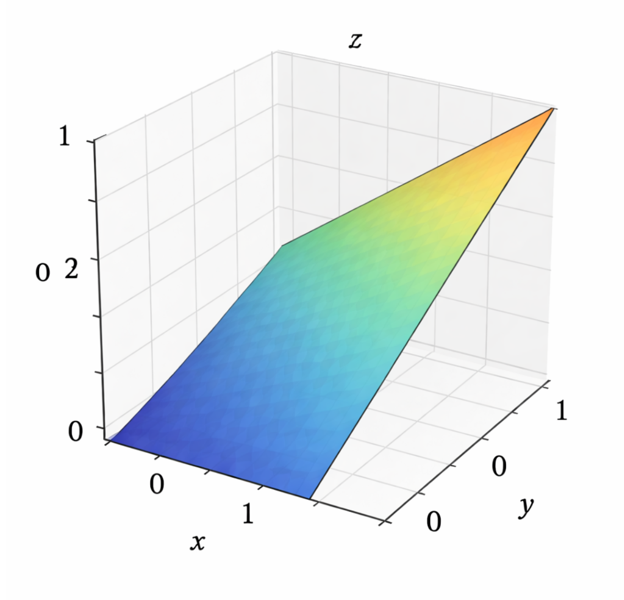


Solution to part c: Because $x^2 + y^2 \leq 1$, we know any point (x, y, z) in the region must lie on or inside the circle of radius 1 centered on the z-axis in a horizontal plane $z = k$. We are also given $2 \leq z \leq 4$, so the given inequalities illustrate the portion of the solid circular cylinder of radius 1, situated between the planes $z = 2$ and $z = 4$.



Problem 1.1.3. *Describe and sketch the surface in $\mathbb{R}^3$ represented by the equation $z = x$.*

Solution: The equation $z = x$ relates the z-coordinate with the x-coordinate for all points in this surface. Thus, we have $\{(x, y, x) \,|\, x \in \mathbb{R}, y \in \mathbb{R}\}$. This is an oblique plane that intersects the xz-plane in the line $z = x$, $y = 0$. The portion of the plane lying in the first octant is shown below.



Distance and Spheres In three dimensions, the distance between two points P_1 and P_2 is given by

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

Problem 1.1.4. *Find the distance between the two points $P(2, -1, 3)$ and $Q(3, 0, 4)$.*

Solution: The distance is equal to

$$|PQ| = \sqrt{(3-2)^2 + (0-(-1))^2 + (4-3)^2} = \sqrt{1+1+1} = \boxed{\sqrt{3}}$$

Let's move on to the last topic of this section.

Definition 1.1.1. *A **sphere** with radius r and center $C(h, k, l)$ is defined as the set of all points $P(x, y, z)$ with a fixed distance of r from C.*

Therefore, P is on the sphere if and only if $|PC| = r$, so

$$\sqrt{(x-h)^2 + (y-k)^2 + (z-l)^2} = r$$

The standard-form equation of a sphere with center $C(h, k, l)$ and radius r is

$$(x-h)^2 + (y-k)^2 + (z-l)^2 = r^2$$

If C is the origin, or $(0, 0, 0)$, then the equation of the sphere is

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

Problem 1.1.5. *Find the equation of the sphere with center $(3, -1, 2)$ passing through the point $(4, 1, 0)$.*

Solution: The radius r of the sphere is the distance between the center and any point on the circle, so

$$r = \sqrt{(4-3)^2 + (1-(-1))^2 + (0-2)^2} = \sqrt{1+4+4} = \sqrt{9} = 3$$

So the equation of the sphere is

$$(x-3)^2 + (y+1)^2 + (z-2)^2 = 3^2 \therefore \boxed{(x-3)^2 + (y+1)^2 + (z-2)^2 = 9}$$

Problem 1.1.6. *Show that $x^2 + y^2 + z^2 + 4x - 6y + 2z + 6 = 0$ represents a sphere, and find its center and radius.*

Solution: If we complete the square, we can rewrite this equation into the standard form equation for a sphere:

$$(x^2 + 4x + 4) + (y^2 - 6y + 9) + (z^2 + 2z + 1) = -6 + 4 + 9 + 1$$

$$(x+2)^2 + (y-3)^2 + (z+1)^2 = 8$$

We observe that this is the equation of a sphere with center $\boxed{(-2, 3, -1)}$ and radius $r = \sqrt{8} = \boxed{2\sqrt{2}}$.

1.2 Vectors: In Essence

From a physics perspective, there are many quantities that cannot be described merely by its scale, or magnitude. For example, the velocity of a moving object can only expressed if we specify both the speed and the direction of motion for the object. Other examples include force, acceleration, and displacement.

Definition 1.2.1. *A* ***vector*** *indicates a quantity with both magnitude and direction.*

Vectors are often represented by an arrow or a directed line segment. The length of the arrow represents the magnitude of the vector and the direction of the arrow corresponds to the direction of the vector. We denote vectors by using boldface notation ($\mathbf{v}$) or arrow notation ($\overrightarrow{v}$). In this book, we will primarily use boldface notation.

For instance, if a particle moves from an **initial point** A to a **terminal point** B, so $\mathbf{v} = \overrightarrow{AB}$. Suppose we have another vector $\mathbf{u} = \overrightarrow{CD}$ with same length and direction, even if it is in a different direction. Thus $\mathbf{u}$ and $\mathbf{v}$ are equivalent, and $\mathbf{u} = v$. The **zero vector**, denoted by $\mathbf{0}$, has length 0. It is also the only vector that does not have a specified direction.

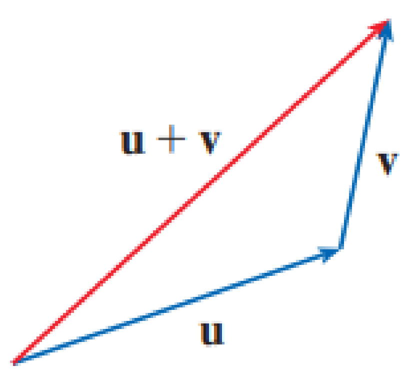
Oftentimes, it is useful to combine vectors. If a particle moves from A to B, it has displacement vector $\overrightarrow{AB}$, then changes direction and moves from B to C, with displacement vector $\overrightarrow{BC}$, then the effect of these displacements is that the particle has ultimately moved from A to C. The resulting displacement vector $\overrightarrow{AC}$ is called the *sum* of the two vectors $\overrightarrow{AB}$ and $\overrightarrow{BC}$.

$$\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$$

In general, if we start with vectors $\mathbf{u}$ and $\mathbf{v}$, we first place $\mathbf{v}$ so that its tail coincides with the tip of $\mathbf{u}$ and define the sum of $\mathbf{u}$ and $\mathbf{v}$.

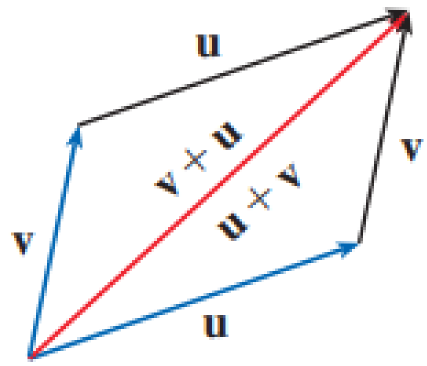
Definition 1.2.2. *If $\mathbf{u}$ and $\mathbf{v}$ are two vectors positioned so the initial point of $\mathbf{v}$ is at the terminal point of $\mathbf{u}$, then the* ***sum*** *$\mathbf{u} + \mathbf{v}$ is the vector from the initial point of $\mathbf{u}$ to the terminal point of $\mathbf{v}$.*

This definition illustrates the **Triangle Law**, which is demonstrated below.



Another way we can combine vectors is to start off with the same vectors $\mathbf{u}$ and $\mathbf{v}$, and draw a copy of $\mathbf{v}$ with the same initial point as $\mathbf{u}$. This gives us a partially-formed parallelogram. If we complete the polygon, then we can observe that $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$, and thus we can find the sum of $\mathbf{u}$

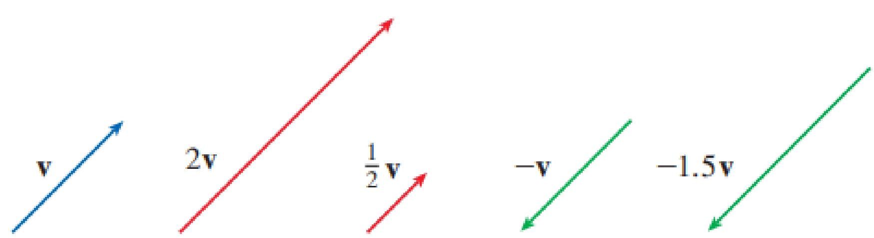
and **v** in this manner. If we place **u** and **v** with common initial point, then $\mathbf{u} + \mathbf{v}$ lies along the diagonal of the parallelogram with **u** and **v** as adjacent sides. This is called the **Parallelogram Law**.



We now define multiplication of a vector **v** by a real-valued c. In this context, we say c is a **scalar** to distinguish it from a vector.

Definition 1.2.3. *If c is a scalar and $\boldsymbol{v}$ is a vector, then the* ***scalar multiple*** *$c\boldsymbol{v}$ is the vector whose length is $|c|$ times the length of $\boldsymbol{v}$ and whose direction is the same as $\boldsymbol{v}$ if $c > 0$ and is opposite to $\boldsymbol{v}$ if $c < 0$. If $c = 0$ or $\boldsymbol{v} = \boldsymbol{0}$ then $c\boldsymbol{v} = \boldsymbol{0}$.*

This definition is illustrated below. Real numbers basically work like scalars, hence the name. Also, notice how two nonzero vectors are **parallel** if they are scalar multiples of one another. The special case $-\mathbf{v} = (-1)\mathbf{v}$ involves a vector with the same length as **v** but pointing in the opposite direction. It is called the **negative** of **v**.



We explore the **difference** of two vectors as $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$. For any two vectors **u** and **v**, we can construct the difference $\mathbf{u} - \mathbf{v}$ by first drawing the negative of **v**, or $-\mathbf{v}$, then adding it to **u** using the Parallelogram Law. Alternatively, we have $\mathbf{v} + (\mathbf{u} - \mathbf{v}) = \mathbf{u}$, so we can also arrive at the result using the Triangle Law. You will find that if **u** and **v** share the same starting point, then the difference $\mathbf{u} = \mathbf{v}$ connects the tips of **v** and **u**.

Component Form We will later see that it is convenient to introduce a coordinate system that can describe vectors in an algebraic fashion. If we place the initial point of a vector **a** at the origin of a rectangular coordinate system, then the terminal point of **a** has coordinates (a_1, a_2) or (a_1, a_2, a_3), depending on whether the system is two- or three-dimensional, respectively. These coordinates are called **components** of **a**, and are expressed in what's called **component form**:

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

Make sure to not confuse the vector $\langle a_1, a_2\rangle$ with the ordered pair (a_1, a_2) which represents a point in the plane.

Let O be the origin and let P be any point in $\mathbb{R}^3$. Thus, the vector representation is $\mathbf{a} = \overrightarrow{OP} = \langle a_1, a_2, a_3\rangle$. If we generalize to consider a directed line segment $\overrightarrow{AB}$ with initial and terminal points $A(x_1, y_1, z_1)$ and $B(x_2, y_2, z_2)$, respectively, then we must have $x_1 + a_1 = x_2$, $y_1 + a_2 = y_2$, and $z_1 + a_3 = z_2$, and so

$$\mathbf{a} = \langle a_1, a_2, a_3\rangle = \langle x_2 - x_1, y_2 - y_1, z_2 - z_1\rangle$$

where $\mathbf{a} = \overrightarrow{AB}$.

Problem 1.2.1. *Find the vector represented by the directed line segment with initial point* $A(2, -3, 4)$ *and* $B(-3, -2, 1)$.

Solution: The vector corresponding to $\overrightarrow{AB}$ is given by

$$\mathbf{a} = \langle -3 - 2, -2, -(-3), 1 - 4\rangle = \boxed{\langle -5, 1, -3\rangle}$$

The **magnitude** or **length** of a vector is the length of any possible representations of it, and is denoted by the symbol $|\mathbf{v}|$ or $||\mathbf{v}||$. If we apply the distance formula to compute the length of segment OP, we have

$$\boxed{|\mathbf{a}| = \sqrt{a_1^2 + a_2^2} \qquad \text{or} \qquad |\mathbf{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}}$$

depending on whether the vector is two- or three-dimensional.

Right now we are faced with the question as to how we can add, subtract, and scalar-wise multiply vectors algebraically/

Algebraic vector operations are performed component-wise. If

$$\mathbf{a} = \langle a_1, a_2\rangle \quad \text{and} \quad \mathbf{b} = \langle b_1, b_2\rangle,$$

then their sum is

$$\mathbf{a} + \mathbf{b} = \langle a_1 + b_1,\ a_2 + b_2\rangle.$$

Similarly, vector subtraction is given by

$$\mathbf{a} - \mathbf{b} = \langle a_1 - b_1,\ a_2 - b_2\rangle.$$

If c is a scalar, then scalar multiplication of $\mathbf{a}$ is

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

Problem 1.2.2. *If* $\boldsymbol{a} = \langle 4, 0, 3\rangle$ *and* $\boldsymbol{b} = \langle -2, 1, 5\rangle$*, find* $|\boldsymbol{a}|$ *and the vectors* $\boldsymbol{a} + \boldsymbol{b}$*,* $\boldsymbol{a} - \boldsymbol{b}$*,* $3\boldsymbol{b}$*, and* $2\boldsymbol{a} + 5\boldsymbol{b}$.

Solution: We simply apply the previous rationale, and the problem is trivial.

$$\begin{aligned}\|\mathbf{a}\| &= \sqrt{4^2+0^2+3^2} \\ &= \sqrt{25} \\ &= \boxed{5}\end{aligned}$$

$$\begin{aligned}\mathbf{a}+\mathbf{b} &= \langle 4,0,3\rangle + \langle -2,1,5\rangle \\ &= \langle 4+(-2), 0+1, 3+5\rangle \\ &= \boxed{\langle 2,1,8\rangle}\end{aligned}$$

$$\begin{aligned}\mathbf{a}-\mathbf{b} &= \langle 4,0,3\rangle - \langle -2,1,5\rangle \\ &= \langle 4-(-2), 0-1, 3-5\rangle \\ &= \boxed{\langle 6,-1,-2\rangle}\end{aligned}$$

$$\begin{aligned}3\mathbf{b} &= 3\langle -2,1,5\rangle \\ &= \langle 3(-2), 3(1), 3(5)\rangle \\ &= \boxed{\langle -6,3,15\rangle}\end{aligned}$$

$$\begin{aligned}2\mathbf{a}+5\mathbf{b} &= 2\langle 4,0,3\rangle + 5\langle -2,1,5\rangle \\ &= \langle 8,0,6\rangle + \langle -10,5,25\rangle \\ &= \boxed{\langle -2,5,31\rangle}\end{aligned}$$

We call V_2 the set of all two-dimensional vectors and V_3 the set of all three-dimensional vectors. Generally, we consider the set V_n as the set of all n-dimensional vectors, which serve as ordered n-tuples of the form

$$\mathbf{a} = \langle a_1, a_2, \ldots, a_n\rangle$$

where $a_1, a_2, \ldots, a_n$ are real numbers and are called the components of $\mathbf{a}$. Addition and scalar multiplication are defined for vectors in V_n the same way as we have done so for $n = 2$ and $n = 3$.

Definition 1.2.4. ***Properties of Vectors***

If $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are vectors in V_n and $c, d \in \mathbb{R}$ are scalars, then

1. $\boldsymbol{a}+\boldsymbol{b} = \boldsymbol{b}+\boldsymbol{a}$
2. $\boldsymbol{a}+(\boldsymbol{b}+\boldsymbol{c}) = (\boldsymbol{a}+\boldsymbol{b})+\boldsymbol{c}$
3. $\boldsymbol{a}+\boldsymbol{0} = \boldsymbol{a}$
4. $\boldsymbol{a}+(-\boldsymbol{a}) = \boldsymbol{0}$
5. $c(\boldsymbol{a}+\boldsymbol{b}) = c\boldsymbol{a}+c\boldsymbol{b}$
6. $(c+d)\boldsymbol{a} = c\boldsymbol{a}+d\boldsymbol{a}$

7. $(cd)\boldsymbol{a} = c(d\boldsymbol{a})$
8. $1\boldsymbol{a} = \boldsymbol{a}$

It is very easy to verify these eight properties of vectors either algebraically or geometrically. We will demonstrate the proof of the first property for commutative vector addition with $n = 2$:

$$\begin{aligned}\mathbf{a} + \mathbf{b} &= \langle a_1, a_2 \rangle + \langle b_1, b_2 \rangle = \langle a_1 + b_1, a_2 + b_2 \rangle \\ &= \langle b_1 + a_1, b_2 + a_2 \rangle = \langle b_1, b_2 \rangle + \langle a_1, a_2 \rangle \\ &= \mathbf{b} + \mathbf{a}\end{aligned}$$

and the proof is complete. □

We now consider three vectors in the set of three-dimensional vectors V_3 that play a very important role. These are

$$\boxed{\mathbf{i} = \langle 1, 0, 0 \rangle \qquad \mathbf{j} = \langle 0, 1, 0 \rangle \qquad \mathbf{k} = \langle 0, 0, 1 \rangle}$$

The vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are referred to as **standard basis vectors**. They all have magnitude 1 and point in the directions of the positive x-, y-, and z-axes. In two dimensions, we only consider $\mathbf{i} = \langle 1, 0 \rangle$ and $\mathbf{j} = \langle 0, 1 \rangle$.

Consider the set V_3. If $\mathbf{a} \in V_3$ can be expressed in its component form $\langle a_1, a_2, a_3 \rangle$, then we write

$$\begin{aligned}\mathbf{a} &= \langle a_1, a_2, a_3 \rangle = \langle a_1, 0, 0 \rangle + \langle 0, a_2, 0 \rangle + \langle 0, 0, a_3 \rangle \\ &= a_1 \langle 1, 0, 0 \rangle + a_2 \langle 0, 1, 0 \rangle + a_3 \langle 0, 0, 1 \rangle \\ \mathbf{a} &= a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}\end{aligned}$$

So we observe that *any* vector in V_3 can be written in terms of the standard basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$. In the case of V_2, we use the standard bases $\mathbf{i}$ and $\mathbf{j}$.

Problem 1.2.3. *If $\boldsymbol{a} = \boldsymbol{i} + 2\,\boldsymbol{j} - 3\,\boldsymbol{k}$ and $\boldsymbol{b} = 4\,\boldsymbol{i} + 7\,\boldsymbol{k}$, express the vector $2\boldsymbol{a} + 3\boldsymbol{b}$ in terms of $\boldsymbol{i}$, $\boldsymbol{j}$, and $\boldsymbol{k}$.*

Solution: There isn't much to do here except apply the properties of vectors.

$$\begin{aligned}2\mathbf{a} + 3\mathbf{b} &= 2(\mathbf{i} + 2\,\mathbf{j} - 3\,\mathbf{k}) + 3(4\,\mathbf{i} + 7\,\mathbf{k}) \\ &= 2\,\mathbf{i} + 4\,\mathbf{j} - 6\,\mathbf{k} + 12\,\mathbf{i} + 21\,\mathbf{k} \\ &= \boxed{14\,\mathbf{i} + 4\,\mathbf{j} + 15\,\mathbf{k}}\end{aligned}$$

A **unit vector** is the name for a vector with magnitude 1. Therefore, the standard basis vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ would fit this definition. In general, if a vector $\mathbf{a}$ is not the zero vector, then the unit vector in the same direction as $\mathbf{a}$ is

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

To quickly verify, let $c = 1/|\mathbf{a}|$. Then $\mathbf{u} = c\mathbf{a}$ and c is a positive scalar, so $\mathbf{u}$ has the same direction as $\mathbf{a}$. Also

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

Problem 1.2.4. *Find the unit vector in the direction of the vector* $\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$.

Solution: The length of this vector is

$$|\mathbf{i} - \mathbf{j} + \mathbf{k}| = \sqrt{1^2 + (-1)^2 + 1^2} = \sqrt{3}$$

so the unit vector with same direction is $\boxed{\mathbf{u} = \frac{1}{\sqrt{3}}(\mathbf{i} - \mathbf{j} + \mathbf{k})}$.

Vector Applications Beyond just mathematics, vectors are highly useful in physics and the engineering fields. In Chapter 2, we will discuss how they can describe the motion of objects in space. For now we will consider a basic example: forces. Force is considered a vector quantity because it has both magnitude (usually expressed in pounds or newtons) and direction. If several forces act on an object at once, then the **resultant force** is represented by the sum of these vectors, an application of vector addition.

Problem 1.2.5. *A* 100*-lb weight hangs from two wires arranged in a system shown below. Calculate the tension forces* $\boldsymbol{T}_1$ *and* $\boldsymbol{T}_2$ *in the wires, as well as the magnitudes of these tensions.*

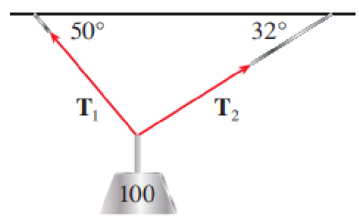


Image Credit: Calculus, Early Transcendentals

Solution: First we express the tension vectors in terms of their horizontal and vertical components.

$$\mathbf{T}_1 = -|\mathbf{T}_1| \cos 50^\circ \, \mathbf{i} + |\mathbf{T}_1| \sin 50^\circ \, \mathbf{j}$$
$$\mathbf{T}_2 = |\mathbf{T}_2| \cos 32^\circ \, \mathbf{i} + |\mathbf{T}_2| \sin 32^\circ \, \mathbf{j}$$

The resultant vector $\mathbf{T}_1 + \mathbf{T}_2$ should cancel out the weight $\mathbf{w} = 100\,\mathbf{j}$ of the object, so we have

$$(-|\mathbf{T}_1| \cos 50^\circ + |\mathbf{T}_2| \cos 32^\circ)\, \mathbf{i} + (|\mathbf{T}_1| \sin 50^\circ + |\mathbf{T}_2| \sin 32^\circ)\, \mathbf{j} = 100\,\mathbf{j}$$

and matching the vector components we get

$$-|\mathbf{T}_1| \cos 50^\circ + |\mathbf{T}_2| \cos 32^\circ = 0$$
$$|\mathbf{T}_1| \sin 50^\circ + |\mathbf{T}_2| \sin 32^\circ = 100$$

We'll solve the first equation for $|\mathbf{T}_2|$ and substitute into the second to get

$$|\mathbf{T}_1| \sin 50^\circ + \frac{|\mathbf{T}_1| \cos 50^\circ}{\cos 32^\circ} \sin 32^\circ = 100$$

$$|\mathbf{T}_1| \left(\sin 50^\circ + \cos 50^\circ \frac{\sin 32^\circ}{\cos 32^\circ} \right) = 100$$

So the magnitudes of the tensions $|\mathbf{T}_1|$ and $|\mathbf{T}_2|$ are

$$|\mathbf{T}_1| = \frac{100}{\sin 50^\circ + \tan 32^\circ \cos 50^\circ} \approx \boxed{85.64 \text{ lb}}$$

$$|\mathbf{T}_2| = \frac{|\mathbf{T}_1| \cos 50^\circ}{\cos 32^\circ} \approx \boxed{64.91 \text{ lb}}$$

and so the tension vectors are

$$\mathbf{T}_1 \approx \boxed{-55.05\,\mathbf{i} + 65.60\,\mathbf{j}}$$
$$\mathbf{T}_2 \approx \boxed{55.05\,\mathbf{i} + 34.40\,\mathbf{j}}$$

Let's branch out even farther in real life. When flying in wind, a plane's *true course* or *track* is the direction of the resultant velocity vector which is obtained by summing the velocity vectors of the plane and the wind. The magnitude of this resultant vector is called the plane's *ground speed.* Similarly, when a boat navigates through the water, it follows a true course in the direction of the resultant of the velocity vectors of the boat as well as the conventional current of the water.

Problem 1.2.6. *A woman launches a boat from the south shore of a straight river that flows directly west at* 4 *mi/h. She wants to land at the point directly across on the opposite shore. If the speed of the boat (relative to the water) is* 8 *mi/h, in what direction should she steer the boat in order to arrive at the desired landing point?*

Solution: Let the coordinate axes be chosen so that the positive x-direction points east and the positive y-direction points north.

The velocity of the river is

$$\mathbf{v}_r = -4\,\mathbf{i} \quad (\text{mi/h})$$

The boat's velocity relative to the water has constant magnitude 8 mi/h. Let θ be the angle the boat is steered east of north. Then

$$\mathbf{v}_{bw} = 8\big(\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}\big)$$

The resultant velocity vector, or the boat's velocity relative to the ground, is given by

$$\mathbf{v}_{bg} = \mathbf{v}_{bw} + \mathbf{v}_r$$

Substituting,

$$\mathbf{v}_{bg} = \big(8\sin\theta - 4\big)\mathbf{i} + 8\cos\theta\,\mathbf{j}$$

To land directly across the river, the net east-west displacement must be zero. Therefore, the x-component of $\mathbf{v}_{bg}$ must vanish:

$$8\sin\theta - 4 = 0$$

Solving,

$$\sin\theta = \frac{1}{2} \quad \Longrightarrow \quad \theta = 30^\circ$$

Thus the woman should steer the boat in the direction pointing $\boxed{30^\circ \text{ northeast}}$.

1.3 Dot Product

So far we have worked with adding two vectors and multiplying vectors by scalars. Now the question arises as to whether it is possible to multiply two vectors and get a useful scalar result? One such product is the dot product, which will be the topic of this section. The other product is the cross product, which is discussed in the next section.

Dot Product of Two Vectors Let **a** and **b** be two vectors. To find the dot product, we multiply their corresponding components and add.

Definition 1.3.1. *If $\boldsymbol{a}$ and $\boldsymbol{b}$ are two vectors in V_3 and $\boldsymbol{a} = \langle a_1, a_2, a_3 \rangle$ and $\boldsymbol{b} = \langle b_1, b_2, b_3 \rangle$, then the* ***dot product*** *of $\boldsymbol{a}$ and $\boldsymbol{b}$ is the number*

$$\boldsymbol{a} \cdot \boldsymbol{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

Again, the dot product of two vectors is a real number, not a vector. For this reason, it is also known as the **scalar product**. Although Definition 1.3.1 is given for three-dimensional vectors (vectors in V_3), it can also be given for two-dimensional vectors (V_2) in a similar fashion:

$$\langle a_1, a_2 \rangle \cdot \langle b_1, b_2 \rangle = a_1 b_1 + a_2 b_2$$

Problem 1.3.1. *Find the dot product of $\boldsymbol{a} = 2\,\boldsymbol{i} - 3\,\boldsymbol{j}$ and $\boldsymbol{b} = -\boldsymbol{i} + 5\,\boldsymbol{j}$.*

Solution: We simply know $\mathbf{a} = \langle a_1, a_2 \rangle = \langle 2, -3 \rangle$ and $\mathbf{b} = \langle b_1, b_2 \rangle = \langle -1, 5 \rangle$. Applying Definition 1.3.1, the dot product is

$$\begin{aligned} \mathbf{a} \cdot \mathbf{b} &= a_1 b_1 + a_2 b_2 \\ &= (2)(-1) + (-3)(5) \\ &= \boxed{-17} \end{aligned}$$

The dot product, being a scalar, obeys many of the laws that hold for ordinary products of real numbers. They are summarized below.

Definition 1.3.2. ***Properties of the Dot Product***

If $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are vectors in V_3 and c is a scalar, then

1. $\boldsymbol{a} \cdot \boldsymbol{a} = |\boldsymbol{a}|^2$
2. $\boldsymbol{a} \cdot \boldsymbol{b} = \boldsymbol{b} \cdot \boldsymbol{a}$
3. $\boldsymbol{a} \cdot (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \cdot \boldsymbol{b} + \boldsymbol{a} \cdot \boldsymbol{c}$
4. $(c\boldsymbol{a}) \cdot \boldsymbol{b} = c(\boldsymbol{a} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot (c\boldsymbol{b})$
5. $\boldsymbol{0} \cdot \boldsymbol{a} = 0$

Proof. (Properties #1 and #3).

The first property is basically saying $\mathbf{a} \cdot \mathbf{a} = \langle a_1, a_2, a_3 \rangle \cdot \langle a_1, a_2, a_3 \rangle$. Applying Definition 1.3.1, we have

$$\mathbf{a} \cdot \mathbf{a} = a_1^2 + a_2^2 + a_3^2 = |\mathbf{a}|^2$$

and the proof is complete. □

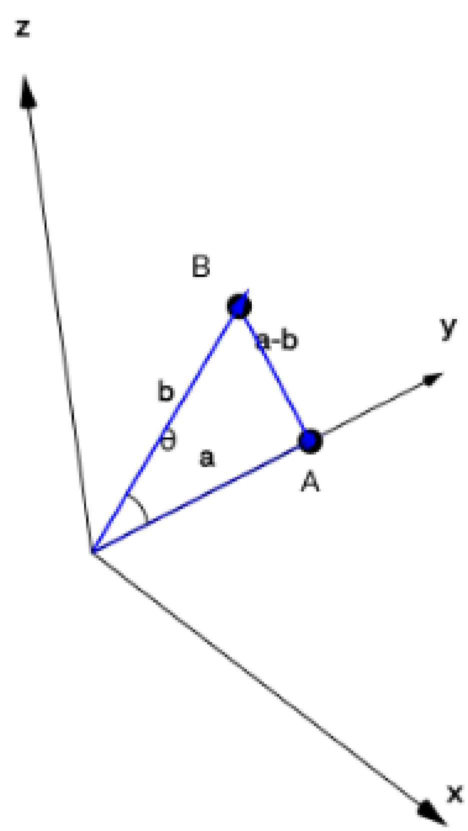
The third property involves some algebraic manipulation. Follow the steps shown below.

$$\begin{aligned}
\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) &= \langle a_1, a_2, a_3 \rangle \cdot \langle b_1 + c_1, b_2 + c_2, b_3 + c_3 \rangle \\
&= a_1(b_1 + c_1) + a_2(b_2 + c_2) + a_3(b_3 + c_3) \\
&= (a_1 b_1 + a_2 b_2 + a_3 b_3) + (a_1 c_1 + a_2 c_2 + a_3 c_3) \\
&= \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}
\end{aligned}$$

and the proof is complete. □

Try to prove the other properties yourself!

The dot product $\mathbf{a} \cdot \mathbf{b}$ has a geometric interpretation; particularly in terms of the angle θ between the two vectors. We let $\mathbf{a}$ and $\mathbf{b}$ start at the origin, with $0 \leq \theta \leq \pi$. Notice that if $\mathbf{a}$ and $\mathbf{b}$ are parallel vectors, then $\theta = 0$ or $\theta = \pi$.



Physicists use the following theorem to define their version of the dot product.

Theorem 1.3.1. *If θ is the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, then the dot product is given by*

$$\boldsymbol{a} \cdot \boldsymbol{b} = |\boldsymbol{a}||\boldsymbol{b}| \cos\theta$$

Proof. Let O be the origin. If we apply the Law of Cosines to triangle OAB, as shown above, we have

$$|AB|^2 = |OA|^2 + |OB|^2 - 2|OA||OB| \cos\theta$$

Observe that $|OA| = |\mathbf{a}|$, $|OB| = |\mathbf{b}|$, and $|AB| = |\mathbf{a} - \mathbf{b}|$, so the Law of Cosines equation becomes

$$|\mathbf{a} - \mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos\theta$$

Using properties of the dot product given in Definition 1.3.2, the left hand side of this equation becomes

$$\begin{aligned} |\mathbf{a} - \mathbf{b}|^2 &= (\mathbf{a} - \mathbf{b}) \cdot (\mathbf{a} - \mathbf{b}) \\ &= \mathbf{a} \cdot \mathbf{a} - \mathbf{a} \cdot \mathbf{b} - \mathbf{b} \cdot \mathbf{a} + \mathbf{b} \cdot \mathbf{b} \\ &= |\mathbf{a}|^2 - 2\mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2 \end{aligned}$$

We have

$$|\mathbf{a}|2 - 2\mathbf{a} \cdot \mathbf{b} + |\mathbf{b}|^2 = |\mathbf{a}|^2 + |\mathbf{b}|^2 - 2|\mathbf{a}||\mathbf{b}| \cos\theta$$

$$-2\mathbf{a} \cdot \mathbf{b} = -2|\mathbf{a}||\mathbf{b}| \cos\theta \therefore \mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}| \cos\theta$$

and the proof is complete. □

Problem 1.3.2. *If two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ have respective lengths of* 3 *and* 7*, and the angle between them is $\pi/4$, find $\boldsymbol{a} \cdot \boldsymbol{b}$.*

Solution: We know the values of $|\mathbf{a}|$ and $|\mathbf{b}|$. Also, $\cos\frac{\pi}{4} = \frac{\sqrt{2}}{2}$, so our answer is

$$3 \cdot 7 \cdot \frac{\sqrt{2}}{2} = \boxed{\frac{21\sqrt{2}}{2}}$$

Theorem 1.3.1 also has a nice corollary which allows us to determine the angle between two vectors given their dot product and lengths:

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

Problem 1.3.3. *Find the angle between the vectors $\boldsymbol{a} = \langle 2, 3, -4 \rangle$ and $\boldsymbol{b} = \langle 6, -1, 5 \rangle$.*

Solution: We find the lengths of the vectors.

$$|\mathbf{a}| = \sqrt{2^2 + 3^2 + (-4)^2} = \sqrt{29} \qquad |\mathbf{b}| = \sqrt{6^2 + (-1)^2 + 5^2} = \sqrt{62}$$

The dot product of the two vectors is

$$\mathbf{a} \cdot \mathbf{b} = (2)(6) + (3)(-1) + (-4)(5) = -11$$

So from the corollary to Theorem 1.3.1, we have

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}||\mathbf{b}|} = \frac{-11}{\sqrt{1798}}$$

So the angle between **a** and **b** is

$$\theta = \cos^{-1}\left(\frac{-11}{\sqrt{1798}}\right) \approx \boxed{1.83 \text{ (or } 103.04^\circ)}$$

Given that **a** and **b** are two nonzero vectors, we can call them **perpendicular** or **orthogonal** if the angle between them is $\theta = \pi/2$. Also, the $\mathbf{0}$ vector is considered to be perpendicular to all vectors.

Problem 1.3.4. *Show that* $2\,\boldsymbol{i} + 2\,\boldsymbol{j} - \boldsymbol{k}$ *is perpendicular to* $5\,\boldsymbol{i} - 4\,\boldsymbol{j} + 2\,\boldsymbol{k}$*.*

Solution: We need to show that the dot product of these two vectors is zero.

$$(2\,\mathbf{i} + 2\,\mathbf{j} - \mathbf{k})\cdot(5\,\mathbf{i} - 4\,\mathbf{j} + 2\,\mathbf{k}) = (2)(5) + (2)(-4) + (-1)(2) = 0$$

and thus the vectors are orthogonal.

Let θ be the angle between the vectors **a** and **b**. When θ is *acute* $(0 \le \theta < \frac{\pi}{2})$, we have $\cos\theta > 0$, so the dot product $\mathbf{a}\cdot\mathbf{b}$ is positive. This indicates that the vectors point generally in the same direction.

When θ is a *right angle* $(\theta = \frac{\pi}{2})$, $\cos\theta = 0$, and therefore $\mathbf{a}\cdot\mathbf{b} = 0$. In this case, the vectors are perpendicular and have no directional alignment.

When θ is *obtuse* $(\frac{\pi}{2} < \theta \le \pi)$, $\cos\theta < 0$, so the dot product $\mathbf{a}\cdot\mathbf{b}$ is negative. This reflects that the vectors point generally in opposite directions.

Thus, the dot product $\mathbf{a}\cdot\mathbf{b}$ measures how closely two vectors align in direction: it is positive for acute angles, zero for right angles, and negative for obtuse angles.

In the extreme cases, if **a** and **b** point in exactly the same direction, then $\theta = 0$, $\cos\theta = 1$, and

$$\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|.$$

If they point in exactly opposite directions, then $\theta = \pi$, $\cos\theta = -1$, and $\mathbf{a}\cdot\mathbf{b} = -|\mathbf{a}||\mathbf{b}|$.

Direction Angles and Direction Cosines We start this topic with a basic definition.

Definition 1.3.3. *The* ***direction angles*** *of a nonzero vector* $\boldsymbol{a}$ *are the angles* α*,* β*, and* γ *on the interval* $[0,\pi]$ *that* $\boldsymbol{a}$ *makes with the positive* x*-,* y*-, and* z*-axes, respectively.*

The cosines of these direction angles, $\cos\alpha$, $\cos\beta$, and $\cos\gamma$, are called the **direction cosines** of the vector **a**. Using the corollary to Theorem 1.3.1, it is easy to list the three direction cosines, using the standard basis vectors **i**, **j**, and **k**:

$$\cos\alpha = \frac{\mathbf{a}\cdot\mathbf{i}}{|\mathbf{a}||\mathbf{i}|} \quad \cos\beta = \frac{\mathbf{a}\cdot\mathbf{j}}{|\mathbf{a}||\mathbf{j}|} \quad \cos\gamma = \frac{\mathbf{a}\cdot\mathbf{k}}{|\mathbf{a}||\mathbf{k}|}$$

which simplifies down to

$$\cos\alpha = \frac{a_1}{|\mathbf{a}|} \quad \cos\beta = \frac{a_2}{|\mathbf{a}|} \quad \cos\gamma = \frac{a_3}{|\mathbf{a}|}$$

It can also be shown that $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$. Therefore,

$$\begin{aligned}\mathbf{a} = \langle a_1, a_2, a_3\rangle &= \langle |\mathbf{a}|\cos\alpha, |\mathbf{a}|\cos\beta, |\mathbf{a}|\cos\gamma\rangle \\ &= |\mathbf{a}|\langle\cos\alpha, \cos\beta, \cos\gamma\rangle\end{aligned}$$

and so

$$\frac{\mathbf{a}}{|\mathbf{a}|} = \langle\cos\alpha, \cos\beta, \cos\gamma\rangle$$

In other words, the direction cosines of **a** are the components of the unit vector **u** in the direction of **a**, where $\mathbf{u} = \frac{\mathbf{a}}{|\mathbf{a}|}$.

Problem 1.3.5. *Find the direction cosines and direction angles of the vector* $\boldsymbol{a} = \langle 2, 1, 4\rangle$.

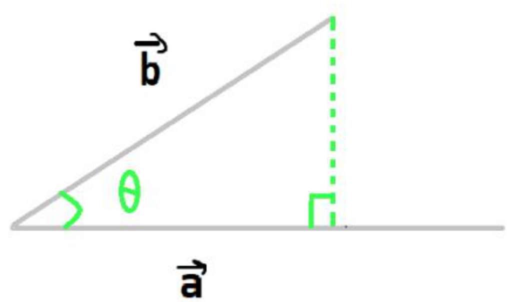
Solution: The length of the vector is $|\mathbf{a}| = \sqrt{2^2 + 1^2 + 4^2} = \sqrt{21}$, so the direction cosines are

$$\boxed{\cos\alpha = \frac{2}{\sqrt{21}} \quad \cos\beta = \frac{1}{\sqrt{21}} \quad \cos\gamma = \frac{4}{\sqrt{21}}}$$

and so taking the inverse cosine gives

$$\boxed{\alpha = \cos^{-1}\left(\frac{2}{\sqrt{21}}\right) \approx 64.1^\circ \quad \beta = \cos^{-1}\left(\frac{1}{\sqrt{21}}\right) \approx 77.4^\circ \quad \gamma = \cos^{-1}\left(\frac{4}{\sqrt{21}}\right) \approx 29.2^\circ}$$

Projections Let's consider two vectors **a** and **b**, and let them share an angle θ between them. Let **b** make a projection onto **a**. To conceptualize this more clearly, assume two sticks are aligned in vector position. Then we place a torch in on condition over **b**. Then we can see a shadow along the first stick (vector **a**) that is produced by the projection of the second stick (vector **b**).



The portion of the vector **b** that lies on **a** is called the **vector projection** of **b** onto **a** and is denoted by $\text{proj}_{\mathbf{a}}\mathbf{b}$. It can help to view this as the "shadow" of **b**.

The **scalar projection** of **b** onto **a** is the signed magnitude of the vector projection, which is really $|\mathbf{b}|\cos\theta$, where θ is between **a** and **b**. Observe that $\text{comp}_{\mathbf{a}}\mathbf{b} < 0$ if $\pi/2 \le \theta \le \pi$. The equation

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

shows that the dot product of **a** and **b** can be interpreted as the length of **a** multiplied by the scalar projection of **b** onto **a**. Since

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

the component of **b** along **a** can be found by taking the dot product of **b** with the unit vector in the direction of **a**. Summarizing these ideas, we have

Scalar projection of **b** onto **a**: $\quad \text{comp}_{\mathbf{a}}\mathbf{b} = \dfrac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|}$

Vector projection of **b** onto **a**: $\quad \text{proj}_{\mathbf{a}}\mathbf{b} = \left(\dfrac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|}\right)\dfrac{\mathbf{a}}{|\mathbf{a}|} = \dfrac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|^2}\,\mathbf{a}$

Notice that the vector projection of **b** onto **a** is really just the scalar projection multiplied by the unit vector in the direction of **a**.

Problem 1.3.6. *Find the scalar and vector projections of* $\boldsymbol{b} = \boldsymbol{i} - \boldsymbol{j} + 2\,\boldsymbol{k}$ *onto* $\boldsymbol{a} = 3\,\boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$*.*

Solution: We need to determine the length of **a**, or $|\mathbf{a}|$:

$$|\mathbf{a}| = \sqrt{3^2 + (-1)^2 + 1^2} = \sqrt{11}$$

Then the scalar projection of **b** onto **a** is

$$\text{comp}_{\mathbf{a}}\mathbf{b} = \frac{\mathbf{a}\cdot\mathbf{b}}{|\mathbf{a}|} = \frac{(1)(3) + (-1)(-1) + (2)(1)}{\sqrt{11}} = \boxed{\frac{6}{\sqrt{11}}}$$

The vector projection is the scalar projection multiplied by the unit vector in the direction of vector **a**:

$$\text{proj}_{\mathbf{a}}\mathbf{b} = \frac{6}{\sqrt{11}}\frac{\mathbf{a}}{|\mathbf{a}|} = \frac{6}{11}\mathbf{a} = \boxed{\left\langle \frac{18}{11}, -\frac{6}{11}, \frac{6}{11} \right\rangle}$$

Dot Product Applications: Work In physics, vector projections are used in calculating the quantity of work.

Definition 1.3.4. *The work W done by a constant force F in moving an object through a distance d is given by*

$$W = Fd$$

Now suppose that the constant force is a vector **F** pointing at an angle with respect to the horizontal. If the force moves an object from the point of origin to another location at a horizontal distance of $|\mathbf{D}|$, where **D** is the displacement vector, then we have

$$W = (|\mathbf{F}|\cos\theta)|\mathbf{D}|$$

and Theorem 1.3.1 gives

$$W = |\mathbf{F}||\mathbf{D}|\cos\theta = \mathbf{F}\cdot\mathbf{D}$$

Thus the work done by a constant force **F** on an object is the dot product, $\mathbf{F}\cdot\mathbf{D}$, where **D** is the displacement vector.

Problem 1.3.7. *A wagon is pulled at a distance of* 100 *m along a horizontal path by a constant force of* 70 *N. The handle of the wagon is held at an angle* $35°$ *above the horizontal. Find the work done on the wagon.*

Solution: Let $\mathbf{F}$ and $\mathbf{D}$ be the force and displacement vectors. It's easy to see that the angle between the two vectors is $\theta = 35°$. When approaching problems like these, first try drawing out the scenario, so you don't risk taking the complementary angle, or $90 - 35 = 65°$. Then the work done is

$$\begin{aligned} W &= \mathbf{F} \cdot \mathbf{D} = |\mathbf{F}||\mathbf{D}| \cos 35° \\ &= (70)(100) \cos 35° \approx 5734\,\mathrm{N} \cdot \mathrm{m} \\ &= \boxed{5734\,\mathrm{J}} \end{aligned}$$

Problem 1.3.8. *A force is given by a vector* $\boldsymbol{F} = 3\,\boldsymbol{i} + 4\,\boldsymbol{j} + 7\,\boldsymbol{k}$ *and moves a particle from the point* $P(1, 0, 0)$ *to* $Q(2, 4, 3)$*. Find the work done on the particle.*

Solution: The displacement vector is $\mathbf{D} = \overrightarrow{PQ} = \langle 1, 4, 3\rangle$, so we take the dot product with $\mathbf{F} = \langle 3, 4, 7\rangle$ to get

$$\begin{aligned} W &= \mathbf{F} \cdot \mathbf{D} = \langle 3, 4, 7\rangle \cdot \langle 1, 4, 3\rangle \\ &= 3 + 16 + 21 = \boxed{40} \end{aligned}$$

If the unit of length is meters and the force is given in newtons, then the work done is 40 J.

1.4 Cross Product

In mathematics, we encounter pairs of nonzero vectors frequently. It is then useful to find a nonzero vector that is orthogonal to both of them. In this section, we explore an operation, the cross product, that produces such a vector.

Cross Product of Two Vectors Consider two vectors $\mathbf{a}$ in $\mathbf{b}$ in V_3. We have $\mathbf{a} = \langle a_1, a_2, a_3\rangle$ and $\mathbf{b} = \langle b_1, b_2, b_3\rangle$. We wish to find a vector $\mathbf{c} = \langle c_1, c_2, c_3\rangle$ that is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$. By the definition of the dot product, we must have $\mathbf{a} \cdot \mathbf{c} = 0$ and $\mathbf{b} \cdot \mathbf{c} = 0$. So we obtain the equations

$$a_1c_1 + a_2c_2 + a_3c_3 = 0$$

$$b_1c_1 + b_2c_2 + b_3c_3 = 0$$

We can multiply the first equation by b_3 and the second by a_3 and subtract the two in order to eliminate c_3.

$$(a_1b_3 - a_3b_1)c_1 + (a_2b_3 - a_3b_2)c_2 = 0$$

This is in the form $pc_1 + qc_2 = 0$, for which the obvious solution is $c_1 = q$ and $c_2 = -p$. Therefore one possible solution c_1, c_2 is

$$c_1 = a_2b_3 - a_3b_2 \qquad c_2 = a_3b_1 - a_1b_3$$

If we substitute into the original system of equations, we can find

$$c_3 = a_1b_2 - a_2b_1$$

Thus the vector **c** perpendicular to **a** and **b** is

$$\langle c_1, c_2, c_3\rangle = \langle a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1\rangle$$

Definition 1.4.1. *The resulting vector* ***c*** *is called the* ***cross product*** *of* ***a*** *and* ***b*** *and is denoted by* $\boldsymbol{a} \times \boldsymbol{b}$*.*

Notice that the cross product is a vector, compared to the scalar that is a dot product. Therefore, it is also called the **vector product**. Also, $\mathbf{a} \times \mathbf{b}$ is defined only when $\mathbf{a}, \mathbf{b} \in V_3$.

It is easier to internalize Definition 1.4.1 if we consider matrix determinants. Let's consider the determinant of **order 2**.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

For **order 3**, we can define the determinant in terms of second-order determinants.

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

Observe that each term on the right hand side of the equation involves a number a_i in the first row of the determinant, multiplied by the second-order determinant obtained from the left side by deleting the row and column in which a_i appears. Also, notice the minus sign in the second term.

If we now rewrite Definition 1.4.1 using second-order determinants and the standard basis vectors, we know the cross product of the vectors $\mathbf{a} = a_1\,\mathbf{i} + a_2\,\mathbf{j} + a_3\,\mathbf{k}$ and $\mathbf{b} = b_1\,\mathbf{i} + b_2\,\mathbf{j} + b_3\,\mathbf{k}$ is

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \mathbf{i} - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \mathbf{j} + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \mathbf{k} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

The use of determinant notation is probably the easiest way to memorize the computation of cross products.

Problem 1.4.1. *If* $\boldsymbol{a} = \langle 1, 3, 4\rangle$ *and* $\boldsymbol{b} = \langle 2, 7, -5\rangle$*, compute the cross product* $\boldsymbol{a} \times \boldsymbol{b}$*.*

Solution: We compute the third order determinant to find the cross product.

$$\begin{aligned} \mathbf{a} \times \mathbf{b} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 3 & 4 \\ 2 & 7 & -5 \end{vmatrix} \\ &= \begin{vmatrix} 3 & 4 \\ 7 & -5 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 4 \\ 2 & -5 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 3 \\ 2 & 7 \end{vmatrix} \mathbf{k} \\ &= (-15 - 28)\,\mathbf{i} - (-5 - 8)\,\mathbf{j} + (7 - 6)\,\mathbf{k} = \boxed{-43\,\mathbf{i} + 13\,\mathbf{j} + \mathbf{k}} \end{aligned}$$

Problem 1.4.2. *Show that for any vector* $\boldsymbol{a} \in V_3$ *we must have* $\boldsymbol{a} \times \boldsymbol{a} = \boldsymbol{0}$.

Since $\mathbf{a}$ is in the set V_3, we can express it as $\mathbf{a} = \langle a_1, a_2, a_3 \rangle$. Thus

$$\begin{aligned} \mathbf{a} \times \mathbf{a} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ a_1 & a_2 & a_3 \end{vmatrix} \\ &= (a_2a_3 - a_3a_2)\,\mathbf{i} - (a_1a_3 - a_3a_1)\,\mathbf{j} + (a_1a_2 - a_2a_1)\,\mathbf{k} \\ &= 0\,\mathbf{i} - 0\,\mathbf{j} + 0\,\mathbf{k} = \mathbf{0} \end{aligned}$$

and the proof is complete. □

Cross Product Properties The most important reason as to why we constructed the cross product $\mathbf{a} \times \mathbf{b}$ was to find a vector perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. We'll examine this fact in the theorem below.

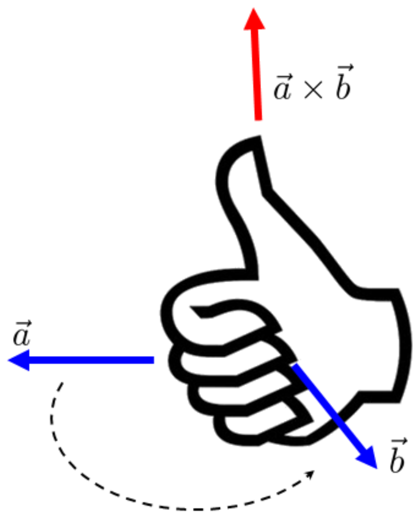
Theorem 1.4.1. *The vector* $\boldsymbol{a} \times \boldsymbol{b}$ *is orthogonal to both* $\boldsymbol{a}$ *and* $\boldsymbol{b}$*, where* $\boldsymbol{a}$ *and* $\boldsymbol{b}$ *are vectors in* V_3.

Proof. To show that $\mathbf{a} \times \mathbf{b}$ is orthogonal to $\mathbf{a}$, we need to check the following dot product:

$$\begin{aligned} (\mathbf{a} \times \mathbf{b}) \cdot \mathbf{a} &= \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} a_1 - \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} a_2 + \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} a_3 \\ &= a_1(a_2b_3 - a_3b_2) - a_2(a_1b_3 - a_3b_1) + a_3(a_1b_2 - a_2b_1) \\ &= a_1a_2b_3 - a_1a_3b_2 - a_1a_2b_3 + a_2a_3b_1 + a_1a_3b_2 - a_2a_3b_1 \\ &= 0 \end{aligned}$$

Similar computations can be applied to show $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{b} = 0$. Therefore, the proof is complete. □

If we treat vectors $\mathbf{a}$ and $\mathbf{b}$ as directed line segments with the same initial point, then Theorem 1.4.1 implies that the cross product $\mathbf{a} \times \mathbf{b}$ points in a direction perpendicular to the plane through $\mathbf{a}$ and $\mathbf{b}$. The actual direction of $\mathbf{a} \times \mathbf{b}$ is given by the *right-hand-rule*; if the fingers of your right hand curl in the direction of rotation at an angle $\theta \in [0, \pi]$ from $\mathbf{a}$ to $\mathbf{b}$, then your thumb points in the direction of the cross product $\mathbf{a} \times \mathbf{b}$.



Now that we can determine the direction of $\mathbf{a} \times \mathbf{b}$, we consider its most important geometric property: length. This is given by the theorem below.

Theorem 1.4.2. *If θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$ with $0 \leq \theta \leq \pi$, then the length of the cross product $\boldsymbol{a} \times \boldsymbol{b}$ is given by*

$$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta$$

Proof. We will use the definition of cross product and the properties of vector length.

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

Taking the square roots of both sides, and recognizing that $\sqrt{\sin^2\theta} = \sin\theta$ because $\sin\theta \geq 0$ on $[0, \pi]$, we have

$$|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}||\mathbf{b}| \sin\theta$$

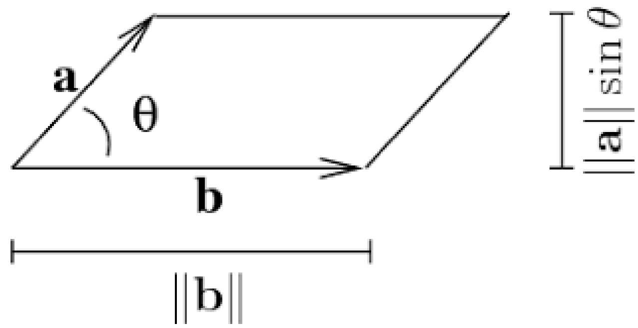
and the proof is complete. ☐

There is a nice corollary that follows from the above theorem, specifically that two vectors $\mathbf{a}$ and $\mathbf{b}$ are parallel if and only if the cross product is the zero vector:

$$\mathbf{a} \times \mathbf{b} = \mathbf{0}$$

Specifically, the vectors are parallel if the angle θ between them is 0 or π. This is because $\sin 0 = 0$ and $\sin\pi = 0$, so $|\mathbf{a} \times \mathbf{b}| = 0$ and thus $\mathbf{a} \times \mathbf{b} = \mathbf{0}$.

Let's move on. Since vectors are determined exclusively by magnitude and direction, we can now consider two vectors **a** and **b** that are *non-parallel*. We know that the magnitude is given by $|\mathbf{a}\times\mathbf{b}| = |\mathbf{a}||\mathbf{b}|\sin\theta$ and the direction is determined by the right-hand rule. Actually, the geometric interpretation of Theorem 1.4.2 can be understood by observing the figure shown below.



If **a** and **b** are treated as directed line segments with the same starting point, then they determine a parallelogram with base $|\mathbf{b}|$ and height $|\mathbf{a}|\sin\theta$, and thus area

$$\boxed{A = |\mathbf{b}|\,(|\mathbf{a}|\sin\theta) = |\mathbf{a}\times\mathbf{b}|}$$

and now we have a nice method of interpreting the magnitude of a cross product.

Problem 1.4.3. *Find a vector perpendicular to the plane that passes through the points* $P(1,4,6)$, $Q(2,5,-1)$, *and* $R(1,-1,1)$.

Solution: The cross product $\overrightarrow{PQ}\times\overrightarrow{PR}$ is orthogonal to both segments $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ and and is also orthogonal to the plane determined by the points P, Q, and R. The vector representations for the directed line segments are

$$\overrightarrow{PQ} = (-2-1)\,\mathbf{i} + (5-4)\,\mathbf{j} + (-1-6)\,\mathbf{k} = -3\,\mathbf{i} + \mathbf{j} - 7\,\mathbf{k}$$

$$\overrightarrow{PR} = (1-1)\,\mathbf{i} + (-1-4)\,\mathbf{j} + (1-6)\,\mathbf{k} = -5\,\mathbf{j} - 5\,\mathbf{k}$$

We now compute the cross product of these two vectors:

$$\begin{aligned}\overrightarrow{PQ}\times\overrightarrow{PR} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -3 & 1 & -7 \\ 0 & -5 & -5 \end{vmatrix} \\ &= (-5-35)\mathbf{i} - (15-0)\mathbf{j} + (15-0)\mathbf{k} \\ &= \boxed{-40\mathbf{i} - 15\mathbf{j} + 15\mathbf{k}}\end{aligned}$$

So the vector $\langle -40, -15, 15\rangle$ is perpendicular to the specified plane. Note that any nonzero scalar multiple of this vector, such as $\langle -8, -3, 3\rangle$ is also perpendicular to the plane.

Problem 1.4.4. *Find the area of the* ***triangle*** *with the same vertices P, Q, and R as in the previous problem.*

Solution: From the previous problem we know that $\overrightarrow{PQ} \times \overrightarrow{PR} = \langle -40, -15, 15\rangle$. The area of the parallelogram determined by the vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ is the length of the cross product:

$$\left|\overrightarrow{PQ} \times \overrightarrow{PR}\right| = \sqrt{(-40)^2 + (-15)^2 + 15^2} = 5\sqrt{82}$$

Intuitively speaking, the area of ΔPQR should be half the area of the parallelogram, that is, $\boxed{\frac{5}{2}\sqrt{82}}$.

If we apply Theorems 1.4.1 and 1.4.2 to the standard basis vectors, taking $\theta = \pi/2$ as they are all mutually orthogonal, we obtain the following six equalities:

1. $\mathbf{i} \times \mathbf{j} = \mathbf{k}$
2. $\mathbf{j} \times \mathbf{k} = \mathbf{i}$
3. $\mathbf{k} \times \mathbf{i} = \mathbf{j}$
4. $\mathbf{j} \times \mathbf{i} = -\mathbf{k}$
5. $\mathbf{k} \times \mathbf{j} = -\mathbf{i}$
6. $\mathbf{i} \times \mathbf{k} = -\mathbf{j}$

Observe that $\mathbf{i} \times \mathbf{j} \neq \mathbf{j} \times \mathbf{i}$, so cross product is not commutative. Also,

$$\mathbf{i} \times (\mathbf{i} \times \mathbf{j}) = \mathbf{i} \times \mathbf{k} = -\mathbf{j}$$

whereas

$$(\mathbf{i} \times \mathbf{i}) \times \mathbf{j} = \mathbf{0} \times \mathbf{j} = \mathbf{0}$$

so the associative law for multiplication is not valid for cross products; in general,

$$(\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})$$

However, some algebraic laws *are* valid for cross product. They are summarized in the definition below.

Definition 1.4.2. *If $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$ are vectors in V_3 and $c \in \mathbb{R}$ is a scalar then*

1. $\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$
2. $(c\boldsymbol{a}) \times \boldsymbol{b} = c(\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a} \times (c\boldsymbol{b})$
3. $\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}$
4. $(\boldsymbol{a} + \boldsymbol{b}) \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{c} + \boldsymbol{b} \times \boldsymbol{c}$
5. $\boldsymbol{a} \cdot (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}$
6. $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}$

It is very easy to prove these properties if you write each vector in component form and simply apply the definition of the cross product. We will prove the fifth property and the remaining can be left as exercises for you.

Proof.

Let $\mathbf{a} = \langle a_1, a_2, a_3\rangle$, $\mathbf{b} = \langle b_1, b_2, b_3\rangle$, and $\mathbf{c} = \langle c_1, c_2, c_3\rangle$. We have

$$\begin{aligned}\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) &= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1)\\ &= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1\\ &= (a_2b_3 - a_3b_2)c_1 + (a_3b_1 - a_1b_3)c_2 + (a_1b_2 - a_2b_1)c_3\\ &= (\mathbf{a}\times\mathbf{b})\cdot\mathbf{c}.\end{aligned}$$

and the proof is complete. □

The Triple Product In the fifth property we just proved, the product $\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})$ is called the **scalar triple product** of the vectors **a**, **b**, and **c**. It is called a triple product because it involves three vectors, and it is scalar because of the dot product. We can express this value as a determinant:

$$\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c}) = \begin{vmatrix} a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}$$

The geometric significance of this triple product is to consider the figure shown below, which represents a parallelepiped determined by the three vectors **a**, **b**, and **c**.

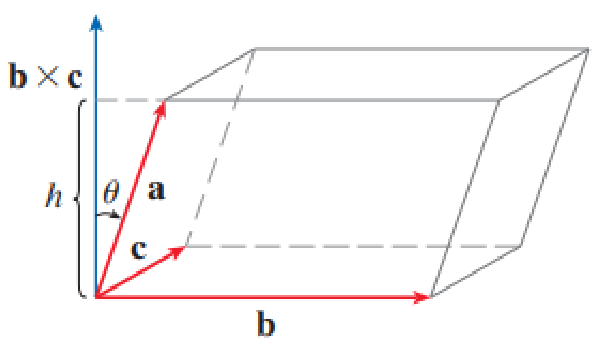


Image Credit: Calculus, Early Transcendentals

We can see that the area of the base (parallelogram) is given by $A = |\mathbf{b}\times\mathbf{c}|$. Let θ be the angle between the vectors **a** and $\mathbf{b}\times\mathbf{c}$, then the height h of the parallelepiped is $h = |\mathbf{a}||\cos\theta|$. We need to put the absolute value just in case $\theta > \pi/2$. Thus the volume of the parallelepiped is

$$\boxed{V = Ah = |\mathbf{b}\times\mathbf{c}||\mathbf{a}||\cos\theta| = |\mathbf{a}\cdot(\mathbf{b}\times\mathbf{c})|}$$

We have a nice fact that follows from the above formula: if the volume of the parallelepiped determined by **a**, **b**, and **c** is 0, then the vectors must lie in the same plane, i.e. they are *coplanar*.

Problem 1.4.5. *Show that the vectors* $\boldsymbol{a} = \langle 1, 4, -7\rangle$*,* $\boldsymbol{b} = \langle 2, -1, 4\rangle$*, and* $\boldsymbol{c} = \langle 0, -9, 18\rangle$ *are coplanar.*

Solution: We simply compute the scalar triple product and check if it is equal to zero.

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \begin{vmatrix} 1 & 4 & -7 \\ 2 & -1 & 4 \\ 0 & -9 & 18 \end{vmatrix}$$

$$= 1\begin{vmatrix} -1 & 4 \\ -9 & 18 \end{vmatrix} - 4\begin{vmatrix} 2 & 4 \\ 0 & 18 \end{vmatrix} - 7\begin{vmatrix} 2 & -1 \\ 0 & -9 \end{vmatrix}$$

$$= 1(18) - 4(36) - 7(-18) = 0.$$

Therefore, the volume of the parallelepiped determined by **a**, **b**, and **c** is 0, so the vectors are coplanar. □

Let's revisit Definition 1.4.2. The sixth property involves $\mathbf{a} \times (\mathbf{b} \times \mathbf{c})$, which is called the **vector triple product** of the vectors **a**, **b**, **c**. This property will arise again in Chapter 2 when we derive Kepler's First Law of planetary motion.

Cross Product Applications: Torque Cross products have the most applications in the physics field. Consider a force **F** acting on a rigid body at a point described by position vector **r**. Suppose the rigid body is a wrench. If we tighten a bolt by applying a force to the wrench, we get this turning effect. This effect has a name−**torque**−and is defined to be the cross product of the position and force vectors:

$$\tau = \mathbf{r} \times \mathbf{F}$$

Torque measures the tendency of a body to rotate about the origin (or generally, its axis of rotation). The direction of the torque vector indicates the axis of rotation. According to Theorem 1.4.2, the magnitude of the torque vector is

$$|\tau| = |\mathbf{r} \times \mathbf{F}| = |\mathbf{r}||\mathbf{F}| \sin\theta$$

where θ is the angle between the position and force vectors. Note that the only component of **F** that can cause a rotation is the one perpendicular to **r**, that is, $|\mathbf{F}| \sin\theta$.

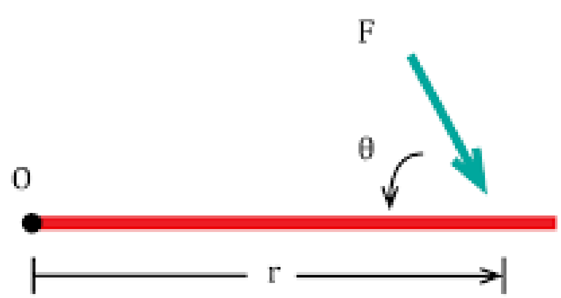


Image Credit: University of Guelph (Ontario) Physics Department

Connecting the previous discussion, we can say that the magnitude of the torque is equal to the area of the parallelogram determined by the vectors $\mathbf{r}$ and $\mathbf{F}$.

Problem 1.4.6. *A bolt is tightened by applying a 40-N force to a 0.25-m wrench, where the angle of the force vector relative to the wrench is 75°. Find the torque magnitude about the center of the bolt.*

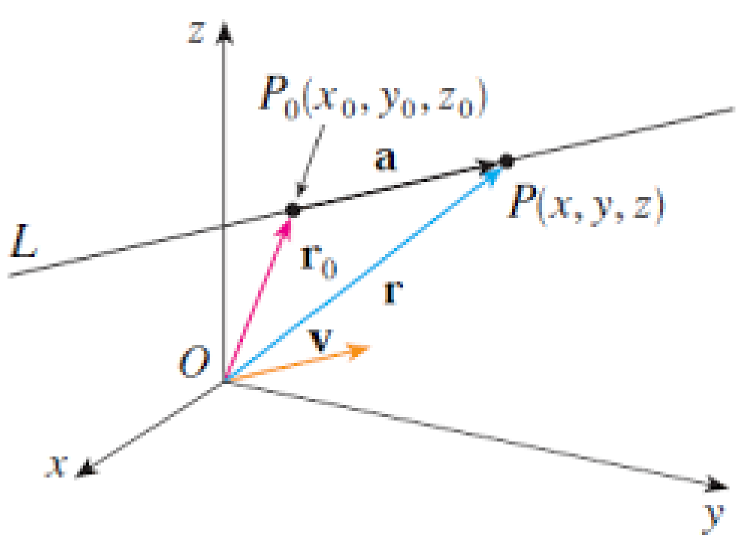
Solution: The magnitude of the torque vector is equal to the magnitude of the cross product of $\mathbf{r}$ and $\mathbf{F}$, or $|\mathbf{r} \times \mathbf{F}|$. We can easily infer that $\theta = 75^\circ$ is the desired angle, i.e. it is located between the position and force vectors.

$$\begin{aligned} |\tau| &= |\mathbf{r} \times \mathbf{F}| = |\mathbf{r}||\mathbf{F}| \sin 75^\circ = (0.25)(40) \sin 75^\circ \\ &= 10 \sin 75^\circ \approx \boxed{9.66\,\mathrm{N} \cdot \mathrm{m}} \end{aligned}$$

1.5 Equations of Lines and Planes

What are Lines? In elementary algebra, a line in the xy-plane is determined when a point on the line and its direction (or inclination) are given. The equation of the line can then be written using point-slope form.

Likewise, in three-dimensional space, a line L can be determined with a point $P_0(x_0, y_0, z_0)$ on L and a direction for L, which is described by a vector $\mathbf{v}$ parallel to the line. Let $P(x, y, z)$ be an arbitrary point on L and let $\mathbf{r}_0$ and $\mathbf{r}$ be the position vectors for P_0 and P, respectively. In other words, they have representations of $\overrightarrow{OP_0}$ and $\overrightarrow{OP}$. If $\mathbf{a}$ is the vector with representation $\overrightarrow{P_0P}$, as shown below, then the Triangle Law for vector addition (see section 1.2) gives $\mathbf{r} = \mathbf{r}_0 + \mathbf{a}$.



Since $\mathbf{a}$ and $\mathbf{v}$ are parallel vectors, there must exist a scalar t such that $\mathbf{a} = t\mathbf{v}$. Thus

$$\boxed{\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}}$$

This is called the **vector equation** of L. Each value of the *parameter* t gives the position vector $\mathbf{r}$ of a point on L. As t varies, the line is traced out by the tip of the vector $\mathbf{r}$.

If $\mathbf{v}$ which gives the direction of L is given in component form, i.e. $\mathbf{v} = \langle a, b, c\rangle$, then $t\mathbf{v} = \langle ta, tb, tc\rangle$. We also have $\mathbf{r} = \langle x, y, z\rangle$ and $\mathbf{r}_0 = \langle x_0, y_0, z_0\rangle$, so the vector equation becomes

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

Two vectors are equivalent if and only if their corresponding components are equal, so we have three scalar equations:

$$x = x_0 + at \quad y = y_0 + bt \quad z = z_0 + ct$$

for all values $t \in \mathbb{R}$. These equations are called **parametric equations** of the L through the point $P_0(x_0, y_0, z_0)$ and parallel to $\mathbf{v} = \langle a, b, c\rangle$. Each value of t gives a unique point (x, y, z) on the line.

Problem 1.5.1. *Answer the following questions.*
(a) Find a vector equation and parametric equations for the line that passes through the point $(5, 1, 3)$ *and is parallel to the vector* $\boldsymbol{i} + 4\,\boldsymbol{j} - 2\,\boldsymbol{k}$.
(b) Find any two other points on the line.

Solution to part a: We have $\mathbf{r}_0 = \langle 5, 1, 3\rangle = 5\,\mathbf{i} + \mathbf{j} + 3\,\mathbf{k}$ and $\mathbf{v} = \mathbf{i} + 4\,\mathbf{j} - 2\,\mathbf{k}$ so the vector equation of the line is

$$\mathbf{r} = (5\,\mathbf{i} + \mathbf{j} + 3\,\mathbf{k}) + t(\mathbf{i} + 4\,\mathbf{j} - 2\,\mathbf{k})$$

$$\therefore \boxed{\mathbf{r} = (5 + t)\,\mathbf{i} + (1 + 4t)\,\mathbf{j} + (3 - 2t)\,\mathbf{k}}$$

and the parametric equations are

$$\boxed{x = 5 + t \quad y = 1 + 4t \quad z = 3 - 2t}$$

Solution to part b: Different parameter values t will give us different points on the line. Let's choose $t = 1$. This gives $x = 6$, $y = 5$, and $z = 1$, so $\boxed{(6, 5, 1)}$ is one point on the line. Now let $t = 2$. This gives $x = 7$, $y = 9$, and $z = -1$, so another point on the line is $\boxed{(7, 9, -1)}$.

Remark. The vector equations and parametric equations of a line are not unique. If we change either the point or choose a different parallel vector, then the equations will change. You can observe this for yourself.

In general for a vector $\mathbf{v} = \langle a, b, c\rangle$ that describes the direction of a line L, then the elements of the set $\{a, b, c\}$ are called the **direction numbers** of L.

We can also describe the equation of L by eliminating the parameter from the set of parametric equations. Provided that $a, b, c \neq 0$, we have

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

If we equate all results by the transitive property, we have

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

These equations are called the **symmetric equations** of L. Note that because a, b, and c are nonzero, they all appear in the denominator. However, even if one of them is zero, we can still eliminate the parameter. If for instance $a = 0$, then the symmetric equations would be

$$x = x_0 \quad \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

indicating that L lies in the vertical plane $x = x_0$, where x_0 is a number.

Problem 1.5.2. *Answer the following questions.*
(a) Find parametric and symmetric equations of the line that passes through the points $A(2, 4, -3)$ *and* $B(3, -1, 1)$.
(b) At what point does the line intersect the xy-plane?

Solution to part a: Notice that we are not explicitly given a vector parallel to the line, but we can construct a vector $\mathbf{v}$ with representation $\overrightarrow{AB}$ which is also parallel to the line by intuition.

$$\mathbf{v} = \langle 3 - 2, -1 - 4, 1 - (-3)\rangle = \langle 1, -5, 4\rangle$$

Therefore the direction numbers are $\{a, b, c\} = \{1, -5, 4\}$. We take the point P_0 as $(2, 4, -3)$, so the parametric equations are

$$\boxed{x = 2 + t \quad y = 4 - 5t \quad z = -3 + 4t}$$

and symmetric equations are

$$\boxed{\frac{x - 2}{1} = \frac{y - 4}{-5} = \frac{z + 3}{4}}$$

Solution to part b: The line intersects the plane xy-plane when $z = 0$. From the parametric equations we have $z = -3 + 4t = 0$, so $t = \frac{3}{4}$. Using this value of t, we have $x = 2 + \frac{3}{4} = \frac{11}{4}$ and $y = 4 - 5\left(\frac{3}{4}\right) = \frac{1}{4}$. Thus our answer is $\boxed{\left(\frac{11}{4}, \frac{1}{4}, 0\right)}$.

In general, we observe that the direction numbers of a line L through the points $P_0(x_0, y_0, z_0)$ and $P_1(x_1, y_1, z_1)$ are $x_1 - x_0$, $y_1 - y_0$, and $z_1 - z_0$, so the symmetric equations are

$$\frac{x - x_0}{x_1 - x_0} = \frac{y - y_0}{y_1 - y_0} = \frac{z - z_0}{z_1 - z_0}$$

Many times we need a description of not an entire line, but usually just a line segment. Let's look back at problem 1.5.2. Consider the line segment AB: if we put $t = 0$ in the parametric equations, we get the point $(2, 4, -3)$ and if we put $t = 1$, we get $(3, -1, 1)$. Therefore AB is described by the parametric equations

$$x = 2 + t \quad y = 4 - 5t \quad z = -3 + 4t \quad 0 \leq t \leq 1$$

or by the vector equation

$$\mathbf{r}(t) = \langle 2 + t, 4 - 5t, -3 + 4t\rangle \quad 0 \leq t \leq 1$$

The vector equation of a line through the tip of the vector $\mathbf{r}_0$ in the direction of a vector $\mathbf{v}$ is given by $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$. If the line also passes through the tip of $\mathbf{r}_1$, then $\mathbf{v} = \mathbf{r}_1 - \mathbf{r}_0$ and thus the vector equation is

$$\boxed{\mathbf{r} = \mathbf{r}_0 + t(\mathbf{r}_1 - r\mathbf{r}_0) = (1-t)\mathbf{r}_0 + t\mathbf{r}_1}$$

on the parameter interval $0 \le t \le 1$.

Problem 1.5.3. *Show that the two lines L_1 and L_2 with parametric equations*

$$L_1: \quad x = 1 + t \quad y = -2 + 3t \quad z = 4 - t$$

$$L_2: \quad x = 2s \quad y = 3 + s \quad z = -3 + 4s$$

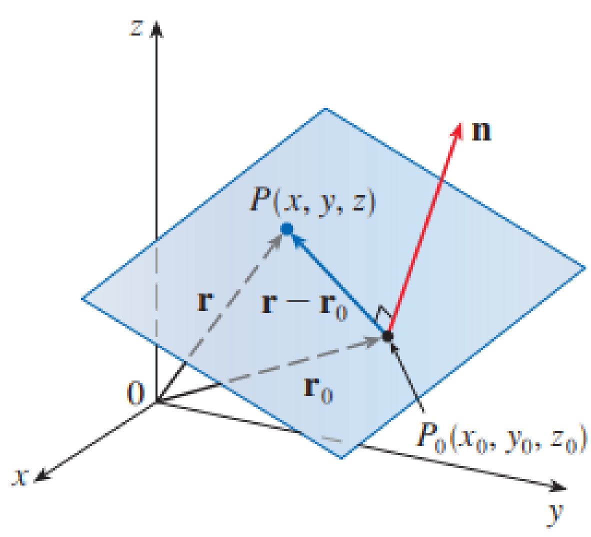
*are **skew**, i.e. they do not intersect and are not parallel (therefore do not lie in the same plane).*

Solution: It is not too difficult to show that L_1 and L_2 are not parallel because the corresponding vectors $\langle 1, 3, -1\rangle$ and $\langle 2, 1, 4\rangle$ are not parallel because the components are not proportional. Also, if L_1 and L_2 intersect, then there would be values of $t, s \in \mathbb{R}$ such that

$$\begin{aligned} 1 + t &= 3s \\ -2 + 3t &= 3 + s \\ 4 - t &= -3 + 4s \end{aligned}$$

Solving the first two equations yields $t = \frac{11}{5}$ and $s = \frac{8}{5}$, and these values don't satisfy the third equation. Therefore, there are no values $t, s \in \mathbb{R}$ so that all three equations are satisfied. Therefore, L_1 and L_2 don't intersect. Since they fail both criteria, they are skew, and the proof is complete. □

Planes In space, lines are determined by a point and direction. Meanwhile, planes are a bit of a challenge to describe. A single vector parallel to the plane is not enough to convey its "direction." However, a vector that is *perpendicular* to the plane absolutely specifies its direction. Thus a plane in space is determined by a point $P_0(x_0, y_0, z_0)$ and a vector $\mathbf{n}$ that is orthogonal to the plane. This vector is also called a **normal vector**. Let $P(x, y, z)$ be an arbitrary point in the plane, and let $\mathbf{r}_0$ and $\mathbf{r}$ be the respective position vectors of the points P_0 and P. Then $\mathbf{r} - \mathbf{r}_0$ has representation $\overrightarrow{P_0P}$.



The normal vector is orthogonal to every vector in the given plane, particularly $\mathbf{r} - \mathbf{r}_0$, therefore the dot product is zero:

$$\boxed{\mathbf{n} \cdot (\mathbf{r} - \mathbf{r}_0) = 0}$$

which can be rewritten as

$$\mathbf{n} \cdot \mathbf{r} = \mathbf{n} \cdot \mathbf{r}_0$$

The boxed equation is referred to as the **vector equation** of the plane. To obtain a scalar equation, however, we write $\mathbf{n} = \langle a, b, c\rangle$, $\mathbf{r} = \langle x, y, z\rangle$, and $\mathbf{r}_0 = \langle x_0, y_0, z_0\rangle$ and the vector equation becomes

$$\langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0$$

Let us expand the left side of the equation to get the scalar equation of the plane through $P_0(x_0, y_0, z_0)$ with normal vector $\mathbf{n} = \langle a, b, c\rangle$:

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

Problem 1.5.4. *Find an equation of the plane passing through the point* $(2, 4, -1)$ *with normal vector* $\boldsymbol{n} = \langle 2, 3, 4\rangle$. *Also, find the intercepts.*

Solution: The scalar equation of the plane is given by

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

which simplifies to $\boxed{2x + 3y + 4z = 12}$.

To find the intercepts we set the other two variables equal to zero at a time. For instance, if $y = z = 0$ the x-intercept can be found to be $\boxed{x = 6}$. Using a similar process, the y- and z-intercepts are $\boxed{y = 4}$ and $\boxed{z = 3}$, respectively.

The equation of a plane can also be written as

$$\boxed{ax + by + cz + d = 0}$$

where $d = -(ax_0 + by_0 + cz_0)$. The above equality is called the **linear equation** of the plane in x, y, and z.

Problem 1.5.5. *Find an equation of the plane that passes through the points* $P(1, 3, 2)$, $Q(3, -1, 6)$, *and* $R(5, 2, 0)$.

Solution: Let's choose two vectors **a** and **b** which correspond to $\overrightarrow{PQ}$ and $\overrightarrow{PR}$, respectively. It is important that both vectors have the same starting point; in this case, we have point P. Also

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

Both these vectors lie in the plane so the cross product $\mathbf{a} \times \mathbf{b}$ is orthogonal to the plane and is considered as the normal vector. Thus

$$\mathbf{n} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2 & -4 & 4 \\ 4 & -1 & -2 \end{vmatrix} = 12\,\mathbf{i} + 20\,\mathbf{j} + 14\,\mathbf{k}$$

With the point $P(1,3,2)$ and the normal vector $\mathbf{n}$ the equation of the plane becomes

$$12(x-1)+20(y-3)+14(z-2)=0$$

or in simplified form, we have $\boxed{6x+10y+7z=50}$.

Problem 1.5.6. *Find the point at which the line with parametric equations $x=2+3t$, $y=-4t$, $z=5+t$ intersects the plane $4x+5y-2z=18$.*

Solution: Substitute the expressions for x, y, and z from the parametric equations into the equation of the plane:

$$4(2+3t)+5(-4t)-2(5+t)=18$$

This reduces to $-10t=20$, so $t=-2$. Therefore, the point of intersection occurs at the parameter value $t=2$. Then $x=2+3(-2)=-4$, $y=-4(-2)=8$, and $z=5-2=3$, so the point of intersection is $\boxed{(-4,8,3)}$.

Parallel Planes We say two planes are **parallel** if their normal vectors are parallel. For example, the planes $x+2y-3z=4$ and $2x+4y-6z=9$ are parallel because their normal vectors are parallel, i.e. $\mathbf{n}_1=\langle 1,2,-3\rangle$, $\mathbf{n}_2=\langle 2,4,-6\rangle$ and $\mathbf{n}_2=2\mathbf{n}_1$. If the two planes are not parallel, then they intersect in a straight line and the angle between the two planes is equal to the acute angle between their normal vectors.

Problem 1.5.7. *Answer the following questions.*
(a) Find the angle between the planes $x+y+z=1$ and $x-2y+3z=1$.
(b) Find symmetric equations for the line of intersection L of these two planes.

Solution to part a: The normal vectors of the planes are

$$\mathbf{n}_1=\langle 1,1,1\rangle \qquad \mathbf{n}_2=\langle 1,-2,3\rangle$$

and so if θ is the angle between these vectors, then the corollary to Theorem 1.3.1 gives

$$\begin{aligned}\cos\theta&=\frac{\mathbf{n}_1\cdot\mathbf{n}_2}{|\mathbf{n}_1||\mathbf{n}_2|}=\frac{1(1)+1(-2)+1(3)}{\sqrt{1+1+1}\sqrt{1+4+9}}=\frac{2}{\sqrt{42}}\\ \theta&=\cos^{-1}\left(\frac{2}{\sqrt{42}}\right)\approx\boxed{72^\circ}\end{aligned}$$

Solution to part b: We first need to find a point on L. Let's find the point where the line intersects the xy-plane by setting $z=0$ in both planes' equations. This gives us $x+y=1$ and $x-2y=1$, and the solution to this system is $x=1$, $y=0$. So we know the point $(1,0,0)$ lies on L.

Now we observe that L lies in both planes, so it is perpendicular to the normal vectors. Thus, a vector $\mathbf{v}$ parallel to L is given by

$$\mathbf{v}=\mathbf{n}_1\times\mathbf{n}_2=\begin{vmatrix}\mathbf{i}&\mathbf{j}&\mathbf{k}\\1&1&1\\1&-2&3\end{vmatrix}=5\,\mathbf{i}-2\,\mathbf{j}-3\,\mathbf{k}$$

and so the symmetric equations of L are written as

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

1.6 Overview of Cylinders and Quadric Surfaces

In sections 1.1 and 1.5, we talked about two different types of surfaces: spheres and planes, respectively. In this section, we talk about two more surfaces, namely cylinders and quadric surfaces.

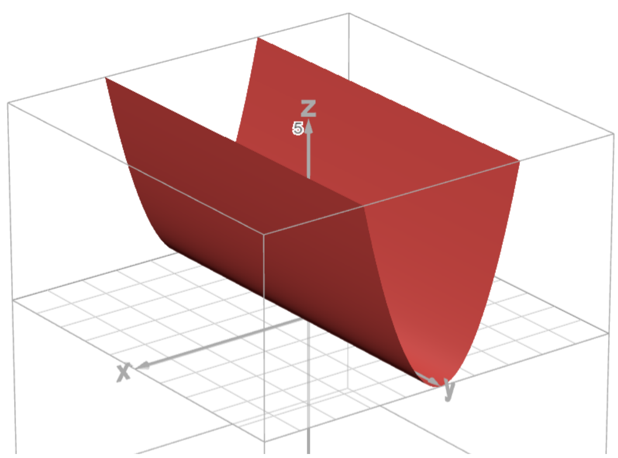
For any surface, it is useful to determine the curves of intersection of the surface with planes parallel to the coordinate planes, prior to sketching its graph. Such curves are called **traces**, or cross-sections, of the surface.

Cylinders

Definition 1.6.1. *A **cylinder** is a surface consisting of all lines, called **rulings**, parallel to a given line and pass through a plane curve.*

Problem 1.6.1. *Sketch the graph of the surface $z = x^2$.*

Solution: First and foremost, the equation of the graph does not involve the variable y. This means any vertical plane $y = k$ (parallel to the xz-plane) intersects the graph in a curve with equation $z = x^2$. Thus these vertical traces are parabolas. Moving these parabolas in the direction of the y-axis yields the **parabolic cylinder** shown below.



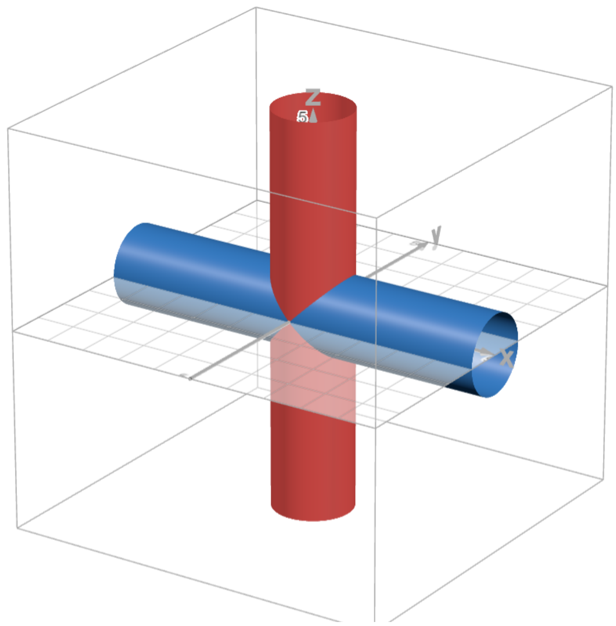
Remark. Generally, if one of the variables x, y, and z is missing from the equation of a surface, then the surface is a cylinder.

Definition 1.6.2. *Identify each surface.*

(a) $x^2 + y^2 = 1$
(b) $y^2 + z^2 = 1$

Solution to part a: The variable z is missing from the equation of this surface. Thus, $x^2 + y^2 = 1$, for $z = k$, so the $x^2 + y^= 1$ represents a circle of radius 1 in the plane $z = k$. The surface $x^2 + y^2 = 1$ is a $\boxed{\text{circular cylinder}}$ symmetric about the z-axis, shown in red.

Solution to part b: In this case x is missing. Using similar intuition, the surface $y^2 + z^2 = 1$ is a $\boxed{\text{circular cylinder}}$ symmetric about the x-axis, shown in blue.



Remark. When dealing with surfaces, it is important to realize that an equation like $x^2 + y^2 = 1$ represents a cylinder and not a circle. The trace of the cylinder $x^2 + y^2 = 1$ is the circle with equations $x^2 + y^2 = 1, z = 0$.

Quadric Surfaces

Definition 1.6.3. *A* ***quadric surface*** *is the graph of a second-degree equation in three variables x, y, and z.*

Generally, the equation is

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fxz + Gx + Hy + Iz + J = 0$$

where $\{A, B, \ldots, J\}$ are constants, but processes of rotation and translation can simplify the equation into one of two *standard forms*

$$Ax^2 + By^2 + Cz^2 + J = 0 \quad \text{or} \quad Ax^2 + By^2 + Iz = 0$$

Quadric surfaces are the counterparts in 3D of the conic sections in the 2D plane.

Definition 1.6.4. *Use traces to sketch the quadric surface with equation $x^2 + \frac{y^2}{9} + \frac{z^2}{4} = 1$.*

Solution: First substitute $z = 0$, to find the trace in the xy-plane as $x^2 + \frac{y^2}{9} = 1$, which is an ellipse. In general, the horizontal trace in the plane $z = k$ is

$$x^2 + \frac{y^2}{9} = 1 - \frac{k^2}{4} \quad z = k$$

which is an ellipse under the constraint $k^2 < 4$, that is, $-2 < x < 2$. In terms of absolute value, if $|k| = 2$, then the trace consists of a single point and the trace is empty for $|k| > 2$.

Similarly, vertical traces parallel to the yz- and xz-planes are ellipses:

$$\frac{y^2}{9} + \frac{z^2}{9} = 1 - k^2 \quad x = k \quad -1 < k < 1$$
$$x^2 + \frac{z^2}{4} = 1 - \frac{k^2}{9} \quad y = k \quad -3 < k < 3$$

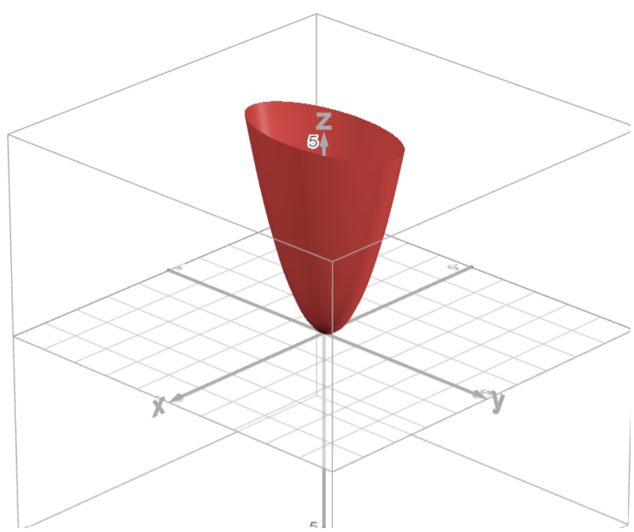
The below sketch demonstrates an **ellipsoid**, because all of its traces are ellipses. Also, the surface is symmetric with each of the coordinate planes, because the equation involves only even powers of x, y, and z.

Problem 1.6.2. *Use traces to sketch $z = 4x^2 + y^2$.*

Solution: Setting $x = 0$ gives $z = y^2$, so the intersection of the surface with the yz-plane is a parabola. If we instead let $x = k$ (a constant), then $z = y^2 + 4k^2$, which shows that slicing the surface with any plane parallel to the yz-plane produces an upward-opening parabola.

Likewise, when $y = k$, the resulting trace is $z = 4x^2 + k^2$, which is also an upward-opening parabola. If we fix $z = k$, the horizontal cross-sections satisfy $4x^2 + y^2 = k$, which represents a family of ellipses for $k > 0$.

By understanding the shapes of these traces, we can sketch the surface as shown below. Since the surface has both elliptical and parabolic cross-sections, the quadric surface given by $z = 4x^2 + y^2$ is called an **elliptic paraboloid**.



Problem 1.6.3. *Sketch the surface $\frac{x^2}{4} + y^2 - \frac{z^2}{4} = 1$.*

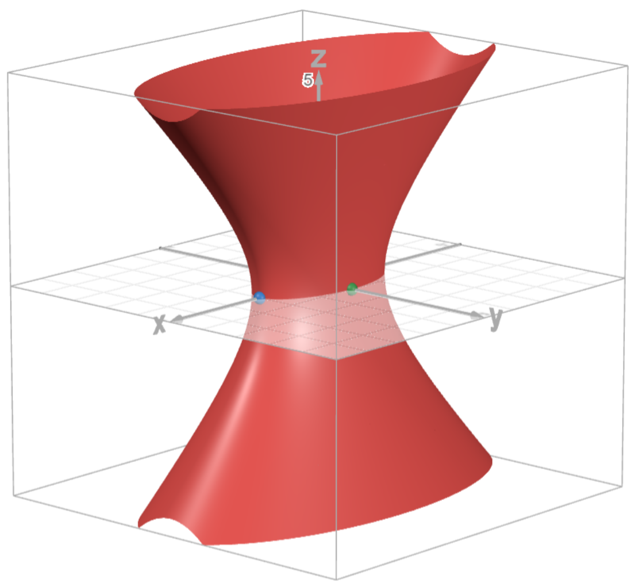
Solution: The trace in the plane $z = k$ gives the ellipse

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

but the traces in the xz- and yz-planes are the hyperbolas

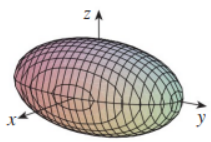
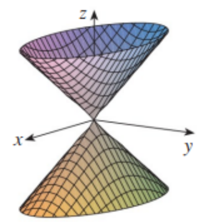
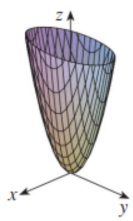
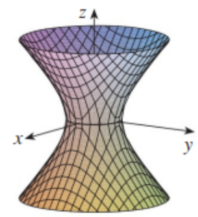
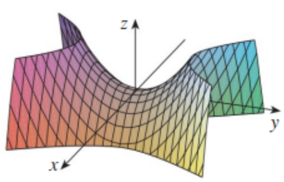
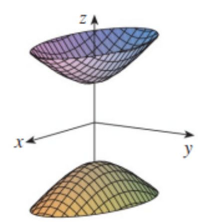
$$\frac{x^2}{4} - \frac{z^2}{4} = 1 \quad y = 0 \qquad \text{and} \qquad y^2 - \frac{z^2}{4} = 1 \quad x = 0$$

This surface is called a **hyperboloid of one sheet** and is sketched below.



The idea of using traces to draw a surface is employed in three-dimensional graphing software. In most such software, traces in the vertical planes $x = k$ and $y = k$ are drawn for equally spaced values of k.

The below table, courtesy of Calculus: Early Transcendentals, shows computer software graphs of the six basic types of quadric surfaces in standard form. For simplicity, all surfaces are symmetric to the z-axis. If symmetric about a different axis, then the equation can be adjusted accordingly.

Surface	Equation	Surface	Equation
Ellipsoid 	$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ All traces are ellipses. If $a = b = c$, the ellipsoid is a sphere.	**Cone** 	$\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces in the planes $x = k$ and $y = k$ are hyperbolas if $k \neq 0$ but are pairs of lines if $k = 0$.
Elliptic Paraboloid 	$\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$ Horizontal traces are ellipses. Vertical traces are parabolas. The variable raised to the first power indicates the axis of the paraboloid.	**Hyperboloid of One Sheet** 	$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$ Horizontal traces are ellipses. Vertical traces are hyperbolas. The axis of symmetry corresponds to the variable whose coefficient is negative.
Hyperbolic Paraboloid 	$\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$ Horizontal traces are hyperbolas. Vertical traces are parabolas. The case where $c < 0$ is illustrated.	**Hyperboloid of Two Sheets** 	$-\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ Horizontal traces in $z = k$ are ellipses if $k > c$ or $k < -c$. Vertical traces are hyperbolas. The two minus signs indicate two sheets.

Problem 1.6.4. *Identify and sketch the surface* $4x^2 - y^2 + 2z^2 + 4 = 0$.

Solution: First we put the equation in standard form by dividing by -4:

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

If we observe the above table, we see that this surface represents a **hyperboloid of two sheets**, the only difference being the axis of symmetry is the y-axis. The traces in the xy- and yz-planes are the hyperbolas

$$-x^2 + \frac{y^2}{4} = 1 \quad z = 0 \qquad \text{and} \qquad \frac{y^2}{4} - \frac{z^2}{2} = 1 \quad x = 0$$

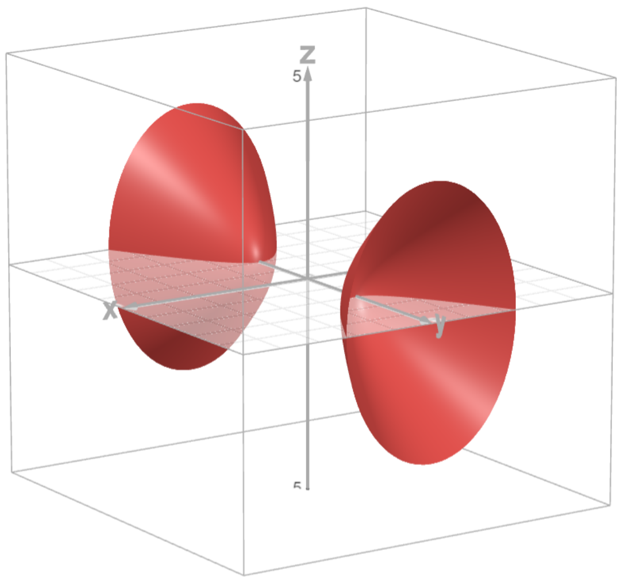
The surface has no trace in the xz-plane, but traces in the vertical planes $y = k$ for $|k| > 2$ are the ellipses

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

which can be written as

$$\frac{x^2}{\left(\frac{k^2}{4} - 1\right)} + \frac{z^2}{2\left(\frac{k^2}{4} - 1\right)} = 1 \qquad y = k$$

These traces are used to sketch the graph shown below.



Problem 1.6.5. *Classify the quadric surface* $x^2 + 2z^2 - 6x - y + 10 = 0$.

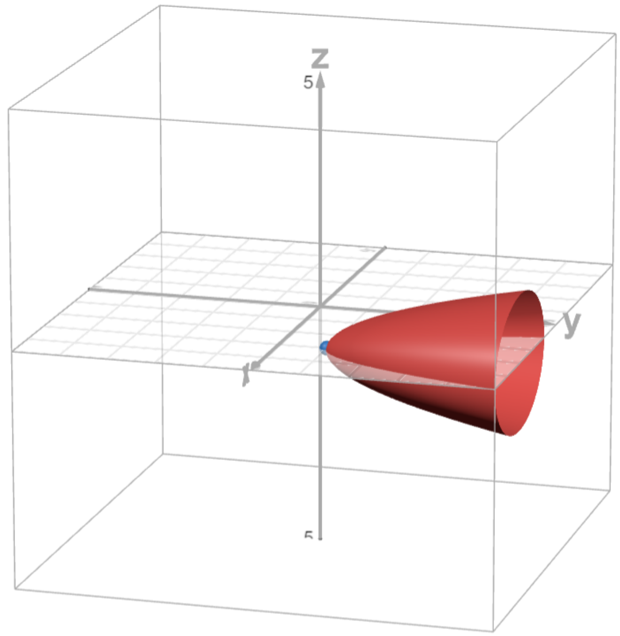
Solution: Completing the square gives

$$y - 1 = (x - 3)^2 + 2z^2$$

This surface represents an **elliptic paraboloid**, as predicted by the table. Here, however, the paraboloid is symmetric about the y-axis, and it is shifted because its vertex is located at the point $(3, 1, 0)$. For $k > 1$, the traces $y = k$ in the plane are the ellipses of the form

$$(x - 3)^2 + 2z^2 = k - 1 \qquad y = k$$

In the xy-plane, the trace is the parabola with equation $y = 1 + (x - 3)^2, z = 0$. The paraboloid and its vertex is shown below.


z
5
y
x
5

2 Vector Functions

In earlier courses, we have worked with real-valued functions, where each input gives a single number as an output. In this chapter, we will move on to vector-valued functions, which output vectors instead of numbers. These functions let us describe curves and surfaces in three-dimensional space and model the motion of objects as they move through space. We will also use vector functions to help derive Kepler's laws of planetary motion.

2.1 Vector Functions and Space Curves

Let's begin with a discussion on functions. A function is any rule that assigns to each element in a the domain an element in the range.

Definition 2.1.1. *A **vector-valued function**, or a **vector function**, is simply a function whose domain is a set of real numbers and whose range is a set of vectors.*

Specifically, we are interested in vectors $\mathbf{r}$ whose range values are three-dimensional vectors, i.e. the set V_3. For every real number t in the domain of $\mathbf{r}$, there is a unique vector in V_3 and is denoted by $\mathbf{r}(t)$. We call the real-valued functions $f(t)$, $g(t)$, and $h(t)$ the *component functions* of $\mathbf{r}(t)$ and so we can write

$$\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle = f(t)\,\mathbf{i} + g(t)\,\mathbf{j} + h(t)\,\mathbf{k}$$

Here, t denotes the independent variable because it represents time in the majority of applications of vector-valued functions.

Problem 2.1.1. *For the vector function* $\mathbf{r}(t) = \langle t^4, \ln(3-t), \sqrt{2t} \rangle$*, determine the component functions and the domain of* $\mathbf{r}$*.*

Solution: This is a vector in V_3, so there are three vector functions $f(t)$, $g(t)$, and $h(t)$ such that $\mathbf{r}(t) = \langle f(t), g(t), h(t) \rangle$. Thus $f(t) = \boxed{t^4}$, $g(t) = \boxed{\ln(3-t)}$, and $h(t) = \boxed{\sqrt{2t}}$. The expressions t^4, $\ln(3-t)$, and $\sqrt{2t}$ are all defined when $3 - t > 0$ and $t \geq 0$. Therefore, the domain of $\mathbf{r}$ is the interval $\boxed{[0, 3)}$.

Limits and Continuity The **limit** of a vector function $\mathbf{r}$ works similar to real-valued functions. You just need to take the limits of the component functions:

$$\lim_{t \to a} \mathbf{r}(t) = \left\langle \lim_{t \to a} f(t), \lim_{t \to a} g(t), \lim_{t \to a} h(t) \right\rangle$$

provided that the limits of the component functions all exist.

Notice that if we wanted, we could have also applied an ε-δ definition, nevertheless, the limits of vector functions are governed by the same rules as for real-valued functions.

Problem 2.1.2. *Find* $\lim_{t\to 0} \mathbf{r}(t)$, *where* $\mathbf{r}(t) = (1+t^3)\,\mathbf{i} + te^{-2t}\,\mathbf{j} + \frac{\sin 3t}{4t}\,\mathbf{k}$.

Solution: The answer to this problem is the vector whose components are given by the limits of the components of vector $\mathbf{r}$.

$$\begin{aligned}\lim_{t\to 0} \mathbf{r}(t) &= \left[\lim_{t\to 0}(1+t^3)\right]\,\mathbf{i} + \left[\lim_{t\to 0} te^{-2t}\right]\,\mathbf{j} + \left[\lim_{t\to 0}\frac{\sin 3t}{4t}\right]\,\mathbf{k} \\ &= \boxed{\mathbf{i} + \frac{3}{4}\,\mathbf{k}}\end{aligned}$$

We call a vector function $\mathbf{r}$ continuous at a point $t = a$ if

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

If we use similar intuition as with limits, we can know that $\mathbf{r}$ is continuous at a if and only if its component functions are continuous at a.

Defining Space Curves Now we will observe a close connection between continuous vector functions and space curves. Suppose that the component real-valued functions f, g, and h are continuous on any interval I. Call C the set of all points (x, y, z) in space satisfying

$$x = f(t) \qquad y = g(t) \qquad z = h(t)$$

as $t \in I$. This set C is defined as the **space curve**. The above three equations are called *parametric equations* of C and t is called a *parameter*. Consider C as being traced out by a moving particle whose position at any time t is the vector $\langle f(t), g(t), h(t)\rangle$. Let this vector be assigned to $\mathbf{r}$. Then $\mathbf{r}(t)$ is the position vector assigned to the point $P(f(t), g(t), h(t))$ on C. Then any continuous vector function defines a space curve traced out by the tip of the moving position vector.

Problem 2.1.3. *Describe the curve defined by the vector function*

$$\mathbf{r}(t) = \langle 2+t, 3+5t, 4-t\rangle$$

Solution: The first step is to parameterize the curve. The parametric equations are

$$x = 2+t \qquad y = 3+5t \qquad z = 4-t$$

We know from section 1.5 that this curve describes a line passing through the point $(2, 3, 4)$ and parallel to the three-dimensional vector $\langle 1, 5, -1\rangle$. The function could also be written as $\mathbf{r} = \mathbf{r}_0 + t\mathbf{v}$, where $\mathbf{r}_0 = \langle 2, 3, 4\rangle$ and $\mathbf{v} = \langle 1, 5, -1\rangle$. So the equation of the line is $\boxed{\mathbf{r}(t) = \langle 2, 3, 4\rangle + t\langle 1, 5, -1\rangle}$.

We can also represent plane curves in vector notation. For example, consider a curve with parametric equations $x = t^2 - t$ and $y = t + 1$. The vector notation is

$$\mathbf{r}(t) = (t^2 - t)\,\mathbf{i} + (t+1)\,\mathbf{j}$$

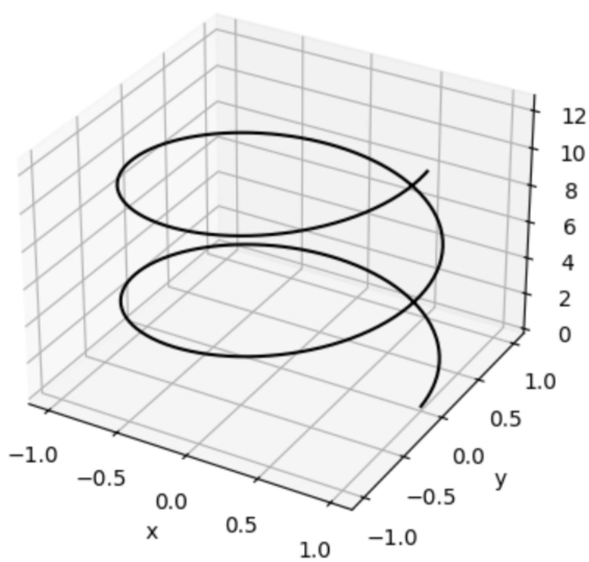
where $\mathbf{i} = \langle 1, 0\rangle$ and $\mathbf{j} = \langle 0, 1\rangle$.

Problem 2.1.4. *Describe the curve with vector equation*

$$\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$$

Solution: We'll need to do a little bit of algebra here. Since $\cos^2 t + \sin^2 t = x^2 + y^2 = 1$ for all values of $t \in I$, where I is some arbitrary parameter interval, we know that the curve must lie on the cylinder $x^2 + y^2 = 1$. The point (x, y, z) lies directly above the point $(x, y, 0)$, which moves counterclockwise around the circle $x^2 + y^2 = 1$, which is the projection of the cylinder in the xy-plane. Also, since $z = t$, we know that the curve spirals upward around the cylinder as t increases.

***Remark*.** The resulting curve is referred to as a **helix**. For the previous problem, it will look something like this:



***Remark*.** The above graph was generated using Python code.

Previously, we were given vector equations for a curve and then asked to provide a geometric description. Now, we will work in reverse: find parametric equations for a curve given its geometric description.

Problem 2.1.5. *Find a vector equation and parametric equations for the line segment conjoining the points $P(1, -2, 3)$ and $Q(4, 0, 5)$.*

Solution: From section 1.5, we know that the vector equation for a line segment conjoining the tips of vectors $\mathbf{r}_0$ and $\mathbf{r}_1$ is

$$\mathbf{r}_1 = (1-t)\mathbf{r}_0 + t\mathbf{r}_1 \qquad 0 \le t \le 1$$

In this case, we take $\mathbf{r}_0 = \langle 1, -2, 3\rangle$ and $\mathbf{r} = \mathbf{r}_1 = \langle 4, 0, 5\rangle$ to obtain the vector equation of the line segment from P to Q as

$$\begin{aligned}\mathbf{r}(t) &= (1-t)\langle 1, -2, 3\rangle + t\langle 4, 0, 5\rangle \\ &= \langle 1 - t + 4t, -2 + 2t + 0, 3 - 3t + 5t\rangle \\ &= \boxed{\langle 1 + 3t, -2 + 2t, 3 + 2t\rangle}\end{aligned}$$

And the corresponding parametric equations are

$$\boxed{x = 1 + 3t \qquad y = -2 + 2t \qquad z = 3 + 2t \qquad 0 \le t \le 1}$$

Problem 2.1.6. *Find a vector function representing the curve of intersection C of the cylinder $x^2 + y^2 = 1$ and the plane $y + z = 2$.*

Solution: We can mentally visualize this scenario. The plane $y + z = 2$ will slice the right circular cylinder $x^2 + y^2 = 1$ into a slanted ellipse, which is C. We will now project C into the xy-plane to obtain the circle $x^2 + y^2 = 1$, $z = 0$. So we can parameterize the projection of C as

$$x = \cos t \qquad y = \sin t \qquad 0 \le t \le 2\pi$$

From the plane $y + z = 2$, we know that the z-coordinate is always $2 - y$, so $z = 2 - y = 2 - \sin t$. Therefore the parametric equations for C are

$$x = \cos t \qquad y = \sin t \qquad z = 2 - \sin t \qquad 0 \le t \le 2\pi$$

and the corresponding vector equation is

$$\boxed{\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + (2 - \sin t)\,\mathbf{k} \qquad 0 \le t \le 2\pi}$$

This vector equation can also be called the *parameterization* of C.

Computer-Aided Design: Space Curves It is simply inevitable that space curves are much more difficult to illustrate by hand than plane curves. To be as accurate as possible, we will often need to use technology. For example, consider the *toroidal spiral*

$$x = (4 + \sin 20t)\cos t \qquad y = (4 + \sin 20t)\,\sin t \qquad z = \cos 20t$$

or the *trefoil knot*

$$x = (2 + \cos 1.5t)\cos t \qquad (2 + \cos 1.5t)\sin t \qquad z = \sin 1.5t$$

It would not be very easy to plot either of these curves by hand. Even when we can use a computer to draw a space curve, optical illusions can still make it difficult to get a good impression of what the curve actually appears.

2.2 Vector Functions: Derivatives and Integrals

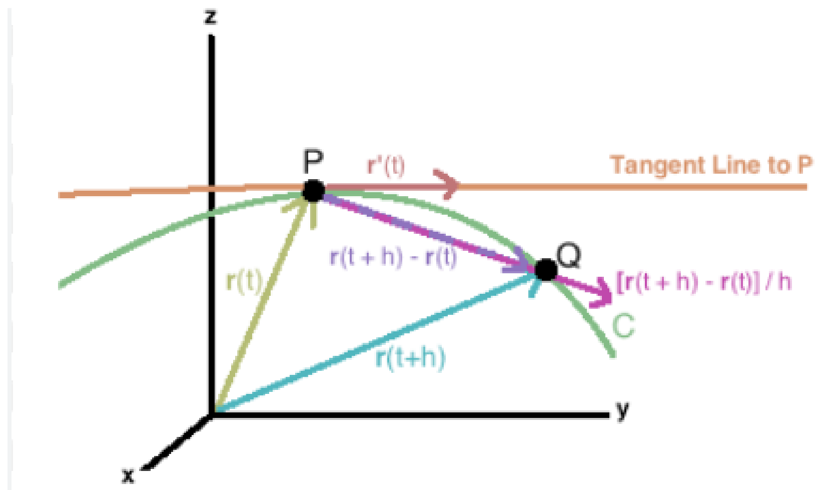
The ultimate application of the material in this chapter is to study planetary motion as well as general movement throughout space. In this section, we pave the way by first understanding the calculus of vector functions.

Derivatives A vector function $\mathbf{r}$ has derivative $\mathbf{r}'$ defined in an almost parallel manner to those for real-valued functions:

$$\boxed{\frac{d\mathbf{r}}{dt} = \mathbf{r}'(t) = \lim_{h\to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}}$$

provided that the limit is existent. The geometric significance of this formula is demonstrated below. Let two points P and Q have respective position vectors $\mathbf{r}(t)$ and $\mathbf{r}(t+h)$. The directed line segment $\overrightarrow{PQ}$ has vector representation of $\mathbf{r}(t+h) - \mathbf{r}(t)$, which is regarded as a *secant vector*.

For $h > 0$, the scalar multiple $(1/h)(\mathbf{r}(t+h) - \mathbf{r}(t))$ is in the same direction of $\mathbf{r}(t+h) - \mathbf{r}(t)$. If we take the limit as $h \to 0$, however, it appears that this vector approaches a vector lying on the tangent line. That vector is $\mathbf{r}'(t)$ and is called the *tangent vector* to the curve defined by $\mathbf{r}$ at the point P, provided that $\mathbf{r}'(t) \neq \mathbf{0}$ and $r'(t)$ is existent (these two conditions imply that the curve C is smooth). The **tangent line** to C at P is defined to be the line through P parallel to the tangent vector $\mathbf{r}'(t)$.



The theorem that follows provides a means to compute the derivative of a vector function $\mathbf{r}$: we simply differentiate each component of the vector.

Theorem 2.2.1. *If $\boldsymbol{r}(t) = \langle f(t), g(t), h(t)\rangle = f(t)\,\boldsymbol{i} + g(t)\,\boldsymbol{j} + h(t)\,\boldsymbol{k}$, where f, g, and h are differentiable functions, then*

$$\boldsymbol{r}'(t) = \langle f'(t), g'(t), h'(t)\rangle = f'(t)\,\boldsymbol{i} + g'(t)\,\boldsymbol{j} + h'(t)\,\boldsymbol{k}$$

Proof. We will use the definition of a derivative.

$$\begin{aligned}
\mathbf{r}'(t) &= \lim_{\Delta t\to 0}\frac{1}{\Delta t}[\mathbf{r}(t+\Delta t)-\mathbf{r}(t)] \\
&= \lim_{\Delta\to 0}\frac{1}{\Delta t}[\langle f(t+\Delta t), g(t+\Delta t), h(t+\Delta t)\rangle - \langle f(t), g(t), h(t)\rangle] \\
&= \lim_{\Delta t\to 0}\left\langle \frac{f(t+\Delta t)-f(t)}{\Delta t}, \frac{g(t+\Delta t)-g(t)}{\Delta t}, \frac{h(t+\Delta t)-h(t)}{\Delta t}\right\rangle \\
&= \left\langle \lim_{\Delta t\to 0}\frac{f(t+\Delta t)-f(t)}{\Delta t}, \lim_{\Delta t\to 0}\frac{g(t+\Delta t)-g(t)}{\Delta t}, \lim_{\Delta t\to 0}\frac{h(t+\Delta t)-h(t)}{\Delta t}\right\rangle \\
&= \langle f'(t),\, g'(t),\, h'(t)\rangle
\end{aligned}$$

A vector that has the same direction as the tangent vector is a unit vector called the **unit tangent vector**:

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

Problem 2.2.1. *Answer the following questions.*
(a) Find the tangent vector to $\boldsymbol{r}(t) = (1-t^2)\,\boldsymbol{i} + \arctan(2t)\,\boldsymbol{k}$.
(b) Find the unit tangent vector to the vector in (a) where $t = 0$.

Solution to part a: According to Theorem 2.2.1 the tangent vector to $\mathbf{r}$ is found by computing the derivatives of all individual components.

$$\mathbf{r}'(t) = -2t\,\mathbf{i} + 0\,\mathbf{j} + \frac{2}{1+4t^2}\,\mathbf{k} = \boxed{-2t\,\mathbf{i} + \frac{2}{1+4t^2}\,\mathbf{k}}$$

Solution to part b: As $\mathbf{r}(0) = \mathbf{i}$ and $\mathbf{r}'(0) = \frac{1}{2}\mathbf{k}$, so the unit tangent vector is

$$\mathbf{T}'(t) = \frac{\frac{1}{2}\,\mathbf{k}}{\sqrt{0+0+\frac{1}{4}}} = \boxed{\mathbf{k}}$$

Problem 2.2.2. *For the curve* $\boldsymbol{r}(t) = \sqrt{t}\,\boldsymbol{i} + (3-t)\,\boldsymbol{j}$, *find* $\boldsymbol{r}'(t)$.

Solution: All we have to do is compute the derivatives of all the components of $\mathbf{r}(t)$. So the answer is

$$\mathbf{r}'(t) = \boxed{\frac{1}{2\sqrt{t}}\,\mathbf{i} - \mathbf{j}}$$

Problem 2.2.3. *Find parametric equations for the tangent line to the helix with parametric equations of*

$$x = 2\cos t \quad y = \sin t \quad z = t$$

at the point $(0, 1, \pi/2)$.

Solution: The vector equation of the helix is $\mathbf{r}(t) = \langle 2\cos t, \sin t, t\rangle$ so

$$\mathbf{r}'(t) = \langle -2\cos t, \cos t, 1\rangle$$

By inspection, we see that the parameter value corresponding to the point $(0, 1, \pi/2)$ is $t = \pi/2$, so the tangent vector there is $\mathbf{r}'(\pi/2) = \langle -2, 0, 1\rangle$. The tangent line is the line through the point $(0, 1, \pi/2)$ parallel to the vector $\langle -2, 0, 1\rangle$, so from section 1.5 we know that the parametric equations for this line are

$$\boxed{x = -2t \quad y = t \quad z = \pi/2 + t}$$

Derivative Rules The following theorem demonstrates the differentiation formulas for vector-valued functions as counterparts to real-valued functions.

Theorem 2.2.2. *If $\boldsymbol{u}$ and $\boldsymbol{v}$ are differentiable vector functions, $c \in \mathbb{R}$ is a scalar, and f is a real-valued function, then:*

1. $\frac{d}{dt}[\boldsymbol{u}(t) + \boldsymbol{v}(t)] = \boldsymbol{u}'(t) + \boldsymbol{v}'(t)$
2. $\frac{d}{dt}[c\boldsymbol{u}(t)] = c\boldsymbol{u}'(t)$
3. $\frac{d}{dt}[f(t)\boldsymbol{u}(t)] = f'(t)\boldsymbol{u}(t) + f(t)\boldsymbol{u}'(t)$
4. $\frac{d}{dt}[\boldsymbol{u}(t) \cdot \boldsymbol{v}(t)] = \boldsymbol{u}'(t) \cdot \boldsymbol{v}(t) + \boldsymbol{u}(t) \cdot \boldsymbol{v}'(t)$
5. $\frac{d}{dt}[\boldsymbol{u}(t) \times \boldsymbol{v}(t)] = \boldsymbol{u}'(t) \times \boldsymbol{v}(t) + \boldsymbol{u}(t) \times \boldsymbol{v}'(t)$
6. $\frac{d}{dt}[\boldsymbol{u}(f(t))] = f'(t)\boldsymbol{u}'(f(t))$

where the last property in this list represents the Chain Rule.

We can prove Formula 4. The rest are left to you as exercises.

Proof. Suppose we have two vector-valued functions

$$\mathbf{u}(t) = \langle f_1(t), f_2(t), f_3(t)\rangle \qquad \mathbf{v} = \langle g_1(t), g_2(t), g_3(t)\rangle$$

Then the dot product is

$$\mathbf{u}(t) \cdot \mathbf{v}(t) = f_1(t)g_1(t) + f_2(t)g_2(t) + f_3(t)g_3(t) = \sum_{i=1}^{3} f_i(t)g_i(t)$$

so applying Product Rule for real-valued functions, we have

$$\begin{aligned}
\frac{d}{dt}\big[\mathbf{u}(t) \cdot \mathbf{v}(t)\big] &= \frac{d}{dt}\sum_{i=1}^{3} f_i(t)g_i(t) = \sum_{i=1}^{3} \frac{d}{dt}\big[f_i(t)g_i(t)\big] \\
&= \sum_{i=1}^{3}\big[f_i'(t)g_i(t) + f_i(t)g_i'(t)\big] \\
&= \sum_{i=1}^{3} f_i'(t)g_i(t) + \sum_{i=1}^{3} f_i(t)g_i'(t) \\
&= \mathbf{u}'(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \mathbf{v}'(t)
\end{aligned}$$

and the proof is complete. □

Formula 4 can also be used to prove the following theorem.

Theorem 2.2.3. *If $|\boldsymbol{r}(t)| = c$, where c is some constant, then $\boldsymbol{r}'(t)$ is orthogonal to $\boldsymbol{r}(t)$ for all values of t.*

Proof. We know from section 1.3 that

$$\mathbf{r}(t) \cdot \mathbf{r}(t) = |\mathbf{r}(t)|^2 = c^2$$

and if c is constant, then c^2 must also be one. Formula 4 of Theorem 2.2.2 gives us

$$0 = \frac{d}{dt}[\mathbf{r}(t) \cdot \mathbf{r}(t)] = \mathbf{r}'(t) \cdot \mathbf{r}(t) + \mathbf{r}(t) \cdot \mathbf{r}'(t) = 2(\mathbf{r}(t) \cdot \mathbf{r}'(t))$$

Therefore $\mathbf{r}(t) \cdot \mathbf{r}'(t) = 0$, so the vectors $\mathbf{r}(t)$ and $\mathbf{r}'(t)$ are proven to be orthogonal. □

Integrals We can explore definite integrals of vector functions $\mathbf{r}(t)$ in pretty much the same manner as for real-valued functions except that the integrand is a vector. We can express the integral of $\mathbf{r}$ in terms of the integrals of its component functions f, g, and h if $\mathbf{r}$ represents a vector in V_3. Keep in mind that we will be using the same Riemann notation that we learned in Calculus I.

$$\begin{aligned}\int_a^b r(t)\,dt &= \lim_{n\to\infty} \sum_{i=1}^{n} r(t_i^*)\Delta t \\ &= \lim_{n\to\infty} \left[\left(\sum_{i=1}^{n} f(t_i^*)\Delta t \right) \mathbf{i} + \left(\sum_{i=1}^{n} g(t_i^*)\Delta t \right) \mathbf{j} + \left(\sum_{i=1}^{n} h(t_i^*)\Delta t \right) \mathbf{k} \right]\end{aligned}$$

so ultimately, the integral can be expressed as

$$\boxed{\int_a^b \mathbf{r}(t)\,dt = \left(\int_a^b f(t)\,dt \right) \mathbf{i} + \left(\int_a^b g(t)\,dt \right) \mathbf{j} + \left(\int_a^b h(t)\,dt \right) \mathbf{k}}$$

Since integration of real-valued functions and vector-valued functions function essentially the same way, we can even extend the Fundamental Theorem of Calculus to continuous vector functions in a parallel process:

$$\int_a^b \mathbf{r}(t)\,dt = \mathbf{R}(t)]_a^b = \mathbf{R}(b) - \mathbf{R}(a)$$

where $\mathbf{R}$ is the antiderivative of $\mathbf{r}$, i.e. $\mathbf{R}'(t) = \mathbf{r}(t)$.

Problem 2.2.4. *Answer the following questions. Let $\boldsymbol{r}(t) = \sin t\,\boldsymbol{i} + 6\,\boldsymbol{j} + 4t\,\boldsymbol{k}$.*
(a) Evaluate $\int \boldsymbol{r}(t)\,dt$.
(b) Evaluate $\int_0^1 \boldsymbol{r}(t)\,dt$.

Solution to part a: All we really need to do is integrate each component function of $\mathbf{r}(t)$ and we are good to go.

$$\begin{aligned}\int \mathbf{r}(t)\,dt &= \left(\int \sin t\,dt\right)\mathbf{i} + \left(\int 6\,dt\right)\mathbf{j} + \left(\int 4t\,dt\right)\mathbf{k}\\ &= \boxed{-\cos t\,\mathbf{i} + 6t\,\mathbf{j} + 2t^2\,\mathbf{k} + \mathbf{C}}\end{aligned}$$

where $\mathbf{C}$ is a "constant" of integration, but it is a vector, not a real number.

Solution to part b: This is a definite integral so we just need to run the antiderivative we found in (a) (dropping the constant vector $\mathbf{C}$, of course) from $t = 0$ to $t = 1$.

$$\begin{aligned}\int_0^1 \mathbf{r}(t)\,dt &= [-\cos t\,\mathbf{i} + 6t\,\mathbf{k} + 2t^2\,\mathbf{k}]_0^1\\ &= [-\cos t]_0^1\,\mathbf{i} + [6t]_0^1\,\mathbf{j} + [2t^2]_0^1\,\mathbf{k}\\ &= \boxed{(1-\cos 1)\,\mathbf{i} + 6\,\mathbf{j} + 2\,\mathbf{k}}\end{aligned}$$

2.3 Arc Length and Curvature

In Calculus II we described a parametric curve with the equations $x = f(t)$ and $y = g(t)$ with $a \le t \le b$ as the limits of lengths approximating polygons. For the case where f and g have continuous first partial derivatives, we have

$$L = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2}\,dt = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt$$

When dealing with three dimensions, however, we define the length of space curve. Let the curve have vector equation $\mathbf{r}(t) = \langle f(t), g(t), h(t)\rangle$ for $a \le t \le b$, or equivalently in parametric form, $x = f(t)$, $y = g(t)$, and $z = h(t)$, given that f', g', and h' are all continuous. As t increases from a to b if the curve is traversed one time, its length is given by

$$\begin{aligned}L &= \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2 + [h'(t)]^2}\,dt\\ &= \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\,dt\end{aligned}$$

Either of these formulas can be put into an even more compact form, i.e.

$$\boxed{L = \int_a^b |\mathbf{r}'(t)|\,dt}$$

where for plane and space curves, respectively, we have $\mathbf{r}(t) = \langle f(t), g(t)\rangle$ and $\mathbf{r}(t) = \langle f(t), g(t), h(t)\rangle$.

Problem 2.3.1. *Find the length of the arc of the circular helix with vector equation* $\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$ *from the points* $(1, 0, 0)$ *to* $(1, 0, 2\pi)$.

Solution: We find the tangent vector

$$\mathbf{r}'(t) = -\sin t\,\mathbf{i} + \cos t\,\mathbf{j} + \mathbf{k}$$

with magnitude

$$|\mathbf{r}'(t)| = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2}$$

Therefore, the arc length of the helix on the interval $0 \leq t \leq 2\pi$ is simply

$$L = \int_0^{2\pi} |\mathbf{r}'(t)|\,dt = \int_0^{2\pi} \sqrt{2}\,dt = \boxed{2\sqrt{2}\,\pi}$$

Keep in mind that there are several possible ways to parameterize a curve C. For example, $\mathbf{r}_1(t) = \langle t, t^2, t^3\rangle$ for $1 \leq t \leq 2$ is functionally equivalent to $\mathbf{r}_2(u) = \langle e^u, e^{2u}, e^{3u}\rangle$ as $t = e^u$. These two expressions are called **parameterizations** of C. In other words, if we were to compute the arc length using either parameterization, we would arrive at the same answer. This is because arc length is a *geometric property* of the curve and is thus independent of parameterization.

The Arc Length Function Let C be a curve with vector function

$$\mathbf{r}(t) = \langle f(t), g(t), h(t)\rangle \qquad a \leq t \leq b$$

where $\mathbf{r}'$ is continuous and C is traversed exactly once as t increases from a to b. The **arc length function** is defined by

$$s(t) = \int_a^t |\mathbf{r}'(u)|\,du = \int_a^t \sqrt{\left(\frac{dx}{du}\right)^2 + \left(\frac{dy}{du}\right)^2 + \left(\frac{dz}{du}\right)^2}\,du$$

If we differentiate both sides of the equation with respect to t, using Part 1 of the Fundamental Theorem of Calculus, we obtain

$$\frac{ds}{dt} = |\mathbf{r}'(t)|$$

It is really useful to parameterize a curve *with respect to arc length* because as stated, arc length is a property that is a part of the curve and does not depend on parameterization. If $\mathbf{r}(t)$ describes a curve that is already expressed in terms of t and $s(t)$ is the arc length of that curve, we can solve for the parameter as a function of arc length, i.e. $t = t(s)$. Therefore, the parameterization of $\mathbf{r}(t)$ with respect to arc length becomes $\mathbf{r}(t) = \mathbf{r}(t(s))$. For example, if $s = 1$, then $\mathbf{r}(t(1))$ represents the position vector of the point 1 unit of length along the curve from an arbitrary initial point.

Problem 2.3.2. *Re-parameterize the circular helix* $\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$ *with respect to arc length measured from the point* $(1, 0, 0)$ *to* $(1, 0, 2\pi)$.

Solution: The initial point, $(1, 0, 0)$, corresponds to a parameter value of $t = 0$. It is trivial to show that

$$\frac{ds}{dt} = |\mathbf{r}'(t)| = \sqrt{2}$$

and so

$$s = s(t) = \int_0^t |\mathbf{r}'(u)|\,du = \int_0^t \sqrt{2}\,du = \sqrt{2}\,t$$

Therefore, we have $t = s/\sqrt{2}$ so we can find the requested parameterization by making this substitution to the original expression for $\mathbf{r}(t)$.

$$\mathbf{r}(t(s)) = \boxed{\cos(s/\sqrt{2})\,\mathbf{i} + \sin(s/\sqrt{2})\,\mathbf{j} + (s/\sqrt{2})\,\mathbf{k}}$$

Curvature We call a parameterization $\mathbf{r}(t)$ smooth if and only if $\mathbf{r}'$ is continuous and $\mathbf{r}'(t) \neq \mathbf{0}$ on any interval I. The curve C is called smooth if and only if its parameterization is smooth. Such curves have no sharp corners or cusps; when the tangent vector turns, the motion is perfectly continuous.

If C is a smooth function with vector function $\mathbf{r}$, then the **unit tangent vector** is defined by

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

and this vector indicates the overall direction of the curve. We will see that as when C is fairly straight, then $\mathbf{T}(t)$ changes direction very slowly but when the curve bends or twists sharply, $\mathbf{T}(t)$ changes direction rapidly. The *curvature* of C at a given point measures the speed at which the curve changes direction at that point. It is defined to be the magnitude of the unit tangent vector with respect to arc length, i.e. since the unit tangent vector has constant magnitude, only changes in direction contribute to the rate of change of $\mathbf{T}$.

Definition 2.3.1. *The* ***curvature*** *of a curve C is given by*

$$\kappa = \left|\frac{d\mathbf{T}}{ds}\right|$$

where $\mathbf{T}$ is the unit tangent vector.

In most cases, however, the curvature is easier to compute when we are working in terms of the parameter t rather than arc length s, so using the Chain Rule, we get

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds}\frac{ds}{dt} \Longrightarrow \kappa = \left|\frac{d\mathbf{T}}{ds}\right| = \left|\frac{d\mathbf{T}/dt}{ds/dt}\right|$$

However, $ds/dt = |\mathbf{r}'(t)|$, so we can express the curvature in an alternate form:

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

Problem 2.3.3. *Determine, in terms of a, the curvature of a circle with radius a.*

Solution: The circle is a curve with parameterization

$$\mathbf{r}(t) = a\cos t\,\mathbf{i} + a\sin t\,\mathbf{j}$$

and the tangent vector $\mathbf{r}'(t) = -a\sin t\,\mathbf{i} + a\cos t\,\mathbf{j}$ has magnitude $|\mathbf{r}'(t)| = a$. The unit tangent vector thus becomes

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} = -\sin t\,\mathbf{i} + \cos t\,\mathbf{j}$$

and the derivative $\mathbf{T}'(t)$ is equal to $-\cos t\,\mathbf{i} - \sin t\,\mathbf{j}$. Thus $|\mathbf{T}'(t)| = 1$ and the curvature of the circle is

$$\kappa(t) = \frac{|\mathbf{T}'(t)|}{|\mathbf{r}'(t)|} = \frac{1}{a}$$

and so the premise of the problem statement is verified. □

Remark. The above result indicates that smaller circles have larger curvatures and vice versa, i.e. the curvature is inversely proportional to the radius of the circle. We can also infer that the curvature of a straight line is 0 because the tangent vector is a fixed constant.

Below shows a more convenient approach to determining the curvature of a curve.

Theorem 2.3.1. *The curvature of a curve with vector function* $\mathbf{r}$ *can be expressed by*

$$\kappa(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3}$$

Proof. Since $\mathbf{T} = \mathbf{r}'/|\mathbf{r}'|$ and $|\mathbf{r}'| = ds/dt$, we can use algebraic manipulation to write the tangent vector as

$$\mathbf{r}' = |\mathbf{r}'||\mathbf{T}| = \frac{ds}{dt}\mathbf{T}$$

The second derivative $\mathbf{r}''$ can be determined by the Product Rule.

$$\mathbf{r}'' = \frac{d^2s}{dt^2}\mathbf{T} + \frac{ds}{dt}\mathbf{T}'$$

Since $\mathbf{T} \times \mathbf{T} = \mathbf{0}$, the relevant cross product is

$$\mathbf{r}' \times \mathbf{r}'' = \left(\frac{ds}{dt}\right)^2 (\mathbf{T} \times \mathbf{T}')$$

Notice that $|\mathbf{T}(t)| = 1$ (a constant) for all $t \in I$, where the curve is defined on I, so the vectors $\mathbf{T}$ and $\mathbf{T}'$ are orthogonal. So using properties of the cross product, we have

$$|\mathbf{r}' \times \mathbf{r}''| \left(\frac{ds}{dt}\right)^2 |\mathbf{T} \times \mathbf{T}'| = \left(\frac{ds}{dt}\right)^2 |\mathbf{T}|\,|\mathbf{T}'| = \left(\frac{ds}{dt}\right)^2 |\mathbf{T}'|$$

$$|\mathbf{T}'| = \frac{|\mathbf{r}' \times \mathbf{r}''|}{\left(\dfrac{ds}{dt}\right)^2} = \frac{|\mathbf{r}' \times \mathbf{r}''|}{|\mathbf{r}'|^2}$$

$$\therefore \kappa = \frac{|\mathbf{T}'|}{|\mathbf{r}'|} = \frac{|\mathbf{r}' \times \mathbf{r}''|}{|\mathbf{r}'|^3}$$

so the proof is complete. □

Problem 2.3.4. *Find the curvature of the twisted cubic* $\mathbf{r}(t) = \langle t, t^2, t^3 \rangle$ *at* $(1, 1, 1)$.

Solution: The required ingredients for this problem are the first and second derivatives of $\mathbf{r}$ and the magnitude of the tangent vector $\mathbf{r}'(t)$.

$$\mathbf{r}'(t) = \langle 1, 2t, 3t^2 \rangle$$
$$\mathbf{r}''(t) = \langle 0, 2, 6t \rangle$$

$$|\mathbf{r}'(t)| = \sqrt{1 + 4t^2 + 9t^4}$$

$$\mathbf{r}'(t) \times \mathbf{r}''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 2t & 3t^2 \\ 0 & 2 & 6t \end{vmatrix} = 6t^2\mathbf{i} - 6t\mathbf{j} + 2\mathbf{k}$$

$$|\mathbf{r}'(t) \times \mathbf{r}''(t)| = \sqrt{36t^4 + 36t^2 + 4} = 2\sqrt{9t^4 + 9t^2 + 1}$$

At the point $(1, 1, 1)$, the corresponding parameter value is $t = 1$, so the curvature is $\kappa(1) = \frac{2\sqrt{19}}{14\sqrt{14}} = \frac{\sqrt{19}}{7\sqrt{14}} = \boxed{\frac{\sqrt{266}}{98}}$.

Special Case for Curvature: Plane Curves Consider a curve in two dimensions with equation $y = f(x)$. The parameter is x and the corresponding vector function is $\mathbf{r}(x) = x\,\mathbf{i} + f(x)\,\mathbf{j}$. Then $\mathbf{r}'(x) = \mathbf{i} + f'(x)\,\mathbf{j}$ and $\mathbf{r}''(x) = \mathbf{i} + f''(x)\,\mathbf{j}$. Since $\mathbf{i} \times \mathbf{j} = \mathbf{k}$ and $\mathbf{j} \times \mathbf{j} = \mathbf{0}$, it follows that $\mathbf{r}'(x) \times \mathbf{r}''(x) = f''(x)\,\mathbf{k}$. Also $|\mathbf{r}'(x)| = \sqrt{1 + [f'(x)]^2}$, so according to Theorem 2.3.1, we have

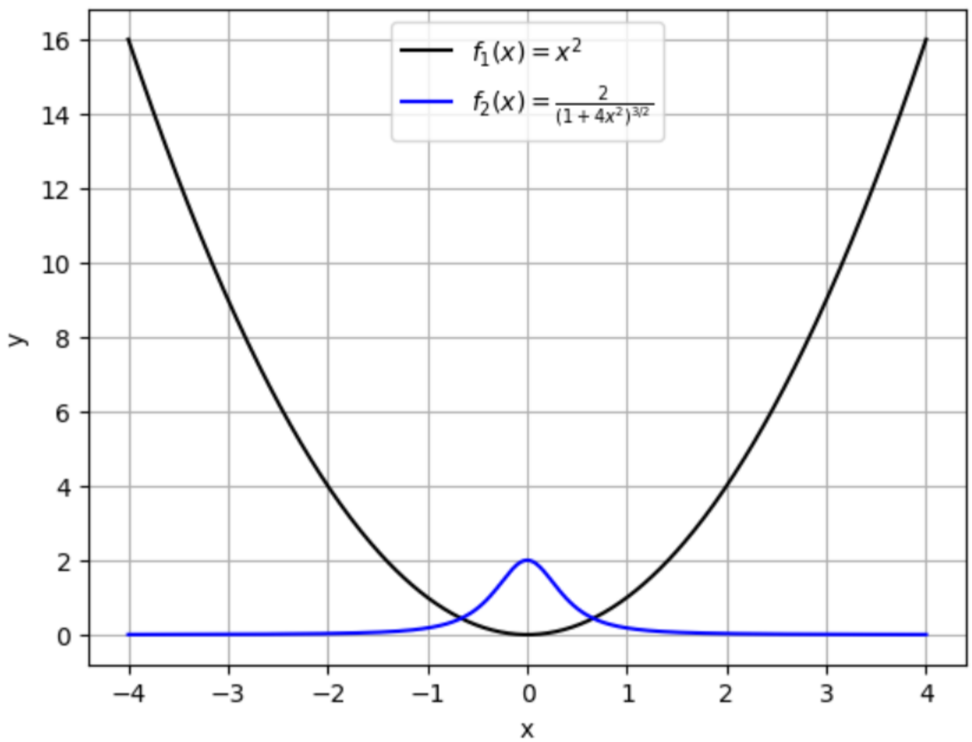
$$\boxed{\kappa(x) = \frac{|f''(x)|}{[1 + (f'(x))^2]^{3/2}}}$$

Problem 2.3.5. *Find the curvature of the parabola $y = x^2$ at the origin and describe the end behavior of the graph of $y = \kappa(x)$ in the xy-plane.*

Solution: We apply the special case formula for determining the curvature of plane curves. We have

$$\kappa(x) = \frac{|y''|}{[1 + (y')^2]^{3/2}} = \frac{2}{(1 + 4x^2)^{3/2}}$$

At the origin, $(0, 0)$, the parameter value is $x = 0$ with curvature $\kappa(0) = \boxed{2}$. Observe from the expression for $\kappa(x)$ that it tends to 0 as x tends to positive or negative infinity. You can observe this behavior in the plot below, which was generated with Python code.



The Normal and Binormal Vectors On a smooth space curve with vector function $\mathbf{r}(t)$, there are numerous vectors orthogonal to the unit tangent vector $\mathbf{T}(t)$. According to Theorem 2.2.3, because $|\mathbf{T}'(t)| = 1$ for all $t \in I$, we have $\mathbf{T}(t) \cdot \mathbf{T}'(t) = 0$, so $\mathbf{T}(t)$ and $\mathbf{T}'(t)$ are orthogonal. It is important to note that although $\mathbf{T}(t)$ is a unit vector, that might not always be the case for $\mathbf{T}'(t)$. However, at any point where $\kappa \neq 0$ we can define the **unit normal vector** as

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

This vector indicates the direction in which the curve turns at each point.

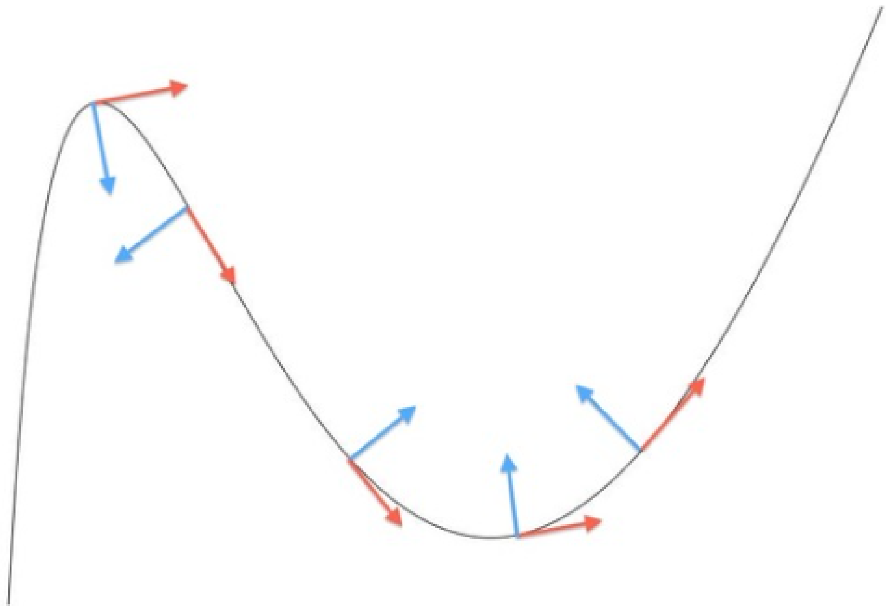


Image Credit: Mathematics LibreTexts

In the above image shows a portion of a curve. The red arrows represent the unit tangent vectors $\mathbf{T}$ and the blue arrows represent the unit normal vectors $\mathbf{N}$. The vector

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

is referred to as the **binormal vector**. As a cross product, it is orthogonal to both $\mathbf{T}(t)$ and $\mathbf{N}(t)$, and is also a unit vector.

Problem 2.3.6. *Find the unit normal and binormal vectors for the circular helix* $\mathbf{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$.

Solution: We first compute the components required for the unit normal vector.

$$\mathbf{r}'(t) = -\sin t\,\mathbf{i} + \cos t\,\mathbf{j} + \mathbf{k} \Longrightarrow |\mathbf{r}'(t)| = \sqrt{2}$$

So the unit tangent vector is

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|} = \frac{1}{\sqrt{2}}(-\sin t\,\mathbf{i} + \cos t\,\mathbf{j} + \mathbf{k}$$

Proceeding further, we obtain

$$\mathbf{T}'(t) = \frac{1}{\sqrt{2}}(-\cos t\,\mathbf{i} - \sin t\,\mathbf{j}) \Longrightarrow |\mathbf{T}'(t)| = \frac{1}{\sqrt{2}}$$

So the unit normal vector is

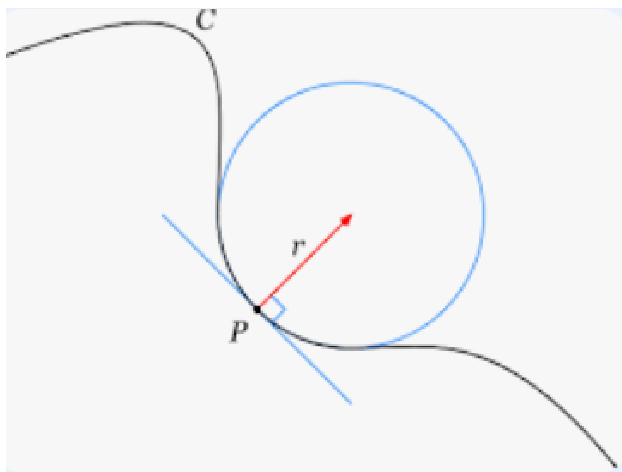
$$\mathbf{N}(t) = \frac{\mathbf{T}'(t)}{|\mathbf{T}'(t)|} = -\cos t\,\mathbf{i} - \sin t\,\mathbf{j} = \boxed{\langle -\cos t, -\sin t, 0\rangle}$$

This indicates that the unit normal at any point on the helix is horizontal and points toward the z-axis. We now compute the binormal vector:

$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t) = \frac{1}{\sqrt{2}} \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin t & \cos t & 1 \\ -\cos t & -\sin t & 0 \end{vmatrix} = \boxed{\frac{1}{\sqrt{2}} \langle \sin t, -\cos t, 1 \rangle}$$

The plane determined by vectors $\mathbf{N}$ and $\mathbf{B}$ at a point P on the curve C is called the *normal plane*. It consists of all lines that are orthogonal to the tangent vector $\mathbf{T}$. Meanwhile, the plane determined by vectors $\mathbf{T}$ and $\mathbf{N}$ is called the *osculating plane*. It is the plane that comes closest to maintaining the part of the curve near P.

The **osculating circle** of C at P is the circle in the osculating plane that passes through the point P with radius $1/\kappa$ and centered a distance of $1/\kappa$ from P along the unit normal $\mathbf{N}$. The center of this circle is called the *center of curvature* at P.

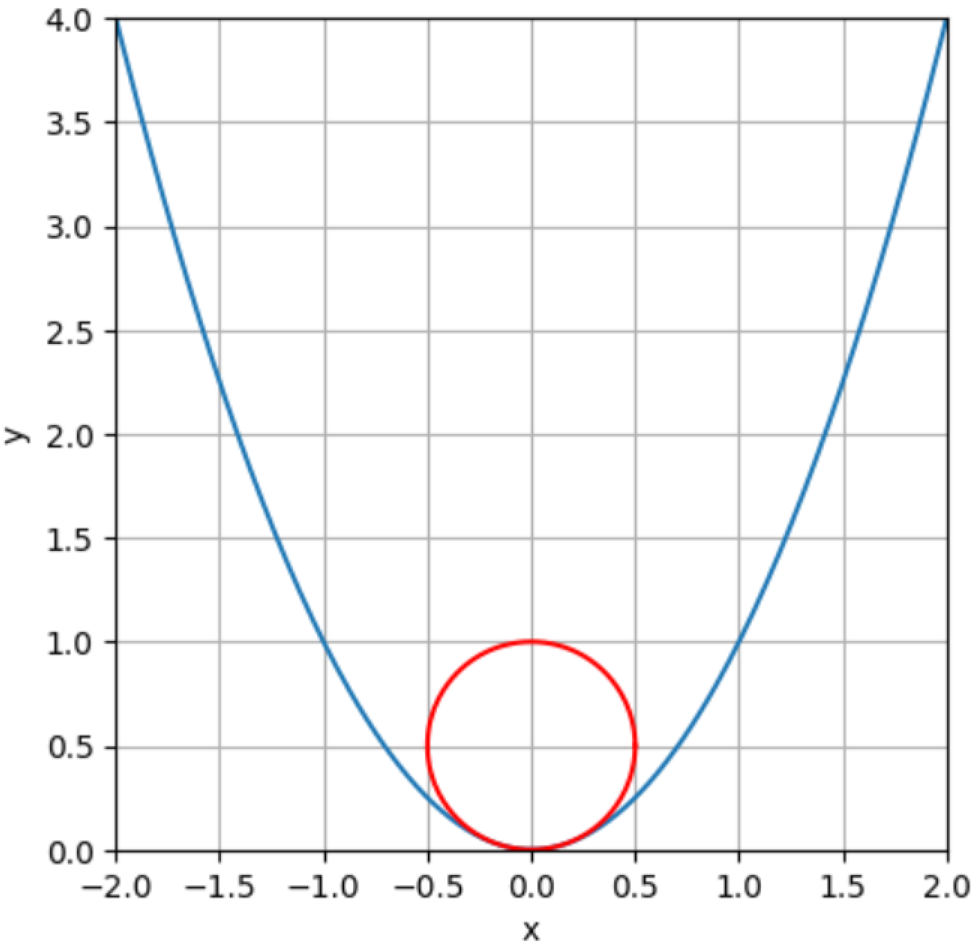


Problem 2.3.7. *Find and graph the osculating circle of the parabola $y = x^2$ at the origin.*

Solution: We know from problem 2.3.5 that the curvature of this plane curve at the origin is $\kappa(0) = 2$ so the radius of the osculating circle at this point is $\frac{1}{2}$. Also, the tangent vector is horizontal at the origin so the normal vector is vertical. Moving a distance of $\frac{1}{2}$ along the normal vector $\mathbf{N} = \langle 0, 1 \rangle$, we obtain the center of curvature as $\left(0, \frac{1}{2}\right)$ so the equation of the osculating circle is

$$\boxed{x^2 + \left(y - \frac{1}{2}\right)^2 = \frac{1}{4}}$$

The circle is graphed (using Python code) below.



Torsion We now know that the curvature $\kappa = |d\mathbf{T}/ds|$ at a point P on some curve C describes how tightly the curve can "bend." Since $\mathbf{T}$ is perpendicular to the normal plane, the quantity $d\mathbf{T}/ds$ indicates how the normal plane changes as P moves along C. We also know that $\mathbf{B}$ is perpendicular to the osculating plane, so the quantity $d\mathbf{B}/ds$ indicates how the osculating plane changes as P moves along C.

It is known that $d\mathbf{B}/ds$ is parallel to unit normal vector $\mathbf{N}$. In other words, there exists some scalar quantity τ such that

$$\frac{d\mathbf{B}}{ds} = -\tau\mathbf{N}$$

The scalar τ is called the *torsion* of C at P. Let's take the dot product with $\mathbf{N}$ of each side of the equation, using the fact that $\mathbf{N} \cdot \mathbf{N} = 1$, we obtain the following definition.

Definition 2.3.2. *The* ***torsion*** *of a curve with respect unit normal and binormal vectors* $\mathbf{N}$ *and* $\mathbf{B}$ *is*

$$\tau = -\frac{d\mathbf{B}}{ds} \cdot \mathbf{N}$$

As with curvature, it is easier to express the result in terms of the parameter t rather than s, so we apply the Chain Rule.

$$\frac{d\mathbf{B}}{dt} = \frac{d\mathbf{B}}{ds}\frac{ds}{dt} \Longrightarrow \frac{d\mathbf{B}}{ds} = \frac{d\mathbf{B}/dt}{ds/dt} = \frac{\mathbf{B}'(t)}{|\mathbf{r}'(t)|}$$

Applying Definition 2.3.2, the torsion of a curve in terms of the parameter t is

$$\boxed{\tau(t) = -\frac{\mathbf{B}'(t)\cdot\mathbf{N}(t)}{|\mathbf{r}'(t)|}}$$

Problem 2.3.8. *Find the torsion of the circular helix* $\mathbf{r}(t) = \langle \cos t, \sin t, t\rangle$.

Solution: In a previous exercise, we determined $ds/dt = |\mathbf{r}'(t)| = \sqrt{2}$, $\mathbf{N}(t) = \langle -\cos t, -\sin t, 0\rangle$, and $\mathbf{B}(t) = (1/\sqrt{2})\langle \sin t, -\cos t, 1\rangle$. Then $\mathbf{B}'(t) = (1/\sqrt{2})\langle \cos t, \sin t, 0\rangle$ so the torsion of this curve is

$$\tau(t) = -\frac{\mathbf{B}'(t)\cdot\mathbf{N}(t)}{|\mathbf{r}'(t)|} = -\frac{1}{2}\langle \cos t, \sin t, 0\rangle \cdot \langle -\cos t, -\sin t, 0\rangle = \boxed{\frac{1}{2}}$$

There is another formula that is often more convenient to compute torsion, only in terms of the vector function $\mathbf{r}(t)$ and its derivatives.

$$\tau(t) = \frac{[\mathbf{r}'(t)\times\mathbf{r}''(t)]\cdot\mathbf{r}'''(t)}{|\mathbf{r}'(t)\times\mathbf{r}''(t)|^2}$$

Problem 2.3.9. *Calculate the torsion of the curve* $\mathbf{r}(t) = \langle e^t, e^{-t}, t\rangle$ *at the point* $(1,1,0)$.

Solution: We have $\mathbf{r}'(t) = \langle e^t, -e^{-t}, 1\rangle$, $\mathbf{r}''(t) = \langle e^t, e^{-t}, 0\rangle$, and $\mathbf{r}'''(t) = \langle e^t, -e^{-t}, 0\rangle$. The requested cross product is

$$\mathbf{r}'(t)\times\mathbf{r}''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ e^t & -e^{-t} & 1 \\ e^t & e^{-t} & 0 \end{vmatrix} = -e^{-t}\,\mathbf{i} + e^t\,\mathbf{j} + 2\,\mathbf{k}$$

The magnitude squared is

$$\left|\mathbf{r}'(t)\times\mathbf{r}''(t)\right|^2 = e^{-2t} + e^{2t} + 4$$

Then, the torsion of the curve for any value $t \in I$ is

$$\begin{aligned}\tau(t) &= \frac{(\mathbf{r}'(t)\times\mathbf{r}''(t))\cdot\mathbf{r}'''(t)}{|\mathbf{r}'(t)\times\mathbf{r}''(t)|^2} \\ &= \frac{\langle -e^{-t}, e^t, 2\rangle\cdot\langle e^t, -e^{-t}, 0\rangle}{e^{-2t}+e^{2t}+4} \\ &= \frac{-1-1+0}{e^{-2t}+e^{2t}+4} = -\frac{2}{e^{-2t}+e^{2t}+4}\end{aligned}$$

The point $(1,1,0)$ corresponds to a parameter value of $t = 0$, so the torsion is

$$\tau(0) = -\frac{2}{1+1+4} = \boxed{-\frac{1}{3}}$$

2.4 Motion in Space

Earlier, we talked about tangent and normal vectors and the concept of curvature. These metrics can be used in physics to study the motion of an object−primarily governed by velocity and acceleration−along a space curve (three dimensions). In this chapter, we will follow in Newton's footsteps by using the aforementioned methods to derive Kepler's First Law of planetary motion.

Velocity, Speed, and Acceleration Suppose a particle is moving through space and its position as a function of time t is described by the vector $\mathbf{r}(t)$. Notice that for small values of h, i.e. $h \to 0$ the vector

$$\frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}$$

approximates the direction of the particle's motion along the curve $\mathbf{r}(t)$. The magnitude, $|\mathbf{r}(t)|$, measures the size of the displacement vector per unit time. The above vector represents the average velocity over a time interval of length $\Delta t = (t+h) - t = h$ and so the limit as $h \to 0$ is the **velocity vector**, $\mathbf{v}(t)$, at time t:

$$\boxed{\mathbf{v}(t) = \lim_{h \to 0} \frac{\mathbf{r}(t+h) - \mathbf{r}(t)}{h}}$$

Based on the definition of the derivative, the velocity vector is also the tangent vector and points in the direction of the tangent line, so $\mathbf{v}(t) = \mathbf{r}'(t)$. The **speed** of the particle at a given time t is given by the magnitude of the velocity vector, or $|\mathbf{v}(t)| = |\mathbf{r}'(t)|$. If we relate this to arc length from section 2.3, we have

$$|\mathbf{v}'(t)| = |\mathbf{r}'(t)| = \frac{ds}{dt} = \text{rate of change of distance with respect to time}$$

Analogous with one-dimensional motion, the **acceleration** of the particle is defined as the derivative of velocity (or the *second* derivative of position):

$$\mathbf{a}(t) = \mathbf{v}'(t) = \mathbf{r}''(t)$$

Problem 2.4.1. *An object moves in a plane with position vector* $\boldsymbol{r}(t) = t^2\,\boldsymbol{i} + t^3\,\boldsymbol{j}$. *Find its velocity, speed, and acceleration when* $t = 1$.

Solution: We need the first and second derivatives of $\mathbf{r}(t)$ in order to find $\mathbf{v}(t)$ and $\mathbf{a}(t)$, respectively.

$$\mathbf{v}(t) = 2t\,\mathbf{i} + 3t^2\,\mathbf{j} \qquad \mathbf{a}(t) = 2\,\mathbf{i} + 6t\,\mathbf{j}$$

and now plugging in $t = 1$, we get the velocity as

$$\mathbf{v}(1) = 2(1)\,\mathbf{i} + 3(1)^2\,\mathbf{j} = \boxed{2\,\mathbf{i} + 3\,\mathbf{j}}$$

and the acceleration as

$$\mathbf{a}(1) = 2\,\mathbf{i} + 6(1)\,\mathbf{j} = \boxed{2\,\mathbf{i} + 6\,\mathbf{j}}$$

We know that speed is the magnitude of velocity, so at $t = 1$ we have

$$|\mathbf{v}(1)| = \sqrt{2^2 + 3^2} = \boxed{\sqrt{13}}$$

Recall the integration of vector functions we saw in section 2.2. These can be applied to particle motion, when finding position vectors when velocity or acceleration vectors are known, as in the next example.

Problem 2.4.2. *A moving particle starts at an initial position vector $\boldsymbol{r}(0) = \langle 1, 0, 0\rangle$ with initial velocity $\boldsymbol{v}(0) = \boldsymbol{i} - \boldsymbol{j} + \boldsymbol{k}$. The acceleration is given by the vector function $\boldsymbol{a}(t) = 4t\,\boldsymbol{i} + 6t\,\boldsymbol{j} + \boldsymbol{k}$. Find its velocity and position as functions of t.*

Solution: Since $\mathbf{a}(t) = \mathbf{v}'(t)$ we can take the integral of the acceleration vector with respect to t.

$$\begin{aligned}\mathbf{v}(t) &= \int \mathbf{a}(t)\,dt = \int (4t\,\mathbf{i} + 6t\,\mathbf{j} + \mathbf{k})\,dt \\ &= 2t^2\,\mathbf{i} + 3t^2\,\mathbf{j} + t\,\mathbf{k} + \mathbf{C}\end{aligned}$$

In order to determine the value of the constant vector $\mathbf{C}$, we use the fact that $\mathbf{v}(0) = \mathbf{i} - \mathbf{j} + \mathbf{k}$. Using the initial condition, we can find that $\mathbf{v}(0) = \mathbf{C}$ so $\mathbf{C} = \mathbf{i} - \mathbf{j} + \mathbf{k}$ and

$$\begin{aligned}\mathbf{v}(t) &= 2t^2\,\mathbf{i} + 3t^2\,\mathbf{j} + t\,\mathbf{k} + \mathbf{i} - \mathbf{j} + \mathbf{k} \\ &= \boxed{(2t^2+1)\,\mathbf{i} + (3t^2-1)\,\mathbf{j} + (t+1)\,\mathbf{k}}\end{aligned}$$

Since $\mathbf{v}(t) = \mathbf{r}'(t)$, we have

$$\begin{aligned}\mathbf{r}(t) &= \int \mathbf{v}(t)\,dt \\ &= \int [(2t^2+1)\,\mathbf{i} + (3t^2-1)\,\mathbf{j} + (t+1)\,\mathbf{k}]\,dt \\ &= \left(\frac{2}{3}t^3 + t\right)\mathbf{i} + \left(t^3 - t\right)\mathbf{j} + \left(\frac{1}{2}t^2 + t\right)\mathbf{k} + \mathbf{D}\end{aligned}$$

Setting $t = 0$, we find that $\mathbf{D} = \mathbf{r}(0) = \mathbf{i}$ so the position as a function of time is

$$\boxed{\mathbf{r}(t) = \left(\frac{2}{3}t^3 + t + 1\right)\mathbf{i} + (t^3 - t)\,\mathbf{j} + \left(\frac{1}{2}t^2 + t\right)\mathbf{k}}$$

In general, we can apply integration of vector-valued functions to recover velocity when acceleration is known and position when velocity is known, so

$$\mathbf{v}(t) = \mathbf{v}(t_0) + \int_{t_0}^{t} \mathbf{a}(u)\,du \qquad \mathbf{r}(t) = \mathbf{r}(t_0) + \int_{t_0}^{t} \mathbf{v}(u)\,du$$

In a real-world context, if there exists some force acting on an object, then its acceleration can be found from *Newton's Second Law of Motion.* If at any time t, a force $\mathbf{F}(t)$ acts on an object of mass m producing an acceleration $\mathbf{a}(t)$, then we have

$$\mathbf{F}(t) = m\mathbf{a}(t)$$

Problem 2.4.3. *A particle of mass m moves in a circular path with constant angular speed ω and position vector $\boldsymbol{r}(t) = a\cos\omega t\,\boldsymbol{i} + a\sin\omega t\,\boldsymbol{j}$. Find the force acting on the object.*

Solution: To find the force, we need to know the acceleration of the particle. The first step is to find the velocity:

$$\mathbf{v}(t) = \mathbf{r}'(t) = -a\omega\sin\omega t\,\mathbf{i} + a\omega\sin\omega t\,\mathbf{j}$$

We take the derivative once again to obtain acceleration:

$$\mathbf{a}(t) = \mathbf{v}'(t) = -a\omega^2\cos\omega t\,\mathbf{i} - a\omega^2\sin\omega t\,\mathbf{j}$$

Thus according to Newton's Second Law of Motion, the force on the particle is

$$\mathbf{F}(t) = m\mathbf{a}(t) = \boxed{-m\omega^2(a\cos\omega t\,\mathbf{i} + a\sin\omega t\,\mathbf{j})}$$

Remark. Notice that $\mathbf{F}(t) = -m\omega^2\mathbf{r}(t)$. This equation demonstrates that the force points in the direction opposite to the radial/position vector $\mathbf{r}(t)$. In physics, such a force is called a *centripetal*, or *center-seeking* force.

Projectile Motion

Problem 2.4.4. *Suppose a projectile is fired with an angle of elevation α and initial velocity $\mathbf{v}_0$. Assume air resistance is negligible and that all external forces are merely due to gravity, find the position function $\mathbf{r}(t)$ for the projectile. What value α maximizes the horizontal distance traveled, or the range of the projectile?*

Solution: We can set up coordinate axes such that the initial position of the projectile is the origin. Since the force due to gravity points straight down, we have

$$\mathbf{F} = m\mathbf{a} = -mg\,\mathbf{j}$$

where $|\mathbf{g}| = 9.8\,\mathrm{m/s^2}$. Thus

$$\mathbf{a} = -g\,\mathbf{j}$$

Since $\mathbf{a} = \mathbf{v}'(t)$, we have $\mathbf{v}(t) = -gt\,\mathbf{j} + \mathbf{v}_0$. To obtain position, we integrate again:

$$\mathbf{r}(t) = -\frac{1}{2}gt^2\,\mathbf{j} + t\mathbf{v}_0 + \mathbf{D}$$

However, $\mathbf{D} = \mathbf{r}(0) = \mathbf{0}$, so the position vector is given by

$$\mathbf{r}(t) = -\frac{1}{2}gt^2\,\mathbf{j} + t\mathbf{v}_0$$

If we let $|\mathbf{v}_0| = v_0$, as the initial speed of the projectile, then

$$\mathbf{v}_0 = v_0\cos\alpha,\mathbf{i} + v_0\sin\alpha\,\mathbf{j}$$

so the position vector becomes

$$\mathbf{r}(t) = (v_0\cos\alpha)t\,\mathbf{i} + \left[(v_0\sin\alpha) - \frac{1}{2}gt^2\right]\mathbf{j}$$

so the parametric equations of this curve are

$$x = (v_0 \cos \alpha)t \qquad y = (v_0 \sin \alpha)t - \frac{1}{2}gt^2$$

The range of the projectile, d, is the value of x that occurs when $y = 0$. Plugging into the parametric equations implies $t = 0$ or $t = (2v_0 \sin \alpha)/g$. We choose the second value because $t > 0$ after the initial launch of the projectile. Therefore

$$d = x = (v_0 \cos \alpha)\frac{2v_0 \sin \alpha}{g} = \frac{v_0^2(2 \sin \alpha \cos \alpha)}{g} = \frac{v_0^2 \sin 2\alpha}{g}$$

The maximum value of the sin function is 1, so clearly, d is maximized when $\sin 2\alpha = 1$, i.e. $\boxed{\alpha = 45^\circ}$.

The Tangential and Normal Components of Acceleration In most cases of particle motion, it is widely considered useful to resolve the acceleration vector into two components. Specifically, one for each of two directions: the tangent and the normal. If we express v as the magnitude of the velocity, i.e. $v = |\mathbf{v}|$, then

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

so $\mathbf{v} = v\mathbf{T}$. Now we will differentiate both sides of the equation with respect to t to obtain acceleration as a function of time.

$$\mathbf{a} = \mathbf{v}' = v'\mathbf{T} + v\mathbf{T}'$$

Let's apply one of the equations for curvature from section 2.3.

$$\kappa = \frac{|\mathbf{T}'|}{|\mathbf{r}'|} = \frac{|\mathbf{T}'|}{v} \therefore |\mathbf{T}'| = \kappa v$$

We also defined the unit normal vector in the previous section, as $\mathbf{N} = \mathbf{T}/|\mathbf{T}'|$ and substituting for $|\mathbf{T}'|$ with the above, we have

$$\mathbf{T}' = |\mathbf{T}'|\mathbf{N} = \kappa v\mathbf{N}$$

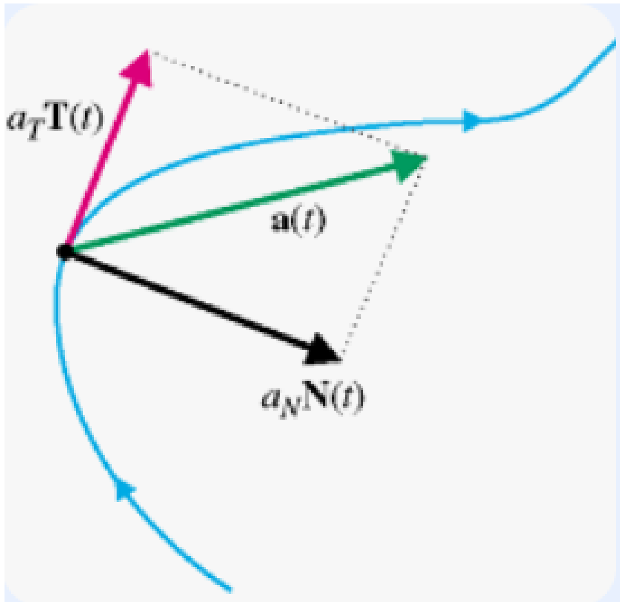
and the acceleration vector is

$$\boxed{\mathbf{a} = v'\mathbf{T} + \kappa v^2\mathbf{N}}$$

Let the tangential and normal components of the acceleration vector be given by a_T and a_N, respectively. Then

$$\mathbf{a} = a_T\mathbf{T} + a_N\mathbf{N}$$

where $a_T = v'$ and $a_N = \kappa v^2$.



It can sometimes be more desirable to express these two components of acceleration that depend only upon the position vector, i.e. expressions for $\mathbf{r}$, $\mathbf{r}'$, and $\mathbf{r}''$. We start by computing the dot product $\mathbf{v} \cdot \mathbf{a}$, letting $\mathbf{v} = v\mathbf{T}$.

$$\begin{aligned}\mathbf{v} \cdot \mathbf{a} &= v\mathbf{T} \cdot (v'\mathbf{T} + \kappa v^2 \mathbf{N}) \\ &= vv'\mathbf{T} \cdot \mathbf{T} + \kappa v^3 \mathbf{T} \cdot \mathbf{N} \\ &= vv'\end{aligned}$$

Note that $\mathbf{T} \cdot \mathbf{T} = 1$ because $\mathbf{T}$ is unit vector (has magnitude 1) and $\mathbf{T} \cdot \mathbf{N} = 0$ because the unit tangent and unit normal vectors are orthogonal. Therefore, the tangential component is

$$a_T = v' = \frac{\mathbf{v} \cdot \mathbf{a}}{v} = \frac{\mathbf{r}'(t) \cdot \mathbf{r}''(t)}{|\mathbf{r}'(t)|}$$

and using the curvature formula given in the previous section, the normal component is

$$a_N = \kappa v^2 = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|^3} |\mathbf{r}'(t)|^2 = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{|\mathbf{r}'(t)|}$$

Problem 2.4.5. *A particle moves with position function* $\mathbf{r}(t) = \langle 1, t, t^2 \rangle$*. Find the tangential and normal components of its acceleration.*

Solution: Let's find $\mathbf{r}'(t)$ and $\mathbf{r}''(t)$ first. We have

$$\mathbf{r}'(t) = \langle 0, 1, 2t \rangle$$

and $\mathbf{r}''(t) = \langle 0, 0, 2 \rangle$. The magnitude of the tangent vector is $|\mathbf{r}'(t)| = \sqrt{0^2 + 1^2 + 4t^2} = \sqrt{1 + 4t^2}$. Therefore, the tangential component of acceleration is

$$a_T = \frac{\langle 0, 1, 2t \rangle \cdot \langle 0, 0, 2 \rangle}{\sqrt{1 + 4t^2}} = \boxed{\frac{4t}{\sqrt{1 + 4t^2}}}$$

Next, we need to determine the cross product $\mathbf{r}'(t) \times \mathbf{r}''(t)$.

$$\mathbf{r}'(t) \times \mathbf{r}''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 1 & 2t \\ 0 & 0 & 2 \end{vmatrix} = 2\,\mathbf{i} = \langle 2, 0, 0 \rangle$$

Thus the normal component is

$$a_N = \frac{|\langle 2,0,0\rangle|}{\sqrt{1+4t^2}} = \boxed{\frac{2}{\sqrt{1+4t^2}}}$$

Kepler's Laws of Planetary Motion One of the greatest accomplishments in the intersection of astrophysics and calculus is the three laws of planetary motion. In the early 17^{th} century, Johannes Kepler formulated the following three laws:

1. A planet's orbit around the sun is elliptical in shape, with the Sun located at one focus.
2. The line that conjoins the sun and a planet sweeps out equal areas in equal time intervals.
3. The square of the period of revolution of a planet is directly proportional to the cube of the length of the major axis of its orbit.

In this section, we prove Kepler's First Law. Since the gravitational force exerted by the sun on the planet is exponentially larger than the forces exerted by other celestial objects, we can safely ignore all other bodies when we consider a system of the sun and a single planet revolving around it. Our coordinate system involves the sun placed at the origin and the planet in question has position vector $\mathbf{r} = \mathbf{r}(t)$. The velocity of the planet is $\mathbf{v} = \mathbf{r}'$ and the acceleration is $\mathbf{a} = \mathbf{r}''$. We use Newton's Second Law of Motion as well as his Law of Gravitation to get

$$\mathbf{F} = m\mathbf{a} \qquad \mathbf{F} = -\frac{GMm}{r^3}\,\mathbf{r} = -\frac{GMm}{r^2}\,\mathbf{u}$$

where $\mathbf{F}$ is the gravitational force on the planet, m and M are the respective masses of the planet and the sun, G is the universal gravitation constant, $r = |\mathbf{r}|$, and $\mathbf{u} = (1/r)\mathbf{r}$ is the unit vector in the direction of $\mathbf{r}$. If we equate the expressions for $\mathbf{F}$, we find

$$\mathbf{a} = -\frac{GM}{r^3}\,\mathbf{r}$$

so it is clear that $\mathbf{a}$ and $\mathbf{r}$ are parallel vectors, therefore $\mathbf{r} \times \mathbf{a} = \mathbf{0}$. Recall the fifth formula from Theorem 2.2.2 regarding the derivative of a cross product. We write

$$\frac{d}{dt}(\mathbf{r} \times \mathbf{v}) = \mathbf{r}' \times \mathbf{v} + \mathbf{r} \times \mathbf{v}' = \mathbf{v} \times \mathbf{v} + \mathbf{r} \times \mathbf{a} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Therefore $\mathbf{r} \times \mathbf{v} = \mathbf{h}$ where $\mathbf{h}$ is some constant vector. We assume that $\mathbf{h} \neq \mathbf{0}$, because then $\mathbf{r}$ and $\mathbf{v}$ are not parallel vectors. In other words, $\mathbf{r} = \mathbf{r}(t)$ is perpendicular to $\mathbf{h}$ for all values of t, so the planet always lies in the plane perpendicular to $\mathbf{h}$, and so the planet's orbit is a plane curve.

Proof.

Rewrite the vector $\mathbf{h}$ as follows:

$$\begin{aligned}\mathbf{h} &= \mathbf{r} \times \mathbf{v} = \mathbf{r} \times \mathbf{r}' = r\mathbf{u} \times (r\mathbf{u})' \\ &= r\mathbf{u} \times (r\mathbf{u}' + r'\mathbf{u}) = r^2(\mathbf{u} \times \mathbf{u}') + rr'(\mathbf{u} \times \mathbf{u}) \\ &= r^2(\mathbf{u} \times \mathbf{u}')\end{aligned}$$

Then using properties of the cross product, we obtain

$$\begin{aligned}\mathbf{a} \times \mathbf{h} &= -\frac{GM}{r^2}\,\mathbf{u} \times (r^2\,\mathbf{u} \times \mathbf{u}') = -GM\,\mathbf{u} \times (\mathbf{u} \times \mathbf{u}') \\ &= -GM\left[(\mathbf{u} \cdot \mathbf{u}')\mathbf{u} - (\mathbf{u} \cdot \mathbf{u})\mathbf{u}'\right]\end{aligned}$$

However, with $\mathbf{u}$ being a unit vector, we know $\mathbf{u} \cdot \mathbf{u} = |\mathbf{u}|^2 = 1$ and since $|\mathbf{u}(t)| = 1$ it follows that

$$\mathbf{u} \times \mathbf{u}' = 0 \therefore \mathbf{a} \times \mathbf{h} = GM\,\mathbf{u}'$$

so we have

$$(\mathbf{v} \times \mathbf{h})' = \mathbf{v}' \times \mathbf{h} + \mathbf{v} \times \mathbf{h}' = \mathbf{a} \times \mathbf{h} = GM\,\mathbf{u}'$$

We can integrate both sides of this equation to get

$$\mathbf{v} \times \mathbf{h} = GM\,\mathbf{u} + \mathbf{c}$$

where $\mathbf{c}$ is a constant vector.

At this point, it makes sense to choose a coordinate system where the standard basis $\mathbf{k}$ points in the same direction as $\mathbf{h}$, so the planet's orbit is located in the xy-plane. Since both $\mathbf{v} \times \mathbf{h}$ and $\mathbf{u}$ are perpendicular to $\mathbf{h}$, the integrated equation shows that $\mathbf{c}$ lies in the xy-plane. This means we can choose the x- and y-axes so that the standard basis vector $\mathbf{i}$ lies in the direction of $\mathbf{c}$.

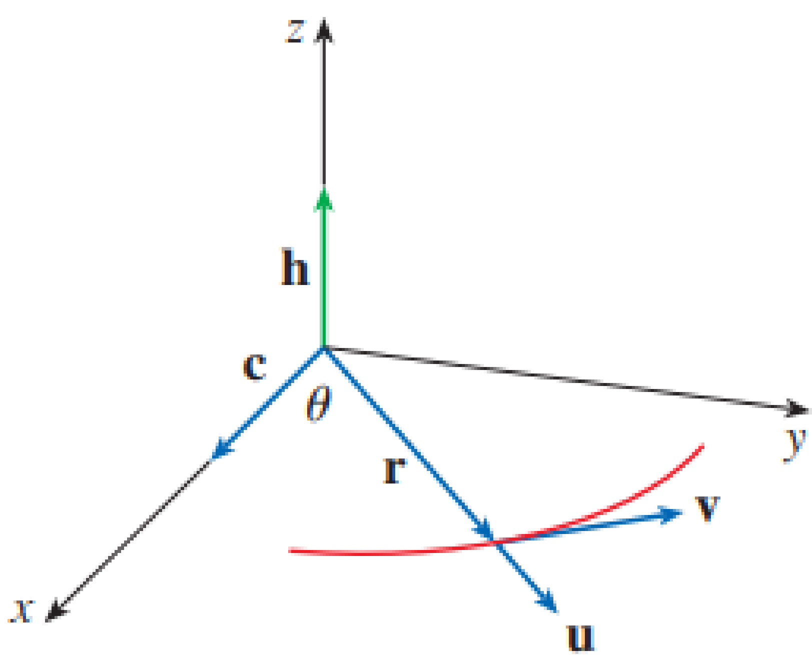


Image Credit: Calculus, Early Transcendentals

Now let θ be the angle between $\mathbf{c}$ and $\mathbf{r}$. The planet has polar coordinates of (r, θ). Therefore

$$\begin{aligned} \mathbf{r} \cdot (\mathbf{v} \times \mathbf{h}) &= \mathbf{r} \cdot (GM\,\mathbf{u} + \mathbf{c}) = GM\,\mathbf{r} \cdot \mathbf{u} + \mathbf{r} \cdot \mathbf{c} \\ &= GM\,r\,\mathbf{u} \cdot \mathbf{u} + |\mathbf{r}||\mathbf{c}| \cos\theta = GM\,r + rc\cos\theta \end{aligned}$$

where $c = |\mathbf{c}|$. Then

$$r = \frac{\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h})}{GM + c\cos\theta} = \frac{1}{GM} \frac{\mathbf{r} \times (\mathbf{v} \times \mathbf{h})}{1 + e\cos\theta}$$

where $e = c/(GM)$. However, if $h = |\mathbf{h}|$ we have

$$\mathbf{r} \cdot (\mathbf{v} \times \mathbf{h}) = (\mathbf{r} \times \mathbf{v}) \cdot \mathbf{h} = \mathbf{h} \cdot \mathbf{h} = |\mathbf{h}|^2 = h^2$$

and so

$$r = \frac{h^2/(GM)}{1 + e\cos\theta} = \frac{eh^2/c}{1 + e\cos\theta}$$

If we write $d = h^2/c$, we obtain

$$r = \frac{ed}{1 + e\cos\theta}$$

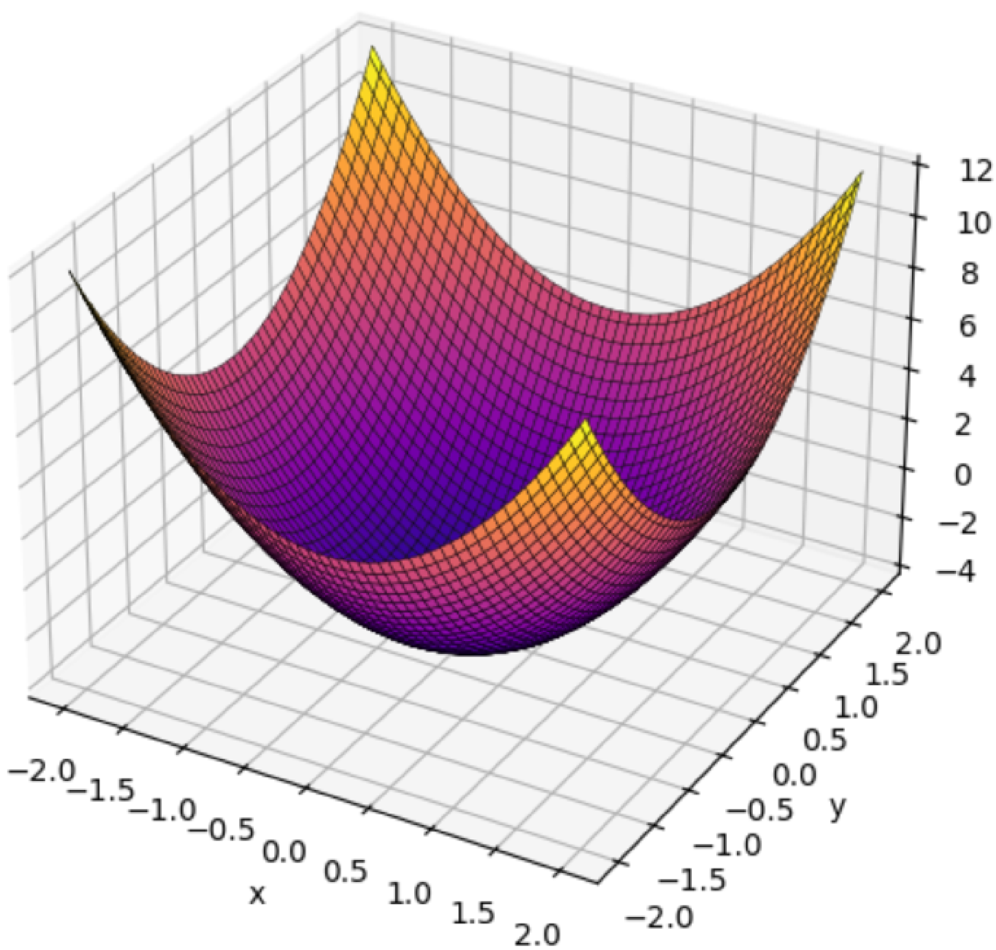
This is the polar equation of a conic section with focus at the origin and eccentricity of e. This is convenient because we know that the planet's orbit is a closed curve and because it has an eccentricity $e > 0$, the conic must be an ellipse, and the proof of Kepler's First Law is complete. □

3 Partial Derivatives

Up until now, we have focused on functions of a single variable. But in real life, many things depend on two or more variables. In this chapter, we will explore functions of several variables and see how the ideas from differential calculus can be extended to handle them.

3.1 Working with Multivariable Functions

In this section, we discuss some basic principles regarding functions of more than one variable. Particularly, we should recall that graphs of functions of two variables, $z = f(x, y)$, are treated as surfaces in three-dimensional space. For example, consider the elliptic paraboloid with equation $z = 2x^2 + 2y^2 - 4$, graphed using Python code.



This surface is an example of a *quadric* surface. We saw several of these in Chapter 1. We will be seeing quadric surfaces fairly regularly later on in this course. Another type of graph that will appear frequently in this course is the graph of a plane. We use a standard convention for graphing planes that makes them easier to draw and visualize. The equation of a plane is given by

$$ax + by + cz = d$$

If we solve for z, we can write this plane in the form of a surface, with

$$z = f(x, y) = Ax + By + D$$

Graphing a plane is generally easy. We find its intercepts with the three coordinate axes and connect them to form a triangle. This triangle represents a portion of the plane and provides a good visualization of its overall shape. For example, consider the plane given by the equation

$$f(x, y) = 12 - 3x - 4y$$

When it comes to graphing, it is probably easier to describe the plane as

$$z = 12 - 3x - 4y \quad \Longrightarrow \quad 3x + 4y + z = 12$$

The idea of the intersection points with each of the x-, y-, and z-axes is based on the fact that two of the other coordinates are zeroed out. For instance, to find the intersection of the plane with the x-axis, we set $y = z = 0$. So the three intersection points are:

$$\begin{aligned} x\text{-axis} &: (4, 0, 0) \\ y\text{-axis} &: (0, 3, 0) \\ z\text{-axis} &: (0, 0, 12) \end{aligned}$$

We can also consider graphs of functions in the form $w = f(x, y, z)$, which would represent four-dimensional surfaces. Obviously, we cannot do this, but it is interesting to consider.

Domain of Multivariable Functions Now, we move on to the subject of the domains of multivariable functions. Recall that for single-variable functions, i.e. $y = f(x)$, the domain consists of all values x such that we could generate a real value for y. Now, if we think about it, this means that the domain of a function of a single variable is an interval (or intervals) of values from the number line, or one dimensional space.

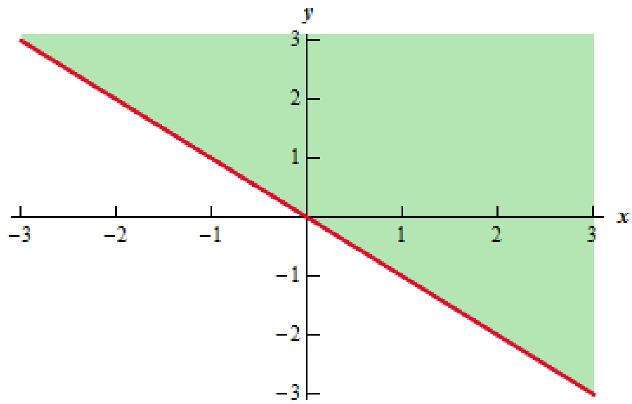
Thinking in two dimensions, the domain of two-variable functions $z = f(x, y)$ are two-dimensional regions consisting of all pairs (x, y) that we could generate a real value for z.

Problem 3.1.1. *Determine the domain of each of the following.*
(a) $f(x, y) = \sqrt{x + y - 1}$
(b) $f(x, y) = \ln(9 - x^2 - 9y^2)$

Solution to part a: In this case, we know that we cannot take the square root of a negative number, so we must require

$$x + y - 1 \geq 0 \Longrightarrow \boxed{x + y \geq 1}$$

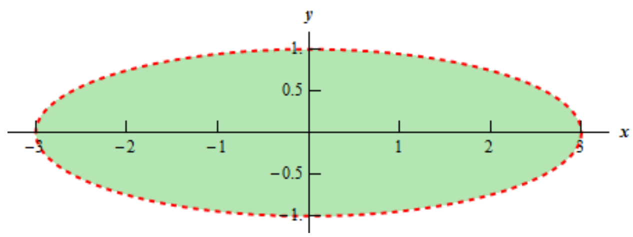
Here is a sketch of the region which illustrates the domain.



Solution to part b: We know that we cannot take the logarithm of a number which is negative or zero. Thus we must require

$$9 - x^2 - 9y^2 > 0 \Longrightarrow \boxed{\frac{x^2}{9} + y^2 < 1}$$

This inequality indicates that our domain is interior to the ellipse shown below.



Now, let's move onto functions of three variables, i.e. $w = f(x, y, z)$. Their domains will be regions in three-dimensional space.

Problem 3.1.2. *Determine the domain of the function* $f(x, y, z) = \frac{1}{\sqrt{x^2+y^2+z^2-16}}$.

Solution: In this case, we need to prevent taking the square root of a negative number as well as zero division. These conditions require

$$x^2 + y^2 + z^2 - 16 > 0 \Longrightarrow x^2 + y^2 + z^2 > 16$$

So the domain for this function is the set of all points that lies completely outside a sphere of radius 4 centered at the origin.

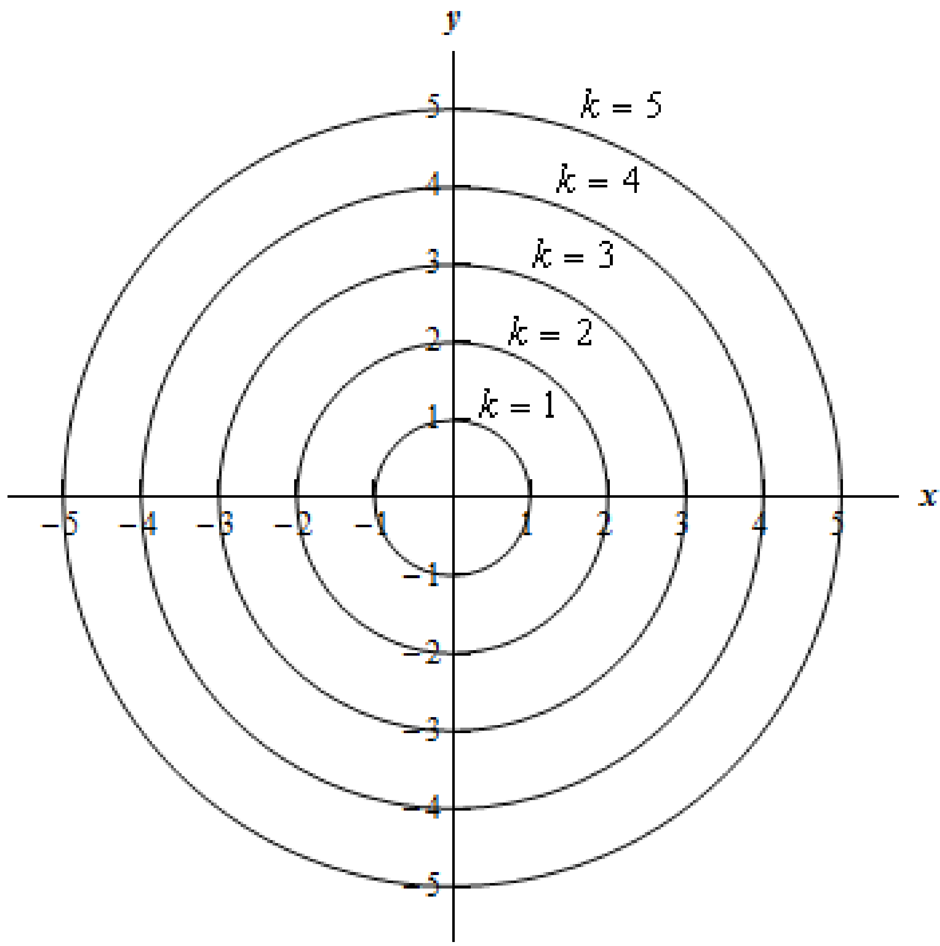
A Brief Discussion on Level or Contour Curves For a function $z = f(x, y)$, the *level curves* or *contour curves* are the two-dimensional curves we obtain by setting $z = k$, where k can be any real number. Hence, the equations of the level curves are $f(x, y) = k$. Note that sometimes the equation can be in the forms $f(x, y, z) = 0$ or $f(x, y, k) = 0$. You've probably seen contour curves before. If you've ever seen an elevation map, this is nothing more than the contour curves describing the function that describes the elevation of land in a particular area.

Problem 3.1.3. *Identify the level curves of* $f(x, y) = \sqrt{x^2 + y^2}$.

Solution: Recall from section 1.7 on quadric surfaces that the surface given by $f(x, y) = \sqrt{x^2 + y^2}$ is the upper half of the right circular cone. The level or contour curves of the function can be found by taking

$$k = \sqrt{x^2 + y^2} \implies x^2 + y^2 = k^2$$

where $k \in \mathbb{R}$. In this case, the level curves of $f(x, y)$ are circles of radius k centered at the origin.

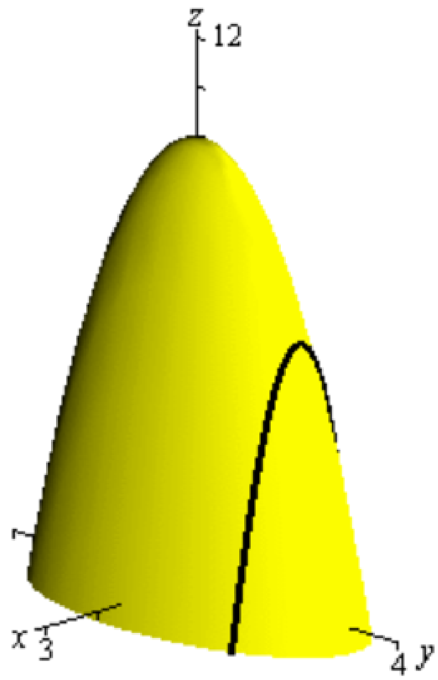


We can think of contours here as the intersection of the surface and the plane. For functions of the form $f(x, y, z)$, we observe *level surfaces*. These are given by $f(x, y, z) = k$, with $k \in \mathbb{R}$.

Traces The last topic of this section will cover *traces*. These are in many ways similar to contours/level curves. Specifically, traces of surfaces are curves that represent the intersection of the surface and the planes $x = a$ or $y = b$.

Problem 3.1.4. *Sketch the traces of* $f(x, y) = 10 - 4x^2 - y^2$ *for the planes* $y = 2$.

Solution: We have $z = f(x, y)$ and with $y = 2$, we have $z = f(x, 2) = 10 - 4x^2 - (2)^2$. This results in the parabola $\boxed{z = 6 - 4x^2}$.



3.2 Limits and Continuity of Multivariable Functions

Let's start this section with the discussion of the behavior of two functions

$$f(x,y) = \frac{\sin(x^2+y^2)}{x^2+y^2} \quad \text{and} \quad g(x,y) = \frac{x^2-y^2}{x^2+y^2}$$

as both x and y approach 0, i.e. the point (x,y) approaches the origin.

Convince yourself that as (x,y) approaches $(0,0)$, the values of $f(x,y)$ are approaching 1 whereas the values of $g(x,y)$ are not approaching any particular number. It turns out that these guesses based on numerical evidence are correct, so we have

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2} = 1 \quad \text{and} \quad \lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2} = \text{DNE}$$

In general, the limit notation is

$$\lim_{(x,y)\to(0,0)} f(x,y) = L$$

and indicates that the values of $f(x,y)$ approach the number L as the point (x,y) approaches the point (a,b) (within the domain of f). In other words, if we make the values of (x,y) as close to (a,b) as possible, we get a similar effect regarding the values of $f(x,y)$ and L.

Proving the Non-Existence of a Limit Let's revisit single-variable calculus for a bit. If we let x approach a value a, there are only two possible directions, i.e. one from the left and one from the right. We also know that if $\lim_{x\to a^-} f(x) \neq \lim_{x\to a^+} f(x)$, then $\lim_{x\to a} f(x)$ does not exist. However, the situation is not that easy when dealing with functions of two variables. This is because we can let (x, y) approach (a, b) from an infinite number of directions in any manner as long as (x, y) stays within the domain of f.

Intuitively, the distance between $f(x, y)$ and L can be made arbitrarily small by placing (x, y) as close to (a, b) as physically possible (but not letting the distance reach 0). Note that this notion only refers to the distance, not the direction. Therefore, if the limit exists, then $f(x, y)$ must approach the same limit *no matter how* (x, y) approaches (a, b). Therefore, a systematic way to demonstrate that $\lim (x, y) \to (a, b) f(x, y)$ does not exist is to find different paths of approach along which the function has different limiting values.

Problem 3.2.1. *Show that the limit* $\lim_{(x,y)\to(0,0)} \frac{x^2-y^2}{x^2+y^2}$ *does not exist.*

We have $f(x, y) = (x^2 - y^2)/(x^2 + y^2)$. First let's approach $(0, 0)$ along the x-axis. On this path we have $y = 0$ so the function becomes $f(x, 0) = x^2/x^2 = 1$ for all $x \neq 0$ and thus

$$f(x, y) \to 1 \quad \text{as} \quad (x, y) \to (0, 0) \quad \text{along the } x-\text{axis}$$

We can now approach along the y-axis by setting $x = 0$. Then $f(0, y) = -y^2/y^2 = -1$ for all $x \neq 0$, so

$$f(x, y) \to -1 \quad \text{as} \quad (x, y) \to (0, 0) \quad \text{along the } y-\text{axis}$$

Since f has two different limiting values as (x, y) approaches $(0, 0)$ along two different paths, the given limit does not exist. □

Problem 3.2.2. *If* $f(x, y) = \frac{xy^2}{x^2+y^4}$, *does* $\lim_{(x,y)\to(0,0)} f(x, y)$ *exist?*

Solution: Let's try to save some time by letting $(x, y) \to (0, 0)$ along any line passing through the origin. If the line is not the y-axis, we have $y = mx$, where m is the slope.

$$f(x, y) = f(x, mx) = \frac{x(mx)^2}{x^2 + (mx)^4} = \frac{x^2m^3}{x^2 + m^4x^4} = \frac{m^2x}{1 + m^4x^2}$$

so $f(x, y) \to 0$ as $(x, y) \to (0, 0)$ across the line $y = mx$. If we let $x = 0$, we get the same result as $(x, y) \to (0, 0)$ along that line. Thus f has the limiting value along every line passing through the origin. But that does not help us determine that the given limit is 0. For instance, let $(x, y) \to (0, 0)$ along the parabola $x = y^2$, so

$$f(x, y) = f(y^2, y) = \frac{y^2 \cdot y^2}{(y^2)^2 + y^4} = \frac{y^4}{2y^4} = \frac{1}{2}$$

so $f(x, y) \to \frac{1}{2}$ as $(x, y) \to (0, 0)$ along $x = y^2$. Since different paths lead to different limiting values, the limit does not exist. □

Properties of Limits Just as for functions of one variable, the calculation of limits for functions of two variables can be greatly simplified by the use of properties of limits. The conventional limit laws from single-variable calculus can be extended to functions of two variables. Assuming that the indicated limits exist, we can state these laws as follows:

1. The limit of a sum is the sum of the limits.
2. The limit of a difference is the difference of the limits.
3. The limit of a constant multiplied by a function is the constant multiplied by the limit of the function.
4. The limit of a product is the product of the limits.
5. The limit of a quotient is the quotient of the limits, assuming the limit of the denominator is nonzero.

A *polynomial function* of two variables is defined as a summation of terms of the form cx^my^n where m and n are nonnegative integers. A *rational function* is defined as the quotient of two polynomial functions. For example, $p(x,y) = x^4+5xy^3-2x^2y$ is a polynomial function, whereas $q(x,y) = \frac{2xy+1}{x^2-y^2}$ is a rational function.

Problem 3.2.3. *Evaluate* $\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y)$.

Solution: Since $f(x,y) = x^2y^3 - x^3y^2 + 3x + 2y$ is a polynomial function, we can compute the limit via direct substitution.

$$\lim_{(x,y)\to(1,2)} (x^2y^3 - x^3y^2 + 3x + 2y) = 1^2 \cdot 2^3 - 1^3 \cdot 2^2 + 3 \cdot 1 + 2 \cdot 2 = \boxed{11}$$

Problem 3.2.4. *Evaluate* $\lim_{(x,y)\to(0,0)} \frac{x^2y+1}{x^3y^2-2x}$ *if it exists.*

Solution: The function $f(x,y) = (x^2y+1)/(x^3y^2-2x)$ is a rational function and the point $(-2,3)$ is in its domain, as the denominator, x^3y^2-2x is not zero at that point. So once again, we can evaluate the limit using direct substitution:

$$\lim_{(x,y)\to(-2,3)} \frac{x^2y+1}{x^3y^2-2x} = \frac{(-2)^2(3)+1}{(-2)^3(3)^2-2(-2)} = \boxed{-\frac{13}{68}}$$

We will see that the Squeeze Theorem also holds for functions of two variables.

Problem 3.2.5. *Find* $\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$ *if it exists.*

Solution: First, we know that the quantities $3x^2$ and x^2+y^2 are positive for all $x, y \in \mathbb{R}$. Hence, we can write

$$\left|\frac{3x^2y}{x^2+y^2}\right| = \frac{3x^2|y|}{x^2+y^2} \leq 3|y|$$

and so we have the bounds

$$-3|y| \leq \frac{3x^2y}{x^2+y^2} \leq 3|y|$$

Now $|y| \to 0$ as $y \to 0$, so $\lim_{(x,y)\to(0,0)}(-3|y|) = 0$ and $\lim_{(x,y)\to(0,0)}(3|y|) = 0$ using the third limit law. Therefore, the Squeeze Theorem forces

$$\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2} = \boxed{0}$$

Continuity Recall from single-variable calculus that it was extremely easy to evaluate limits of continuous functions, i.e. we could apply direct substitution, because the core criterion for continuity was $\lim_{x\to a} f(x) = f(a)$. We can apply the same principle in functions of two variables.

Definition 3.2.1. *A function f of two variables is **continuous** at a point (a, b) if*

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

Also, f is said to be continuous on D if and only if the function is continuous at every point $(a, b) \in D$.

More literally, the idea of continuity is that if the point (x, y) changes by s small amount, the corresponding function value $f(x, y)$ also changes by a small amount. Geometrically, this means a surface that is the graph of a continuous function will have no holes or breaks.

Problem 3.2.6. *Where is the function $f(x, y) = \frac{x^2-y^2}{x^2+y^2}$ discontinuous?*

Solution: The function f is discontinuous at the point $\boxed{(0, 0)}$ because it is not defined there. Since f is a rational function, it is continuous only on its domain D, where $D = \{(x, y) \,|\, (x, y) \neq (0, 0)\}$.

Just as with functions of a single variable, composition provides a method for combining two continuous functions. Specifically, if f is a continuous function of two variables and g is a continuous function of a single variable defined on the range of f, then the composite function $h = g \circ f$, defined by

$$h(x, y) = g(f(x, y)),$$

is also continuous.

Problem 3.2.7. *Where is the function $h(x, y) = e^{-(x^2+y^2)}$ continuous?*

Solution: The function $f(x, y) = x^2 + y^2$ is a polynomial function and is continuous everywhere on $\mathbb{R}^2$. Meanwhile, the function $g(t) = e^{-t}$ is continuous for $t \in \mathbb{R}^2$, so the composition

$$h(x, y) = g(f(x, y)) = e^{-(x^2+y^2)}$$

is continuous on $\boxed{\mathbb{R}^2}$.

Problem 3.2.8. *Where is the function* $h(x, y) = \arctan(y/x)$ *continuous?*

Solution: The function $f(x, y) = y/x$ is a rational function and is continuous everywhere on $\mathbb{R}^2$ except for the line $x = 0$. Meanwhile, the function $g(t) = \arctan t$ is continuous for $t \in \mathbb{R}$, so the composition

$$h(x, y) = g(f(x, y)) = \arctan(y/x)$$

is continuous on $\boxed{\text{the subset of } \mathbb{R}^2 \text{ excluding the line } x = 0}$.

Functions of Three or More Variables Everything we have done so far involving limits and continuity can be extended to functions of three or more variables. In particular, the limit notation becomes

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

and indicates that the values of $f(x, y, z)$ approach the number L as the point (x, y, z) approaches the point (a, b, c), where $(a, b, c) \in E$, the domain of f. Also, the function f is continuous at (a, b, c) if and only if

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

Problem 3.2.9. *Where is the function* $f(x, y, z) = \frac{1}{x^2+y^2+z^2-1}$ *continuous?*

Solution: We have a rational function $f(x, y, z) = \frac{1}{x^2+y^2+z^2-1}$ of three variables and is therefore continuous at every point in $\mathbb{R}^3$ except where $x^2 + y^2 + z^2 = 1$. In other words, f is continuous on $\boxed{\text{the subset of } \mathbb{R}^3 \text{ excluding the unit sphere}}$.

3.3 Partial Derivatives: In Essence

Introduction to Partial Derivatives In general, if f is a function of two variables x and y, we can control the variability of these values to make some interesting conclusions. Suppose we place a restriction that y is fixed, i.e. $y = b$, where b is a constant. Then we really are considering a single-variable function $g(x) = f(x, b)$. If g has a derivative at a, then $g'(a)$ is called the **partial derivative of f with respect to x** at (a, b) and is denoted by $f_x(a, b)$. Thus we have $f_x(a, b) = g'(a)$ where $g(x) = f(x, b)$. According to the definition of a derivative, we have

$$g'(a) = \lim_{h\to 0} \frac{g(a + h) - g(a)}{h}$$

and so the partial derivative of f with respect to x is

$$\boxed{f_x(a, b) = \lim_{h\to 0} \frac{f(a + h, b) - f(a, b)}{h}}$$

Likewise, if we fix $x = a$ and allow y to vary, we get a function $G(y) = f(a, y)$ so the partial derivative of f with respect to y is

$$\boxed{f_y(a, b) = \lim_{h\to 0} \frac{f(a, b + h) - f(a, b)}{h}}$$

If we allow (a, b) to vary, then f_x and f_y become functions of two variables $f_x(x, y)$ and $f_y(x, y)$. There are also many alternative notations for partial derivatives. For example, f_x can be written as f_1 or $D_1 f$ (indicates differentiation with respect to the first variable) or $\partial f/\partial x$ (but this cannot be interpreted as a ratio of differentials).

In order to compute partial derivatives, all we need to do is remember that the partial derivative with respect to x is the same as the ordinary derivative of the single-variable function g obtained by keeping y fixed, and vice versa. Let's apply this rule in practice.

Problem 3.3.1. *If $f(x, y) = x^3 + x^2y^3 - 2y^2$, find $f_x(2, 1)$ and $f_y(2, 1)$.*

Solution: To find f_x, we hold y constant and differentiate with respect to x. This yields $f_x(x, y) = 3x^2 + 2xy^3$. Plugging in $(2, 1)$ yields $3 \cdot 2^2 + 2 \cdot 2 \cdot 1^3 = \boxed{16}$. On the other hand, holding x constant and differentiating with respect to y yields $f_y(x, y) = 3x^2y^2 - 4y$, so plugging in $(1, 1)$ yields $3 \cdot 2^2 \cdot 1^2 - 4 \cdot 1 = \boxed{8}$.

Problem 3.3.2. *If $f(x, y) = \cos\left(\frac{x}{1+y}\right)$, calculate f_x and f_y.*

Solution: This problem will require us to use the Chain Rule for single-variable functions.

$$\frac{\partial f}{\partial x} = -\sin\left(\frac{x}{1+y}\right) \cdot \frac{\partial}{\partial x}\left(\frac{x}{1+y}\right) = \boxed{-\sin\left(\frac{x}{1+y}\right) \cdot \frac{1}{1+y}}$$
$$\frac{\partial f}{\partial y} = -\sin\left(\frac{x}{1+y}\right) \cdot \frac{\partial}{\partial y}\left(\frac{x}{1+y}\right) = \boxed{\sin\left(\frac{x}{1+y}\right) \cdot \frac{x}{(1+y)^2}}$$

Interpreting Partial Derivatives Now, we want to develop a geometric interpretation for partial derivatives. Recall that the equation $z = f(x, y)$ represents a surface S, which is the graph of f. If $f(a, b) = c$, then the point $P(a, b, c)$ lies on S If we fix $y = b$, we restrict our attention to the curve C_1 in which the vertical plane intersects S. Similarly, the vertical plane $x = a$ intersects S in a curve C_2. Both curves pass through the point P.

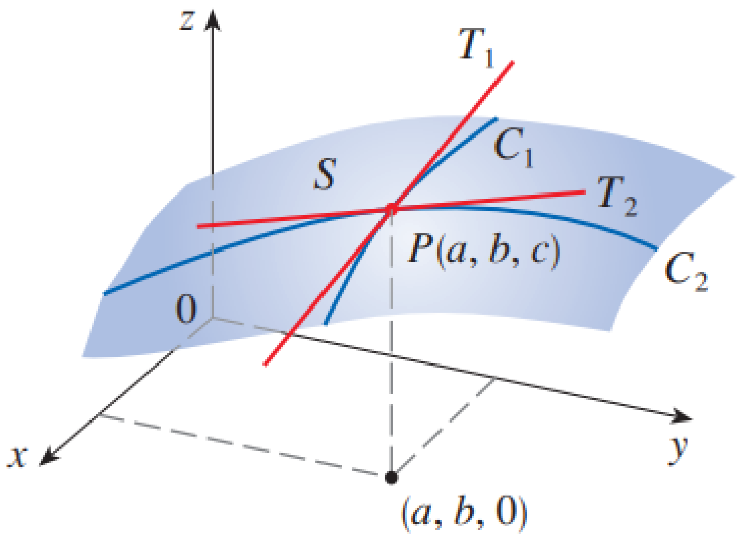


Image Credit: Calculus, Early Transcendentals

Observe that C_1 represents the graph of the function $g(x) = f(x, b)$. Therefore, the slope of its tangent line T_1 at the point P is given by $g'(a) = f_x(a, b)$.

Similarly, C_2 is the graph of the function $G(y) = f(a, y)$, and the slope of its tangent line T_2 at P is $G'(b) = f_y(a, b)$.

Hence, the partial derivatives $f_x(a, b)$ and $f_y(a, b)$ have a geometric meaning: they correspond to the slopes of the tangent lines at the point $P(a, b, c)$ to the cross-sectional curves C_1 and C_2 obtained by intersecting the surface S with the planes $y = b$ and $x = a$, respectively.

Problem 3.3.3. *If $f(x, y) = 4 - x^2 - 2y^2$, find $f_x(1, 1)$ and $f_y(1, 1)$ and interpret them.*

Solution: The partial derivatives are $f_x(x, y) = -2x$ and $f_y(x, y) = -4y$. At the point $(1, 1)$, we have $f_x(1, 1) = \boxed{-2}$ and $f_y(1, 1) = \boxed{-4}$. The graph of f is the paraboloid with equation $z = f(x, y) = 4 - x^2 - 2y^2$ and the vertical plane $y = 1$ intersects it in the parabola $z = 2 - x^2$, $y = 1$. The slope of the tangent line to this parabola at the point $(1, 1, f(1)) = (1, 1, 1)$ is $f_x(1, 1) = -2$. Likewise, the plane $x = 1$ intersects the paraboloid in the parabola $z = 3 - 2y^2$, $x = 1$, and the slope of the tangent line at $(1, 1, 1)$ is $f_y(1, 1) = -4$.

The key idea here is that we can interpret partial derivatives as *rates of change.* If $z = f(x, y)$, then $\partial z/\partial x$ represents the rate of change of z with respect to x when y is fixed. Likewise, $\partial z/\partial y$ represents the rate of change of z with respect to y when x is fixed.

Problem 3.3.4. *The body mass index of a person with mass m and height h is defined by the function $B(m, h) = \frac{m}{h^2}$. Calculate the partial derivatives of B for a young man with $m = 64$ kg and $h = 1.68$ m and interpret these values.*

Solution: We find the partial derivative with respect to m via treating h as a constant.

$$B_m(m, h) = \frac{\partial}{\partial m}\left(\frac{m}{h^2}\right) = \frac{1}{h^2}$$

so $B_m(64, 1.68) = \frac{1}{(1.68)^2} \approx \boxed{0.35\,(\text{kg/m}^2)/\text{kg}}$. This is the rate at which the man's BMI increases with respect to his weight when he weighs 64 kg and his height is 1.68 m. Suppose his weight increases by 1 kg, but his height is unchanged, so his BMI will increase by approximately $B(64, 1.68) \approx 22.68$ by about 0.35.

Now, let m be constant. The partial derivative with respect to h is

$$B_h(m, h) = \frac{\partial}{\partial h}\left(\frac{m}{h^2}\right) = m\left(-\frac{2}{h^3}\right) = -\frac{2m}{h^3}$$

so $B_h(64, 1.68) = -\frac{2\cdot 64}{(1.68)^3} \approx \boxed{-27\,(\text{kg/m}^2)/\text{m}}$. This represents the rate at which the man's BMI increases with respect to his height when he weighs 64 kg and his height is 1.68 m. So if the man is still growing and his weight stays unchanged while his height increases by a small amount, say 1 cm, then his BMI will *decrease* by about $27(0.01) = 0.27$.

Problem 3.3.5. *Find z_x and z_y if z is defined implicitly as a function of x and y for the equation*

$$x^3 + y^3 + z^3 + 6xyz + 4 = 0$$

Then evaluate these partial derivatives at the point $(-1, 1, 2)$.

Solution: To find z_x, we need to differentiate z with respect to x, ensuring we treat y as a constant and z as a function of x.

$$3x^2 + 3z^2\frac{\partial z}{\partial x} + 6yz + 6xy\frac{\partial z}{\partial x} = 0$$

Solving this equation for $\frac{\partial z}{\partial x}$, we obtain

$$\frac{\partial z}{\partial x} = \boxed{-\frac{x^2 + 2yz}{z^2 + 2xy}}$$

Similarly, implicit differentiation with respect to y gives

$$\frac{\partial z}{\partial y} = \boxed{-\frac{y^2 + 2xz}{z^2 + 2xy}}$$

The point $(-1, 1, 2)$ satisfies the equation $x^3 + y^3 + z^3 + 6xyz + 4 = 0$, so it lies on the surface defined by $z = f(x, y)$. At the point $(-1, 1, 2)$ we have $z_x = \boxed{-\frac{5}{2}}$ and $z_y = \boxed{\frac{3}{2}}$.

Functions of Three or More Variables Partial derivatives can also be expressed for functions of three variables more. For example, if f is a function of three variables x, y, and z, then the partial derivative with respect to x is

$$f_x(x, y, z) = \lim_{h \to 0} \frac{f(x + h, y, z) - f(x, y, z)}{h}$$

This is obtained by regarding y and z as constants and differentiating $f(x, y, z)$ with respect to only x. The partial derivative f_x can be interpreted as the rate of change of f with respect to x

when y and z are held fixed. Unfortunately, we cannot interpret this geometrically because the graph of f lies in four-dimensional space.

In general, if u is a function of n variables, where $u = f(x_1, x_2, \ldots, x_n)$, the partial derivative with respect to the ith variable is given by

$$\frac{\partial u}{\partial x_i} = \lim_{h\to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n)}{h}$$

and also with notations like $\frac{\partial u}{\partial x_i} = \frac{\partial f}{\partial x_i} = f_{x_i} = f_i = D_i f$.

Problem 3.3.6. *Find f_x, f_y, and f_z if $f(x, y, z) = e^{xy} \ln z$.*

Solution: To find f_x, we hold y and z constant and differentiate with respect to x. This yields $f_x = \boxed{ye^{xy} \ln z}$. Now, for f_y, we hold x and z constant to obtain $f_y = \boxed{xe^{xy} \ln z}$. You can tell where this is going for the case of f_z. We get $f_z = \boxed{\dfrac{e^{xy}}{z}}$.

Higher Order Partial Derivatives Notice that if f is a function of two variables, then its partial derivatives f_x and f_y also have the same property. This means we can also consider *their* partial derivatives $(f_x)_x$, $(f_x)_y$, $(f_y)_x$, and $(f_y)_y$. These are called the **second-order partial derivatives** of f. If $z = f(x, y)$, we have the following notation.

$$\begin{aligned}
(f_x)_x &= f_{xx} = f_{11} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 z}{\partial x^2} \\
(f_x)_y &= f_{xy} = f_{12} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y\, \partial x} = \frac{\partial^2 z}{\partial y\, \partial x} \\
(f_y)_x &= f_{yx} = f_{21} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x\, \partial y} = \frac{\partial^2 z}{\partial x\, \partial y} \\
(f_y)_y &= f_{yy} = f_{22} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = \frac{\partial^2 z}{\partial y^2}
\end{aligned}$$

Therefore, the notation f_{xy} indicates that we differentiate first with respect to x and then with respect to y. With f_{yx}, the order of differentiation is swapped.

Problem 3.3.7. *Find the second partial derivatives for the function $f(x, y) = x^3 + x^2y^3 - 2y^2$.*

Solution: The first step is to compute the first-order partial derivatives

$$f_x(x, y) = 3x^2 + 2xy^3 \qquad f_y(x, y) = 3x^2y^2 - 4y$$

Therefore, we have

$$f_{xx} = \frac{\partial}{\partial x}(3x^2 + 2xy^3) = \boxed{6x + 2y^3} \qquad f_{xy} = \frac{\partial}{\partial y}(3x^2 + 2xy^3) = \boxed{6xy^2}$$

$$f_{yx} = \frac{\partial}{\partial x}(3x^2y^2 - 4y) = \boxed{6xy^2} \qquad f_{yy} = \frac{\partial}{\partial y}(3x^2y^2 - 4y) = \boxed{6x^2y - 4}$$

Remark. In the above problem, $f_{xy} = f_{yx}$. This is not a mere coincidence. It turns out that the mixed partial derivatives f_{xy} and f_{yx} are always equivalent when the function f and the first and second-order partial derivatives are continuous in a region.

The following theorem, discovered by French mathematician Alexis Clairaut, summarizes the conditions in which we can assert $f_{xy} = f_{yx}$.

Theorem 3.3.1. *Suppose f is defined on a disk D that contains the point (a, b). If f_{xy} and f_{yx} are both continuous on D, then Clairaut's Theorem states that $f_{xy}(a, b) = f_{yx}(a, b)$ for all $(a, b) \in D$.*

Proof. For small values of h, where $h \neq 0$, consider the difference

$$\Delta(h) = [f(a+h, b+h) - f(a+h, b)] - [f(a, b+h) - f(a, b)]$$

If we let $g(x) = f(x, b+h) - f(x, b)$, then

$$\Delta(h) = g(a+h) - g(a)$$

According to Mean Value Theorem, there is some value c between a and $a+h$ such that

$$g(a+h) - g(a) = g'(c)h = h[f_x(c, b+h) - f_x(c, b)]$$

Applying the Mean Value Theorem again to f_x, we get some number d between b and $b+h$ such that

$$f_x(c, b+h) - f_x(c, b) = f_{xy}(c, d)h$$

Combining the equations, we obtain

$$\Delta(h) = h^2 f_{xy}(c, d)$$

If $h \to 0$, then $(c, d) \to (a, b)$, so the continuity of f_{xy} at (a, b) guarantees

$$\lim_{h \to 0} \frac{\Delta(h)}{h^2} = \lim_{(c,d) \to (a,b)} f_{xy}(c, d) = f_{xy}(a, b)$$

Similarly, we write

$$\Delta(h) = [f(a+h, b+h) - f(a, b+h)] - [f(a+h, b) - f(a, b)]$$

Applying the Mean Value Theorem twice and the continuity of f_{yx} at (a, b), we obtain

$$\lim_{h \to 0} \frac{\Delta(h)}{h^2} = f_{yx}(a, b)$$

From the above derivations, it follows that $f_{xy} = f_{yx}$, so the proof is complete. □

Problem 3.3.8. *Calculate f_{xxyz} for $f(x, y, z) = \sin(3x + yz)$.*

Solution: We simply apply step-by-step partial differentiation. First, $f_x = 3\cos(3x+yz)$. Then $f_{xx} = (f_x)_x = -9\sin(3x+yz)$. Then $f_{xxy} = (f_{xx})_y = -9z\cos(3x+yz)$. Finally, $f_{xxyz} = (f_{xxy})_z = \boxed{-9\cos(3x+yz) + 9yz\sin(3x+yz)}$.

Partial Differential Equations (PDEs) Partial derivatives frequently occur in *partial differential equations* (PDEs) that express a multitude of physical laws. For example, the differential equation below

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

is called **Laplace's equation**, named after mathematician Pierre Laplace. The solutions of this equation are grouped in a family called **harmonic functions**, and they play a major role in problems involving heat conduction, fluid flow, and electric potential.

Problem 3.3.9. *Prove that the function $u(x, y) = e^x \sin y$ is a solution of Laplace's equation.*

Solution: We just need to compute the second-order partial derivatives of u.

$$u_x = e^x \sin y \qquad u_y = e^x \cos y$$

$$u_{xx} = e^x \sin y \qquad u_{yy} = -e^x \sin y$$

So $u_{xx} + u_{yy} = e^x \sin y - e^x \sin y = 0$, so $u(x, y)$ is a solution of Laplace's equation and the proof is complete. □

The **wave equation**, given by

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}$$

is used to describe the motion of a waveform, e.g. ocean waves, sound waves, electromagnetic waves, etc. For instance, if $u(x, t)$ represents the the displacement of a vibrating violin string at time t and distance x from one end of the string, then $u(x, t)$ satisfies the wave equation. The constant a is based on the string density and force of tension within it.

Problem 3.3.10. *Prove that the function $u(x, t) = \sin(x - at)$ satisfies the wave equation.*

Solution: Similarly as verifying Laplace's equation, we just need to compute the second order partial derivatives.

$$u_x = \cos(x - at) \qquad u_t = -a \cos(x - at)$$

$$u_{xx} = -\sin(x - at) \qquad u_{tt} = -a^2 \sin(x - at) = a^2 u_{xx}$$

Since $u_{tt} = a^2 u_{xx}$, the function $u(x, t) = \sin(x - at)$ is a solution to the wave equation and the proof is complete. □

PDEs involving three variables are also very important in science and engineering. The three-dimensional variant of Laplace's equation is

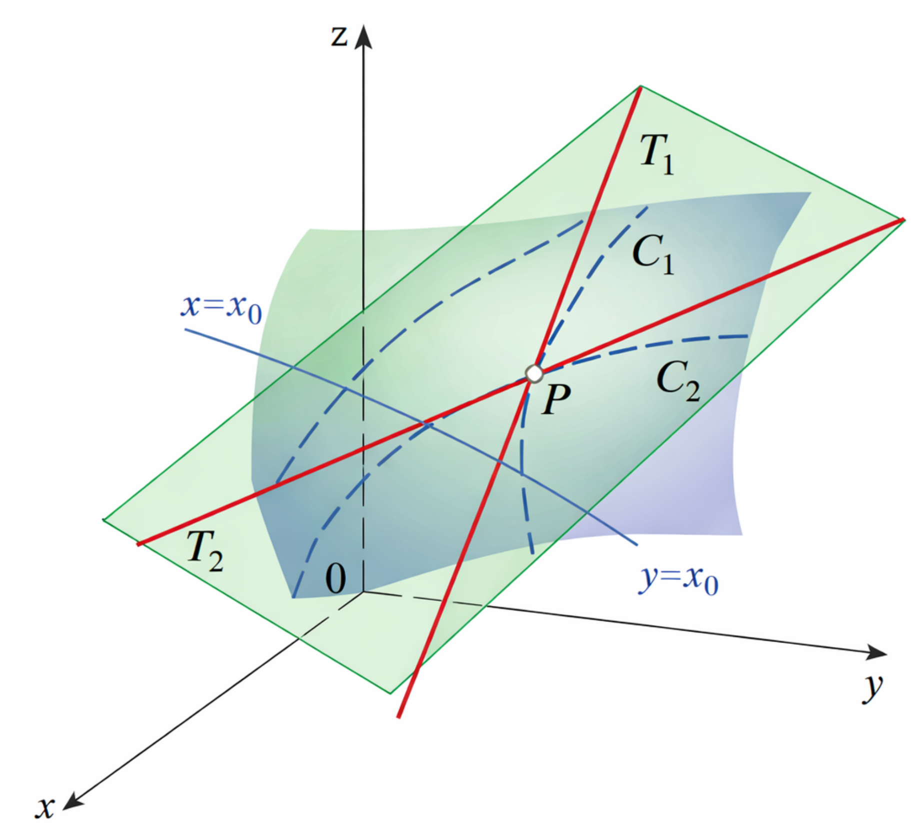
$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

and one such application is in geophysics. Suppose u represents the magnetic field strength at a point (x, y, z). The magnetic field strength indicates the distribution of iron-rich materials and reflects different types of rocks and fault locations.

3.4 Tangent Planes and Linear Approximations

A central concept in single-variable calculus is that when we examine a differentiable function very close to a point, its graph begins to look like its tangent line, allowing the function to be approximated by a linear expression. This section extends that idea to three dimensions. When we focus closely on a point on a surface defined by a differentiable function of two variables, the surface increasingly resembles a plane, its tangent plane, so the function can be approximated by a linear function of two variables. We also generalize the concept of differentials to functions involving two or more variables.

Tangent Planes Suppose a surface S has equation $z = f(x, y)$, where f has continuous first-order partial derivatives f_x and f_y. Let $P(x_0, y_0, z_0)$ be an arbitrary point on S, and let C_1 and C_2 be the two curves obtained by intersecting the vertical planes $y = y_0$ and $x = x_0$ with the surface S. Then P lies on both C_1 and C_2. Call T_1 and T_2 the tangent lines to C_1 and C_2, respectively, at point P. Then the tangent plane to the surface S at the point P is defined to be the plane that contains both tangent lines T_1 and T_2.



We can think of the tangent plane to S at P as the set of all possible tangent lines to P to curves lying in S and passing through P. Also, we know from section 1.5 that the equation of any plane passing through the point $P(x_0, y_0, z_0)$ has closed form

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

Let's do some algebraic manipulation here. Divide the equation by C and let $a = -A/C$ and $b = -B/C$. This yields a new equation

$$z - z_0 = a(x - x_0) + b(y - y_0)$$

If the above equation represents the tangent plane at P, then the intersection with the plane $y = y_0$ gives the tangent line T_1. Setting $y = y_0$ causes the $b(y - y_0)$ term to zero out, leaving

$$z - z_0 = a(x - x_0)$$

and this is the point-slope form equation of a line with slope a. But we know from section 3.3 this is just equal to $f_x(x_0, y_0)$. Therefore, $a = f_x(x_0, y_0)$. Similarly, setting $x = x_0$, it follows that $z - z_0 = b(y - y_0)$, which represents T_2, where $b = f_y(x_0, y_0)$. This brings us to the following definition.

Definition 3.4.1. *Suppose f has continuous partial derivatives. The* ***equation of the tangent plane*** *to the surface $z = f(x, y)$ at the point $P(x_0, y_0, z_0)$ is*

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)$$

Problem 3.4.1. *Find the tangent plane to the elliptic paraboloid $z = 2x^2 + y^2$ at the point $(1, 1, 3)$.*

Solution: Let $z = f(x, y) = 2x^2 + y^2$. We have

$$\begin{aligned} f_x(x, y) &= 4x \quad & f_y(x, y) &= 2y \\ f_x(1, 1) &= 4 \quad & f_y(1, 1) &= 2 \end{aligned}$$

Also, $z_0 = f(1, 1) = 3$, so the equation of the tangent plane is

$$z - 3 = 4(x - 1) + 2(y - 1) \implies \boxed{z = 4x + 2y - 3}$$

Linear Approximations In problem 3.4.1 we found that an equation of the tangent plane to the graph of $f(x, y) = 2x^2 + y^2$ at $(1, 1, 3)$ is $z = 4x + 2y - 3$. Therefore, the linear function

$$L(x, y) = 4x + 2y - 3$$

is a good approximation of the function value when (x, y) is in the vicinity of $(1, 1)$. The function L is called the *linearization* of f at $(1, 1)$ and $f(x, y) \approx 4x + 2y - 3$ is called the *linear approximation* of f at $(1, 1)$. For instance, at the point $(1.1, 0.95)$, the linear approximation is

$$f(1.1, 0.95) \approx 4(1.1) + 2(0.95) - 3 = 3.3$$

which is pretty close to the true value of $f(1.1, 0.95) = 2(1.1)^2 + (0.95)^2 = 3.3225$. But if we chose a point relatively far from $(1, 1)$, e.g. $(2, 3)$, our approximation is no longer suitable. For instance, $L(2, 3) = 11$ whereas $f(2, 3) = 17$.

We already know that the equation of the tangent plane to the graph of $f(x, y)$ at the point $(a, b, f(a, b))$ is

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

The linear function $L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$ is the linearization of f at (a, b) and the approximation $f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$ is the linear approximation of the tangent plane approximation of f at (a, b).

Now suppose for $z = f(x, y)$, x changes by an amount Δx from a to $a + \Delta x$ and y changes by an amount Δy from b to $b + \Delta y$. So the corresponding **increment** of z is

$$\Delta z = f(a + \Delta x, b + \Delta y) - f(a, b)$$

We now discuss the differentiability of multivariable functions. For a function $z = f(x, y)$, if the partial derivatives f_x and f_y exist in the vicinity of (a, b) and are continuous at (a, b), then f is differentiable at said point.

Problem 3.4.2. *Show that $f(x, y) = xe^{xy}$ is differentiable at $(1, 0)$ and find the linearization at that point. Then use it to approximate $f(1.1, -0.1)$.*

Solution: The partial derivatives are

$$\begin{aligned} f_x(x, y) &= e^{xy} + xye^{xy} \qquad & f_y(x, y) &= x^2 e^{xy} \\ f_x(1, 0) &= 1 & f_y(1, 0) &= 1 \end{aligned}$$

We observe that f_x and f_y are both continuous functions, so f is differentiable. The linearization is

$$\begin{aligned} L(x, y) &= f(1, 0) + f_x(1, 0)(x - 1) + f_y(1, 0)(y - 0) \\ &= 1 + 1(x - 1) + 1 \cdot y = x + y \end{aligned}$$

Therefore, the corresponding linear approximation is $xe^{xy} \approx x + y$, so we have $f(1.1, -0.1) \approx 1.1 - 0.1 = \boxed{1}$.

Differentials For single-variable differential functions $y = f(x)$, we defined the differential dx to be an independent variable, i.e. the differential of y became

$$dy = f'(x)\, dx$$

Consider the increment Δy and the differential dy, i.e. Δy represents the change in height of the curve $y = f(x)$ and dy represents the change in height of the tangent line when x changes by a small amount Δx.

Definition 3.4.2. *For a differentiable function of two variables $z = f(x, y)$, the* ***differentials*** *dx and dy are independent variables, and the differential dz, also known as the* ***total differential****, is*

$$dz = f_x(x, y)\, dx + f_y(x, y)\, dy = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy$$

Keep in mind that some persons in academia may use df in place of dz. Let $dx = \Delta x = x - a$ and $dy = \Delta y = y - b$. The differential of z becomes

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

so the linear approximation can be written as

$$f(x, y) \approx f(a, b) + dz$$

Problem 3.4.3. *Answer the following questions.*
(a) If $z = f(x, y) = x^2 + 3xy - y^2$, find the differential dz.
(b) If x changes from 2 to 2.05 and y changes from 3 to 2.96, compare the values of Δz and dz.

Solution to part a: The differential dz is given by Definition 3.4.2:

$$dz = \frac{\partial z}{\partial x}\, dx + \frac{\partial z}{\partial y}\, dy = \boxed{(2x + 3y)\, dx + (3x - 2y)\, dy}$$

Solution to part b: Let $x = 2$, $dx = \Delta x = 0.05$, $y = 3$, and $dy = \Delta y = -0.04$. Then

$$dz = [2(2) + 3(3)](0.05) + [3(2) - 2(3)](-0.04) = 0.65$$

Now, the increment of z is

$$\begin{aligned}\Delta z &= f(2.05, 2.96) - f(2, 3)\\ &= [(2.05)^2 + 3(2.05)(2.96) - (2.96)^2] - [2^2 + 3(2)(3) - 3^2]\\ &= 0.6449\end{aligned}$$

Clearly, $\boxed{\Delta z \approx dz}$, but it is easier to calculate the latter.

Problem 3.4.4. *The base radius and height of a right circular cone are measured as 10 cm and 25 cm, respectively, with possible measurement error of ε cm in each.*
(a) Use differentials to estimate the maximum error in the calculated volume of the cone.
(b) What is the estimated maximum error in volume if the radius and height are measured with errors up to 0.1 cm?

Solution to part a: A cone with base radius r and height h has volume $V = \pi r^2 h/3$. The differential of V is

$$dV = \frac{\partial V}{\partial r}\, dr + \frac{\partial V}{\partial h}\, dh = \frac{2\pi rh}{3}\, dr + \frac{\pi r^2}{3}\, dh$$

Since each error is at most ε, we have $|\Delta r| \leq \varepsilon$ and $|\Delta h| \leq \varepsilon$. To estimate the largest error in calculating the volume, we take the largest error in measuring r and h. This results in $dr = \varepsilon$ and $dh = \varepsilon$. Since $r = 10$ and $h = 25$, we get

$$\Delta V \approx dV = \frac{500\pi}{3}\varepsilon + \frac{100\pi}{3}\varepsilon = 200\pi\varepsilon$$

Thus the maximum error in the calculated volume is about $\boxed{200\pi\varepsilon\,\text{cm}^3}$.

Solution to part b: If the largest error in each measurement is $\varepsilon = 0.1$ cm, then $dV = 200\pi(0.1) \approx 63$, so the estimated error in volume is about $\boxed{63\,\text{cm}^3}$.

Functions of Three or More Variables Everything we have discussed so far, e.g. linearization, differentiability of functions, and differentials can be generalized to functions of three or more variables. The linear approximation can be represented by

$$f(x, y, z) \approx f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c)$$

and the linearization $L(x, y, z)$ is illustrated by the right side of the expression.

If $w = f(z, y, z)$, the **increment** of w is

$$\Delta w = f(x + \Delta x, y + \Delta y, z + \Delta z) - f(x, y, z)$$

the differential dw is defined in terms of the differentials dx, dy, and dz, which are independent variables. Then

$$dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy + \frac{\partial w}{\partial z}\,dz$$

Problem 3.4.5. *The dimensions of a rectangular box are measured to be* 75 *cm,* 60 *cm, and* 40 *cm, and each measurement is correct to within* ε *cm.*
(a) Use differentials to estimate the largest possible error when the volume of the box is calculated from these measurements.
(b) What is the estimated maximum error in the calculated volume if the measured dimensions are correct to within 0.2 *cm?*

Solution to part a: Let the dimensions of the rectangular box with x, y, and z. Then the volume of the box is $V = xyz$ and so the differential dV is

$$dV = \frac{\partial V}{\partial x}\,dx + \frac{\partial V}{\partial y}\,dy + \frac{\partial V}{\partial z}\,dz = yz\,dx + xz\,dy + xy\,dz$$

It is known that $|\Delta x| \le \varepsilon$, $|\Delta y| \le \varepsilon$, and $|\Delta z| \le \varepsilon$. To estimate the largest error in the volume, we set $dx = \varepsilon$, $dy = \varepsilon$, and $dz = \varepsilon$, along with $x = 75$, $y = 60$, and $z = 40$.

$$\Delta V \approx dV = (60)(40)\varepsilon + (75)(40)\varepsilon + (75)(60)\varepsilon = \boxed{9900\varepsilon}$$

In other words, the maximum error in the calculated volume is about 9900 times larger than the error in each individual measurement.

Solution to part b: If the largest error in each individual measurement is $\varepsilon = 0.2$ cm, then we have $dV = 9900(0.2) = 1980$, so an error of only 0.2 cm in measuring each dimension could lead to an error of approximately $\boxed{1980\,\text{cm}^3}$ in the calculated volume.

3.5 Chain Rule

By now we all understand how to apply the chain rule for single-variable functions, with sufficient experience in Calculus I and prior practice exercises in this book. Now, we will extend the chain rule to more complicated scenarios. But first, we will review the notation for the chain rule of multivariable functions.

Most people are already familiar with the following notation:

$$F(x) = f(g(x)) \qquad F'(x) = f'(g(x)) \cdot g'(x)$$

However, there is also an alternate notation that is not used much in Calculus I and/or Calculus II. Here it is:

$$y = f(x) \quad \text{and} \quad x = g(t) \quad \text{then} \quad \frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}$$

Note that the derivative $\frac{dy}{dt}$ simply refers to the fact that if we were to plug in x (a function of t) then y would automatically be a function of t. What may help in remembering this form of the chain rule is acknowledging that the two derivatives on the right side are basically fractions, with the dx terms canceling out to get equivalent terms on both sides.

Chain Rule Equations As with many topics in multivariable calculus, there are several different formulas that depend on the number of variables we are dealing with. Let's start with functions of two variables, i.e. $z = f(x, y)$. There are still many different possibilities we can look at.

Case 1. Let $z = f(x, y)$, where $x = g(t)$ and $y = h(t)$. How can we compute $\frac{dz}{dt}$?

Well, this is pretty analogous to the standard chain rule in Calculus I. In this case, we simply compute an ordinary derivative because z is a function of t provided that we substitute for x and y. We have

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

So, basically what we're doing here is differentiating f with respect to each variable in it and then multiplying each of these by the derivative of said variable with respect to t. The final step is to then add everything up.

Problem 3.5.1. *If* $z = xe^{xy}$, $x = t^2$, *and* $y = \frac{1}{t}$, *compute* $\frac{dz}{dt}$.

Solution: We first apply the formula directly:

$$\begin{aligned}\frac{dz}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\ &= (e^{xy} + yxe^{xy})(2t) + x^2e^{xy}\left(-\frac{1}{t^2}\right) \\ &= 2t(e^{xy} + yxe^{xy}) - \frac{1}{t^2}x^2e^{xy}\end{aligned}$$

Technically, we computed the derivative, but we should substitute the functions of t for x and y because the derivative expression already contains t terms. This gives us

$$\frac{dz}{dt} = 2t(e^t + te^t) - \frac{1}{t^2} \cdot t^4e^t = \boxed{2te^t + t^2e^t}$$

Before we move onto the second case, consider the special scenario where

$$z = f(x, y) \qquad y = g(x)$$

The Chain Rule for $\frac{dz}{dx}$ becomes

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

where $\frac{dx}{dx} = \frac{d}{dx}(x) = 1$.

Problem 3.5.2. *If $z = x\ln(xy) + y^3$ and $y = \cos(x^2+1)$, find $\frac{dz}{dx}$.*

Solution: We simply apply the above formula, but let's be careful with Product Rule.

$$\begin{aligned}\frac{dz}{dx} &= \left(\ln(xy) + x\frac{y}{xy}\right) + \left(x\frac{x}{xy} + 3y^2\right)\left(-2x\sin(x^2+1)\right)\\ &= \ln\left(x\cos(x^2+1)\right) + 1 - 2x\sin(x^2+1)\left(\frac{x}{\cos(x^2+1)} + 3\cos^2(x^2+1)\right)\\ &= \boxed{\ln\left(x\cos(x^2+1)\right) + 1 - 2x^2\tan(x^2+1) - 6x\sin(x^2+1)\cos^2(x^2+1)}\end{aligned}$$

Case 2. We have $z = f(x,y)$ with $x = g(s,t)$ and $y = h(s,t)$. We wish to compute the partial derivatives $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.

In this case we would have to substitute in for x and y and obtain z as a function of s and t so for both scenarios, here is the Chain Rule:

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

Here is a quick example.

Problem 3.5.3. *Suppose $z = e^{2r}\sin(3\theta)$, $r = st - t^2$, and $\theta = \sqrt{s^2+t^2}$. Calculate the partial derivatives $\frac{\partial z}{\partial s}$ and $\frac{\partial z}{\partial t}$.*

Solution: The Chain Rule for $\frac{\partial z}{\partial s}$ is

$$\begin{aligned}\frac{\partial z}{\partial s} &= \left(2e^{2r}\sin(3\theta)\right)(t) + \left(3e^{2r}\cos(3\theta)\right)\frac{s}{\sqrt{s^2+t^2}}\\ &= \boxed{t\left(2e^{2(st-t^2)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3se^{2(st-t^2)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}}\end{aligned}$$

Now, for $\frac{\partial z}{\partial t}$, we have

$$\begin{aligned}\frac{\partial z}{\partial t} &= \left(2e^{2r}\sin(3\theta)\right)(s-2t) + \left(3e^{2r}\cos(3\theta)\right)\frac{t}{\sqrt{s^2+t^2}}\\ &= \boxed{(s-2t)\left(2e^{2(st-t^2)}\sin\left(3\sqrt{s^2+t^2}\right)\right) + \frac{3te^{2(st-t^2)}\cos\left(3\sqrt{s^2+t^2}\right)}{\sqrt{s^2+t^2}}}\end{aligned}$$

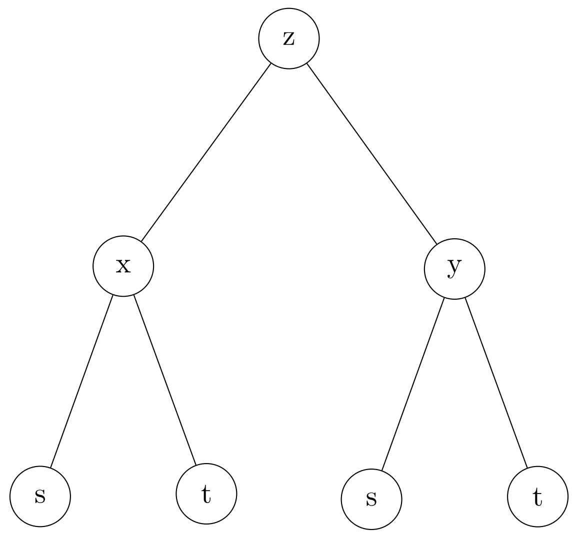
Now let's generalize the Chain Rule. Suppose that z is a function of n variables, namely, $x_1, x_2, \ldots, x_n$ and each variable is a function of m variables, $t_1, t_2, \ldots, t_m$. For any variable t_i, for $i = 1, 2, \ldots, n$ we have the following:

$$\boxed{\frac{\partial z}{\partial t_i} = \frac{\partial z}{\partial x_1}\frac{\partial x_1}{\partial t_i} + \frac{\partial z}{\partial x_2}\frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial z}{\partial x_n}\frac{\partial x_n}{\partial t_i}}$$

This looks like a lot, and honestly, it is. Our focus should not be on trying to mindlessly memorize this. A more systematic approach is to construct a **tree diagram**. This will give us the correct chain rule for any situation. Let's revisit the chain rule formula for the case of $\frac{\partial z}{\partial s}$, where $z = f(x, y)$, $x = g(s, t)$, and $y = h(s, t)$. We know

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

And here is the tree diagram.



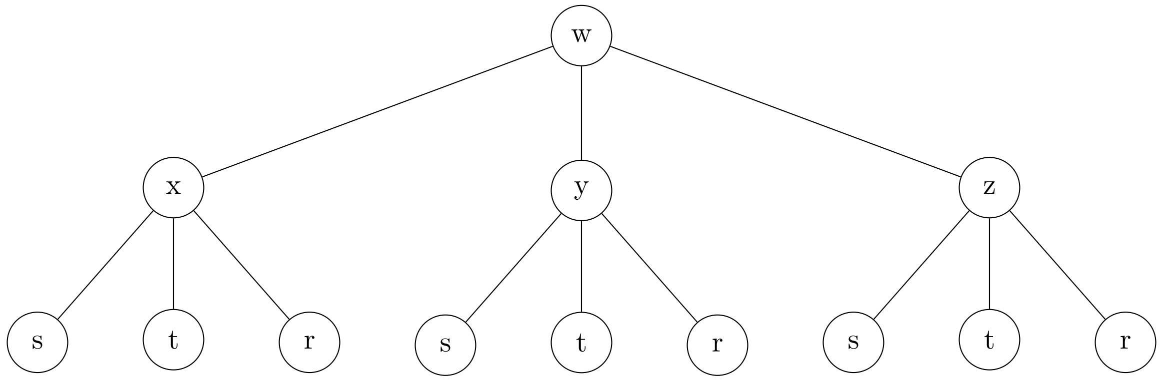
To interpret this tree, let us start at the top (root) with the original function and branch out from that point. The first set of branches represents the variables of the function. From these components, a further set of branches provides the variables that both x and y depend on. We connect each letter with a line and each line represents a partial derivative as shown. Note that the letter in the numerator of the partial derivative is the upper "node" of the tree and the letter in the denominator of the partial derivative is the lower "node" of the tree.

Specifically, to obtain the Chain Rule, we start at the bottom and for each branch ending with the variable we wish to take the derivative with respect to, we move up the tree until we hit the top multiplying the derivatives that we see along that set of branches. Once we've done this for each branch that ends at s, we add everything up to obtain the answer in that scenario.

Remark. Sometimes when trees get too large or messy, we do not explicitly include the derivatives. If you understand which derivative belongs on each branch, you'll be fine.

Problem 3.5.4. *Let $w = f(x, y, z)$, $x = g_1(s, t, r)$, $y = g_2(s, t, r)$, and $z = g_3(s, t, r)$. Find the partial derivative $\frac{\partial w}{\partial r}$.*

Solution: We will construct a tree diagram.



It is trivial to see that the requested partial derivative will be

$$\frac{\partial w}{\partial r} = \boxed{\frac{\partial f}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial r} + \frac{\partial f}{\partial z}\frac{\partial z}{\partial r}}$$

Implicit Differentiation The last topic of this section will be a revisit to implicit differentiation, a technique we saw in Calculus I. When we apply multivariable chain rule, it becomes very easy to perform implicit differentiation in higher-order calculus problems.

Consider a function of the form $F(x, y) = 0$. If it is not initially in this form, we can simply rearrange the terms. Here, y is defined implicitly as a function of x, i.e. $y = y(x)$. In Calculus I, we would be asked to compute $\frac{dy}{dx}$ and if we recall, it could be a pretty messy process. However, using the Chain Rule defined in this section, we can make implicit differentiation a much more pleasant process. Let's start by differentiating both sides of the equation with respect to x. To do that, we apply the chain rule on the left and the right side will intuitively differentiate to zero:

$$F_x + F_y\frac{dy}{dx} = 0 \Longrightarrow \frac{dy}{dx} = -\frac{F_x}{F_y}$$

As shown above, we simply need to solve for $\frac{dy}{dx}$, assuming $F_y \neq 0$ and now we have a very nice formula for implicit differentiation.

Problem 3.5.5. *Find $\frac{dy}{dx}$ for the implicit function $x\cos(3y) + x^3y^5 = 3x - e^{xy}$.*

Solution: We need to get a zero on the right side of the equation, and this is simple:

$$x\cos(3y) + x^3y^5 - 3x + e^{xy} = 0$$

We have $F(x, y) = x\cos(3y) + x^3y^5 - 3x + e^{xy}$ so we simply use

$$\frac{dy}{dx} = -\frac{F_x}{F_y}$$

to obtain the answer. We have

$$F_x = -3x\sin(3y) + 5x^3y^4 + xe^{xy} \quad \text{and} \quad F_y = \cos(3y) + 3x^2y^5 - 3 + ye^{xy}$$

so the final answer is

$$\frac{dy}{dx} = \boxed{-\frac{\cos(3y) + 3x^2y^5 - 3 + ye^{xy}}{-3x\sin(3y) + 5x^3y^4 + xe^{xy}}}$$

We can also handle more complicated functions, i.e. those of the form $F(x, y, z) = 0$. Let's assume $z = f(x, y)$ and that we wish to determine partial derivatives $\frac{\partial z}{\partial x}$ and/or $\frac{\partial z}{\partial y}$.

For $\frac{\partial z}{\partial x}$ we need to differentiate both sides with respect to x, and remember that we treat y as a constant.

$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial x} \ + \ \frac{\partial F}{\partial y}\frac{\partial y}{\partial x} \ + \ \frac{\partial F}{\partial z}\frac{\partial z}{\partial x} \ = \ 0$$

Obviously, $\frac{\partial x}{\partial x} = 1$ and with y as a constant, we have $\frac{\partial y}{\partial x} = 0$. Plugging in and solving for $\frac{\partial z}{\partial x}$, we obtain

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z}$$

If we adjust the logic slightly, we can also show

$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z}$$

Let's end the section with one final practice problem.

Problem 3.5.6. *For the implicit function $x^2\sin(2y - 5z) = 1 + y\cos(6zx)$, find the first partials z_x and z_y.*

Solution: Let's first get everything on one side:

$$F(x, y, z) = x^2\sin(2y - 5z) - 1 - y\cos(6zx) = 0$$

All we need to do is apply the formulas we derived above.

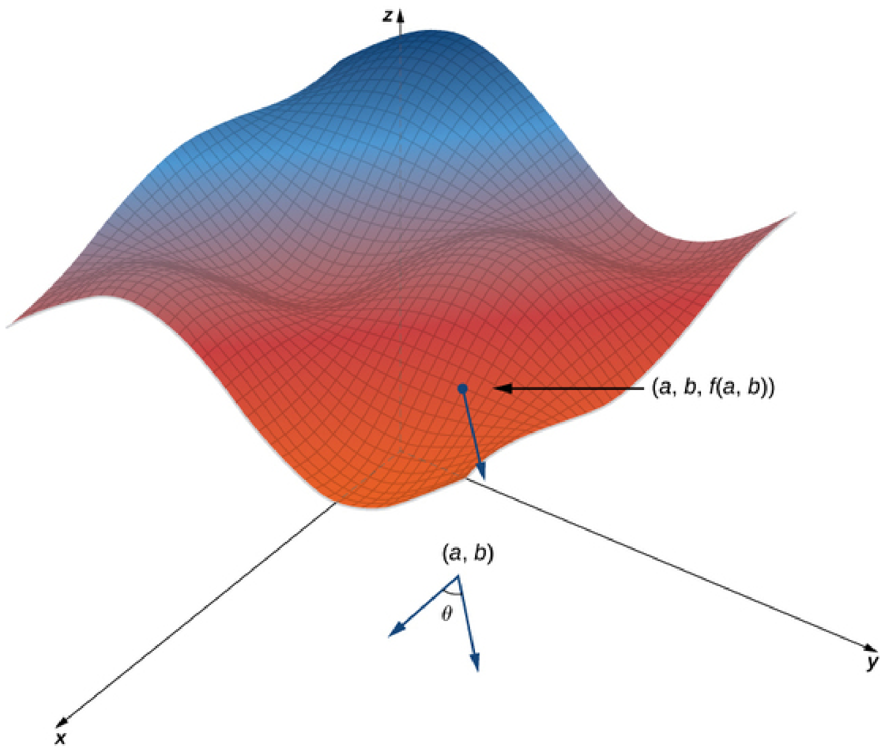
$$z_x = \frac{\partial z}{\partial x} = \boxed{-\frac{2x\sin(2y - 5z) + 6yz\sin(6zx)}{-5x^2\cos(2y - 5z) + 6yx\sin(6zx)}}$$

$$z_y = \frac{\partial z}{\partial y} = \boxed{-\frac{2x^2\cos(2y - 5z) - \cos(6zx)}{-5x^2\cos(2y - 5z) + 6yx\sin(6zx)}}$$

3.6 The Gradient Vector and Directional Derivatives

Suppose $z = f(x, y)$ is a function of two variables. It has two first partial derivatives, $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$. These derivatives correspond to each of the independent variables and can be interpreted as instantaneous rates of change (that is, as slopes of a tangent line). Specifically, $\frac{\partial z}{\partial x}$ represents the slope of a tangent line passing through a given point on the surface $z = f(x, y)$, assuming the tangent line is parallel to the x-axis. Similarly, $\frac{\partial z}{\partial y}$ represents the slope of that tangent line parallel to the y-axis. In this section, we consider the possibility of a tangent line parallel to neither axis.

Directional Derivatives A surface has equation $z = f(x, y)$. Given a point (a, b) in the domain of f, we choose a direction to travel from that point. The direction angle, denoted by θ, measures the indicated direction, taken relative to the counterclockwise direction in the xy-plane. At the positive x-axis, the starting value is $\theta = 0$. Suppose we travel some distance h; the direction in which we travel is given by the unit vector $\mathbf{u} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$. Thus, the z-coordinate of the second point on the graph is $z = f(a + h\cos\theta, b + h\sin\theta)$.



We can calculate the slope of the secant line by dividing the difference in z-values by the length of the line segment connecting the two points in the domain. The slope is

$$m = \frac{f(a + h\cos\theta, b + h\sin\theta) - f(a, b)}{h}$$

Intuitively, we can find the slope of the tangent line in the same direction, if we take the limit as $h \to 0$.

Definition 3.6.1. *Let $z = f(x, y)$ be a function of two variables with a domain of D. Let $(a, b) \in D$ and define a unit vector $\mathbf{u} = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$. The **directional derivative** of f in the direction of $\mathbf{u}$ is given by*

$$D_{\mathbf{u}} f(a, b) = \lim_{h \to 0} \frac{f(a + h\cos\theta, b + h\sin\theta) - f(a, b)}{h}$$

assuming that the limit exists.

Keep in mind that the point (a, b) is randomly chosen from the domain D, so we can use Definition 3.6.1 to find the directional derivative as a function of x and y.

$$D_{\mathbf{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + h\cos\theta, y + h\sin\theta) - f(x, y)}{h}$$

Problem 3.6.1. *Let $\theta = \arccos(3/5)$. Calculate the directional derivative $D_{\mathbf{u}}f(x, y)$ of $f(x, y) = x^2 - xy + 3y^2$ in the direction of $\mathbf{u} = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$. at the point $(-1, 2)$.*

Solution: We know that $\cos\theta = \frac{3}{5}$ and θ is acute, so it follows that

$$\sin\theta = \sqrt{1 - \left(\frac{3}{5}\right)^2} = \sqrt{\frac{16}{25}} = \frac{4}{5}$$

Taking $f(x, y) = x^2 - xy + 3y^2$, we first calculate $f(x + h\cos\theta, y + h\sin\theta)$.

$$\begin{aligned} f(x + h\cos\theta, y + h\sin\theta) &= (x + h\cos\theta)^2 - (x + h\cos\theta)(y + h\sin\theta) + 3(y + h\sin\theta)^2 \\ &= x^2 + 2xh\cos\theta + h^2\cos^2\theta - xy - xh\sin\theta - yh\cos\theta \\ &- h^2\sin\theta\cos\theta + 3y^2 + 6yh\sin\theta + 3h^2\sin^2\theta \\ &= x^2 + 2xh\left(\frac{3}{5}\right) + \frac{9h^2}{25} - \frac{4xh}{5} - \frac{3yh}{5} - \frac{12h^2}{25} + 3y^2 + \frac{6yh}{5} + \frac{21yh}{5} \\ &= x^2 - xy + 3y^2 + \frac{2xh}{5} + \frac{9h^2}{5} + \frac{21yh}{5} \end{aligned}$$

We take $a = x$ and $b = y$ according to Definition 3.6.1 to obtain

$$\begin{aligned} D_{\mathbf{u}}f(x, y) &= \lim_{h\to 0} \frac{f(x + h\cos\theta, y + h\sin\theta) - f(x, y)}{h} \\ &= \lim_{h\to 0} \frac{(x^2 - xy + 3y^2 + \frac{2xh}{5} + \frac{9h^2}{5} + \frac{21yh}{5}) - (x^2 - xy + 3y^2)}{h} \\ &= \lim_{h\to 0} \frac{\frac{2xh}{5} + \frac{9h^2}{5} + \frac{21yh}{5}}{h} \\ &= \lim_{h\to 0} \left(\frac{2x}{5} + \frac{9h}{5} + \frac{21y}{5}\right) \\ &= \frac{2x + 21y}{5} \end{aligned}$$

Finally, we substitute $x = -1$ and $y = 2$ to calculate $D_{\mathbf{u}}f(-1, 2)$:

$$D_{\mathbf{u}}f(-1, 2) = \frac{2(-1) + 21(2)}{5} = \frac{-2 + 42}{5} = \boxed{8}$$

An easier approach to calculating directional derivatives that involves partial derivatives is outlined in the following theorem.

Theorem 3.6.1. *Let $z = f(x, y)$ be a function of two variables x and y. Assume that f has continuous partial derivatives f_x and f_y. The directional derivative of f in the direction of $\mathbf{u} = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$ is given by*

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta$$

Proof. Applying Definition 3.6.1, the directional derivative of f in the direction of the unit vector $\mathbf{u} = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$ at a point (x_0, y_0) in the domain of f can be written as

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}$$

Suppose $x = x_0 + t\cos\theta$, $y = y_0 + t\sin\theta$, and define $g(t) = f(x, y)$. Since f_x and f_y are both continuous and existent, we can use the chain rule of two variables to calculate $g'(t)$.

$$g'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta$$

Also, if $t = 0$, then $x = x_0$ and $y = y_0$ so

$$g'(0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta$$

so using the definition of the derivative, we know

$$g'(0) = \lim_{t\to 0}\frac{g(t) - g(0)}{t} = \lim_{t\to 0}\frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t}$$

Thus, $D_{\mathbf{u}}f(x_0, y_0) = f_x(x_0, y_0)\cos\theta + f_y(x_0, y_0)\sin\theta$. Since the point (x_0, y_0) is an arbitrary point from the domain of f, this formula holds true for all such points in the domain where f_x and f_y exist. So we have

$$D_{\mathbf{u}}f(x, y) = f_x(x, y)\cos\theta + f_y(x, y)\sin\theta$$

and the proof is complete. □

Problem 3.6.2. *Find the directional derivative* $D_{\mathbf{u}}f(x, y)$ *of* $f(x, y) = 3x^2y - 4xy^3 + 3y^2 - 4x$ *in the direction of the unit vector* $\mathbf{u} = \left(\cos\frac{\pi}{3}\right)\mathbf{i} + \left(\sin\frac{\pi}{3}\right)\mathbf{j}$. *What is* $D_{\mathbf{u}}f(3, 4)$*?*

Solution: The first step here is to calculate the partial derivatives, $f_x(x, y)$ and $f_y(x, y)$.

$$f_x(x, y) = 6xy - 4y^3 - 4$$
$$f_y(x, y) = 3x^2 - 12xy^2 + 6y$$

Also, $\theta = \frac{\pi}{3}$, so the directional derivative at any point (x, y) is given by

$$D_{\mathbf{u}}f(x, y) = \frac{(6xy - 4y^3 - 4)(1)}{2} + \frac{(3x^2 - 12xy^2 + 6y)\sqrt{3}}{2}$$

and so the directional derivative at the point $(3, 4)$ is

$$D_{\mathbf{u}}f(3, 4) = \frac{72 - 256 - 4}{2} + \frac{(27 - 572 + 24)\sqrt{3}}{2} = \boxed{-94 - \frac{525\sqrt{3}}{2}}$$

Gradient Notice that the expression $f_x(x, y)\cos\theta + f_y(x, y)\sin\theta$ can be written as the dot product of two vectors. The first vector is defined as $\nabla f(x, y) = f_x(x, y)\,\mathbf{i} + f_y(x, y)\,\mathbf{j}$ and the second is defined as $\mathbf{u} = (\cos\theta)\,\mathbf{i} + (\sin\theta)\,\mathbf{j}$. Thus, we can write the directional derivative as

$$D_{\mathbf{u}}f(x, y) = \nabla f(x, y)\cdot\mathbf{u}$$

The vector ∇f is called the *gradient* of the function $f(x, y)$, and it is read as "**del** f."

Problem 3.6.3. *Find the gradient* $\nabla f(x, y)$ *for the function* $f(x, y) = \sin 3x\cos 3y$.

Solution: We just need to find the partials f_x and f_y. These are:

$$f_x(x, y) = 3\cos 3x\cos 3y$$
$$f_y(x, y) = -3\sin 3x\sin 3y$$

so the gradient is

$$\begin{aligned}\nabla f(x,y) &= f_x(x,y)\,\mathbf{i} + f_y(x,y)\,\mathbf{j}\\ &= \boxed{(3\cos 3x\cos 3y)\,\mathbf{i} - (3\sin 3x\sin 3y)\,\mathbf{j}}\end{aligned}$$

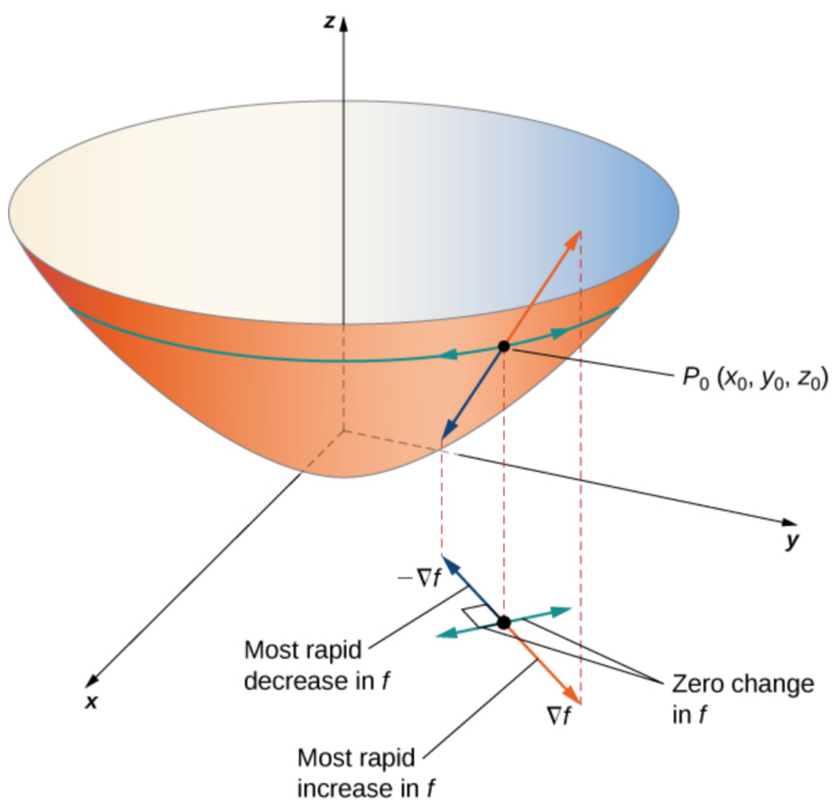
The gradient vector has some interesting properties. These are based on the properties of the dot product. Specifically, if the angle between two vectors $\mathbf{a}$ and $\mathbf{b}$ is φ, then $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\varphi$. Thus, if the angle between $\nabla f(x_0, y_0)$ and $\mathbf{u}$ is φ, we have

$$D_{\mathbf{u}}f(x_0,y_0) = \nabla f(x_0,y_0)\cdot\mathbf{u} = |\nabla f(x_0,y_0)||\mathbf{u}|\cos\varphi = |\nabla f(x_0,y_0)|\cos\varphi$$

Notice that the $|\mathbf{u}|$ disappears because $\mathbf{u}$ is a unit vector with magnitude 1. Therefore, the directional derivative is equal to the magnitude of the gradient vector multiplied by $\cos\varphi$, where $-1 \leq \cos\varphi \leq 1$.

Definition 3.6.2. *Suppose the function $z = f(x,y)$ is differentiable at the point (x_0, y_0). There are three cases.*

I. If $\nabla f(x_0,y_0) = \mathbf{0}$, then $D_{\mathbf{u}}f(x_0,y_0) = 0$ for any unit vector $\mathbf{u}$.
II. If $\nabla f(x_0,y_0) \neq \mathbf{0}$, then $D_{\mathbf{u}}f(x_0,y_0)$ is maximized when $\mathbf{u}$ points in the same direction of $\nabla f(x_0,y_0)$. The maximum value is equal to $|\nabla f(x_0,y_0)|$.
III. If $\nabla f(x_0,y_0) \neq \mathbf{0}$, then $D_{\mathbf{u}}f(x_0,y_0)$ is minimized when $\mathbf{u}$ points in the opposite direction from $\nabla f(x_0,y_0)$. The minimum value is equal to $-|\nabla f(x_0,y_0)|$.



Problem 3.6.4. *Find the direction for which the directional derivative of $f(x,y) = 4x - xy + 2y^2$ at the point $(-2,3)$ is maximized. What is the maximum value?*

Solution: We need to find the gradient of f at a point (x, y) to solve this problem.

$$\begin{aligned}\nabla f(x,y) &= f_x(x,y)\,\mathbf{i} + f_y(x,y)\,\mathbf{j} \\ &= (4-y)\,\mathbf{i} + (-x+4y)\,\mathbf{j}\end{aligned}$$

We evaluate the value of the gradient at $(-2, 3)$.

$$\nabla f(-2,3) = \boxed{\mathbf{i} + 14\,\mathbf{j}}$$

To find the unit vector which travels in the same direction as $\nabla f(-2, 3)$, we normalize the gradient vector to get

$$\mathbf{u} = \frac{\nabla f(-2,3)}{|\nabla f(-2,3)|} = \frac{1}{\sqrt{197}}\,\mathbf{i} + \frac{14}{\sqrt{197}}\,\mathbf{j} = \frac{\sqrt{197}}{197}\,\mathbf{i} + \frac{14\sqrt{197}}{197}\,\mathbf{j}$$

which gives an angle of $\theta = \sin^{-1}\left(\frac{14\sqrt{197}}{197}\right) \approx 1.499\,\text{rad}$. Thus, the maximum value of the directional derivative is $|\nabla f(-2,3)| = \boxed{\sqrt{197}}$.

Gradients and Level Curves Recall from Chapter 2 that if a curve is defined parametrically by the pair $x(t)$ and $y(t)$, then the vector $x'(t)\,\mathbf{i} + y'(t)\,\mathbf{j}$ is tangent to the curve for every parameter value t in the domain. Now suppose $z = f(x, y)$ is differentiable for x and y, and (x_0, y_0) is in its domain. Also, let $x_0 = x(t_0)$ and $y_0 = y(t_0)$ for some t and consider the level curve $f(x, y) = k$. Call $g(t) = f(x(t), y(t))$ and the derivative $g'(t)$ on the level curve is given by the Chain Rule:

$$g'(t) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t)$$

But $g'(t) = 0$ because $g(t) = k$, a constant. We have

$$f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) = 0$$

But this is just the dot product

$$\nabla f(x,y) \cdot \langle x'(t), y'(t)\rangle = 0$$

which implies these two vectors are orthogonal. However, the second vector is tangent to the level curve, which implies the gradient must be normal to the level curve, which gives rise to the following definition.

Definition 3.6.3. *Suppose the function $z = f(x, y)$ has continuous first-order partial derivatives in an open disk centered at a point (x_0, y_0). If $\nabla f(x, y) \neq \mathbf{0}$, then the gradient is normal to the level curve of f at (x_0, y_0).*

Problem 3.6.5. *For $f(x, y) = 2x^2 - 3xy + 8y^2 + 2x - 4y + 4$, find a tangent vector to the level curve at the point $(-2, 1)$.*

Solution: We first determine the gradient vector $\nabla f(x, y)$.

$$f_x(x,y) = 4x - 3y + 2 \quad f_y = -3x + 16y - 4 \Longrightarrow \nabla f(x,y) = (4x - 3y + 2)\,\mathbf{i} + (-3x + 16y - 4)\,\mathbf{j}$$

At the point $(-2, 1)$, the gradient vector is

$$\nabla f(-2,1) = (4(-2) - 3(1) + 2)\,\mathbf{i} + (-3(-2) + 16(1) - 4)\,\mathbf{j} = -9\,\mathbf{i} + 18\,\mathbf{j}$$

This gradient vector is orthogonal to the level curve at the point $(-2, 1)$. We can obtain a tangent vector by reversing the components and multiplying either by -1. (this should bring back memories from perpendicular slope in Algebra I!). For example, a valid tangent vector is $\boxed{-18\,\mathbf{i} - 9\,\mathbf{j}}$.

Three Dimensional Gradients and Directional Derivatives We can take our discussion on gradients and directional derivatives in two dimensions and extend to a third dimension. For instance, if $f(x, y, z)$ is a function of three variables, where f_x, f_y, and f_z all exist, then the gradient $\nabla f(x, y, z)$ is given by

$$\nabla f(x, y, z) = f_x(x, y, z)\,\mathbf{i} + f_y(x, y, z)\,\mathbf{j} + f_z(x, y, z)\,\mathbf{k}$$

Problem 3.6.6. *Find the gradient of the function* $f(x, y) = 5x^2 - 2xy + y^2 - 4yz + z^2 + 3xz$.

Solution: We just need to calculate the partial derivatives f_x, f_y, and f_z and then apply the definition of the gradient vector. We have

$$f_x = 10x - 2y + 3z \quad f_y = -2x + 2y - 4z \quad f_z = 3x - 4y + 2z$$

Thus the gradient vector is

$$\begin{aligned} \nabla f(x, y, z) &= f_x(x, y, z)\,\mathbf{i} + f_y(x, y, z)\,\mathbf{j} + f_z(x, y, z)\,\mathbf{k} \\ &= \boxed{(10x - 2y + 3z)\,\mathbf{i} + (-2x + 2y - 4z)\,\mathbf{j} + (3x - 4y + 2z)\,\mathbf{k}} \end{aligned}$$

As expected, we can generalize the directional derivative to functions of three variables. To do this, we need to have a vector with three components. This vector is called a unit vector, and the components are called directional cosines (recall from section 1.3). Initially fix a three-dimensional unit vector $\mathbf{u}$ at the origin; it forms three different angles with the positive x-, y-, and z-axes. Let's call them α, β, and γ, respectively. Then, intuitively, the directional cosines are given by $\cos\alpha$, $\cos\beta$, and $\cos\gamma$, and they form the components of $\mathbf{u}$. Additionally, because $\mathbf{u}$ is a unit vector, we must have $\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$.

Definition 3.6.4. *Suppose $w = f(x, y, z)$ is a function of three variables defined on a domain D. Choose an arbitrary point $(x_0, y_0, z_0) \in D$ and let $\mathbf{u} = \cos\alpha\,\mathbf{i} + \cos\beta\,\mathbf{j} + \cos\gamma\,\mathbf{k}$ be a unit vector. The directional derivative of f in the direction of $\mathbf{u}$ is given by*

$$D_{\mathbf{u}} f(x_0, y_0, z_0) = \lim_{t \to 0} \frac{f(x_0 + t\cos\alpha, y_0 + t\cos\beta, z_0 + t\cos\gamma) - f(x_0, y_0, z_0)}{t}$$

assuming that the limit exists.

Using similar rationale as in Theorem 3.6.1, we can develop the formula for the directional derivative of a function of three variables using the gradient vector.

$$D_{\mathbf{u}} f(x, y, z) = \nabla f(x, y, z) \cdot \mathbf{u} = f_x(x, y, z)\cos\alpha + f_y(x, y, z)\cos\beta + f_z(x, y, z)\cos\gamma$$

where the three angles α, β, and γ determine $\mathbf{u}$. In practice, however, we may not always be given a unit vector. In such cases, we take the vector given, then normalize it to obtain the unit vector in the desired direction.

Problem 3.6.7. *Calculate $D_{\mathbf{v}} f(1, -2, 3)$ in the direction of $\mathbf{v} = \langle -1, 2, 2\rangle$ for the function $f(x, y, z) = 5x^2 - 2xy + y^2 - 4yz + z^2 + 3xz$.*

Solution: The first step here is to normalize the vector $\mathbf{v}$ so we can obtain the unit vector $\mathbf{u}$. The magnitude of $\mathbf{v}$ is

$$|\mathbf{v}| = \sqrt{(-1)^2 + 2^2 + 2^2} = \sqrt{9} = 3$$

so the unit vector is $\mathbf{u} = \frac{\mathbf{v}}{|\mathbf{v}|} = -\frac{1}{3}\,\mathbf{i} + \frac{2}{3}\,\mathbf{j} + \frac{2}{3}\,\mathbf{k}$, in the same direction as $\mathbf{v}$. Thus, we have $\cos\alpha = -\frac{1}{3}$, $\cos\beta = \frac{2}{3}$, and $\cos\gamma = \frac{2}{3}$. Now, we compute the partial derivatives of f.

$$f_x(x, y, z) = 10x - 2y + 3z \quad f_y(x, y, z) = -2x + 2y - 4z \quad f_z(x, y, z) = -4y + 2z + 3x$$

and now substitute these values and the direction cosines into the formula based on the dot product of $\nabla f(x, y, z)$ and $\mathbf{u}$:

$$\begin{aligned} D_{\mathbf{v}}f(x, y, z) &= f_x(x, y, z)\cos\alpha + f_y(x, y, z)\cos\beta + f_z(x, y, z)\cos\gamma \\ &= (10x - 2y + 3z)\left(-\frac{1}{3}\right) + (-2x + 2y - 4z)\left(\frac{2}{3}\right) + (-4y + 2z + 3x)\left(\frac{2}{3}\right) \\ &= -\frac{10x}{3} + \frac{2y}{3} - \frac{3z}{3} - \frac{4x}{3} + \frac{4y}{3} - \frac{8z}{3} - \frac{8y}{3} + \frac{4z}{3} + \frac{6x}{3} \\ &= -\frac{8x}{3} - \frac{2y}{3} - \frac{7z}{3} \end{aligned}$$

Finally, to calculate $D_{\mathbf{v}}$ at $(1, -2, 3)$, we substitute $x = 1$, $y = -2$, and $z = 3$.

$$\begin{aligned} D_{\mathbf{v}}f(1, -2, 3) &= -\frac{8(1)}{3} - \frac{2(-2)}{3} - \frac{7(3)}{3} \\ &= -\frac{8}{3} + \frac{4}{3} - \frac{21}{3} \\ &= \boxed{-\frac{25}{3}} \end{aligned}$$

3.7 Finding Extreme Values

We know that in calculus, one of the main uses of ordinary derivative is to calculate minimum and maximum values, i.e. extreme values. In this section, we will explore how to use partial derivatives to locate maxima and minima of multivariable functions. For example, how can we maximize the volume of a box without a lid if we have a fixed amount of cardboard to work with? Questions like these are what we aim to find the solutions to in this section.

Definition 3.7.1. *A function of two variables has a* ***local maximum*** *at (a, b) if $f(x, y) \leq f(a, b)$ when (x, y) is near (a, b). The quantity $f(a, b)$ is called a* ***local maximum value****. On the other hand, if $f(x, y) \geq f(a, b)$ when (x, y) is near (a, b), then f has a* ***local minimum*** *at (a, b) and $f(a, b)$ is a* ***local minimum value****.*

We also know that, for single-variable functions, if f has a local maximum or minimum at c, and if $f'(c)$ exists, then $f'(c) = 0$. The following theorem states a similar result for functions of two variables.

Theorem 3.7.1. *If f has a local maximum or minimum value at (a, b), and the first-order partial derivatives of f exist there, then $f_x(a, b) = 0$ and $f_y(a, b) = 0$.*

Proof. Let $g(x) = f(x, b)$. If f has a local extreme point at (a, b), then g has a local maximum or minimum value at a, so $g'(a) = 0$. But $g'(a) = f_x(a, b)$ and so $f_x(a, b) = 0$. Similarly, if we construct a function $G(y) = f(a, y)$, we will obtain $f_y(a, b) = 0$. □

A point (a, b) is called a **critical**, or stationary point of f if $f_x(a, b) = 0$ and $f_y(a, b) = 0$, or if one of these partial derivatives does not exist. Theorem 3.7.1 states that if f has no local maximum or minimum at (a, b), then (a, b) is a critical point of f. However, *not all* critical points can give rise to maxima and minima.

Problem 3.7.1. *Find the extreme values of $f(x, y) = y^2 - x^2$.*

Solution: We compute the appropriate partial derivatives. Since $f_x = -2x$ and $f_y = 2y$, the only valid critical point is $(0, 0)$. For points on the x-axis, $y = 0$, so $f(x, y) = -x^2 < 0$ (if $x \neq 0$). However, for points on the y-axis, we have $x = 0$ so $f(x, y) = y^2 > 0$ (if $y \neq 0$). Thus, every disk with center $(0, 0)$ contains points where f takes on both positive and negative values. Thus $f(0, 0) = 0$ cannot be an extreme value for f, so f has $\boxed{\text{no extreme value}}$.

Recall that for a function of one variable, a critical number c satisfying $f'(c) = 0$ can correspond to a local maximum, a local minimum, or neither.

A similar idea applies to functions of two variables. If (a, b) is a critical point of a function f, meaning that

$$f_x(a, b) = 0 \quad \text{and} \quad f_y(a, b) = 0,$$

then $f(a, b)$ may be a local maximum, a local minimum, or neither. In the case where it is neither, the point (a, b) is called a **saddle point**. The term comes from the shape the surface can take near such a point. Although the graph near a saddle point does not have to look exactly like a saddle, the surface crosses its tangent plane at that point.

To determine whether a function has a local maximum, local minimum, or saddle point at a critical point, we use a test introduced at the end of this section. This test is analogous to the Second Derivative Test for functions of one variable.

Definition 3.7.2. *Suppose the second partial derivatives of f are continuous on a disk centered at (a, b), and $f_x(a, b) = f_y(a, b) = 0$. Let*

$$D = D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2$$

(a) If $D > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local minimum.
(b) If $D < 0$ and $f_{xx}(a, b) < 0$, then $f(a, b)$ is a local maximum.
(c) If $D < 0$, then (a, b) is a saddle point of f.

Remark 1. If $D = 0$, the test is inconclusive.
Remark 2. To remember this formula, it is better to write D as a determinant:

$$D = \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = f_{xx}f_{yy} - (f_{xy})^2$$

Problem 3.7.2. *Find the local maximum and minimum values and saddle points of $f(x, y) = x^4 + y^4 - 4xy + 1$.*

Solution: We find the first-order partial derivatives f_x and f_y.

$$f_x = 4x^3 - 4y \qquad f_y = 4y^3 - 4x$$

The critical points occur where both partials are zero.

$$x^3 - y = 0 \qquad y^3 - x = 0$$

We can substitute $y = x^3$ from the first equation into the second one. This gives

$$0 = x^9 - x = x(x^8 - 1) = x(x^4 - 1)(x^4 + 1) = x(x^2 - 1)(x^2 + 1)(x^4 + 1)$$

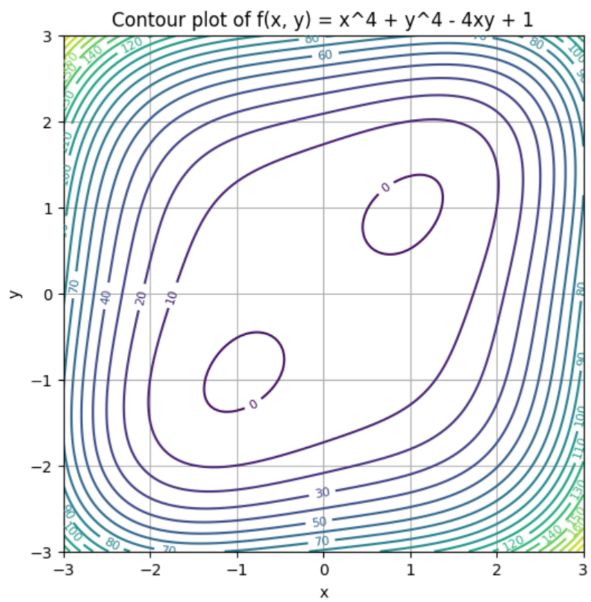
so there are three real solutions: $x = 0, 1, -1$. The three critical points of $f(x, y)$ are $(0, 0)$, $(1, 1)$, and $(-1, -1)$.

Now, we calculate the second-order partial derivatives and $D(x, y)$:

$$f_{xx} = 12x^2 \qquad f_{xy} = -4 \qquad f_{yy} = 12y^2$$

$$D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 144x^2y^2 - 16$$

Since $D(0, 0) = -16 < 0$, the point $(0, 0, f(0, 0)) = \boxed{(0, 0, 1)}$ is a saddle point. Since $D(1, 1) = 128 > 0$ and $f_{xx}(1, 1) = 12 > 0$, we see that if $\boxed{f(1, 1) = -1}$ is a local minimum. Similarly, $D(-1, -1) = 128 > 0$ and $f_{xx}(-1, -1) = 12 > 0$, so $\boxed{f(-1, -1) = -1}$ is also a local minimum. The contour graph of f is shown below, built using Python code.



Problem 3.7.3. *Find the shortest distance from the point* $(1, 0, -2)$ *to the plane* $x + 2y + z = 4$.

Solution: The distance from any point (x, y, z) to the point $(1, 0, -2)$ can be found using the Distance Formula.

$$d = \sqrt{(x-1)^2 + y^2 + (z+2)^2}$$

but if (x, y, z) lies on the plane $x + 2y + z = 4$, we can replace z with $4 - x - 2y$ so we have $d = \sqrt{(x-1)^2 + y^2 + (6 - x - 2y)^2}$. In order to minimize d, we can minimize the simpler expression for d^2.

$$d^2 = f(x, y) = (x-1)^2 + y^2 + (6 - x - 2y)^2$$

By solving the system of equations

$$f_x = 2(x-1) - 2(6 - x - 2y) = 4x + 4y - 14 = 0,$$

$$f_y = 2y - 4(6 - x - 2y) = 4x + 10y - 24 = 0,$$

we find that the only critical point is $\left(\frac{11}{6}, \frac{5}{3}\right)$. Since $f_{xx} = 4, \quad f_{xy} = 4, \quad$ and $\quad f_{yy} = 10$, the discriminant is

$$D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 4(10) - 4^2 = 40 - 16 = 24 > 0.$$

Because $D > 0$ and $f_{xx} > 0$, the Second Derivative Test tells us that

$$f\left(\frac{11}{6}, \frac{5}{3}\right)$$

is a local minimum. In fact, this minimum is absolute, since there must be a point on the given plane closest to $(1, 0, -2)$. If $x = \frac{11}{6}$ and $y = \frac{5}{3}$, then the distance is

$$d = \sqrt{(x-1)^2 + y^2 + (6 - x - 2y)^2} = \sqrt{\left(\frac{5}{6}\right)^2 + \left(\frac{5}{3}\right)^2 + \left(\frac{5}{6}\right)^2} = \frac{5}{6}\sqrt{6}$$

Therefore, the shortest distance from $(1, 0, -2)$ to the plane $x + 2y + z = 4$ is $\boxed{\frac{5}{6}\sqrt{6}}$.

Problem 3.7.4. *A rectangular box without a lid is to be made using* $12\,m^2$ *of cardboard. Find the maximum volume of such a box.*

Solution: Call x, y, and z the length, width, and height of the box, respectively, in meters. The volume of the box is

$$V = xyz$$

We can reduce V into a function of only two variables because of the constraint

$$2xz + 2yz + xy = 12$$

so $z = (12 - xy)/[2(x + y)]$, so the expression for V becomes

$$V = xy\frac{12 - xy}{2(x+y)} = \frac{12xy - x^2y^2}{2(x+y)}$$

The partial derivatives are

$$\frac{\partial V}{\partial x} = \frac{y^2(12 - 2xy - x^2)}{2(x+y)^2} \qquad \frac{\partial V}{\partial y} = \frac{x^2(12 - 2xy - y^2)}{2(x+y)^2}$$

If V is a maximum, then $\partial V/\partial x = \partial V/\partial y = 0$, but $x = 0$ or $y = 0$ implies $V = 0$. So we need to solve the equations

$$12 - 2xy - x^2 = 0 \qquad 12 - 2xy - y^2 = 0$$

These equations imply $x^2 = y^2$ and so $x = y$, because in this problem, $x > 0$ and $y > 0$. Making this substitution, we obtain $12 - 3x^2 = 0$, so $x = 2$, $y = 2$, and $z = (12 - 2 \cdot 2)/[2(2 + 2)] = 1$. We could just use the Second Derivative Test, but alternatively, the physical nature of this problem suggests that there must be an absolute maximum volume which has to occur at a critical point of V. So the maximum volume of the box is $V = 2 \cdot 2 \cdot 1 = \boxed{4\,\text{m}^3}$.

Absolute Maximum and Minimum Values Just as for single-variable functions, the absolute extrema for a function f are the largest and smallest values that the function achieves on its domain.

Definition 3.7.3. *Let (a, b) be a point in the domain D of a function f of two variables. Then $f(a, b)$ is the*

- ***absolute maximum*** *value of f on D if $f(a, b) \geq f(x, y)$ for all $(x, y) \in D$.*
- ***absolute minimum*** *value of f on D if $f(a, b) \leq f(x, y)$ for all $(x, y) \in D$.*

For single-variable functions, the Extreme Value Theorem says that if f is continuous on a closed interval $[a, b]$, then f has an absolute minimum value and an absolute maximum value. We found these by evaluating f at not just the critical points, but also the endpoints a and b of the interval.

For functions of two variables, the process is similar. Instead of a closed interval with two endpoints, we use a **closed set** in $\mathbb{R}^2$ that contains all boundary points of $f(x,y)$. For instance, the disk

$$D = \{(x,y)\,|\,x^2+y^2 \leq 1\}$$

contains all points on or inside the circle $x^2+y^2=1$. It is a closed set because it contains all the boundary points. But if even one point on the boundary curve were excluded, this set would *not* be closed. On the other hand, a **bounded set** in $\mathbb{R}^2$ is one that is contained within some disk. In other words, it is finite. So we have the counterpart of the Extreme Value Theorem in two dimensions below.

Theorem 3.7.2. *If f is continuous on a closed, bounded set $D \in \mathbb{R}^2$, then f attains an absolute maximum value $f(x_1,y_1)$ and an absolute minimum value $f(x_2,y_2)$ at some points (x_1,y_1) and (x_2,y_2) in D.*

We can also extend the strategy involving the closed interval. In order to find the absolute maximum and minimum values of a continuous function f on a closed, bounded set D, we apply three steps:

1. Find the values of f at the critical points of f in D.
2. Find the extreme values of f on the boundary of D.
3. The largest of the values from steps 1 and 2 is the absolute maximum value; the smallest of these values is the absolute minimum value.

Problem 3.7.5. *Find the absolute maximum and minimum values of the function $f(x,y) = x^2 - 2xy + 2y$ on the rectangle $D = \{(x,y)\,|\,0 \leq x \leq 3, 0 \leq y \leq 2\}$.*

Solution: Since f is a polynomial function, it is continuous on the closed, bonded rectangle D, so the Extreme Value Theorem guarantees both an absolute maximum and an absolute minimum. The first step is to find the critical points. These occur when

$$f_x = 2x - 2y = 0 \quad \text{and } f_y = -2x + 2 = 0$$

so the only critical point is $(1,1)$. This point is in D and the value of f at this point is $f(1,1) = 1$. The next step is to look at the values of f on the boundary of D, which consists of the four line segments L_1 (from $(0,0)$ to $(3,0)$), L_2 (from $(3,0)$ to $(3,2)$), L_3 (from $(3,2)$ to $(0,2)$), and L_4 ($(0,2)$ to $(0,0)$). On L_1 we have $y = 0$ and

$$f(x,0) = x^2 \qquad 0 \leq x \leq 3$$

This is an increasing function of x, so the minimum value is $f(0,0) = 0$ and the maximum value is $f(3,0) = 9$. On L_2 we have $x = 3$ and

$$f(3,y) = 9 - 4y \qquad 0 \leq y \leq 2$$

This is a decreasing function of y, so the maximum value is $f(3,0) = 9$ and the minimum value is $f(3,2) = 1$. On L_3 we have $y = 2$ and

$$f(x,2) = x^2 - 4x + 4 \qquad 0 \leq x \leq 3$$

Since $f(x,2)$ is simply $(x-2)^2$, the minimum value of this function occurs at $f(2,2) = 0$ and the maximum value is $f(0,2) = 4$. Finally, on L_4 we have $x = 0$ and

$$f(0,y) = 2y \qquad 0 \leq y \leq 2$$

with maximum value of $f(0,2) = 4$ and minimum value $f(0,0) = 0$. Thus, on the overall boundary D, the minimum value of f is $\boxed{0}$ and the maximum value of f is $\boxed{4}$.

Proof of the Second Derivative Test We will close this section by proving the Second Derivatives Test.

Proof. Let's compute the second-order directional derivative of f in the direction of the unit vector $\mathbf{u} = \langle h, k \rangle$. The first-order derivative is simple to compute.

$$D_{\mathbf{u}} f = f_x h + f_y k$$

To find the second-order directional derivative, we apply the differentiation a second time.

$$\begin{aligned} D_u^2 f = D_u(D_u f) &= \frac{\partial}{\partial x}(D_u f)\, h + \frac{\partial}{\partial y}(D_u f)\, k \\ &= (f_{xx}h + f_{yx}k)h + (f_{xy}h + f_{yy}k)k \\ &= f_{xx}h^2 + 2f_{xy}hk + f_{yy}k^2. \end{aligned}$$

If we compute the square in this expression, we yield

$$D_{\mathbf{u}}^2 f = f_{xx}\left(h + \frac{f_{xy}}{f_{xx}}\, k\right)^2 + \frac{k^2}{f_{xx}}\left(f_{xx}f_{yy} - f_{xy}^2\right)$$

We are given that $f_{xx}(a,b) > 0$ and $D(a,b) > 0$. Since f_{xx} and

$$D = f_{xx}f_{yy} - f_{xy}^2$$

are continuous functions, there exists a disk B centered at (a,b) with radius $\delta > 0$ such that $f_{xx}(x,y) > 0$ and $D(x,y) > 0$ for all $(x,y) \in B$. From the above equation, it follows that $D_{\mathbf{u}}^2 f(x,y) > 0$ for every $(x,y) \in B$. Hence, if C is the curve obtained by intersecting the graph of f with the vertical plane through $P(a,b,f(a,b))$ in the direction of $\mathbf{u}$, then C is concave upward on an interval of length 2δ. Since this holds for every direction $\mathbf{u}$, restricting (x,y) to lie in B implies that the graph of f lies above its horizontal tangent plane at P. Therefore, $f(x,y) \geq f(a,b)$ for all $(x,y) \in B$. This shows that $f(a,b)$ is a local minimum, and the proof is complete. □

3.8 Method of Lagrange Multipliers

In the previous section, we optimized a function, that is, we found its absolute extrema, over a region that included its boundary. Identifying candidate optimal points in the interior of the region is generally straightforward: we simply locate the critical points and evaluate the function there. In contrast, as the examples showed, determining candidate optimal points along the boundary is often a much longer and more cumbersome process.

In this section, we introduce another method for optimizing a function subject to one or more constraints. These constraints may arise from equations describing the boundary of a region, although we will not focus on such cases here. This method applies to general constraints and does not depend on how the constraints are obtained.

If we wish to *optimize* (locate maximum and minimum values of) a function $f(x,y,z)$ subject to a constraint $g(x,y,z) = k$, this process is actually simple, although the process can still be somewhat overwhelming. Again, the constraint may be the equation that describes the boundary of a region or it may not be.

Lagrange Multipliers We seek to solve the following system of equations:

$$\begin{aligned} \nabla f(x,y,z) &= \lambda \nabla g(x,y,z) \\ g(x,y,z) &= k \end{aligned}$$

Next, we plug in all solutions of the form (x, y, z) into f and identify the minimum and maximum values, provided that they exist and vector $\nabla g \neq \mathbf{0}$ at the point.

Here, the constant λ is defined as the **Lagrange Multiplier**.

If you look closely, notice that the system of equations shown above actually consists of *four* equations, just written in a simpler form. This is by the definition of the gradient vector. Writing in terms of its components, we have

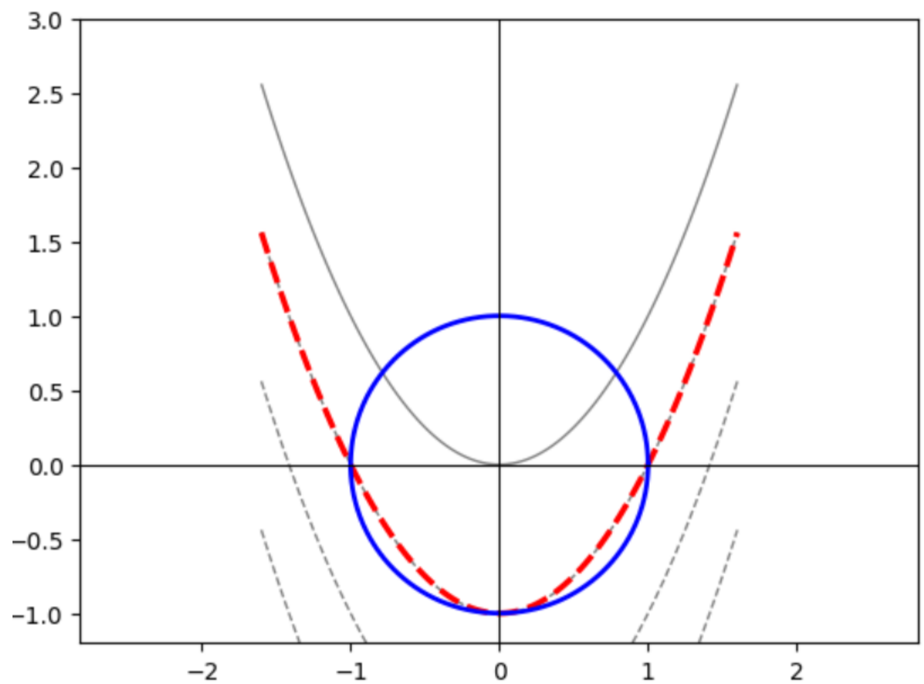
$$\langle f_x, f_y, f_z \rangle = \lambda \langle g_x, g_y, g_z \rangle = \langle \lambda g_x, \lambda g_y, \lambda g_z \rangle$$

We know from Chapter 1 that two vectors are equivalent if and only if their individual components are equal. So

$$f_x = \lambda g_x \quad f_y = \lambda g_y \quad f_z = \lambda g_z$$

Along with the constraint $g(x, y, z) = k$, we have four equations with four unknowns x, y, z, and λ. Note that for functions of two variables, we will only have three equations with the three unknowns x, y, and λ.

To justify the above formulas, let's consider the minimum and maximum values of the function $f(x, y) = 8x^2 - 2y$ subject to the constraint $x^2 + y^2 = 1$. Later in this section, we will show that the minimum value of the function is -2 which occurs at $(0, 1)$ and the maximum value of the function is 8.125 which occurs at $\left(-\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$ and $\left(\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$. Here's a sketch of both the constraint and $f(x, y) = k$ for various values k.



Note that the solution to the system must lie somewhere on the graph of the constraint, i.e. $x^2 + y^2 = 1$. Because we are looking to optimize the function, the point (x, y) must occur when the graph of $f(x, y) = k$ intersects the graph of the constraint when k is either the minimum or maximum value of $f(x, y)$.

Specifically, we can see that the graph of $f(x, y) = -2$, i.e. the graph of the minimum value of $f(x, y)$, barely touches the graph of the constraint at the point $(0, 1)$. In fact, those two graphs are tangent, intersecting in exactly one point.

This implies that the normal vectors of these graphs must be parallel, i.e. the two normal vectors must be scalar multiples of each other. Hence we have derived

$$\nabla f(x,y,z) = \lambda \nabla g(x,y,z)$$

for some scalar ∇.

Also, keep in mind that if k is smaller than the minimum value of $f(x,y)$, the graph $f(x,y) = k$ does not intersect the graph of the constraint and thus it is not possible for the function to take a value of k at a point to satisfy the constraint. This is likewise for the case where k is larger than the maximum value of $f(x,y)$.

Again, by observation, we see that the level curve $f(x,y) = 8.125$ will just touch the graph of the constraint at two points. We already know that these two points are $\left(-\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$ and $\left(\frac{3\sqrt{7}}{8}, -\frac{1}{8}\right)$.

The same logic applies for values of k greater or less than 8.125 as it did earlier. While the 2D justification uses curves, the logic scales to higher dimensions. In 3D, for instance, we work with surfaces rather than lines. At a maximum or minimum, the objective and constraint surfaces become tangent (parallel), meaning their normal vectors/gradients align.

Problem 3.8.1. *Find the dimensions of the box with maximum volume given that the total surface area is* $64\,cm^2$.

Solution: Since we are optimizing volume, our function is

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

Also, the surface area must be a constant value, 64. So we have

$$2xy + 2xz + 2yz = 64 \Longrightarrow xy + xz + yz = 32$$

Thus, the constraint function is given by

$$g(x,y,z) = xy + xz + yz$$

Obviously, the main function $f(x,y,z) = xyz$ will not have any minimum or maximum values unless we place restrictions on the variables. The only sensible restriction we have so far is that each of x, y, and z must be positive. This, of course, instantly means that the function does have a minimum, zero, even though this is a silly value as it also means we pretty much don't have a box. It does however mean that we know $f(x,y,z)$ indeed has a minimum value.

Now let us check for a maximum. $f(x,y,z)$ alone would not have a maximum if we allow x, y, and z to grow unbounded. This is why we have a constraint function $g(x,y,z) = 32$. Using the Method of Lagrange Multipliers, we have the following four equations we need to solve.

$$(1) \quad yz = \lambda(y+z)$$

$$(2) \quad xz = \lambda(x+z)$$

$$(3) \quad xy = \lambda(x+y)$$

$$(4) \quad xy + xz + yz = 32$$

We will use a systematic approach. Let's multiply equation (1) by x, equation (2) by y, and equation (3) by z. This gives us

$$xyz = \lambda x(y+z) \qquad (5)$$

$$xyz = \lambda y(x+z) \qquad (6)$$

$$xyz = \lambda z(x+y) \qquad (7)$$

Notice that we can set equations (5) and (6) equal to each other. This will give us

$$\begin{aligned} \lambda x(y+z) &= \lambda y(x+z) \\ \lambda(xy+xz) - \lambda(yx+yz) &= 0 \\ \lambda(xz-yz) &= 0 \\ &\implies \lambda = 0 \text{ or } xz = yz \end{aligned}$$

This gives us two possibilities. The first, where $\lambda = 0$, is not possible because then equation (1) would reduce to

$$yz = 0 \implies y = 0 \text{ or } z = 0$$

Since we have a trivial constraint that each of x, y, and z must be positive, we can discount the case where $\lambda = 0$. This leaves us with the second possibility:

$$xz = yz$$

Since $z \neq 0$ we can cancel the z from both sides. This gives us

$$x = y$$

Now, let's set equations (6) and (7) equal. Doing this gives us

$$\begin{aligned} \lambda y(x+z) &= \lambda z(x+y) \\ \lambda(yx+yz-zx-zy) &= 0 \\ \lambda(yx-zx) &= 0 \\ &\implies \lambda = 0 \text{ or } yx = zx \end{aligned}$$

Again, we discount the case where $\lambda = 0$, so we get

$$yx = zx$$

and because $x \neq 0$, we must have

$$z = y$$

If we plug in equations (8) and (9) into equation (4), we get

$$y^2 + y^2 + y^2 = 3y^2 = 32 \implies y = \pm\sqrt{\frac{32}{3}} \approx \pm 3.266$$

However, y must be positive so our solution is

$$x = y = z \approx 3.266$$

So our scenario illustrates a cube. We should be a little careful here. Since we've only got one solution we might be tempted to assume that these are the dimensions that will give the largest volume. Anytime we get a single solution we really need to verify that it is a maximum (or minimum if that is what we are looking for).

However, this is not difficult. The volume at our solution is given by

$$V = f\left(\sqrt{\frac{32}{3}}, \sqrt{\frac{32}{3}}, \sqrt{\frac{32}{3}}\right) = \left(\sqrt{\frac{32}{3}}\right)^3 \approx 34.8376$$

Notice that we already know a maximum of $f(x, y, z)$ will exist and so to verify that this is truly a maximum all we have to do is find another set of dimensions that satisfy our constraint and calculate the volume. If the volume of this new set of dimensions is smaller that the volume above then we know that our solution does give a maximum.

So, let's find a new set of dimensions for the box. The only thing we need to worry about is that they will satisfy the constraint. Outside of that there aren't other constraints on the size of the dimensions. So, we can freely pick two values and then use the constraint to determine the third value.

Let's choose $x = y = 1$, because these numbers are convenient to work with. Plugging these into the constraint gives us

$$1 + x + z = 32 \implies 2z = 31 \implies z = \frac{31}{2}$$

We have

$$V = f\left(1, 1, \frac{31}{2}\right) = \frac{31}{2} = 15.5 < 34.8376$$

So the new dimensions give us a smaller volume. We can convince ourselves that the dimensions that will give a maximum volume of the box are $\boxed{x = y = z \approx 3.266}$.

Problem 3.8.2. *Find the maximum and minimum values of the function $f(x, y) = 5x - 3y$ subject to the constraint $x^2 + y^2 = 136$.*

Solution: Since we are dealing with a function of only two variables, this will be easier than the last problem. Also, it should be clear that the range of possible solutions lies on a disk of radius $\sqrt{136}$, a closed and bounded region, $-\sqrt{136} \leq x, y \leq \sqrt{136}$, and the Extreme Value Theorem guarantees the existence of a minimum and maximum value. We solve the following system:

$$\begin{aligned} 5 &= 2\lambda x \\ -3 &= 2\lambda y \\ x^2 + y^2 &= 136 \end{aligned}$$

Of course, we cannot have $\lambda = 0$ because this would not satisfy the first two equations. Now that we know $\lambda \neq 0$ we can solve the first two equations for x and y:

$$x = \frac{5}{2\lambda} \qquad y = -\frac{3}{2\lambda}$$

Plugging these values into the constraint function, we obtain

$$\frac{25}{4\lambda^2} + \frac{9}{4\lambda^2} = \frac{17}{2\lambda^2} = 136$$

We can solve for the constant λ.

$$\lambda^2 = \frac{1}{16} \implies \lambda = \pm\frac{1}{4}$$

Now we can determine the points that are potential maximums and/or minimums. If $\lambda = -\frac{1}{4}$, we get $x = -10$ and $y = 6$. And if $\lambda = \frac{1}{4}$, we get $x = 10$ and $y = -6$. Finally, we plug these points into the function $f(x, y)$. Since $f(-10, 6) = -68$ and $f(10, -6) = 68$, the minimum and maximum values of the function are $\boxed{-68}$ and $\boxed{68}$, respectively.

Lagrange Multipliers with Two Constraints The final topic we will discuss in this section is what to do if we are faced with a function subject to more than one constraint. Specifically, we will consider two constraints.

In general, let $f(x, y, z)$ be subject to the constraints $g(x, y, z) = c$ and $h(x, y, z) = k$. We need to solve the system

$$\begin{aligned} \nabla f(x, y, z) &= \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \\ g(x, y, z) &= c \\ h(x, y, z) &= k \end{aligned}$$

So it is easy to tell that now we have *two* Lagrange Multipliers, specifically λ and μ.

4 Iterated Integrals

In single-variable calculus, integration let us measure things like area under a curve, distance traveled, or total mass along a line. But the real world isn't just one-dimensional. Processes like charge distributing through a surface, rotational inertia, or center of mass mass in a solid happen in multiple dimensions. To handle this, we use iterated integrals, which break multi-dimensional problems into a series of familiar single-variable integrals. This chapter introduces iterated integrals, shows how to test and interpret them geometrically, and sets the foundation for double and triple integrals, volume calculations, and Fubini's Theorem—so we can navigate higher-dimensional spaces just like the x-axis.

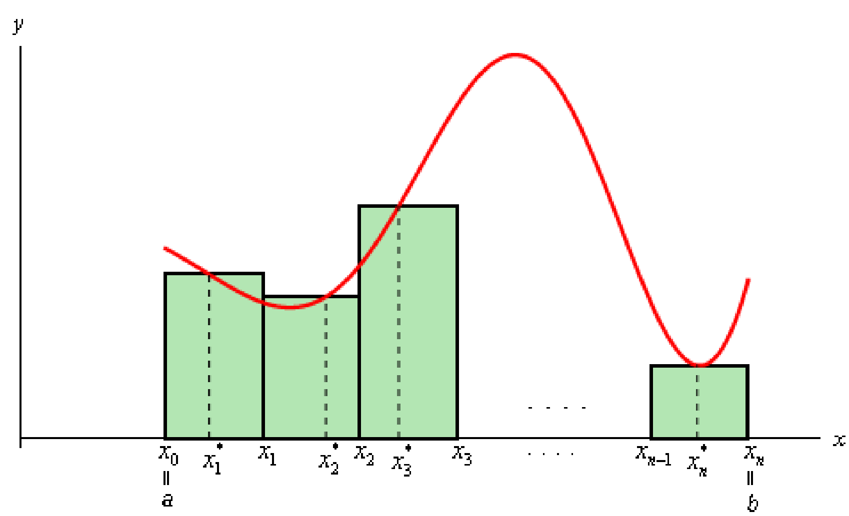
4.1 Double Integrals over Rectangles

Recall from single-variable calculus where we discussed definite integrals for single-variable functions. Specifically, the integral

$$\int_a^b f(x)\,dx$$

involves x values in the closed interval $[a, b]$, or $a \leq x \leq b$. We can also say that we are integrating the function f over the interval $[a, b]$.

When we actually *derived* the definition of the definite integral, we treated it as an area problem. Specifically, we were curious as to what is the area under the curve defined by a function $y = f(x)$. We broke up the interval $a \leq x \leq b$ into sub-intervals of width Δx and chose a point x_i^* from each interval as shown below.



Each rectangle of height $f(x_i^*)$ and its area could then be used to approximate the total area under the curve as follows:

$$A \approx f(x_1^*)\Delta x + f(x_2^*)\Delta x + \cdots + f(x_{n-1}^*)\Delta x + f(x_n^*)\Delta x$$

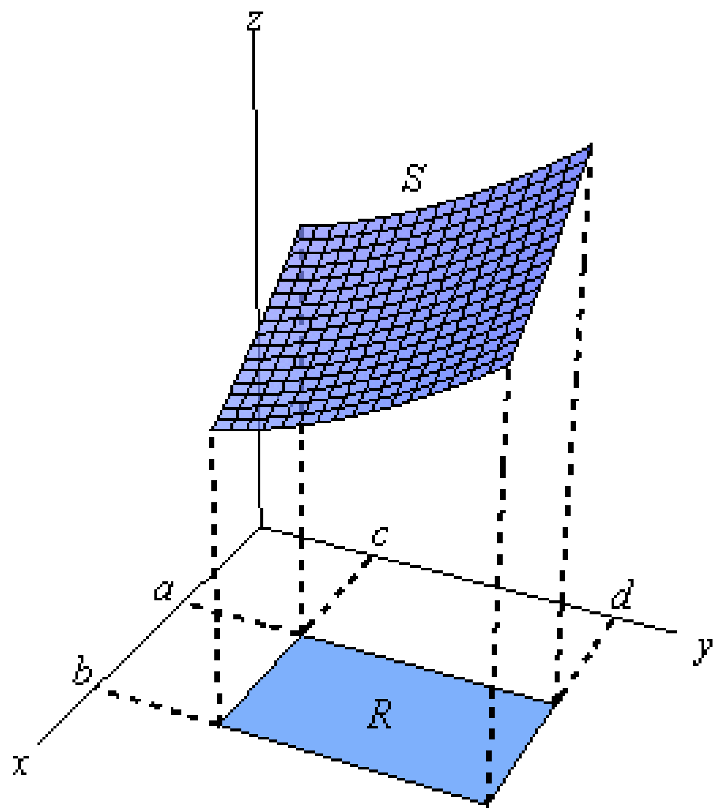
We took the limit as $n \to \infty$ to get the exact area, and thus the definition of the definite integral:

$$\boxed{\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*)\Delta x}$$

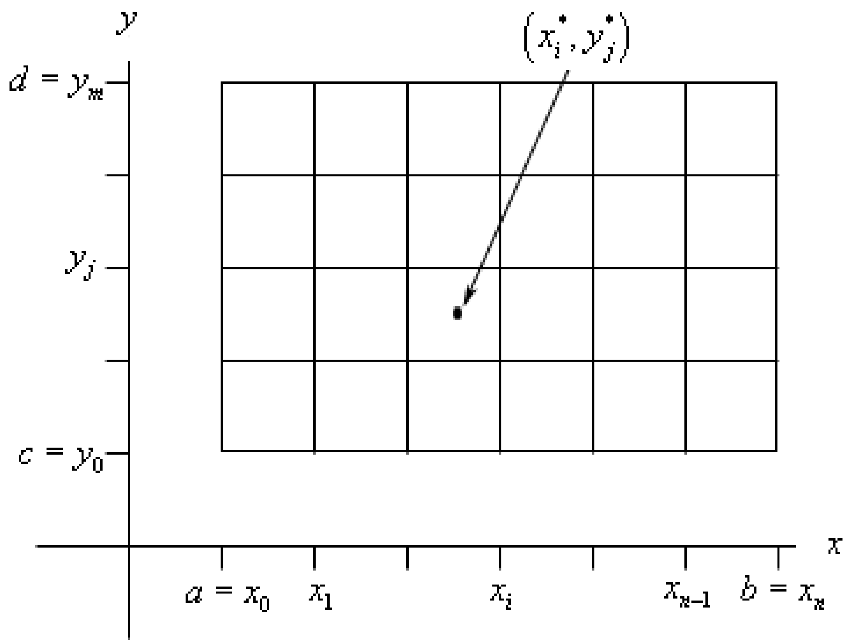
In this section, we integrate a function of two variables, or $f(x, y)$. With one-variable functions, we integrated over an interval (one-dimensional). So when integrating a function of two variables, it makes sense to integrate over a two-dimensional space (a subspace of $\mathbb{R}^2$). In this section, we will assume that the region in $\mathbb{R}^2$ is a rectangle, just to keep things simple:

$$R = [a, b] \times [c, d]$$

In other words, we have two independent intervals $a \le x \le b$ and $c \le y \le d$. We will also assume (initially) that $f(x, y) \ge 0$. Keep in mind this does not always have to be the case. Regardless, consider the graph of S given by graphing $f(x, y)$ over R.



In parallel to finding the area under a curve for one-variable functions, let's now consider the volume of the region under the surface S (and above the xy-plane, as $f(x, y) \ge 0$). In a similar process, divide up the interval $a \le x \le b$ into n sub-intervals and the interval $c \le y \le d$ into m sub-intervals. We then divide R into a series of smaller rectangles and from each of these we choose a point (x_i^*, y_j^*). Below shows a sketch of this process.



Over each of these small rectangles we can construct a rectangular box with with height $f(x_i^*, y_j^*)$. Each base has an area of ΔA, so the the volume of each box is $f(x_i^*, y_j^*)\,\Delta A$. Therefore, the approximate volume under the surface is then

$$V \approx \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i^*, y_j^*)\,\Delta A$$

The reason for the double summation is that we need to add up the volume contributions for both the x and y directions. To increase the approximation's accuracy and ultimately get an exact volume, we let $n, m \to \infty$. Then

$$V = \lim_{n,m\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{m} f(x_i^*, y_j^*)\,\Delta A$$

This should look familiar. This summation is of the same form as the Riemann sum for integrating a single-variable function $f(x)$. The difference is that we are integrating a function of *two* variables over a rectangle. We know that the approximation becomes more accurate as $n, m \to \infty$, so the official definition for the double integral is

$$\boxed{\iint_R f(x,y)\,dA = \lim_{n,m\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} f(x_i^*, y_j^*)\,\Delta A}$$

One interpretation of the double integral of $f(x,y)$ over the rectangle R is the *volume* V under the function $f(x,y)$ and above the xy-plane. In other words,

$$V = \iint_R f(x,y)\,dA$$

Problem 4.1.1. *Estimate the volume of the solid that lies above the square $R = [0,2] \times [0,2]$ and below the elliptic paraboloid $z = 16 - x^2 - 2y^2$. Divide R into four equal squares and choose the sample point to be the upper right corner of each square R_{ij}.*

Solution: We really just need to apply the formula involving the double sum. The paraboloid is the graph of $f(x,y) = 16 - x^2 - 2y^2$ and the area of each square is $\Delta A = 1$. Approximating the volume by the Riemann sum with $m = n = 2$ gives

$$\begin{aligned} V &\approx \sum_{i=1}^{2}\sum_{j=1}^{2} f(x_i, y_j)\,\Delta A \\ &= f(1,1)\,\Delta A + f(1,2)\,\Delta A + f(2,1)\,\Delta A + f(2,2)\,\Delta A \\ &= 13(1) + 7(1) + 10(1) + 4(1) = \boxed{34} \end{aligned}$$

If we increase the number of squares, our approximation becomes better.

Problem 4.1.2. *If R is the rectangular region $\{(x,y)\,|\,-1 \le x \le 1,\ -2 \le y \le 2\}$, evaluate the integral $\iint_R \sqrt{1-x^2}\,dA$.*

Solution: We can't evaluate the double integral directly from the definition alone, but it helps to notice that $\sqrt{1-x^2} \ge 0$. We can, however, compute the integral by interpreting it as a volume. If $z = f(x,y) = \sqrt{1-x^2}$, then $x^2 + z^2 = 1$ and $z \ge 0$, so the double integral represents the volume of the solid S that lies below the circular cylinder $x^2 + z^2 = 1$ and above the rectangle R. Because the cylinder is centered at the origin and the height is two times the radius, the volume of S is the area of a semi-circular base with radius 1 multiplied by the height of the cylinder, which is 2:

$$\iint_R \sqrt{1-x^2}\,dA = \frac{1}{2}\pi(1)^2 \cdot 4 = \boxed{2\pi}$$

The Midpoint Rule As with single integrals, double integrals are also governed by the Midpoint Rule. In this case, the sample point (x_{ij}^*, y_{ij}^*) in R_{ij} is chosen as the center $(\bar{x}_i, \bar{y}_j)$ of R_{ij}. In other words, $\bar{x}_i$ is the midpoint of $[x_{i-1}, x_i]$ and $\bar{y}_j$ is the midpoint of $[y_{i-1}, y_i]$. The rule essentially states

$$\iint_R f(x,y)\,dA \approx \sum_{i=1}^{m}\sum_{j=1}^{n} f(\bar{x}_i, \bar{y}_j)\,\Delta A$$

Problem 4.1.3. *Use the Midpoint Rule with $m = n = 2$ to estimate $\iint_R (x - 3y^2)\,dA$, where $R = \{(x,y)\,|\,0 \le x \le 2,\ 1 \le y \le 2\}$.*

Solution: In order to use the Midpoint Rule with $m = n = 2$, we need to evaluate $f(x,y)$ at the centers of each of the four sub-rectangles. In this case, we have $\bar{x}_1 = \frac{1}{2}$, $\bar{x}_2 = \frac{3}{2}$, $\bar{y}_1 = \frac{5}{4}$, and $\bar{y}_2 = \frac{7}{4}$. The area of each sub-rectangle is $\Delta A = \frac{1}{2}$. Thus

$$\begin{aligned} \iint_R (x - 3y^2)\,dA &\approx \sum_{i=1}^{2}\sum_{j=1}^{2} f(\bar{x}_i, \bar{y}_j)\,\Delta A \\ &= f(\bar{x}_1, \bar{y}_1)\,\Delta A + f(\bar{x}_1, \bar{y}_2)\,\Delta A + f(\bar{x}_2, \bar{y}_1)\,\Delta A + f(\bar{x}_2, \bar{y}_2)\,\Delta A \\ &= \left(-\frac{67}{16}\right)\frac{1}{2} + \left(-\frac{139}{16}\right)\frac{1}{2} + \left(-\frac{51}{16}\right)\frac{1}{2} + \left(-\frac{123}{16}\right)\frac{1}{2} \\ &= \boxed{-\frac{95}{8}} \end{aligned}$$

Remark #1. In problem 4.1.5, we will see that the exact value of this double integral is -12. Our answer, $-\frac{95}{8} = -11.875$, is pretty close to the exact answer. If we increase the number of sub-rectangles and make them smaller, then the approximations approach the exact value of the double integral.

Remark #2. Remember that the interpretation for a double integral as a volume is only applicable to *positive* functions. However, $f(x, y) = x - 3y^2$ is not a positive function, so its integral is not a volume. We will get to problems where we learn how to describe integrals of functions that cannot serve as volumes.

Iterated Integrals Suppose that a two-variable function $f(x, y)$ is integrable on the region $R = [a, b]\times[c, d]$. The notation $\int_c^d f(x, y)\,dy$ means x is fixed and $f(x, y)$ is integrated with respect to y from $y = c$ to $y = d$. This is called *partial integration*, with respect to y. Since $\int_c^d f(x, y)\,dy$ depends on the value of x, it defines a function of x

$$A(x) = \int_c^d f(x, y)\,dy$$

If we integrate A with respect to x from $x = a$ to $x = b$, we get

$$\int_a^b A(x)\,dx = \int_a^b \left[\int_c^d f(x, y)\,dy\right] dx$$

The right-hand-side of this equation demonstrates what is called an **iterated integral**. In practice, the brackets are not included. Therefore

$$\int_a^b \int_c^d f(x, y)\,dy\,dx = \int_a^b \left[\int_c^d f(x, y)\,dy\right] dx$$

meaning that we integrate first with respect to y (holding x fixed) from $y = c$ to $y = d$, and then we integrate the resulting function of x with respect to x from $x = a$ to $x = b$.

Alternatively, the iterated integral

$$\int_c^d \int_a^b f(x, y)\,dx\,dy = \int_c^d \left[\int_a^b f(x, y)\,dx\right] dy$$

means that we first integrate with respect to y (holding x fixed) from $x = a$ to $x = b$, and then we integrate the resulting function of y with respect to y from $y = c$ to $y = d$. Notice in both cases, we operate from the *inside out.*

Problem 4.1.4. *Evaluate the following iterated integrals.*

(a) $\int_0^3 \int_1^2 x^2y\,dy\,dx$
(b) $\int_1^2 \int_0^3 x^2y\,dx\,dy$

Solution to part a: We first integrate with respect to y, holding x constant.

$$\int_1^2 x^2y\,dy = \left[x^2\frac{y^2}{2}\right]_{y=1}^{y=2} = x^2\left(\frac{2^2}{2}\right) - x^2\left(\frac{1^2}{2}\right) = \frac{3}{2}x^2$$

Thus our function A given in the earlier discussion is $A(x) = \frac{3}{2}x^2$ in the context of this problem. We now integrate with respect to x from $x = 0$ to $x = 3$.

$$\int_0^3 \int_1^2 x^2y\,dy\,dx = \int_0^3 \left[\int_1^2 x^2y\,dy\right] dx = \int_0^3 \frac{3}{2}x^2\,dx = \frac{x^3}{2}]_0^3 = \boxed{\frac{27}{2}}$$

Solution to part b: The only difference is that we integrate first with respect to x, treating y as a constant:

$$\int_1^2 \int_0^3 x^2 y \, dx \, dy = \int_1^2 \left[\int_0^3 x^2 y \, dx \right] dy = \int_1^2 \left[\frac{x^3}{3} y \right]_{x=0}^{x=3} dy$$
$$= \int_1^2 9y \, dy = 9 \frac{y^2}{2}]_1^2 = \boxed{\frac{27}{2}}$$

It turns out that we obtained the same answer regardless of whether we integrated with respect to x or y first. It turns out that for all iterated integrals, the order of integration does not matter. (This is similar to Clairaut's Theorem regarding the equality of mixed partial derivatives). The following theorem gives a practical method for evaluating a double integral by expressing it as an iterated integral.

Theorem 4.1.1. *If f is continuous on the rectangle*

$$R = \{(x, y) \,|\, a \leq x \leq b, \, c \leq y \leq d\}$$

then ***Fubini's Theorem*** *gives the following equality:*

$$\iint_R f(x, y) \, dA = \int_a^b \int_c^d f(x, y) \, dy \, dx = \int_c^d \int_a^b f(x, y) \, dx \, dy$$

This is true in general if we assume f is bounded on R, f is discontinuous only on a finite number of smooth curves, and the iterated integrals exist.

Proof. It is too difficult to include the full proof of Fubini's Theorem in this book, but we can give a proper indication as to why it is true for the case of $f(x, y) \geq 0$. If f is a positive function, then the double integral $\iint_R f(x, y) \, dA$ represents the volume V of the solid S lying above R and underneath the surface with equation $z = f(x, y)$. We also have another formula from single-variable calculus used for volume, that is

$$V = \int_a^b A(x) \, dx$$

where $A(x)$ is the area of a cross-section of S in the plane through x perpendicular to the x-axis. Also, $A(x)$ is the area under the curve C whose equation is $z = f(x, y)$, where x is held constant while $c \leq y \leq d$. Thus

$$A(x) = \int_c^d f(x, y) \, dy$$

and so

$$\iint_R f(x, y) \, dA = V = \int_a^b A(x) \, dx = \int_a^b \int_c^d f(x, y) \, dy \, dx$$

Similarly, if we used cross sections perpendicular to the y-axis, we get

$$\iint_R f(x, y) \, dA = \int_c^d \int_a^b f(x, y) \, dx \, dy$$

and the proof is complete. □

Problem 4.1.5. *Evaluate $\iint_R (x - 3y^2) \, dA$, where $R = \{(x, y) \,|\, 0 \leq x \leq 2, \, 1 \leq y \leq 2\}$. Hint: compare your result with the solution to problem* 4.1.3.

Solution 1: Apply Fubini's Theorem, integrating with respect to y first:

$$\iint_R (x - 3y^2)\, dA = \int_0^2 \int_1^2 (x - 3y^2)\, dy\, dx = \int_0^2 \left[xy - y^3\right]_{y=1}^{y=2} dx$$
$$= \int_0^2 (x - 7)\, dx = \frac{x^2}{2} - 7x\Big]_0^2 = \boxed{-12}$$

Solution 2: Here we also apply Fubini's Theorem, this time integrating with respect to x first:

$$\iint_R (x - 3y^2)\, dA = \int_1^2 \int_0^2 (x - 3y^2)\, dx\, dy = \int_1^2 \left[\frac{x^2}{2} - 3xy^2\right]_{x=0}^{x=2} dy$$
$$= \int_1^2 (2 - 6y^2)\, dy = 2y - 2y^3\Big]_1^2 = \boxed{-12}$$

Remark. In problem 4.1.3, we used the Midpoint Rule and obtained a result, -11.875, which is a good approximation for the exact value, -12, guaranteed by Fubini's Theorem.

Problem 4.1.6. *Evaluate* $\iint_R y\sin(xy)\, dA$*, where* $R = [1, 2] \times [0, \pi]$.

Solution: Apply Fubini's Theorem, integrating first with respect to x:

$$\iint_R y\sin(xy)\, dA = \int_0^\pi \int_1^2 y\sin(xy)\, dx\, dy$$

$$= \int_0^\pi y\left[-\frac{1}{y}\cos(xy)\right]_{x=1}^{x=2} dy$$

$$= \int_0^\pi (-\cos 2y + \cos y)\, dy$$
$$= -\frac{1}{2}\sin 2y + \sin y\Big]_0^\pi = \boxed{0}$$

Remark. If we reversed the order of integration, we would get the same result, but if it was done in this problem, we would end up with

$$\iint_R y\sin(xy)\, dA = \int_1^2 \int_0^\pi y\sin(xy)\, dy\, dx$$

but this order of integration is more difficult to deal with. In particular, we would need to apply integration by parts twice. Therefore, it is always best to determine which order gives us a simpler process of integration when calculating double integrals.

Problem 4.1.7. *Calculate the volume of the solid S bounded by the elliptic paraboloid $x^2 + 2y^2 + z = 16$, the three coordinate planes, and the planes $x = 2$ and $y = 2$.*

Solution: The main difficulty here is determining the bounds for our integral, rather than simply setting it up. Three coordinate planes refer to the planes $x = 0$, $y = 0$, and $z = 0$. If we want to express the surface as a function of x and y, we use $z = f(x, y) = 16 - x^2 - 2y^2$. Therefore, S is the solid lying under the surface $z = 16 - x^2 - 2y^2$ and above the square $R = [0, 2] \times [0, 2]$. The volume of S is

$$\begin{aligned} V &= \iint_R (16 - x^2 - 2y^2)\, dA = \int_0^2 \int_0^2 (16 - x^2 - 2y^2)\, dx\, dy \\ &= \int_0^2 \left[16x - \left(\frac{1}{3}\right) x^3 - 2y^2 x\right]_{x=0}^{x=2} dy \\ &= \int_0^2 \left(\frac{88}{3} - 4y^2\right) dy = \left[\frac{88}{3} y - \frac{4}{3} y^3\right]_0^2 \\ &= \boxed{48} \end{aligned}$$

Now let's consider the special case where $z = f(x, y)$ can be expressed as a product $g(x)h(y)$, i.e. a function of x only multiplied by a function of y only. Also, let $R = [a, b] \times [c, d]$. Then Fubini's Theorem gives

$$\iint_R f(x, y)\, dA = \int_c^d \int_a^b g(x)h(y)\, dx\, dy = \int_c^d \left[\int_a^b g(x)h(y)\, dx\right] dy$$

In the inner integral, y is held constant, so $h(y)$ is also constant and we write

$$\int_c^d \left[\int_a^b g(x)h(y)\, dx\right] dy = \int_c^d \left[h(y) \left(\int_a^b g(x)\, dx\right)\right] dy = \int_a^b g(x)\, dx \int_c^d h(y)\, dy$$

since $\int_a^b g(x)\, dx$ is a constant. Therefore, in this special case the double integral of f can be written as the product of two single integrals:

$$\boxed{\iint_R g(x)h(y)\, dA = \int_a^b g(x)\, dx \int_c^d h(y)\, dy \quad \text{where } R = [a, b] \times [c, d]}$$

Average Value In single-variable calculus, we defined the average value of a single-variable function $f(x)$ on an interval $[a, b]$ as

$$f_{avg} = \frac{1}{b - a} \int_a^b f(x)\, dx$$

In a similar fashion, the **average value** of a two-variable function $f(x, y)$ defined on a rectangle R to be

$$f_{avg} = \frac{1}{A(R)} \iint_R f(x, y)\, dA$$

where $A(R)$ is the area of R.
If $f(x, y) \geq 0$, we have

$$A(R) \cdot f_{avg} = \iint_R f(x, y)\, dA$$

which demonstrates the box with base R and height f_{avg} has the same volume as the solid lying under the graph of $z = f(x, y)$.

Problem 4.1.8. *Calculate the average value of $f(x, y) = x^2 + y$ over the rectangle $R = [0, 2] \times [0, 3]$.*

Solution: The area of the region R is $A(R) = 2 \cdot 3 = 6$. The double integral of $f(x, y)$ over the region R is given by

$$\begin{aligned} \iint_R f(x,y)\,dA &= \int_0^3 \int_0^2 f(x,y)\,dy\,dx \\ &= \int_0^3 \int_0^2 (x^2+y)\,dx\,dy \\ &= \int_0^3 \left[\frac{x^3}{3} + xy\right]_{x=0}^{x=2} dx \\ &= \int_0^3 \left(\frac{8}{3} + 2y\right) dy \\ &= \left[\frac{8}{3}y + y^2\right]_0^3 = 17 \end{aligned}$$

So the average value is $f_{avg} = \frac{1}{6} \cdot 17 = \boxed{\frac{17}{6}}$.

4.2 Double Integrals over General Regions

In the previous section we just dipped our feet in the waters of double integrals, by considering rectangular regions. In reality, however, most regions are not rectangular, so we need to consider the double integral

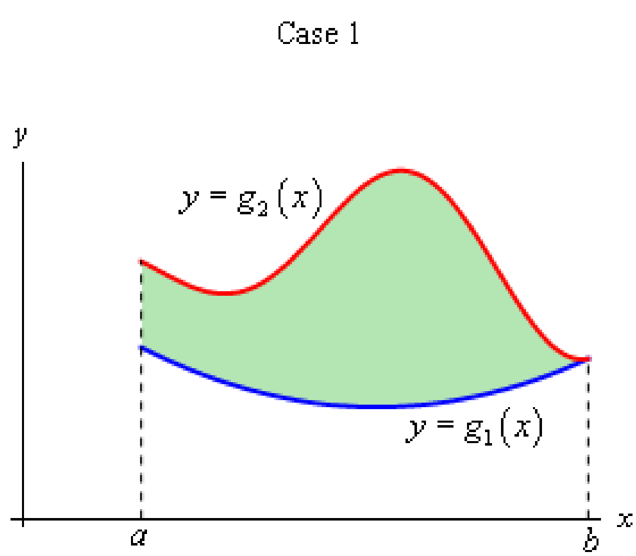
$$\iint_D f(x,y)\,dA$$

where D is any region, or in context, a **general region**.

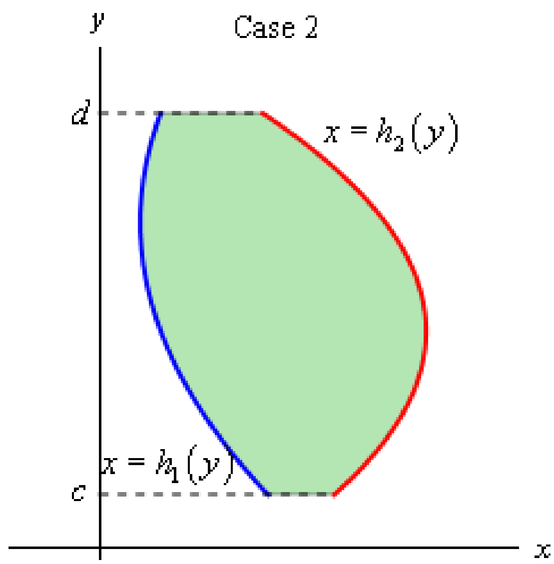
When calculating double integrals over general regions, we often use *set builder notation*, which is a fancy way of expressing the domain of our variables of integration.

There are two cases of regions that we need to consider for double integrals: type I and type II regions.

Type I Regions Define a region D as $D = \{(x, y) \mid a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$. A sketch of a type I region would look like this:



Type II Regions Define a region D as $D = \{(x, y) \mid h_1(y) \le x \le h_2(y),\ c \le y \le d\}$. A sketch of a type II region would look like this:



Double Integral Setup and Properties In the case of type I regions, we set up the double integral as

$$\iint_D f(x, y)\, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y), dy\, dx$$

In the case of type II regions, we set up the double integral as

$$\iint_D f(x, y)\, dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\, dx\, dy$$

Now, let's go through some important properties of double integrals before we jump into practice problems. You'll notice that all of them were really extensions of properties of single integrals.

Theorem 4.2.1. ***Properties***

1. $\iint_D f(x, y) + g(x, y)\, dA = \iint_D f(x, y)\, dA + \iint_D g(x, y)\, dA$
2. $\iint_D cf(x, y)\, dA = c \iint_D f(x, y)\, dA$
3. *If the region D can be split into two separate regions D_1 and D_2, then the integral can be expressed as:*

$$\iint_D f(x, y)\, dA = \iint_{D_1} f(x, y)\, dA + \iint_{D_2} f(x, y)\, dA$$

Now, we can do some practice!

Problem 4.2.1. *Evaluate the double integral $\iint_D e^{\frac{x}{y}}\, dA$, where $D = \{(x, y) \mid 1 \le y \le 2,\ y \le x \le y^3\}$.*

Solution: All we really need to do is apply the formula.

$$\begin{aligned}\iint_D e^{\frac{x}{y}}\, dA &= \int_1^2 \int_y^{y^3} e^{\frac{x}{y}}\, dx\, dy \\ &= \int_1^2 \left(ye^{y^2} - ye^1 \right) dy \\ &= \left(\frac{1}{2}e^{y^2} - \frac{1}{2}y^2 e^1 \right) \Big|_1^2 \\ &= \boxed{\frac{1}{2}e^4 - 2e}\end{aligned}$$

Problem 4.2.2. *Evaluate* $\iint_D 2xy - y^3\, dA$*, where* D *is the plane region bounded by* $y = \sqrt{x}$ *and* $y = x^2$*.*

Solution: The curves $y = \sqrt{x}$ and $y = x^2$ intersect at $x = 0$ and $x = 1$. Also for $0 \le x \le 1$, the curve $y = \sqrt{x}$ is above $y = x^2$. In set builder notation, we have $D = \{(x, y) \,|\, 0 \le x \le 1,\, x^2 \le y \le \sqrt{x}\}$. We can now evaluate the integral.

$$\begin{aligned}\iint_D 2xy - y^3\, dA &= \int_0^1 \int_{x^2}^{\sqrt{x}} 2xy - y^3\, dy\, dx \\ &= \int_0^1 \left(xy^2 - \frac{1}{4}y^4 \right) \Big|_{x^2}^{\sqrt{x}}\, dx \\ &= \int_0^1 \left(\frac{3}{4}x^2 - x^5 + \frac{1}{4}x^8 \right) dx \\ &= \left(\frac{1}{12}x^3 - \frac{1}{6}x^6 + \frac{1}{36}x^9 \right) \Big|_0^1 \\ &= \boxed{-2}\end{aligned}$$

4.3 Double Integrals: Converting to Polar Coordinates

Consider evaluating a double integral like this:

$$\iint_R f(x, y)\, dA$$

where the region R is a disk centered at the origin. Describing this region using rectangular coordinates can be cumbersome, whereas polar coordinates provide a much simpler characterization. More generally, when a region is naturally expressed in polar coordinates, it is often preferable to rewrite the double integral in polar form before evaluating.

Polar Coordinates: A Review Most of you were exposed to the polar coordinate system in your Precalculus course. The *polar coordinates*, (r, θ), of a point are related to its rectangular coordinates by the equations

$$r^2 = x^2 + y^2 \quad x = r\cos\theta \quad y = r\sin\theta$$

As far as regions that are convenient to express in polar coordinates, a circle centered at the origin is a perfect example. In particular, the *unit circle* has the equation $r = 1$, and the angle θ, with respect to the positive x-axis, ranges from 0 to 2π. It is also convenient to apply polar coordinates to regions involving concentric circles and circular sectors.

Double Integrals in Polar Coordinates Any region described in polar coordinates can be thought of as a **polar rectangle**

$$R = \{(r, \theta) \,|\, a \le r \le b,\, \alpha \le \beta\}$$

Let R be a polar rectangle defined by

$$a \le r \le b, \qquad \alpha \le \theta \le \beta.$$

To approximate the double integral $\iint_R f(x, y)\, dA$, partition the interval $[a, b]$ into m equal subintervals $[r_{i-1}, r_i]$ of width $\Delta r = (b - a)/m$ and partition $[\alpha, \beta]$ into n equal subintervals $[\theta_{j-1}, \theta_j]$ of width $\Delta\theta = (\beta - \alpha)/n$. The resulting circles $r = r_i$ and rays $\theta = \theta_j$ subdivide R into small polar rectangles R_{ij}. This process is shown below.

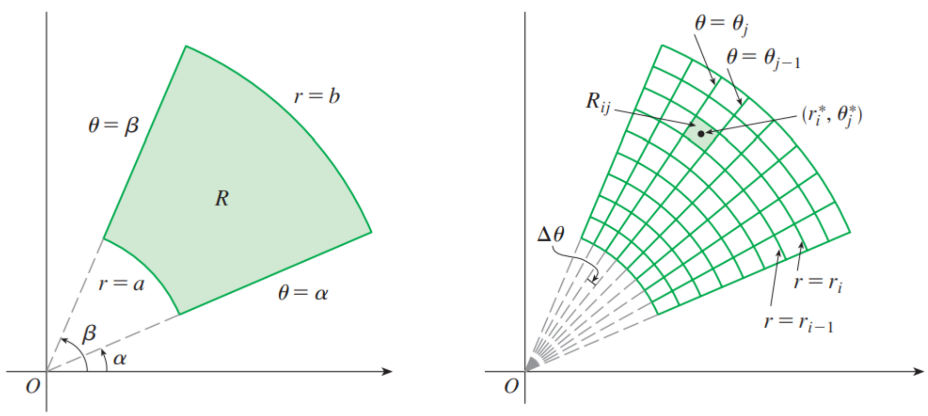


Image Credit: Calculus, Early Transcendentals

The "center" of the polar sub-rectangle

$$R_{ij} = \{(r, \theta) \,|\, r_{i-1} \le r \le r_i,\, \theta_{j-1} \le \theta \le \theta_j\}$$

has polar coordinates of

$$r_i^* = \frac{1}{2}(r_{i-1} + r_i) \qquad \theta_j^* = \frac{1}{2}(\theta_{j-1} + \theta_j)$$

We can compute the area of R_{ij} using the fact that the area of a circular sector with radius r and central angle θ is $\frac{1}{2}r^2\theta$. If we subtract the area of two sectors, each of which has central angle $\Delta\theta = \theta_j - \theta_{j-1}$, the area of R_{ij} is

$$\begin{aligned}\Delta A_i &= \frac{1}{2}r_i^2\,\Delta\theta - \frac{1}{2}r_{i-1}^2\,\Delta\theta = \frac{1}{2}\left(r_i^2 - r_{i-1}^2\right)\Delta\theta \\ &= \frac{1}{2}\left(r_i + r_{i-1}\right)\left(r_i - r_{i-1}\right)\Delta\theta = r_i^*\,\Delta r\,\Delta\theta\end{aligned}$$

Remark. We have defined the integral $\iint_R f(x, y)\, dA$ in terms of ordinary rectangles, it can be shown for continuous f, we can always obtain the same answer using polar rectangles.

The rectangular coordinates of the center of R_{ij} are $(r_i^* \cos\theta_j^*, r_i^* \sin\theta_j^*)$, so the relevant Riemann sum for the conversion to polar coordinates is

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f\left(r_i^*\cos\theta_j^*,\, r_i^*\sin\theta_j^*\right)\,\Delta A_i = \sum_{i=1}^{m}\sum_{j=1}^{n} f\left(r_i^*\cos\theta_j^*,\, r_i^*\sin\theta_j^*\right)\, r_i^*\,\Delta r\,\Delta\theta$$

Suppose $g(r,\theta) = rf(r\cos\theta, r\sin\theta)$, then the above Riemann sum can be written as

$$\sum_{i=1}^{m}\sum_{j=1}^{n} g(r_i^*, \theta_j^*)\,\Delta r\,\Delta\theta$$

which also serves as the Riemann sum for the integral

$$\int_{\alpha}^{\beta}\int_{a}^{b} g(r,\theta)\,dr\,d\theta$$

Therefore we have

$$\begin{aligned}
\iint_R f(x,y)\,dA &= \lim_{m,n\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n} f\big(r_i^*\cos\theta_j^*,\, r_i^*\sin\theta_j^*\big)\ \Delta A_i \\
&= \lim_{m,n\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n} g\big(r_i^*, \theta_j^*\big)\ \Delta r\,\Delta\theta \\
&= \int_{\alpha}^{\beta}\int_{a}^{b} g(r,\theta)\,dr\,d\theta \\
&= \int_{\alpha}^{\beta}\int_{a}^{b} f(r\cos\theta,\, r\sin\theta)\ r\,dr\,d\theta
\end{aligned}$$

In general, we can convert from rectangular coordinates to polar coordinates as long as f is continuous on the polar rectangle given by $0 \le a \le r \le b$ and $\alpha \le \theta \le \beta$, where $0 \le \beta - \alpha \le 2\pi$. We also rewrite the area element as $dA = r\,dr\,d\theta$. Be sure not to forget the additional r factor. A classic method for remembering the area element of a polar rectangle is to consider it as an ordinary rectangle with dimensions of $r\,d\theta$ and dr and therefore has "area" $dA = r\,dr\,d\theta$.

Problem 4.3.1. *Evaluate $\iint_R (3y + 4x^2)\,dA$ where R is the region in the upper-half plane bounded by the circles $x^2 + y^2 = 1$ and $x^2 + y^2 = 4$.*

Solution: We can describe the region in set-builder notation

$$R = \{(x,y)\,|\,y \ge 0,\ 1 \le x^2 + y^2 \le 4\}$$

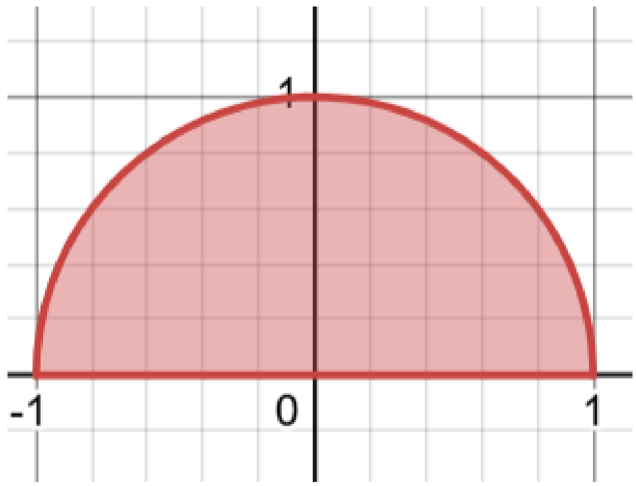
as the upper half-ring, concentric between semicircles of radii 1 and 2. Therefore in polar coordinates R is given by $1 \le r \le 2$, $0 \le \theta \le \pi$. Thus

$$\begin{aligned}
\iint_R (3x + 4y^2)\,dA &= \int_0^{\pi}\int_1^2 \big[3(r\cos\theta) + 4(r\sin\theta)^2\big]\,r\,dr\,d\theta \\
&= \int_0^{\pi}\int_1^2 \big(3r^2\cos\theta + 4r^3\sin^2\theta\big)\ dr\,d\theta \\
&= \int_0^{\pi} \big[r^3\cos\theta + r^4\sin^2\theta\big]_{r=1}^{r=2}\,d\theta \\
&= \int_0^{\pi} \big(7\cos\theta + 15\sin^2\theta\big)\,d\theta \\
&= \int_0^{\pi} \left[7\cos\theta + \frac{15}{2}(1 - \cos 2\theta)\right]d\theta \\
&= \left[7\sin\theta + \frac{15\theta}{2} - \frac{15}{4}\sin 2\theta\right]_0^{\pi} \\
&= \boxed{\frac{15\pi}{2}}
\end{aligned}$$

Problem 4.3.2. *Evaluate the double integral* $\int_{-1}^{1}\int_{0}^{\sqrt{1-x^2}}(x^2+y^2)\,dy\,dx$.

Solution: Here, R is the upper-half of the unit circle, and is described as

$$R = \{(x,y)\,|\,-1 \le x \le 1,\ 0 \le y \le \sqrt{1-x^2}\}$$



Since R is a half-disk, it's more convenient to use polar coordinates:

$$R = \{(r,\theta)\,|\,0 \le \theta \le \pi,\ 0 \le r \le 1\}$$

Therefore we have

$$\begin{aligned}\int_{-1}^{1}\int_{0}^{\sqrt{1-x^2}}(x^2+y^2)\,dy\,dx &= \int_{0}^{\pi}\int_{0}^{1}(r^2)\,r\,dr\,d\theta \\ &= \int_{0}^{\pi}\left[\frac{r^4}{4}\right]_{r=0}^{r=1}d\theta = \frac{1}{4}\int_{0}^{\pi}d\theta = \boxed{\frac{\pi}{4}}\end{aligned}$$

Problem 4.3.3. *Find the volume of the solid bounded by the plane* $z = 0$ *and the paraboloid* $z = 1 - x^2 - y^2$.

Solution: In order to determine the intersection between the plane and the paraboloid, we set $z = 0$ in the equation of the solid, yielding $x^2 + y^2 = 1$. Therefore the solid lies under the paraboloid and above the disk D given by $x^2 + y^2 \le 1$. In polar coordinates, the region is $0 \le r \le 1$, $0 \le \theta \le 2\pi$. Since $1 - x^2 - y^2$ is equivalent to $1 - r^2$ (using the substitution $x^2 + y^2 = r^2$), the volume is

$$\begin{aligned}V &= \iint_D (1-x^2-y^2)\,dA = \int_{0}^{2\pi}\int_{0}^{1}(1-r^2)\,r\,dr\,d\theta \\ &= \int_{0}^{2\pi}d\theta\int_{0}^{1}(r-r^3)\,dr = 2\pi\left[\frac{r^2}{2}-\frac{r^4}{4}\right]_0^1 = \boxed{\frac{\pi}{2}}\end{aligned}$$

We can extend our discussion on double integrals with polar coordinates to more complex regions, similar to the type II rectangular regions we saw in section 4.2. Using the substitutions for converting rectangular to polar coordinates along with the equations for general regions, we obtain the following equation:

$$\boxed{\iint_D f(x,y)\,dA = \int_{\alpha}^{\beta}\int_{h_1(\theta)}^{h_2(\theta)} f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta}$$

if f is continuous on a polar region of the form

$$D = \{(r, \theta) \mid \alpha \le \theta \le \beta,\, h_1(\theta) \le r \le h_2(\theta)\}$$

The following illustration demonstrates this concept.

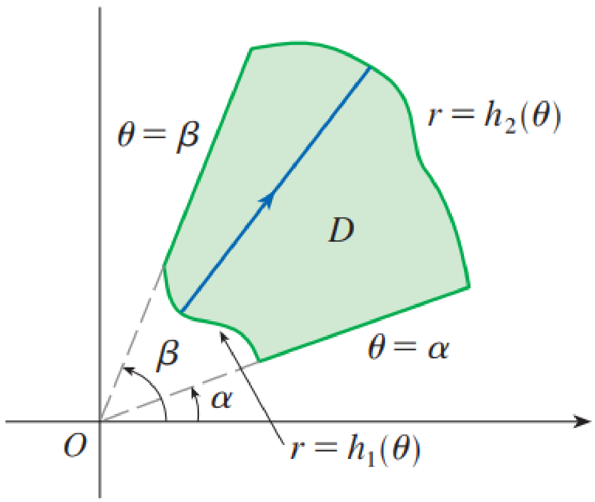


Image Credit: Calculus, Early Transcendentals

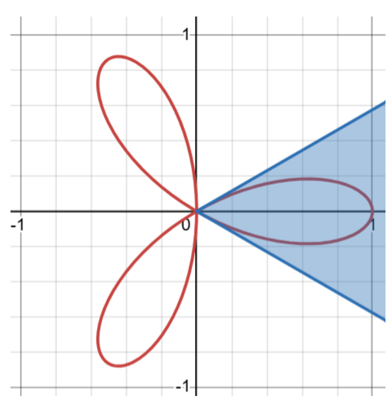
Now let's consider the special case of a generic polar curve $h(\theta)$ from Calculus II. By setting $f(x, y) = 1$, $h_1(\theta) = 0$, and $h_2(\theta) = h(\theta)$, the area of the region D bounded by $\theta = \alpha$, $\theta = \beta$, and $r = h(\theta)$ is

$$\begin{aligned} A(D) &= \iint_D 1\, dA = \int_\alpha^\beta \int_0^{h(\theta)} r\, dr\, d\theta \\ &= \int_\alpha^\beta \left[\frac{r^2}{2}\right]_0^{h(\theta)} d\theta = \int_\alpha^\beta \frac{1}{2}[h(\theta)]^2 d\theta \end{aligned}$$

which is a result we know well.

Problem 4.3.4. *Use a double integral to calculate the area enclosed by one loop of the three-petaled rose* $r = \cos 3\theta$.

Solution: Here is a simple sketch of the region.



If we set $r \geq 0$, then one petal is traced for $\cos 3\theta \geq 0$, so we know $-\pi/6 \leq \theta \leq \pi/6$, because $\cos 3\theta = 0$ for $\theta = \pm\pi/6$. Therefore, a single loop can be expressed as the polar region

$$D = \{(r,\theta) \mid -\pi/6 \leq \theta \leq \pi/6,\, 0 \leq r \leq \cos 3\theta\}$$

So the area is

$$\begin{aligned} A(D) &= \iint_D dA = \int_{-\pi/6}^{\pi/6} \int_0^{\cos 3\theta} r\,dr\,d\theta \\ &= \int_{-\pi/6}^{\pi/6} \left[\frac{1}{2}r^2\right]_0^{\cos 3\theta} d\theta = \frac{1}{2}\int_{-\pi/6}^{\pi/6} \cos^2 3\theta\,d\theta \\ &= \frac{1}{4}\int_{-\pi/6}^{\pi/6} (1+\cos 6\theta)\,d\theta = \frac{1}{4}\left[\theta + \frac{1}{6}\sin 6\theta\right]_{-\pi/6}^{\pi/6} = \boxed{\frac{\pi}{12}} \end{aligned}$$

Problem 4.3.5. *Calculate the volume of the solid that lies under the paraboloid $z = x^2 + y^2$, above the xy-plane, and inside the cylinder $x^2 + y^2 = 2x$.*

Solution: The solid lies above the disk D whose boundary circle has the equation $x^2 + y^2 = 2x$. We can complete the square to get

$$x^2 - 2x + y^2 = 0 \Rightarrow (x^2 - 2x + 1) + y^2 = 0 + 1 \Rightarrow (x-1)^2 + y^2 = 1$$

The boundary circle $x^2 + y^2 = 2x$ becomes $r^2 = 2r\cos\theta$, or $r = 2\cos\theta$. Thus the disk D is given by

$$D = \{(r,\theta) \mid -\pi/2 \leq \pi/2,\, 0 \leq r \leq 2\cos\theta\}$$

so the volume of the solid is

$$\begin{aligned} V &= \iint_D (x^2+y^2)\,dA \int_{-\pi/2}^{\pi/2} \int_0^{2\cos\theta} r^2\,r\,dr\,d\theta = \int_{-\pi/2}^{\pi/2} \left[\frac{r^4}{4}\right]_0^{2\cos\theta} d\theta \\ &= 4\int_{-\pi/2}^{\pi/2} \cos^4\theta\,d\theta = 8\int_0^{\pi/2} \cos^4\theta\,d\theta = 8\int_0^{\pi/2} \left(\frac{1+\cos 2\theta}{2}\right)^2 d\theta \\ &= 2\int_0^{\pi/2} \left[1 + 2\cos 2\theta + \frac{1}{2}(1+\cos 4\theta)\right] d\theta \\ &= 2\left[\frac{3}{2}\theta + \sin 2\theta + \frac{1}{8}\sin 4\theta\right]_0^{\pi/2} = 2\left(\frac{3}{2}\right)\left(\frac{\pi}{2}\right) = \boxed{\frac{3\pi}{2}} \end{aligned}$$

Problem 4.3.6. ***CHALLENGE PROBLEM***. *The Gaussian integral, $\int_{-\infty}^{\infty} e^{-x^2}dx$ is a powerful integral used in many aspects of mathematics and engineering. It is particularly used in probability theory (for normal distributions) and quantum mechanics. Using Fubini's Theorem and polar coordinates, prove that the value of this integral is $\sqrt{\pi}$.*

Solution: We wish to evaluate $I = \int_{-\infty}^{\infty} e^{-x^2}dx$. Since there is no elementary antiderivative for e^{-x^2}, we will use a multivariable approach. First, we consider the result of squaring this integral:

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}dy\right)$$

Don't get confused by the y terms. It's just a dummy variable, which is independent from x, allowing us to form a double integral.

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-x^2}e^{-y^2}\,dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy$$

According to Fubini's Theorem, the product of two convergent single integrals is equal to the corresponding double integral over the plane $\mathbb{R}^2$.

We now convert to polar coordinates. Set

$$x = r\cos\theta \quad y = r\sin\theta \quad r \geq 0 \quad 0 \leq \theta \leq 2\pi$$

We know that $dx\,dy = dA = r\,dr\,d\theta$ and $x^2 + y^2 = r^2$ so the double integral is

$$\begin{aligned}
I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2}\,dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2}\,dy\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(x^2+y^2)}\,dx\,dy \\
&= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,r\,dr\,d\theta = \int_0^{2\pi} d\theta \int_0^{\infty} re^{-r^2}\,dr \\
&= \int_0^{2\pi} d\theta \int_0^{\infty} e^{-u}\,\frac{du}{2} \quad (u = r^2,\, du = 2r\,dr) \\
&= \int_0^{2\pi} d\theta \cdot \frac{1}{2}\int_0^{\infty} e^{-u} = \int_0^{2\pi} d\theta \cdot \frac{1}{2}\left[-e^{-u}\right]_0^{\infty} \\
&= \int_0^{2\pi} d\theta \cdot \frac{1}{2}(0-(-1)) = \int_0^{2\pi} d\theta \cdot \frac{1}{2}\cdot 1 \\
&= \frac{1}{2}\cdot 2\pi = \pi
\end{aligned}$$

We finally take the square root of I^2, to obtain I.

$$I = \sqrt{I^2} = \sqrt{\pi}$$

and the proof is complete. □

4.4 Double Integrals: Applications

The most basic application of double integrals is computing volumes of solids. In this section, we will explore real-world, practical uses of double integrals, such as computing mass, electric charge, center of mass, and moment of inertia. We will also observe the importance of multiple integrals when applied to probability density functions of random variables.

Density and Mass Consider a lamina with variable density. Let it occupy a region D of the xy-plane and its **density** (mass per unit area) at a point (x, y) in D is denoted by $\rho(x, y)$, where ρ is continuous on D. We have

$$\rho(x, y) = \lim \frac{\Delta m}{\Delta A}$$

where Δm and ΔA are the mass and area (respectively) of an infinitesimally small rectangle containing the point (x, y), and the limit is taken as the dimensions of the rectangle approach 0.

In order to find the total mass m of the lamina, we divide a rectangle R into smaller sub-rectangles R_{ij} of equal size, and consider $\rho(x, y) = 0$ outside the boundary of D. Let's choose a point (x_{ij}^*, y_{ij}^*) in R_{ij}. Then the mass of that part of the lamina is approximately $\rho\left(x_{ij}^*, y_{ij}^*\right)\,\Delta A$ where ΔA is the area of R_{ij}. If we find the sum of all such masses to approximate the total mass:

$$m \approx \sum_{i=1}^{k}\sum_{j=1}^{l} \rho\left(x_{ij}^*, y_{ij}^*\right)\,\Delta A$$

Let the number of sub-rectangles increase as $k, l \to \infty$. We then obtain the total *exact* mass of the lamina as

$$\boxed{m = \iint_D \rho\,(x, y)\; dA}$$

In physics, we consider other forms of density that can be treated in a similar manner. For instance, if we distribute electric charge over a region D and the charge density (charge per unit area) is given by $\sigma\,(x, y)$ at a point (x, y) in D, then the total **electric charge** Q is given by the double integral

$$\boxed{Q = \iint_D \sigma\,(x, y)\; dA}$$

Problem 4.4.1. *Charge is distributed over the triangular region D with vertices $(0, 1)$, $(1, 1)$, and $(1, 0)$. Additionally, the charge density at any point (x, y) inside D is given by $\sigma\,(x, y) = xy$, measured in coulombs per square meter (C/m^2). Calculate the total electric charge.*

Solution: If you sketch the description of D in the xy-plane and treat x and y as the independent and dependent variables, respectively, it is very easy to see that as x varies from 0 to 1, y varies from $1 - x$ to 1. Thus, we just need to set up and evaluate the double integral for the charge Q:

$$\begin{aligned} Q &= \iint_D \sigma\,(x, y)\; dA = \int_0^1 \int_{1-x}^1 xy\, dy\, dx = \int_0^1 \left[x\frac{y^2}{2}\right]_{y=1-x}^{y=1} dx = \int_0^1 \frac{x}{2}[1^2 - (1-x)^2]\, dx \\ &= \frac{1}{2}\int_0^1 (2x^2 - x^3)\, dx = \frac{1}{2}\left[\frac{2x^3}{3} - \frac{x^4}{4}\right] = \frac{5}{24} \end{aligned}$$

The total charge is $\boxed{\frac{5}{24}\,\text{C}}$.

Moments and Center of Mass We already know that if we cut up a lamina occupying a region D into sub-rectangles R_{ij}, then the mass of R_{ij} is approximately $\rho\left(x_{ij}^*, y_{ij}^*\right)\,\Delta A$, so we can approximate the moment of R_{ij} with respect to the x-axis by

$$[\rho\left(x_{ij}^*, y_{ij}^*\right)\,\Delta A]y_{ij}^*$$

Add these quantities and take the limit as the number of sub-rectangles approaches infinity. In this way, we obtain the **moment** of the entire lamina **about the x-axis**:

$$\boxed{M_x = \iint_D y\,\rho\,(x, y)\; dA}$$

Similarly, the **moment about the y-axis** is

$$\boxed{M_y = \iint_D x\,\rho\,(x, y)\; dA}$$

Now, we define the center of mass with coordinates $(\bar{x}, \bar{y})$ such that $m\bar{x} = M_y$ and $m\bar{y} = M_x$. Why is this physically significant? Because the lamina behaves as if all its mass is concentrated at this special point. In other words, if you support a lamina at its center of mass, then it will *always* balance horizontally.

Definition 4.4.1. *The coordinates* $(\bar{x}, \bar{y})$ *of the* ***center of mass*** *of a lamina occupying the region* D *and having density function* $\rho(x, y)$ *are*

$$\bar{x} = \frac{M_y}{m} = \frac{1}{m} \iint_D x\,\rho(x, y)\, dA \qquad \bar{y} = \frac{M_x}{m} = \frac{1}{m} \iint_D y\,\rho(x, y)\, dA$$

where the mass m *is known as*

$$m = \iint_D \rho(x, y)\, dA$$

Problem 4.4.2. *The density at any point on a semicircular lamina of radius* a *is proportional to the distance from the center of the circle. Find the center of mass of the lamina.*

Solution: Our best bet would be to position the lamina on the upper half of the circle $x^2 + y^2 = a^2$. Then the distance from any point (x, y) in the region to the center of the circle (the origin) would be $x^2 + y^2$. Our density function is therefore

$$\rho(x, y) = k\sqrt{x^2 + y^2}$$

for some constant k. Because of the density function and the shape of the lamina, it is most convenient to use polar coordinates. Thus, $\sqrt{x^2 + y^2} = r$ and the region D is given by $0 \le r \le a$, $0 \le \theta \le \pi$. We need to first find the mass of the lamina:

$$\begin{aligned} m &= \iint_D \rho(x, y)\, dA = \iint_D k\sqrt{x^2 + y^2}\, dA \\ &= \int_0^\pi \int_0^a (kr)\, r\, dr\, d\theta = k \int_0^\pi d\theta \int_0^a r^2\, dr = k\pi \left[\frac{r^3}{3}\right]_0^a = \frac{k\pi a^3}{3} \end{aligned}$$

Note that the lamina and the density function are both symmetric about the x-axis, so we must have $\bar{x} = 0$. The y-coordinate is given by

$$\begin{aligned} \bar{y} &= \frac{1}{m} \iint_D \rho(x, y)\, dA = \frac{3}{k\pi a^3} \int_0^\pi \int_0^a r\sin\theta\,(kr)\, r\, dr\, d\theta \\ &= \frac{3}{\pi a^3} \int_0^\pi \sin\theta\, d\theta \int_0^a r^3\, dr = \frac{3}{\pi a^3} [-\cos\theta]_0^\pi \left[\frac{r^4}{4}\right]_0^a \\ &= \frac{3}{\pi a^3} \frac{2a^4}{4} = \frac{3a}{2\pi} \end{aligned}$$

Therefore, the center of mass is located at the point $\boxed{(0, 3a/2\pi)}$.

Moment of Inertia The **moment of inertia**, also called the second moment, of a point particle with mass m is given by mr^2, where r is the distance from the particle to the rotating axis. We extend our object of interest to a lamina with density function $\rho(x, y)$ occupying a region D.

Let's divide D into smaller sub-rectangles as expected, and take the limit of the sum as the number of sub-rectangles approaches infinity. The result is the **moment of inertia** of the lamina **about the** x**-axis**:

$$\boxed{I_x = \lim_{m,n\to\infty} \sum_{i=1}^{m} \sum_{j=1}^{n} (y_{ij}^*)^2 \rho\left(x_{ij}^*, y_{ij}^*\right)\, \Delta A = \iint_D y^2 \rho(x, y)\, dA}$$

Similarly, the **moment of inertia about the** y**-axis** is

$$\boxed{I_y = \lim_{m,n\to\infty} (x_{ij}^*)^2 \rho\left(x_{ij}^*, y_{ij}^*\right)\, \Delta A = \iint_D x^2 \rho(x, y)\, dA}$$

Finally, we consider the **moment of inertia about the origin**, also called the **polar moment of inertia**, as the sum of I_x and I_y:

$$\begin{aligned} I_0 &= I_x + I_y \\ &= \boxed{\iint_D (x^2 + y^2)\rho\,(x, y)\ dA} \end{aligned}$$

Problem 4.4.3. *Let D be a homogeneous disk with density $\rho\,(x, y) = \rho$, center the origin, and radius a. Find the moments of inertia I_x, I_y, and I_0.*

Solution: The boundary of D is the circle $x^2 + y^2 = a^2$. In polar coordinates, we have the region $0 \le \theta \le 2\pi$, $0 \le r \le a$. We proceed with the following:

$$\begin{aligned} I_x &= \iint_D y^2 \rho\, dA = \rho \int_0^{2\pi} \int_0^a (r \sin\theta)^2\, r\, dr\, d\theta \\ &= \rho \int_0^{2\pi} \sin^2\theta\, d\theta \int_0^a r^3\, dr = \rho \int_0^{2\pi} \frac{1}{2}(1 - \cos 2\theta)\, d\theta \int_0^a r^3\, dr \\ &= \frac{\rho}{2} \left[\theta - \frac{1}{2}\sin 2\theta\right]_0^{2\pi} \left[\frac{r^4}{4}\right]_0^a = \boxed{\frac{1}{4}\pi\rho a^4} \end{aligned}$$

$$\begin{aligned} I_y &= \iint_D x^2 \rho\, dA = \rho \int_0^{2\pi} \int_0^a (r \cos\theta)^2\, r\, dr\, d\theta \\ &= \rho \int_0^{2\pi} \frac{1}{2}(1 + \cos 2\theta)\, d\theta \int_0^a r^3\, dr = \boxed{\frac{1}{4}\pi\rho a^4} \end{aligned}$$

Note that $I_x = I_y$ due to the symmetry of the problem. We could either use the double integral for I_0, or compute it as

$$\begin{aligned} I_0 &= I_x + I_y \\ &= \frac{1}{4}\pi\rho a^4 + \frac{1}{2}\pi\rho a^4 \\ &= \boxed{\frac{1}{2}\pi\rho a^4} \end{aligned}$$

Let's notice something here. The mass of the disk is

$$m = \text{density} \times \text{area} = \rho\left(\pi a^2\right)$$

so the moment of inertia of the disk about the origin (its axis) can be expressed as

$$I_0 = \frac{\pi\rho a^4}{2} = \frac{1}{2}\left(\rho\pi a^2\right) a^2 = \frac{1}{2} m a^2$$

This equation suggests: if we increase the mass or radius of the disk, its moment of inertia also increases. Let's think about this in a physics perspective; the moment of inertia of a wheel is what makes it difficult to start or stop its rotation, just as a car's mass is the metric which makes it difficult to start or stop its motion.

Definition 4.4.2. *The* ***radius of gyration of a lamina about an axis*** *is the number R such that*

$$mR^2 = I$$

where m and I are the mass and moment of inertia about the given axis, respectively.

Important: If the mass of the lamina were concentrated at a distance R from the axis, then the moment of inertia of this "point mass" would be numerically equal to the moment of inertia of the lamina. In particular, the radius of gyration $\bar{\bar{y}}$ with respect to the x-axis and the radius of gyration $\bar{\bar{x}}$ with respect to the y-axis are given by the below equations:

$$m\bar{\bar{y}}^2 = I_x \qquad m\bar{\bar{x}}^2 = I_y$$

Thus, the point with coordinates $(\bar{\bar{x}}, \bar{\bar{y}})$ is where the mass of the lamina can be concentrated without changing its moment of inertia with respect to the coordinate axes.

Problem 4.4.4. *From the previous problem, find the radius of gyration about the x-axis of the disk.*

Solution: As noted, the mass of the disk is $m = \rho\pi a^2$, so

$$\bar{\bar{y}}^2 = \frac{\frac{1}{4}\pi\rho a^4}{\pi\rho a^2} = \frac{a^2}{4}$$

Therefore, the radius of gyration about the x-axis is $\boxed{\bar{\bar{y}} = a/2}$, which is half the radius of the disk.

Probability Density Consider a pair of continuous random variables X and Y.

Definition 4.4.3. *The* ***joint density function*** *of X and Y is a function f of two variables such that the probability that (X, Y) lies in a region D is*

$$P((X,Y) \in D) = \iint_D f(x,y)\, dA$$

In particular, if D is a rectangle, then the probability that X lies between a and b and Y lies between c and d is

$$P(a \leq X \leq b,\, c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)\, dy\, dx$$

Because probability values are non-negative, and are measured on a scale from 0 to 1, there are two properties that exist for all probability density functions:

$$f(x,y) \geq 0 \qquad \iint_{\mathbb{R}^2} f(x,y)\, dA = 1$$

where the double integral over $\mathbb{R}^2$ is an improper integral defined as the limit of double integrals over expanding circles or squares, and we can write

$$\iint_{\mathbb{R}^2} f(x,y)\, dA = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dy\, dx$$

Problem 4.4.5. *If the joint density function for X and Y is*

$$f(x,y) = \begin{cases} C(x+2y), & \text{if } 0 \leq x \leq 10,\ 0 \leq y \leq 10, \\ 0, & \text{otherwise} \end{cases}$$

find the value of the constant C. Then calculate $P(X \leq 5, Y \geq 3)$.

Solution: We can find the value of C by setting the value of the double integral over $\mathbb{R}^2$ to be 1. Since $f(x,y) = 0$ outside the rectangle $R = [0,10] \times [0,10]$, we get

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dy\, dx = \int_0^{10} \int_0^{10} C(x+2y)\, dy\, dx = C\int_0^{10} [xy + y^2]_{y=0}^{y=10}\, dx$$
$$= C\int_0^{10} (10x + 100)\, dx = 1500C$$

Since this quantity must equal 1, we have

$$1500C = 1 \therefore C = \boxed{\frac{1}{1500}}$$

Now we can compute the probability that X is at most 5 and Y is at least 3:

$$\begin{aligned} P(X \le 5, Y \ge 3) &= \int_{-\infty}^{5}\int_{3}^{\infty} f(x,y)\,dy\,dx = \int_0^5\int_3^{10} \frac{1}{1500}(x+2y)\,dy\,dx \\ &= \frac{1}{1500}\int_0^5 [xy+y^2]_{y=3}^{y=10}\,dx = \frac{1}{1500}\int_0^5 (7x+91)\,dx \\ &= \frac{1}{1500}\cdot\left[\frac{7}{2}x^2 + 91x\right]_0^5 \\ &= \frac{\left(\frac{1085}{2}\right)}{1500} \approx \boxed{0.3617} \end{aligned}$$

Suppose X is a random variable with probability density function $f_1(x)$ and Y is a random variable with probability density function $f_2(y)$. Then X and Y are called **independent random variables** if their joint density function is equal to the product of the individual density functions, or

$$f(x,y) = f_1(x)f_2(y)$$

Note: This topic is not really covered in most introductory calculus classes, so I'll introduce it in terms of one variable before moving on to the multivariable aspect.

We can model waiting times (e.g. in a queue or shopping line) by using the exponential density functions

$$f(t) = \begin{cases} 0, & \text{if } t < 0, \\ \mu^{-1}e^{-t/\mu}, & \text{if } t \ge 0. \end{cases}$$

where μ is the mean waiting time. In the following exercise, we consider a situation with two independent waiting times.

Problem 4.4.6. *The manager of a movie theater determines that the average time moviegoers wait in line to buy a ticket for a film is* 10 *minutes and the average time they wait to buy popcorn is* 5 *minutes. Assuming that the waiting times are independent, find the probability that a moviegoer waits a total of less than* 20 *minutes before taking his or her seat.*

Solution: Since both the waiting time X for the ticket purchase and the waiting time Y in the refreshment line are random variables modeled by exponential density functions, we have

$$f_1(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{10}e^{-x/10} & \text{if } x \ge 0 \end{cases}$$

$$f_2(y) = \begin{cases} 0 & \text{if } y < 0 \\ \frac{1}{5}e^{-y/5} & \text{if } y \ge 0 \end{cases}$$

Since X and Y are independent random variables, the joint density function is

$$f(x,y) = f_1(x)f_2(y) = \begin{cases} \frac{1}{50}e^{-x/10}e^{-y/5} & \text{if } x \ge 0,\ y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

We want the probability that $X + Y < 20$:

$$P(X + Y < 20) = P((X, Y) \in D)$$

where D is the triangular region with vertices $(0, 0)$, $(0, 20)$, and $(20, 0)$. As x varies from 0 to 20, the y-coordinate changes from $y = 0$ to $y = 20 - x$ (this is the equation of the hypotenuse of the triangle!). Thus

$$\begin{aligned}
P(X + Y < 20) &= \iint_D f(x, y)\, dA = \int_0^{20} \int_0^{20-x} \frac{1}{50} e^{-x/10} e^{-y/5}\, dy\, dx \\
&= \frac{1}{50} \int_0^{20} [e^{-x/10}(-5)e^{-y/5}]_{y=0}^{y=20}\, dx = \frac{1}{10} \int_0^{20} e^{-x/10}(1 - e^{(x-20)/5})\, dx \\
&= \frac{1}{10} \int_0^{20} (e^{-x/10} - e^{-4} e^{x/10})\, dx \\
&\approx \boxed{0.7476}
\end{aligned}$$

Almost three-fourths of all moviegoers wait a total of less than 20 minutes before taking their seats.

Expected Values In single-variable calculus, if X is a random variable with probability density function f, then its *mean* is given by

$$\mu = \int_{-\infty}^{\infty} x f(x)\, dx$$

If we have two independent random variables X and Y with joint density function $f(x, y)$, we define the **X-mean** and **Y-mean**, also called the *expected values* of X and Y, to be

$$\mu_1 = \iint_{\mathbb{R}^2} x\, f(x, y)\, dA \qquad \mu_2 = \iint_{\mathbb{R}^2} y\, f(x, y)\, dA$$

Notice how similar these formulas look to the calculations for moment and moment of inertia for a lamina with density ρ?

In the next and final exercise of this section, we will deal with normal distributions. In univariate (one-variable) statistics, a random variable is **normally distributed** if its probability density function has the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$$

where μ and σ are the mean and standard deviation, respectively.

Problem 4.4.7. *A factory produces cylinder-shaped roller bearings that are sold as having diameter* 4.0 *cm and length* 6.0 *cm. Actually the diameters X are normally distributed with mean* 4.0 *cm and standard deviation* 0.01 *cm while the lengths Y are normally distributed with mean* 6.0 *cm and standard deviation* 0.01 *cm. Assuming that X and Y are independent random variables, find the probability that a bearing randomly chosen from the production line has either length or diameter that differs from the mean by more than* 0.02 *cm.*

Solution: We know X and Y are normally distributed with $\mu_1 = 4.0$, $\mu_2 = 6.0$, and $\sigma_1 = \sigma_2 = 0.01$. So the individual density functions for X and Y are

$$f_1(x) = \frac{1}{0.01\sqrt{2\pi}} e^{-(x-4)^2/0.0002} \qquad f_2(y) = \frac{1}{0.01\sqrt{2\pi}} e^{-(y-6)^2/0.0002}$$

Since X and Y are independent, the joint density is the product of the two functions:

$$\begin{aligned}
f(x, y) = f_1(x) f_2(y) &= \frac{1}{0.0002\pi} e^{-(x-4)^2/0.0002} e^{-(y-6)^2/0.0002} \\
&= \frac{5000}{\pi} e^{-5000[(x-4)^2 + (y-6)^2]}
\end{aligned}$$

It is actually easier to first calculate the probability that both X and Y differ from their means by less than 0.02 cm. We can use a calculator to estimate the integral:

$$\begin{aligned} P(3.98 < X < 4.02, 5.98 < Y < 6.02) &= \int_{3.98}^{4.02} \int_{5.98}^{6.02} f(x,y)\,dy\,dx \\ &= \frac{5000}{\pi} \int_{3.98}^{4.02} \int_{5.98}^{6.02} e^{-5000[(x-4)^2+(y-6)^2]}\,dy\,dx \\ &\approx 0.91 \end{aligned}$$

This is the complement of our desired answer, so the probability that either X or Y differs from its mean by more than 0.02 cm is approximately $1 - 0.91 \approx \boxed{0.09}$.

4.5 Surface Area

In this section we will explore the sole application of double integrals, asides from area and volume interpretations. We will explore the concept of surface area of a surface bounded by $z = f(x, y)$ where $(x, y) \in D$, some region in the xy-plane. Make sure you realize this idea is not restricted to just the xy-plane, though!

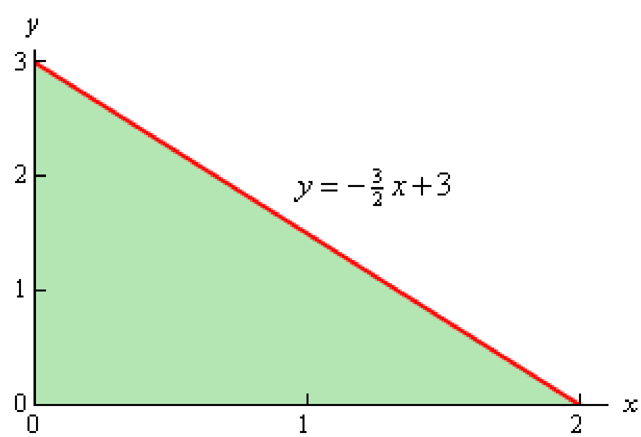
Definition 4.5.1. *The* ***surface area*** *of the surface $z = f(x, y)$ over the region D is given by the double integral*

$$\iint_D \sqrt{[f_x]^2 + [f_y]^2 + 1}\,dA$$

Try these problems below.

Problem 4.5.1. *Find the surface area of the part of the plane $3x + 2y + z = 6$ that lies in the first octant.*

Solution: The first octant is the portion of the xyz-coordinate system in which all three variables are positive. The sketch of the region D is shown below. We can determine this by setting $z = 0$ for the xy-plane, which will give us the equation for the upper bound on the y-coordinate.



In order to use the surface area formula, we need to solve for z in terms of x and y, and also gather the relevant partial derivatives.

$$z = 6 - 3x - 2y \qquad f_x = -3 \qquad f_y = -2$$

The limits defining the region D are

$$0 \le x \le 2 \qquad 0 \le y \le -\frac{3}{2}x + 3$$

Thus, the surface area is then

$$\begin{aligned} S &= \iint_D \sqrt{(-3)^2 + (-2)^2 + 1}\, dA \\ &= \int_0^2 \int_0^{-\frac{3}{2}x+3} \sqrt{14}\, dy\, dx \\ &= \sqrt{14} \int_0^2 \left(-\frac{3}{2}x + 3\right) dx \\ &= \sqrt{14} \left(-\frac{3}{4}x^2 + 3x\right) \Big|_0^2 \\ &= \boxed{3\sqrt{14}} \end{aligned}$$

Problem 4.5.2. *Determine the surface area of the part of $z = xy$ that lies within the cylinder with equation $x^2 + y^2 = 1$.*

Solution: In this case, $z = f(x, y)$ has already been provided to us, so we can find the partial derivatives $f_x = y$ and $f_y = x$. Also, the plane region D inside the provided cylinder will be a disk of radius 1 centered at the origin.

The proper integral for the surface area is then

$$S = \iint_D \sqrt{x^2 + y^2 + 1}\, dA$$

As D is a disk, it is most convenient to use polar coordinates. We proceed with

$$\begin{aligned} S &= \iint_D \sqrt{x^2 + y^2 + 1}\, dA \\ &= \int_0^{2\pi} \int_0^1 r\sqrt{1 + r^2}\, dr\, d\theta \\ &= \int_0^{2\pi} \frac{1}{2}\left(\frac{2}{3}\right)(1 + r^2)^{3/2}\Big|_0^1\, d\theta \\ &= \int_0^{2\pi} \frac{1}{3}\left(2^{3/2} - 1\right) d\theta \\ &= \boxed{\frac{2\pi}{3}\left(2^{3/2} - 1\right)} \end{aligned}$$

4.6 Triple Integrals: In Essence

The concept of integrating over two-dimensional regions was explored in section 4.2: with double integrals. Now, we will extend to three dimensions. As you might expect, we will use a triple integral, and the proper notation (let E be some three-dimensional region) is

$$\iiint_E f(x, y, z)\, dV$$

We can start simple by integrating over some rectangular box

$$B = [a, b] \times [c, d] \times [r, s]$$

Using this notation, the order follows as x, then y, then z. In that case, the triple integral becomes

$$\iiint_B f(x,y,z)\,dV = \int_r^s \int_c^d \int_a^b f(x,y,z)\,dz\,dy\,dx$$

Note: There are six different possible orders to compute the triple integral, and which order you do the integral in will depend upon the function and the order that you feel will be simplest. **We will get the same answer regardless of the order**.

Let's practice!

Problem 4.6.1. *Evaluate* $\iiint_B 8xyz$ *where* $B = [2,3] \times [1,2] \times [0,1]$.

Solution: The standard convention is to do x first, then y, then z, but just to make the point that the order is irrelevant, we'll go z first, then x, then y. Our differential volume element will be $dV = dz\,dx\,dy$.

$$\begin{aligned}
\iiint_B 8xyz\,dV &= \int_1^2 \int_2^3 \int_0^1 8xyz\,dz\,dx\,dy \\
&= \int_1^2 \int_2^3 4xyz^2|_0^1\,dx\,dy \\
&= \int_1^2 \int_2^3 4xy\,dx\,dy \\
&= \int_1^2 2x^2y|_2^3\,dy \\
&= \int_1^2 10y\,dy = \boxed{15}
\end{aligned}$$

Before we discuss broader areas, let's first understand a clear geometric picture of the triple integral. This will help us in some upcoming examples.

Definition 4.6.1. *The volume of the three-dimensional region E is given by the integral,*

$$V = \iiint_E 1\,dV$$

Now, let's talk about the broader three-dimensional spaces. There are three different options for a general area. Here's a drawing of the first option.

Type I Regions Define a region E such that $E = \{(x,y,z) \mid (x,y) \in D, u_1(x,y) \le z \le u_2(x,y)\}$ where $(x,y) \in D$ indicates the point (x,y) lies in the xy-plane region D. The triple integral follows as

$$\iiint_E f(x,y,z)\,dV = \iint_D \left[\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz\right] dA$$

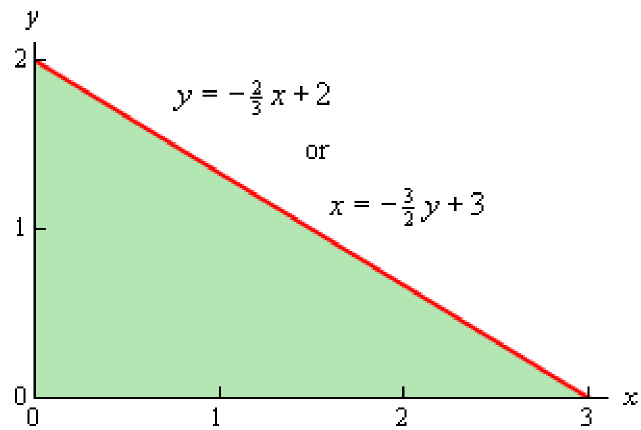
where the double integral can be evaluated through any of the methods we went over previously.

Problem 4.6.2. *Evaluate* $\iiint_E 2x\,dV$ *where E is the region under the plane with equation* $2x + 3y + z = 6$ *lying in the first octant.*

Solution: Before we proceed, we should understand the idea of an *octant*. In 2D, the coordinate system could be divided into $2^2 = 4$ quadrants. Similarly, the 3D coordinate system is divided into $2^3 = 8$ octants. In the first octant, all of x, y, and z are positive. Now, the key detail is identifying the region D in the xy-plane. The easiest way to do this is to actually find the intercepts of the plane. To find the x intercept, set y and z to 0, to find the y intercept, set x and z to 0, and so on.

$$2x + 3(0) + (0) = 6 \therefore (3, 0, 0)$$
$$2(0) + 3y + (0) = 6 \therefore (0, 2, 0)$$
$$2(0) + 3(0) + z = 6 \therefore (0, 0, 6)$$

When we project the surface in the xy-plane, we essentially set the z coordinate to 0. Thus, D becomes the triangle with vertices of $(3, 0)$, $(0, 2)$, and $(0, 0)$. Here's a sketch:



Since we are under the plane and in the first octant, the boundaries for z are

$$0 \le z \le 6 - 2x - 3y$$

We can calculate the double integral over D by using either the following two sets of inequalities:

$$0 \le x \le 3 \qquad 0 \le x \le -\frac{3}{2}y + 3$$

$$0 \le y \le -\frac{2}{3}x + 2 \qquad 0 \le y \le 2$$

Since neither holds any concrete advantage, let's just use the first one.

$$\begin{aligned}
\iiint_E 2x\,dV &= \iint_D \left[\int_0^{6-2x-3y} 2x\,dz\right] dA \\
&= \iint_D 2xz|_0^{6-2x-3y}\,dA \\
&= \int_0^3 \int_0^{-\frac{2}{3}x+2} 2x(6-2x-3y)\,dy\,dx \\
&= \int_0^3 (12xy - 4x^2y - 3xy^2)|_0^{-\frac{2}{3}x+2}\,dx \\
&= \int_0^3 \left(\frac{4}{3}x^3 - 8x^2 + 12x\right) dx \\
&= \left(\frac{1}{3}x^4 - \frac{8}{3}x^3 + 6x^2\right)|_0^3 \\
&= \boxed{9}
\end{aligned}$$

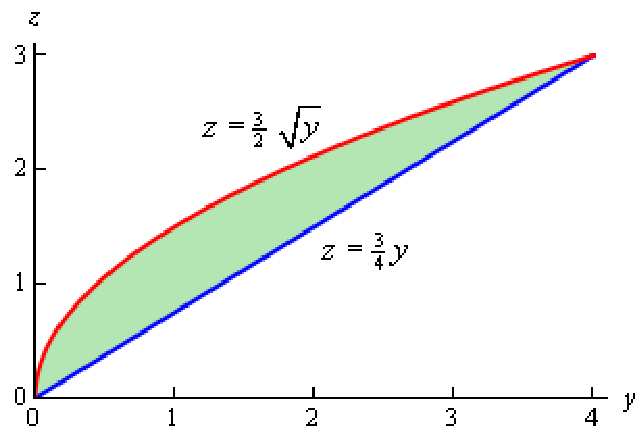
Type II Regions Define a region E such that $E = \{(x, y, z) \,|\, (y, z) \in D, u_1(y, z) \leq x \leq u_2(y, z)\}$ where $(y, z) \in D$ indicates the point (y, z) lies in the yz-plane region D. The triple integral follows as

$$\iiint_E f(x, y, z)\, dV = \iint_D \left[\int_{u_1(y,z)}^{u_2(y,z)} f(x, y, z)\, dx\right] dA$$

Of course, for the double integral, we can choose any method as we wish, including polar coordinates if optimal.

Problem 4.6.3. *Determine the volume of the region that lies behind the plane $x + y + z = 8$ and in front of the region in the yz-plane that is bounded by $z = \frac{3}{2}\sqrt{y}$ and $z = \frac{3}{4}y$.*

Solution: This time, they actually told us the region D, so we don't really need much effort to figure that out. Here's a quick sketch of the curves defining D projected out past the plane:



The limits for each variable are:

$$0 \leq y \leq 4$$
$$\frac{3}{4}y \leq z \leq \frac{3}{2}\sqrt{y}$$
$$0 \leq x \leq 8 - y - z$$

Thus, the volume of the region is simply

$$\begin{aligned}
V = \iiint_E 1\, dV &= \iint_D \left[\int_0^{8-y-z} dx\right] dA \\
&= \int_0^4 \int_{3y/4}^{3\sqrt{y}/2} (8 - y - z)\, dz\, dy \\
&= \int_0^4 \left(8z - yz - \frac{1}{2}z^2\right)\Big|_{3y/4}^{3\sqrt{y}/2} dy \\
&= \int_0^4 \left(12y^{1/2} - \frac{57}{8}y - \frac{3}{2}y^{3/2} + \frac{33}{32}y^2\right) dy \\
&= \left(8y^{3/2} - \frac{57}{16}y^2 - \frac{3}{5}y^{5/2} + \frac{11}{32}y^3\right)\Big|_0^4 = \boxed{\frac{49}{5}}
\end{aligned}$$

Type III Regions Define a region E such that $E = \{(x,y,z)\,|\,(x,z) \in D, u_1(x,z) \leq y \leq u_2(x,z)\}$ where $(x,z) \in D$ indicates the point (x,z) lies in the xz-plane region D. The triple integral follows as

$$\iiint_E f(x,y,z)\,dV = \iint_D \left[\int_{u_1(x,z)}^{u_2(x,z)} f(x,y,z)\,dy\right] dA$$

where we will use either of the two possible orders for integrating over D in the xz-plane or we can use polar coordinates if needed.

Problem 4.6.4. *Evaluate* $\iiint_E \sqrt{3x^2+3z^2}\,dV$ *where* E *is the solid bounded by* $y = 2x^2+2z^2$ *and the plane* $y = 8$.

Solution: The solid E is a paraboloid, with cross sections perpendicular to the base as circles. The region D in the xz-plane can be determined by setting the y-coordinate to 0, essentially, setting $2x^2 + 2z^2 = 8$. Thus, the region D is $x^2 + z^2 = 4$. This equation hints that we should use polar coordinates. Don't think we can only use them in the xy-plane! We can simply translate them here; the equations are

$$x = r\cos\theta \qquad z = r\sin\theta$$

This leads to $x^2 + z^2 = r^2$, and we are now ready for polar coordinates.

Keeping this in mind, we arrive at the following limits of integration:

$$2x^2 + 2z^2 \leq y \leq 8$$
$$0 \leq r \leq 2$$
$$0 \leq \theta \leq 2\pi$$

Thus we obtain

$$\begin{aligned}\iiint_E \sqrt{3x^2+3z^2}\,dV &= \iint_D \left[\int_{2x^2+2z^2}^{8} \sqrt{3x^2+3z^2}\,dy\right] dA \\ &= \iint_D \left(y\sqrt{3x^2+3z^2}\right)\Big|_{2x^2+2z^2}^{8}\,dA \\ &= \iint_D \sqrt{3(x^2+z^2)}\left(8-(2x^2+2z^2)\right)\,dA\end{aligned}$$

In order to compute the double integral, we first convert everything into polar coordinates. Therefore, the integrand becomes

$$\begin{aligned}\sqrt{3(x^2+z^2)}\left(8-(2x^2+2z^2)\right) &= \sqrt{3r^2}\left(8-2r^2\right) \\ &= \sqrt{3}r\left(8-2r^2\right) \\ &= \sqrt{3}\left(8r-2r^3\right)\end{aligned}$$

and we finally evaluate the integral:

$$\begin{aligned}\iiint_E \sqrt{3x^2+3z^2}\,dV &= \iint_D \sqrt{3}\left(8r-2r^3\right)\,dA \\ &= \sqrt{3}\int_0^{2\pi} \left(\frac{8}{3}r^3 - \frac{2}{5}r^5\right)\Big|_0^2 d\theta \\ &= \sqrt{3}\int_0^{2\pi} \frac{128}{15}\,d\theta \\ &= \boxed{\frac{256\sqrt{3}\pi}{15}}\end{aligned}$$

4.7 Triple Integrals: Converting to Cylindrical Coordinates

As we saw in section 4.3, certain curves and regions are easier to work with for polar coordinates. For example, if a point P has Cartesian coordinates (x, y) and polar coordinates (r, θ), then

$$x = r\cos\theta \qquad y = r\sin\theta$$

$$r^2 = x^2 + y^2 \qquad \tan\theta = \frac{y}{x}$$

In three dimensions, we extend this to the *cylindrical coordinate* system, which is similar to polar form and gives us convenient representations of common solids and surfaces.

Cylindrical Coordinate System In this system, a point P is represented by an ordered triple (r, θ, z), where r and θ are the polar coordinates of the projection of P onto the xy-plane (where $z = 0$). and z is the distance from the xy-plane to P. To convert from cylindrical to rectangular coordinates, we use

$$x = r\cos\theta \qquad y = r\sin\theta \qquad z = z$$

and for vice versa, we get

$$r^2 = x^2 + y^2 \qquad \tan\theta = \frac{y}{x} \qquad z = z$$

Here is a point P in cylindrical coordinates:

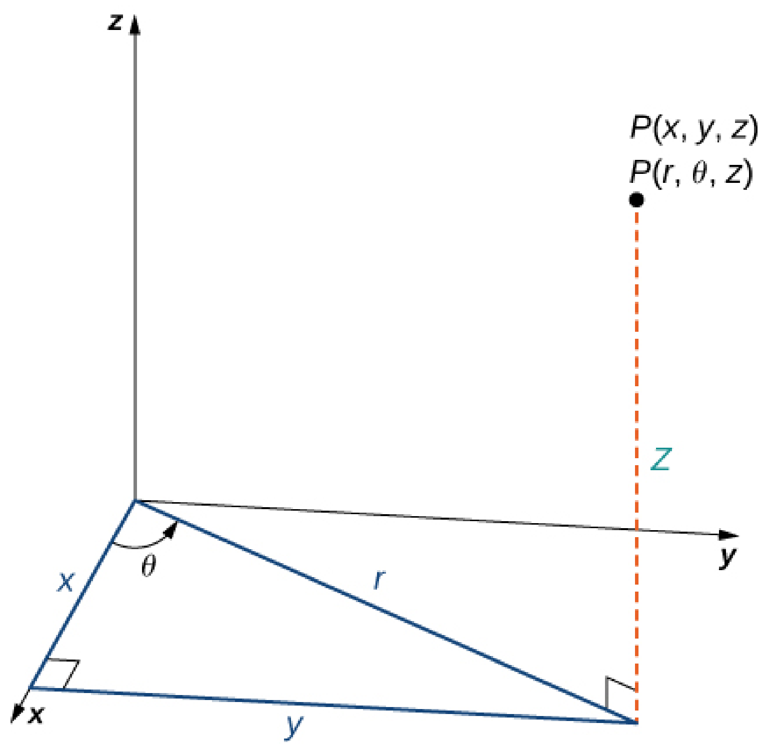


Image Credit: Lumen Learning

Problem 4.7.1. *Find the rectangular (Cartesian) coordinates of the point with cylindrical coordinates* $(2, 2\pi/3, 1)$.

Solution: We will apply $x = r\cos\theta$ and $y = r\sin\theta$ to find the x- and y-coordinates of the point. Also, $z = z$, so this coordinate stays the same.

$$x = 2\cos\frac{2\pi}{3} = 2\left(-\frac{1}{2}\right) = -1$$
$$y = 2\sin\frac{2\pi}{3} = 2\left(\frac{\sqrt{3}}{2}\right) = \sqrt{3}$$
$$z = 1$$

Thus, the point in rectangular coordinates is $\boxed{(-1, \sqrt{3}, 1)}$.

Problem 4.7.2. *Find the cylindrical coordinates of the point with rectangular coordinates* $(3, -3, -7)$.

Solution: Since $x > 0$ and $y < 0$, we note that θ is in quadrant IV of the xy-plane.

$$r = \sqrt{3^2 + (-3)^2} = 3\sqrt{2}$$
$$\tan\theta = \frac{-3}{3} = -1 \therefore \theta = \frac{7\pi}{4} + 2n\pi$$
$$z = -7$$

One set of cylindrical coordinates is $\boxed{(3\sqrt{2}, 7\pi/4, -7)}$. Another could be $(3\sqrt{2}, -\pi/4, -7)$. Similar to polar coordinates, there are infinitely many choices.

It's most useful to apply cylindrical coordinates when solving problems involving symmetry about an axis. In most problems, you will find that the z-axis coincides with the axis of symmetry. For example, the cylinder $x^2 + y^2 = c^2$ is symmetric about the z-axis. If you convert to polar form, and take the square root, then in cylindrical coordinates this results in $r = c$. Additionally, for your information, the equation $\theta = c$ is a vertical plane through the origin and the equation $z = c$ is a horizontal plane perpendicular to the z-axis.

Problem 4.7.3. *Describe the surface whose equation in cylindrical coordinates is* $z = r$.

Solution: This equation says the z value, or height, of each point on the surface is equivalent to r, which is the distance from that point to the z-axis. Intuitively, θ is free to vary. Thus, we can claim that any horizontal trace in the plane $z = k$, for $k > 0$, is a circle with radius k. As k continuously increases, these traces suggest that the surface is a $\boxed{\text{cone}}$. Converting to rectangular coordinates, we find

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

which is indeed, a cone symmetric about the z-axis. □

Triple Integrals in Cylindrical Coordinates Let's say we have a type I region E whose projection is D in the xy-plane is conveniently described in polar coordinates. If f is continuous, then the set-builder notation is

$$E = \{(x, y, z) \,|\, (x, y) \in D,\, u_1(x, y) \le z \le u_2(x, y)\}$$

where D in polar coordinates is

$$D = \{(r, \theta) \,|\, \alpha \le \theta \le \beta,\, h_1(\theta) \le r \le h_2(\theta)\}$$

We obtained this formula in section 4.6 for the triple integral over a type I region:

$$\iiint_E f(x, y, z)\, dV = \iint_D \left[\int_{u_1(x,y)}^{u_2(x,y)} f(x, y, z)\, dz\right] dA$$

But since we can express D in polar coordinates, we combine everything to yield

$$\boxed{\iiint_E f(x,y,z)\,dV = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} \int_{u_1(r\cos\theta, r\sin\theta)}^{u_2(r\cos\theta, r\sin\theta)} f(r\cos\theta, r\sin\theta, z)\, r\,dz\,dr\,d\theta}$$

The above formula explains how to evaluate a triple integral in cylindrical coordinates. Essentially, we convert variables using $x = r\cos\theta$, $y = r\sin\theta$, leaving z as it is, with appropriate limits of integration. We then replace the volume element dV with $r\,dz\,dr\,d\theta$. The below image will explain how you can remember this. It will be most effective to use this formula if E is a solid region that can be described as a cylinder or, more generally, a *cylindrical wedge*, and especially if the function $f(x,y)$ involves the term $x^2 + y^2$.

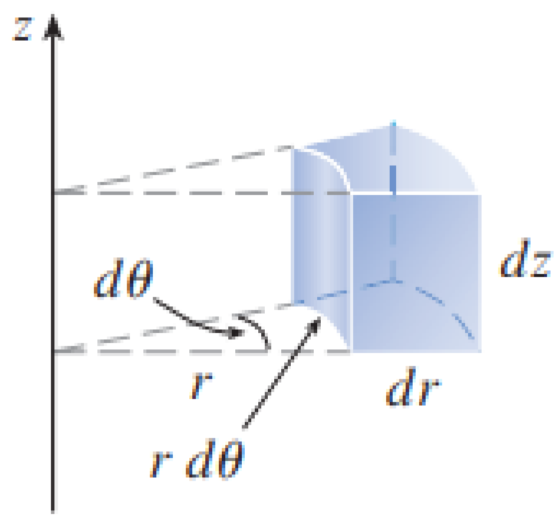


Image Credit: Calculus, Early Transcendentals

Let's do some practice!

Problem 4.7.4. *Evaluate* $\iiint_E x^2\,dV$, *where* E *is the solid lying under the paraboloid* $z = 4 - x^2 - y^2$ *and above the* xy*-plane.*

Solution: The problem statement implies that E is symmetric about the z-axis. Therefore, we will use cylindrical coordinates to solve. We have $z = 4 - x^2 - y^2 = 4 - (x^2 + y^2) = 4 - r^2$ in polar form. The paraboloid intersects the xy-plane at $z = 0$, so $x^2 + y^2 = r^2 = 4$, or equivalently, $r = 2$, so the projection of E onto the xy-plane is the disk $r \le 2$. In set-builder notation,

$$E = \{(r,\theta,z) \,|\, 0 \le \theta \le 2\pi, 0 \le r \le 2, 0 \le z \le 4 - r^2\}$$

so then we get

$$\begin{aligned}\iiint_E x^2\,dV &= \int_0^{2\pi}\int_0^2\int_0^{4-r^2}(r\cos\theta)^2\,r\,dz\,dr\,d\theta\\ &= \int_0^{2\pi}\int_0^2 (r^3\cos^2\theta)(4-r^2)\,dr\,d\theta\\ &= \int_0^{2\pi}\cos^2\theta\,d\theta\int_0^2(4r^3-r^5)\,dr\\ &= \frac{1}{2}\left[\theta+\frac{1}{2}\cos 2\theta\right]_0^{2\pi}\left[r^4-\frac{1}{6}r^6\right]_0^2\\ &= \frac{1}{2}(2\pi)\left(16-\frac{32}{3}\right)\\ &= \boxed{\frac{16}{3}\pi}\end{aligned}$$

Problem 4.7.5. *A solid E lies inside the cylinder $x^2+y^2=1$ to the right of the xz-plane, below the plane $z=4$, and above the paraboloid $z=1-x^2-y^2$. The density at any point is directly proportional to the distance from the axis of the cylinder. Calculate the mass of E.*

Solution: Let's break this down one step at a time. In cylindrical coordinates, the cylinder is $r=1$ and the paraboloid is $z=1-r^2$. Also, because E is restricted to the right of the xz-plane, θ is bounded by $\theta=0$ to $\theta=\pi$, not the full 2π (be careful with this!), so we can write:

$$E=\{(r,\theta,z)\,|\,0\le\theta\le\pi,\,0\le r\le 1,\,1-r^2\le z\le 4\}$$

The distance from any point to the z-axis is $\sqrt{x^2+y^2}$ so the density function is

$$\rho(x,y,z)=k\sqrt{x^2+y^2}=kr$$

where k is an arbitrary constant. The mass of E is then

$$\begin{aligned}m=\iiint_E k\sqrt{x^2+y^2}\,dV &= \int_0^{\pi}\int_0^1\int_{1-r^2}^4 (kr)\,r\,dz\,dr\,d\theta\\ &= \int_0^{\pi}\int_0^1 kr^2\left[4-(1-r^2)\right]\,dr\,d\theta\\ &= k\int_0^{\pi}d\theta\int_0^1(3r^2+r^4)\,dr\\ &= \pi k\left[r^3+\frac{r^5}{5}\right]_0^1\\ &= \boxed{\frac{6\pi k}{5}}\end{aligned}$$

Problem 4.7.6. *Evaluate the triple integral $\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_{\sqrt{x^2+y^2}}^{2}\left(x^2+y^2\right)\,dz\,dy\,dx$.*

Solution: This is the triple integral over the solid region

$$E=\{(x,y,z)\,|\,-2\le x\le 2,\,-\sqrt{4-x^2}\le y\le\sqrt{4-x^2},\,\sqrt{x^2+y^2}\le z\le 2\}$$

The projection of E onto the xy-plane is the disk $x^2 + y^2 \leq 4$. We can see that the lower surface of E is the cone $z = \sqrt{x^2 + y^2}$ and the upper surface is the plane $z = 2$. In cylindrical coordinates, we can describe the region much simpler:

$$E = \{(r, \theta, z) \,|\, 0 \leq \theta \leq 2\pi,\, 0 \leq r \leq 2,\, r \leq z \leq 2\}$$

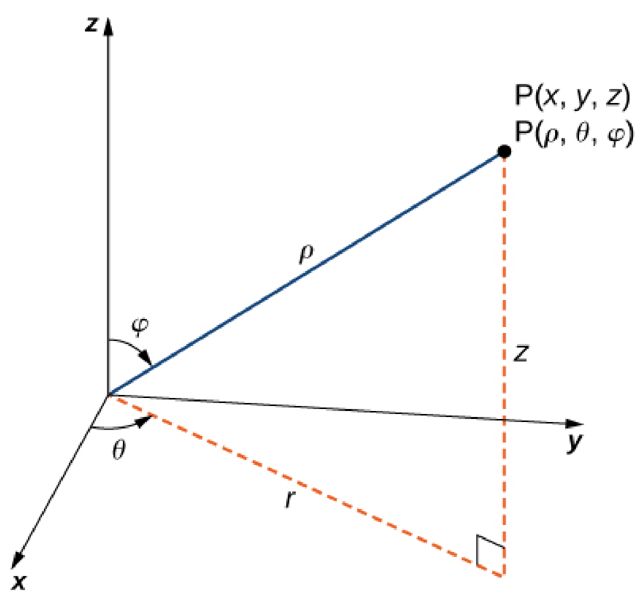
Thus, we have

$$\begin{aligned}
\int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \int_{\sqrt{x^2+y^2}}^{2} (x^2 + y^2)\, dz\, dy\, dx &= \iiint_E (x^2 + y^2)\, dV \\
&= \int_0^{2\pi} \int_0^2 \int_r^2 r^2\, r\, dz\, dr\, d\theta \\
&= \int_0^{2\pi} d\theta \int_0^2 r^3(2 - r)\, dr \\
&= 2\pi \left[\frac{1}{2} r^4 - \frac{1}{5} r^5\right]_0^2 \\
&= \boxed{\frac{16\pi}{5}}
\end{aligned}$$

4.8 Triple Integrals: Converting to Spherical Coordinates

In the previous section, we talked about the cylindrical coordinate system. Another useful 3D system is the *spherical coordinate* system, which is especially with surfaces involving spheres or cones.

Spherical Coordinate System In three dimensional space, or $\mathbb{R}^3$, the *spherical coordinate* system involves a point P specified by its distance ρ away from the origin, the polar angle θ from the positive x-axis (same as the cylindrical coordinate system), and the angle ϕ from the positive z-axis and the line segment OP, as shown below. Note that $\rho > 0$ and $0 \leq \phi \leq \pi$. Spherical coordinates are useful for triple integrals over regions that are symmetric with respect to the origin (this is distinct from cylindrical coordinates in the case of axis-based symmetry), which serves as point O.



Observing the geometry, it is not difficult to connect spherical coordinates with rectangular coordinates:

$$x = \rho \sin\phi \cos\theta \quad y = \rho \sin\phi \sin\theta \quad z = \rho \cos\phi$$

and

$$\rho^2 = x^2 + y^2 + z^2 \quad \tan\theta = \frac{y}{x} \quad \phi = \arccos\left(\frac{z}{\sqrt{x^2+y^2+z^2}}\right)$$

We should also know these relationships to make conversions simpler:

$$r = \rho\sin\phi \quad \theta = \theta \quad z = \rho\cos\phi$$

and

$$\rho = \sqrt{r^2+z^2} \quad \theta = \theta \quad \phi = \arccos\left(\frac{z}{\sqrt{r^2+z^2}}\right)$$

Let's test our knowledge with the following exercises.

Problem 4.8.1. *Convert the point $(0, 2\sqrt{3}, -2)$ from Cartesian to spherical coordinates.*

Solution: We know that $\rho = \sqrt{x^2+y^2+z^2} = \sqrt{0+12+4} = 4$. We also know that $\cos\phi = \frac{z}{p}$ and $\cos\theta = \frac{x}{\rho\sin\phi}$. Substituting known values, we obtain

$$\cos\phi = \frac{-2}{4} = -\frac{1}{2} \therefore \phi = \frac{2\pi}{3}$$
$$\cos\theta = \frac{0}{4\sin\frac{2\pi}{3}} = 0 \therefore \theta = \frac{\pi}{2}$$

Note that we did not choose $\theta = \frac{3\pi}{2}$, because the y-coordinate, $2\sqrt{3}$, is positive. Therefore, the spherical coordinates of the given point are $\boxed{(4, \pi/3, 2\pi/3)}$.

Problem 4.8.2. *Convert the point $(2, \pi/4, \pi/3)$ from spherical to Cartesian coordinates.*

Solution: This conversion should be trivial. We know $x = \rho\sin\phi\cos\theta$, $y = \rho\sin\phi\sin\theta$, and $z = \rho\cos\phi$.

$$x = 2\sin\frac{\pi}{3}\cos\frac{\pi}{4} = 2\left(\frac{\sqrt{3}}{2}\right)\left(\frac{1}{\sqrt{2}}\right) = \sqrt{\frac{3}{2}}$$
$$y = 2\sin\frac{\pi}{3}\sin\frac{\pi}{4} = 2\left(\frac{\sqrt{3}}{2}\right)\left(\frac{1}{\sqrt{2}}\right) = \sqrt{\frac{3}{2}}$$
$$z = 2\cos\frac{\pi}{3} = 2\left(\frac{1}{2}\right) = 1$$

Thus, the rectangular coordinates are $\boxed{(\sqrt{3/2}, \sqrt{3/2}, 1)}$.

The spherical coordinate system is especially useful in problems where there is symmetry about a point, and the origin is placed at this point. For example, the sphere with center the origin and radius c has the simple equation $\rho = c$, hence the name "spherical" coordinates. The graph of the equation $\theta = c$ is a vertical half-plane, and the equation $\phi = c$ represents a half-cone around the z-axis.

Triple Integrals in Spherical Coordinates In the spherical coordinate system, we have a concept of a **spherical wedge**, rather than a rectangular box:

$$E = \{(\rho, \theta, \phi) \,|\, a \le \phi \le b, \alpha \le \theta \le \beta,\, c \le \phi \le d\}$$

where $a \ge 0$, $\beta - \alpha \le 2\pi$, and $d - c \le \pi$. Traditionally, we saw triple integrals as dividing up solids into small rectangular boxes (in section 4.6), the same can be done with divisions into small spherical wedges. Therefore, we divide E into smaller spherical wedges E_{ijk} with equally spaced spheres $\rho = \rho_i$, half-planes

$\theta = \theta_j$, and half-cones $\phi = \phi_k$.

You can visualize E_{ijk} as an "almost" rectangular box with dimensions $\Delta\rho$, $\rho_i\Delta\phi$ (circular arc with radius ρ_i, angle $\Delta\phi$), and $\rho_i \sin\phi_k \Delta\theta$ (circular arc with radius $\rho_i \sin\phi_k$, angle $\Delta\theta$). So an approximation for the volume of E_{ijk} is

$$\Delta V_{ijk} \approx (\Delta\rho)(\rho_i\Delta\phi)(\rho_i \sin\phi_k\Delta\theta) = \rho_i^2 \sin\phi_k \Delta\rho\Delta\theta\Delta\phi$$

Let us propose that for any point $(\bar{\rho_i}, \bar{\theta_j}, \bar{\phi_k})$ in E_{ijk}, the *exact volume* of E_{ijk} is

$$\Delta V_{ijk} = \bar{p_i}^2 \sin\bar{\phi_k}\Delta\rho\Delta\theta\Delta\phi$$

If $(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})$ are the rectangular coordinates of this point, then

$$\iiint_E f(x,y,z)\,dV = \lim_{l,m,n\to\infty} \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} f(x^*_{ijk}, y^*_{ijk}, z^*_{ijk})\,\Delta V_{ijk}$$

$$= \lim_{l,m,n\to\infty} \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\,\bar{\rho_i}^2 \sin\bar{\phi_k}\,\Delta\rho\,\Delta\theta\,\Delta\phi$$

Using the properties of Riemann sums, we arrive at the formula for triple integration involving spherical coordinates.

$$\boxed{\iiint_E f(x,y,z)\,dV = \int_c^d \int_\alpha^\beta \int_a^b f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\,\rho^2\sin\phi\,d\rho\,d\theta\,d\phi}$$

where E is the spherical wedge

$$E = \{(\rho,\theta,\phi)\,|\,a \le \rho \le b, \alpha \le \theta \le \beta, c \le \phi \le d\}$$

We can convert any triple integral from rectangular to spherical coordinates by writing

$$x = \rho\sin\phi\cos\theta \quad y = \rho\sin\phi\sin\theta \quad z = \rho\cos\phi$$

and replacing dV with $\rho^2 \sin\phi \, d\rho \, d\theta \, d\phi$.

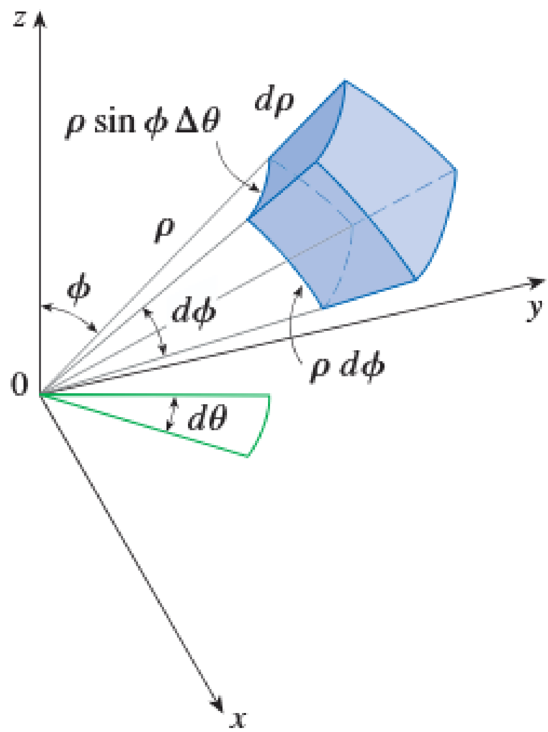


Image Credit: Calculus, Early Transcendentals

Similar to previous sections, we can use general regions. Consider

$$E = \{(\rho, \theta, \phi) \,|\, \alpha \leq \theta \leq \beta, c \leq \phi \leq d, g_1(\theta, \phi) \leq \rho \leq g_2(\theta, \phi)\}$$

as one generic example.

Usually, spherical coordinates are used to compute triple integrals when conical and spherical surfaces form the boundary of the integrating region.

Let's practice!

Problem 4.8.3. *Evaluate* $\iiint_B e^{(x^2+y^2+z^2)^{3/2}} \, dV$, *where* B *is a ball of radius* 2.

Solution: We have

$$B = \{(x, y, z) \,|\, x^2 + y^2 + z^2 \leq 4\}$$

Since the boundary of B is a full sphere, we should use spherical coordinates:

$$B = \{(\rho, \theta, \phi) \,|\, 0 \leq \rho \leq 2, \, 0 \leq \theta \leq 2\pi, \, 0 \leq \phi \leq \pi\}$$

Also, we know $x^2 + y^2 + z^2 = \rho^2$, so

$$\begin{aligned}\iiint_B e^{(x^2+y^2+z^2)^{3/2}}\,dV &= \int_0^{\pi}\int_0^{2\pi}\int_0^{2} e^{(\rho^2)^{3/2}}\rho^2\sin\phi\,d\rho\,d\theta\,d\phi \\ &= \int_0^{\pi}\sin\phi\,d\phi\int_0^{2\pi}d\theta\int_0^{2}\rho^2 e^{p^3}\,d\rho \\ &= [-\cos\phi]_0^{\pi}(2\pi)\left[\frac{1}{3}e^{\rho^3}\right]_0^2 \\ &= \boxed{\frac{4}{3}\pi\left(e^8-1\right)}\end{aligned}$$

Problem 4.8.4. *Calculate the volume of the solid that lies above the cone* $z = \sqrt{x^2+y^2}$ *and below the sphere* $x^2+y^2+z^2 = z$.

Solution: Let's consider the equation of the sphere.

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

The point $(0, 0, 0)$ satisfies this equation. Completing the square, we get

$$x^2 + y^2 + (z^2 - z) = 0 \therefore x^2 + y^2 + \left(z - \frac{1}{2}\right)^2 = \frac{1}{4}$$

This indicates that the sphere passes through the origin and is centered at $(0, 0, \frac{1}{2})$. In spherical coordinates, the equation of the sphere is written as

$$\rho^2 = \rho\cos\phi$$

which implies $\rho = \cos\phi$ as ρ is strictly non-negative.
Meanwhile, the equation of the cone can be written as

$$\rho\cos\phi = \sqrt{\rho^2\sin^2\phi\cos^2\theta + \rho^2\sin^2\phi\sin^2\theta} = \rho\sin\phi$$

In order for $\rho\cos\phi = \rho\sin\phi$, we must have $\phi = \pi/4$. Therefore, the description of the solid region E in spherical coordinates is

$$E = \{(\rho, \theta, \phi)\,|\,0 \le \theta \le 2\pi,\ 0 \le \phi \le \pi/4,\ 0 \le \rho \le \cos\phi\}$$

We now find the volume of E:

$$\begin{aligned}V(E) = \iiint_E dV &= \int_0^{2\pi}\int_0^{\pi/4}\int_0^{\cos\phi}\rho^2\sin\phi\,d\rho\,d\phi\,d\theta \\ &= \int_0^{2\pi}d\theta\int_0^{\pi/4}\sin\phi\left[\frac{\rho^3}{3}\right]_{\rho=0}^{\rho=\cos\phi}d\phi \\ &= \frac{2\pi}{3}\int_0^{\pi/4}\sin\phi\cos^3\phi\,d\phi \\ &= \frac{2\pi}{3}\left[-\frac{\cos^4\phi}{4}\right]_0^{\pi/4} \\ &= \boxed{\frac{\pi}{8}}\end{aligned}$$

Problem 4.8.5. *Let $E \subset \mathbb{R}^3$ be a solid conducting sphere with radius r, centered at the origin. The charge density is non-uniform and given by $\sigma(\theta, \phi) = k \csc\phi$, where k is a positive constant. Calculate the total electric charge, Q, on the sphere.*

Solution: As we know, the applications of double integrals in section 4.4 can also be applied to triple integrals. If in three-dimensional space, we are given a subset E with some charge density σ, then the total electric charge on E must be

$$Q = \iiint_E \sigma \, dV$$

In this case, $\sigma = k \csc\phi$ and $dV = \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$. The region E is described as

$$E = \{(\rho, \theta, \phi) \,|\, 0 \leq \rho \leq r,\ 0 \leq \theta \leq 2\pi,\ 0 \leq \phi \leq \pi\}$$

Thus, the total electric charge on the region E is

$$\begin{aligned}
Q &= \iiint_E k \csc\phi \cdot \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta \\
&= k \int_0^{2\pi} \int_0^{\pi} \int_0^{r} \rho^2 \sin\phi \csc\phi \, d\rho \, d\phi \, d\theta \\
&= k \int_0^{r} \rho^2 \, d\rho \int_0^{\pi} d\phi \int_0^{2\pi} d\theta \\
&= k \cdot \left[\frac{1}{3}\rho^3\right]_0^r \cdot [\phi]_0^{\pi} \cdot [\theta]_0^{2\pi} \\
&= \boxed{\frac{2}{3} k \pi^2 r^3}
\end{aligned}$$

4.9 The Jacobian and Change of Variables

In one-variable calculus, we frequently used a change of variables (a.k.a. substitution) to simplify the process of evaluating an integral. The Substitution Rule involved replacing the rules of x with u, leading to

$$\int_a^b f(x) \, dx = \int_c^d f(g(u)) \, g'(u) \, du$$

where $x = g(u)$, $a = g(c)$, and $b = g(d)$. In other words, the substitution works as

$$\int_a^b f(x) \, dx = \int_c^d f(x(u)) \, \frac{dx}{du} \, du$$

In this section, we discuss how change of variables can help us in evaluating double and triple integrals.

Change of Variables (Double Integrals) We actually have seen the most common method for change of variables regarding double integrals: the conversion to polar coordinates. The new variables, r and θ, are related by the old variables, x and y as

$$x = r\cos\theta \quad y = r\sin\theta$$

so as demonstrated in section 4.3, we found

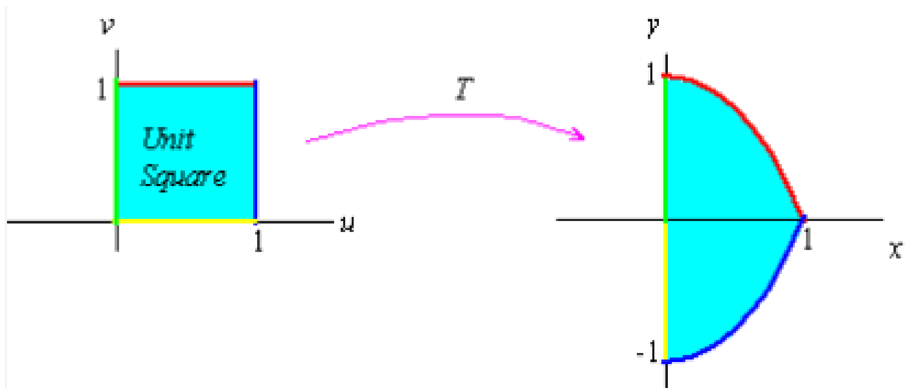
$$\iint_R f(x, y) \, dA = \iint_S f(r\cos\theta,\, r\sin\theta) \, r \, dr \, d\theta$$

where S is the region in the $r\theta$-plane corresponding to the region R in the xy-plane.

Generally, a change of variables is viewed as a **transformation** T from the uv-plane to the xy-plane:

$$T(u, v) = (x, y)$$

where $x = g(u, v)$ and $y = h(u, v)$, or sometimes, $x = x(u, v)$ and $y = y(u, v)$.
For simplicity, we let T be a $\boldsymbol{C^1}$ **transformation**, which means g and h are continuously differentiable, i.e. have continuous first-order partial derivatives. In reality, the transformation T is just a function whose domain and range are subsets of $\mathbb{R}^2$. If $T(u_1, v_1) = (x_1, y_1)$, then (x_1, y_1) is called the **image** of (u_1, v_1). If *no two points* have the same image, then T is called a **one-to-one transformation**. Here's how a transformation T on a region S in the uv-plane works.



T transforms S, the region on the left, into a region R, on the right, from the uv- to the xy-plane. R is called the **image** of S, consisting of the images of all points in S.

Also, if T is one-to-one, then it has an **inverse transformation** T^{-1} from the xy-plane to the uv-plane and we can then solve for u and v in terms of x and y:

$$u = G(x, y) \quad v = H(x, y)$$

Problem 4.9.1. *A transformation T is defined by the equations*

$$x = u^2 - v^2 \quad y = 2uv$$

Find the image of the square $S = \{(u, v) \,|\, 0 \le u \le 1,\, 0 \le v \le 1\}$.

Solution: T basically sends the boundary of the square S in the uv-plane to the boundary of its image, R, in the xy-plane. We start by finding the images of each side of the square S. There are four sides, which we will call S_1, S_2, S_3, and S_4. The first side, S_1, is given by the equation $v = 0$, $0 \le u \le 1$. According to the change of variables, $x = u^2$, $y = 0$, and thus $0 \le x \le 1$. Thus, S_1 is mapped onto the line segment from $(0, 0)$ to $(1, 0)$. The second side, S_2, is $u = 1$, $0 \le v \le 1$. Plugging $u = 1$ in the equations, the change of variables becomes

$$x = 1 - v^2 \qquad y = 2v$$

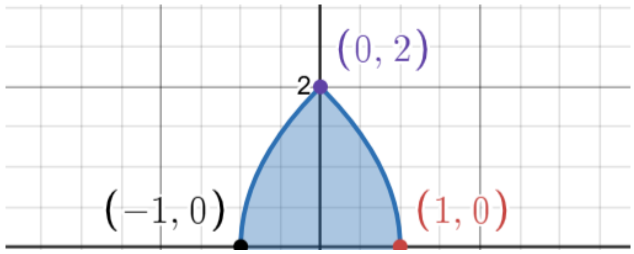
We can solve this system and eliminate v, yielding

$$x = 1 - \frac{y^2}{4} \quad 0 \le x \le 1$$

which is the portion of a parabola. Also, S_3 is given by $v = 1$, $0 \le u \le 1$ whose image is the parabolic arc

$$x = \frac{y^2}{4} - 1 \quad -1 \le x \le 0$$

Finally, the side S_4 is given by $u = 0$, $0 \le v \le 1$ whose image is $x = -v^2$, $y = 0$, for $-1 \le x \le 0$. Also, notice that moving counterclockwise around S maps to moving counterclockwise around the image, R. In conclusion, the image of S is the region R bounded by the x-axis and the two parabolas $x = 1 - \frac{y^2}{4}$ and $x = \frac{y^2}{4} - 1$, shown below.



Let's see how a change of variables can affect a double integral. Start with a small rectangle S in the uv-plane whose lower left corner is the point with coordinates (u_0, v_0) with dimensions of Δu and Δv.
The image of S is a region R in the xy-plane, one of whose boundary points is $T(u_0, v_0) = (x_0, y_0)$. The vector

$$\mathbf{r}(u, v) = g(u, v)\,\mathbf{i} + h(u, v)\,\mathbf{j}$$

is the position vector of the image of the point (u, v) The equation of the lower side of region S is $v = v_0$, given by the vector function $\mathbf{r}(u, v_0)$. The tangent vector to this image curve at (x_0, y_0) is

$$\mathbf{r}_u = g_u(u_0, v_0)\,\mathbf{i} + h_u(u_0, v_0)\,\mathbf{j} = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j}$$

Similarly, the tangent vector at (x_0, y_0) to the left side of S (where $u = u_0$) is

$$\mathbf{r}_v = g_v(u_0, v_0)\,\mathbf{i} + h_v(u_0, v_0)\,\mathbf{j} = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j}$$

Now the key thing is to define two secant vectors $\mathbf{a}$ and $\mathbf{b}$ for small changes Δu and Δv:

$$\mathbf{a} = \mathbf{r}(u_0 + \Delta u, v_0) - \mathbf{r}(u_0, v_0) \qquad \mathbf{b} = \mathbf{r}(u_0, v_0 + \Delta v) - \mathbf{r}(u_0, v_0)$$

These two secant vectors determine a parallelogram that approximates the image region $R = T(S)$. However, we know

$$\mathbf{r}_u = \lim_{\Delta u \to 0} \frac{\mathbf{r}(u_0 + \Delta u, v_0) - \mathbf{r}(u_0, v_0)}{\Delta u}$$

so $\mathbf{r}(u_0 + \Delta u, v_0) - \mathbf{r}(u_0, v_0) \approx \Delta u\,\mathbf{r}_u$ and similarly, $\mathbf{r}(u_0, v_0 + \Delta v) - \mathbf{r}(u_0, v_0) \approx \Delta v\,\mathbf{r}_v$. The approximate area of R is the area of this approximate parallelogram, which we know from section 1.4, is

$$|(\Delta u\,\mathbf{r}_u) \times (\Delta v\,\mathbf{r}_v)| = |\mathbf{r}_u \times \mathbf{r}_v|\,\Delta u\,\Delta v$$

The cross product $\mathbf{r}_u \times \mathbf{r}_v$ is

$$\mathbf{r}_u \times \mathbf{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} \mathbf{k} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} \mathbf{k}$$

The determinant involved in this calculation is referred to as the *Jacobian* of this transformation, and has a special notation.

Definition 4.9.1. *The* ***Jacobian*** *of the transformation T given by $x = g(u, v)$ and $y = h(u, v)$ is*

$$\frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u}\frac{\partial y}{\partial v} - \frac{\partial x}{\partial v}\frac{\partial y}{\partial u}$$

Using this notation, the approximate area of R is given by

$$\Delta A \approx \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \Delta u\, \Delta v$$

where the Jacobian is evaluated at (u_0, v_0).

Next we divide a general region S in the uv-plane into sub-rectangles S_{ij}, and call their images in the xy-plane as R_{ij}.

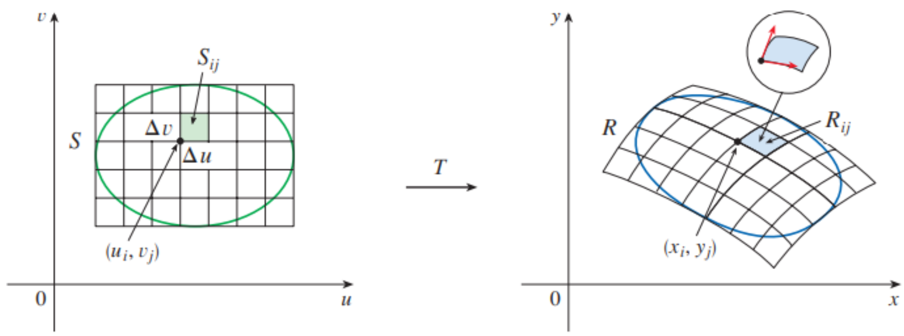


Image Credit: Calculus, Early Transcendentals

Applying the area approximation to each R_{ij}, the approximate double integral of f over the new region, R, is

$$\begin{aligned}\iint_R f(x,y)\,dA &= \sum_{i=1}^{m}\sum_{j=1}^{n} f(x_i, y_j)\,\Delta A \\ &\approx \sum_{i=1}^{m}\sum_{j=1}^{n} f(g(u_i, v_j),\, h(u_i, v_j)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| \Delta u\, \Delta v\end{aligned}$$

This double sum is the Riemann sum for the double integral

$$\iint_S f(g(u,v),\, h(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv$$

This argument suggests the validity of the following theorem.

Theorem 4.9.1. *Suppose that T is a C^1 transformation whose Jacobian is nonzero and T maps a region S in the uv-plane onto a region R in the xy-plane. Suppose f is continuous on R and that R and S are one of type I or II plane regions. Also, assume T is one-to-one, except perhaps on the boundary of S. Then*

$$\iint_R f(x,y)\,dA = \iint_S f(x(u,v),\, y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv$$

This theorem implies we can convert a double integral in x and y into one in u and v using the Jacobian and writing

$$dA = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\, dv$$

With double integrals, the direct application of the Jacobian is the conversion to polar coordinates. In this case, the transformation T takes place from the $r\theta$-plane to the xy-plane, and the equations are

$$x = g(r,\theta) = r\cos\theta \qquad y = h(r,\theta) = r\sin\theta$$

In other words, T maps an ordinary rectangle in the $r\theta$-plane ($\alpha \le \theta \le \beta$, $a \le r \le b$) into a polar rectangle in the xy-plane. The Jacobian of the transformation is given by

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r > 0$$

Thus, according to Theorem 4.9.1, the double integral in polar coordinates is

$$\begin{aligned} \iint_R f(x,y)\,dx\,dy &= \iint_S f(r\cos\theta, r\sin\theta)\left|\frac{\partial(x,y)}{\partial(r,\theta)}\right| dr\,d\theta \\ &= \int_\alpha^\beta \int_a^b f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta \end{aligned}$$

which is the same result derived in section 4.3.

Problem 4.9.2. *Use the change of variables $x = u^2 - v^2$, $y = 2uv$ to evaluate the integral $\iint_R y\,dA$, where R is the region bounded by the x-axis and the parabolas $y^2 = 4 - 4x$ and $y^2 = 4 + 4x$, for $y \ge 0$.*

Solution: We discussed the region R in problem 4.9.1. We discovered that $T(S) = R$, where S was the square $[0,1]\times[0,1]$. Since S is a much simpler region than R, it makes sense to make this change of variables. The Jacobian of the transformation is

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} 2u & -2v \\ 2v & 2u \end{vmatrix} = 4u^2 + 4v^2 > 0$$

Therefore, the value of the double integral is

$$\iint_R y\,dA = \iint_S 2uv\left|\frac{\partial(x,y)}{\partial(u,v)}\right| dA = \int_0^1\int_0^1 (2uv)(4)(u^2+v^2)\,du\,dv$$

$$= 8\int_0^1\int_0^1 (u^3v + uv^3)\,du\,dv = 8\int_0^1 \left[\frac{1}{4}u^4v + \frac{1}{2}u^2v^3\right]_{u=0}^{u=1} dv$$

$$= \int_0^1 (2v + 4v^3)\,dv = [v^2 + v^4]_0^1 = \boxed{2}$$

Problem 4.9.3. *Evaluate $\iint_R e^{(x+y)/(x-y)}\,dA$, where R is the trapezoidal region with vertices $(1,0)$, $(2,0)$, $(0,-2)$, and $(0,-1)$.*

Solution: It is extremely difficult to integrate $e^{(x+y)/(x-y)}$ directly, so we make a change of variables suggested by

$$u = x + y \qquad v = x - y$$

Since this refers to the inverse transformation T^{-1} from the xy-plane to the uv-plane, we need to solve for x and y to get the transformation T from the uv-plane to the xy-plane.

$$x = \frac{1}{2}(u+v) \qquad y = \frac{1}{2}(u-v)$$

The Jacobian of this transformation T is given by

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = -\frac{1}{2}$$

To determine the region S in the uv-plane corresponding to R, we notice the sides of R lie on the following lines:

$$y = 0 \quad x - y = 2 \quad x = 0 \quad x - y = 1$$

Using the equations for x and y in terms of u and v, the image lines in the uv-plane are

$$u = v \quad v = 2 \quad u = -v \quad v = 1$$

If you sketch these lines in the uv-plane, it is easy to see that R is a trapezoidal region with vertices $(1, 1)$, $(2, 2)$, $(-2, 2)$, and $(-1, 1)$. Since

$$S = \{(u, v) \,|\, 1 \le v \le 2,\, -v \le u \le v\}$$

the value of the double integral is

$$\begin{aligned}
\iint_R e^{(x+y)/(x-y)}\, dA &= \iint_S e^{u/v} \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\, dv \\
&= \int_1^2 \int_{-v}^{v} e^{u/v} \left(\frac{1}{2}\right) du\, dv \\
&= \frac{1}{2}\int_1^2 [ve^{u/v}]_{u=-v}^{u=v}\, dv \\
&= \frac{1}{2}\int_1^2 (e - e^{-1})\, v\, dv \\
&= \boxed{\frac{3}{4}(e - e^{-1})}
\end{aligned}$$

Change of Variables (Triple Integrals) Suppose T is a one-to-one transformation that maps a region S in three-dimensional uvw-space onto a region R in xyz-space via the equations

$$x = g(u, v, w) \quad y = h(u, v, w) \quad z = k(u, v, w)$$

The Jacobian of T is given by the 3×3 determinant:

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}$$

Based on Theorem 4.9.1 for double integrals, we can generalize a similar result for transformations involving triple integrals:

$$\boxed{\iint_R f(x, y, z)\, dV = \iiint_S f(x(u, v, w), y(u, v, w), z(u, v, w)) \left|\frac{\partial(x, y, z)}{\partial(u, v, w)}\right| du\, dv\, dw}$$

Let's put this into practice.

Problem 4.9.4. *Derive the formula for converting triple integrals from Cartesian to spherical coordinates.*

Solution: There are three equations to convert from Cartesian coordinates to spherical coordinates:

$$x = \rho \sin\phi \cos\theta \quad y = \rho \sin\phi \sin\theta \quad z = \rho \cos\phi$$

These will serve as our change of variables. The Jacobian of the transformation is

$$\begin{aligned}
\left|\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}\right| &= \begin{vmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{vmatrix} \\
&= \cos\phi \begin{vmatrix} -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \end{vmatrix} - \rho\sin\phi \begin{vmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta \end{vmatrix} \\
&= \cos\phi\left(-\rho^2\sin\phi\cos\phi\sin^2\theta - \rho^2\sin\phi\cos\phi\cos^2\theta\right) \\
&\quad - \rho\sin\phi\left(\rho\sin^2\phi\cos^2\theta + \rho\sin^2\phi\sin^2\theta\right) \\
&= -\rho^2\sin\phi\cos^2\phi - \rho^2\sin\phi\sin^2\phi \\
&= -\rho^2\sin\phi.
\end{aligned}$$

Since $0 \le \phi \le \pi$, we have $\sin\phi \ge 0$. Thus

$$\left|\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}\right| = |-\rho^2\sin\phi| = \rho^2\sin\phi$$

and so the triple integral is

$$\iiint_R f(x,y,z)\,dV = \iiint_S f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\,\rho^2\sin\phi\,d\rho\,d\theta\,d\phi$$

which is the same result in section 4.8.

5 Vector Calculus

Vector calculus lets us study forces, flows, and fields in space. In this chapter, we explore how quantities move and interact−following arrows along paths, measuring how much passes through curves or surfaces, and uncovering the rotations and sources that shape a field. We will extend curves to surfaces, exploring both local behavior and global patterns, and see how complex motions follow simple, elegant principles. By the end, we will be able to understand vector calculus not just as formulas, but as a powerful language for describing motion, geometry, and the world around us.

5.1 Vector Fields: In Essence

Core Idea A vector field is a function whose domain is a set of points in a standard coordinate system (commonly $\mathbb{R}^2$ or $\mathbb{R}^3$) and whose range is a set of vectors in V_2 (or V_3).

Definition 5.1.1. *Let D be a set in the plane region $\mathbb{R}^2$. A **vector field** on $\mathbb{R}^2$ is a function $\boldsymbol{F}$ that assigns to each point (x, y) in D a two-dimensional vector $\boldsymbol{F}(x, y)$.*

The most common (and simple) method of visualizing a vector field is to draw the arrow representing the vector $\mathbf{F}(x, y)$ at some point (x, y). Obviously, this won't be perfect, but it would form a reasonable impression for what $\mathbf{F}$ looks like. Since $\mathbf{F}$ in this case is a two-dimensional vector, we can write it in any of the three component forms:

$$\mathbf{F}(x, y) = P(x, y)\,\mathbf{i} + Q(x, y)\,\mathbf{j} = \langle P(x, y), Q(x, y)\rangle = P\,\mathbf{i} + Q\,\mathbf{j}$$

Note: P and Q are scalar functions of x and y and are occasionally referred to as **scalar fields** to distinguish them from vector fields.

Definition 5.1.2. *Let E be a set in the plane region $\mathbb{R}^3$. A **vector field** on $\mathbb{R}^3$ is a function $\boldsymbol{F}$ that assigns to each point (x, y, z) in E a three-dimensional vector $\boldsymbol{F}(x, y, z)$.*

Notice how similar this statement is to Definition 5.1.1!

Call P, Q, and R the component functions of $\mathbf{F}(x, y, z)$. The relevant expression is

$$\mathbf{F}(x, y, z) = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

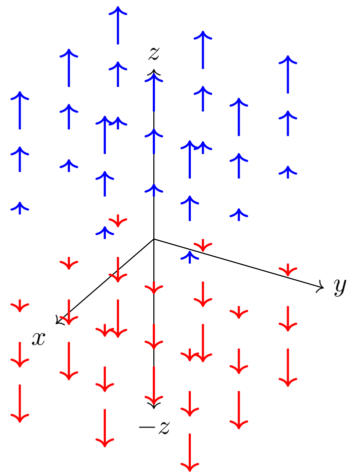
At this point, we must describe the concept of continuity. Any vector field $\mathbf{F}$ is **continuous** if and only if its component functions, or corresponding scalar fields, are also continuous.

Additionally, note that in this book, we will sometimes identify a point (x, y, z) with the position vector $\mathbf{x} = \langle x, y, z\rangle$ and use the notation $\mathbf{F}(\mathbf{x})$ rather than $\mathbf{F}(x, y, z)$, effectively making $\mathbf{F}$ a function that assigns a vector $\mathbf{F}(\mathbf{x})$ to a vector $\mathbf{x}$.

Problem 5.1.1. *Sketch the vector field on $\mathbb{R}^3$ given by $\boldsymbol{F}(x, y, z) = z\,\boldsymbol{k}$.*

Solution: The vector field $\mathbf{F}(x, y, z) = z\,\mathbf{k}$ on $\mathbb{R}^3$ assigns to each point (x, y, z) a vector pointing in the positive (or negative) direction z-direction with length proportional to z. Specifically,

- *At height* $z > 0$*, the vectors point upwards and grow longer as* z *increases.*
- *At height* $z < 0$*, the vectors point downwards with length increasing as* $|z|$ *increases.*
- *On the xy-plane, that is* $z = 0$*, all vectors have zero length.*

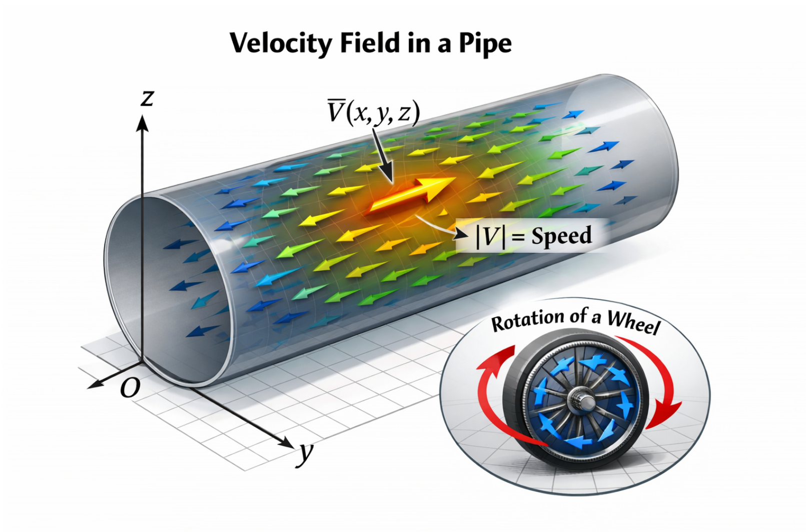


We were able to sketch this vector field because its equation was relatively simple. However, in many cases, three-dimensional vector fields can be impossible to sketch by hand and so we need to use computer software.

Remark. There are various software tools for mathematical modeling and simulations, such as **MATLAB**, **SimuLinks**, **Mathematica**, and many others.

Applications of Vector Fields The mathematical concept of a vector field is extended to describe physical phenomena like velocity, gravity, and charge by assigning a vector to every point in a region of space. This vector represents the magnitude and direction of the physical quantity at that specific point. The field illustrates how a quantity influences space and any object that enters that space.

Imagine a fluid flowing through a medium (usually a simple pipe) and let $\mathbf{V}(x, y, z)$ be the velocity vector at an arbitrary point. By definition of a vector field, $\mathbf{V}$ assigns a vector each point (x, y, z) within the domain E (in this case, the interior of the pipe), hence $\mathbf{V}$ is called a **velocity field** on $\mathbb{R}^3$. At any given point, the speed is given by the length of the corresponding velocity vector.



Velocity fields occur in other areas of physics, as well. The above vector field could be used to describe the rotation of a wheel.

Let's consider another example. Newton's Law of Gravitation states that the magnitude of the gravitational force between two objects with masses m and M is

$$|\mathbf{F}| = \frac{mMG}{r^2}$$

where r is the center-to-center distance between the two objects and G is the gravitational constant. Suppose the object having mass M is located at the origin in $\mathbb{R}^3$. Also, let the position vector of the object with mass m be $\mathbf{x} = \langle x, y, z\rangle$. Thus $r = |\mathbf{x}| \therefore r^2 = |\mathbf{x}|^2$. The gravitational force exerted on the second object points *towards* the origin, so the unit vector in this direction will be

$$-\frac{\mathbf{x}}{|\mathbf{x}|}$$

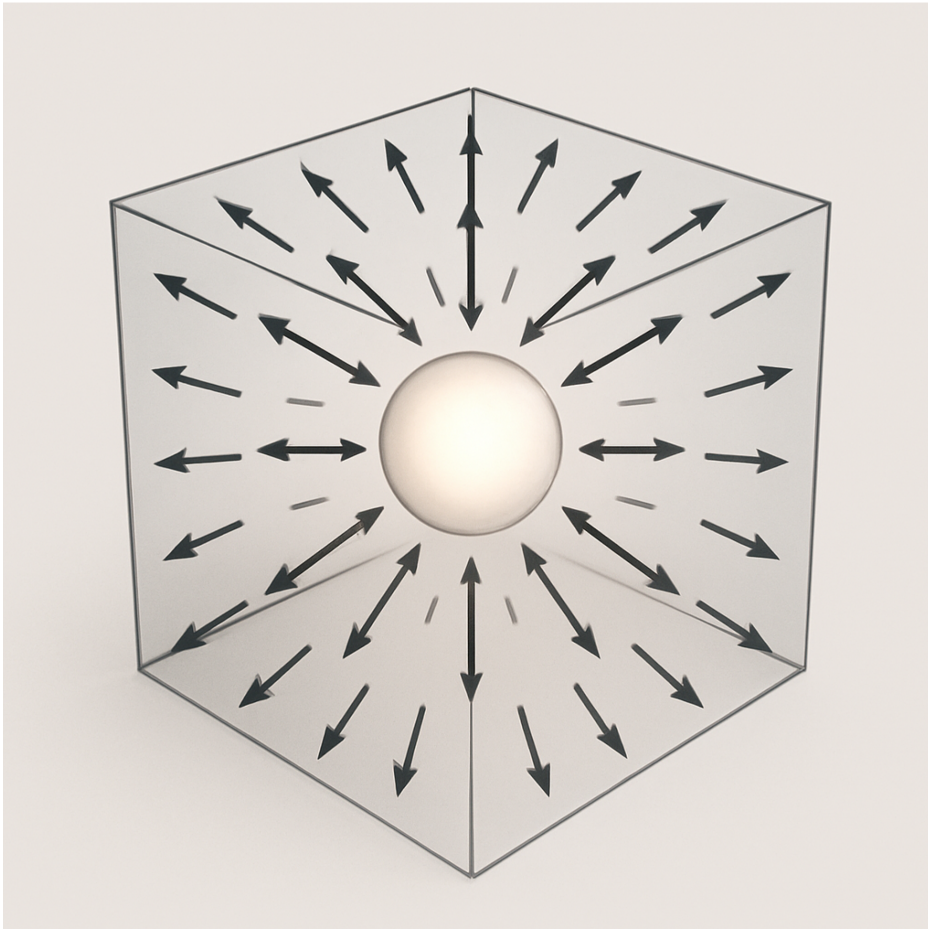
and thus the gravitational force acting on the object at $\mathbf{x} = \langle x, y, z\rangle$ is

$$\mathbf{F}(\mathbf{x}) = -\frac{mMG}{|\mathbf{x}|^3}\,\mathbf{x}$$

The vector field defined above is called a **gravitational field**, because it associates a force vector $\mathbf{F}(\mathbf{x})$ with every point $\mathbf{x}$ in space. Using the fact that $\mathbf{x} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ and $|\mathbf{x}| = \sqrt{x^2+y^2+z^2}$, we can write the component form of the field as

$$\mathbf{F}(x, y, z) = \frac{-mMG}{\sqrt{x^2+y^2+z^2}}\,\mathbf{i} + \frac{-mMG}{\sqrt{x^2+y^2+z^2}}\,\mathbf{j} + \frac{-mMG}{\sqrt{x^2+y^2+z^2}}\,\mathbf{k}$$

Below shows a sketch of this gravitational field in 3D:



Finally, we discuss electrostatic fields. Suppose some electric charge Q is concentrated at the origin. Coulomb's Law states that the electric force $\mathbf{F}(\mathbf{x})$ exerted by this charge on a test particle of charge q located at a point (x, y, z) with corresponding position vector $\mathbf{x} = \langle x, y, z \rangle$ is given by

$$\mathbf{F}(\mathbf{x}) = \frac{\varepsilon q Q}{|\mathbf{x}|^3}\,\mathbf{x}$$

where ε is a constant whose value depends on the units chosen. For like charges, $qQ > 0$ and the force is *repulsive*, and for opposite charges, $qQ < 0$ and the force is *attractive*. Note that gravitational and electrostatic fields are examples of **force fields**.

In practicality, physicists often consider force $\mathbf{F}$ *per unit charge*:

$$\mathbf{E} = \frac{1}{q}\mathbf{F}(\mathbf{x}) = \frac{\varepsilon Q}{|\mathbf{x}|^3}\,\mathbf{x}$$

Then $\mathbf{E}$ is a vector field on $\mathbb{R}^3$ called the **electric field** of Q.

Gradient Vector Fields Let f be a scalar function of two variables, i.e. $f(x, y)$. Recall from our discussion of partial derivatives that the gradient ∇f is defined by

$$\nabla f(x, y) = f_x(x, y)\,\mathbf{i} + f_y(x, y)\,\mathbf{j}$$

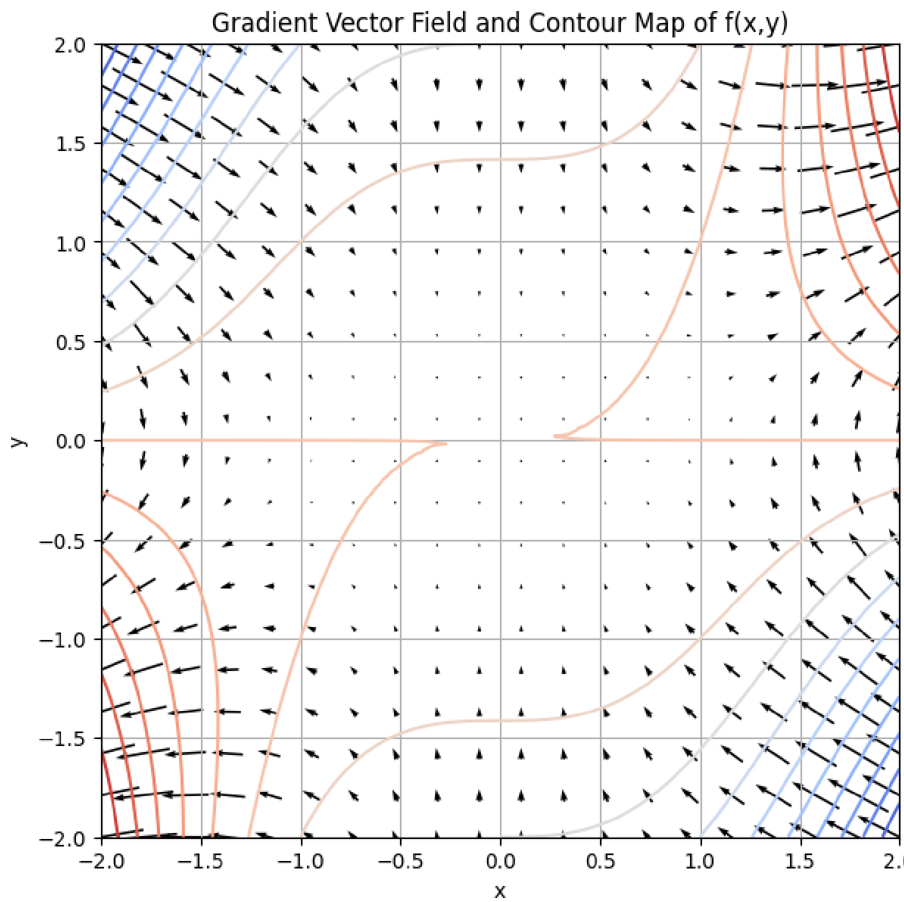
Thus ∇f can also be thought of as a vector field on $\mathbb{R}^2$; specifically, it is a **gradient vector field**. Similar logic applies if f is a scalar function of three variables.

Problem 5.1.2. *Find the gradient vector field of $f(x, y) = x^3y - y^2$. Plot the gradient vector field together with a contour map of f. How are they related?*

Solution: The gradient vector field is simply

$$\nabla f(x, y) = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} = \boxed{3x^2y\,\mathbf{i} + (x^3 - 2y)\,\mathbf{j}}$$

Below shows a contour map of f with the gradient vector field. Notice how the gradient vectors are orthogonal to the level curves, which we would expect based no our knowledge of the gradient vector in section 3.6.



The last thing we need to talk about is **conservative vector fields**. A field $\mathbf{F}$ is called conservative if and only if it is the gradient of some scalar function. In simpler terms, there exists a function f such that $\mathbf{F} = \nabla f$, where f is called a *potential function* for $\mathbf{F}$.

It's important to note that not ALL vector fields are conservative, but from a physics point of view, many such fields do arise. Let's consider the gravitational field $\mathbf{F}$ as previously defined. If

$$f(x, y, z) = \frac{mMG}{\sqrt{x^2 + y^2 + z^2}}$$

then we find

$$\begin{aligned}\nabla f(x,y,z) &= \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} + \frac{\partial f}{\partial z}\,\mathbf{k} \\ &= \frac{-mMG}{(x^2+y^2+z^2)^{3/2}}\,\mathbf{i} + \frac{-mMG}{(x^2+y^2+z^2)^{3/2}}\,\mathbf{j} + \frac{-mMG}{(x^2+y^2+z^2)^{3/2}}\,\mathbf{k} \\ &= \mathbf{F}(x,y,z)\end{aligned}$$

and so the gravitational field is conservative. □

5.2 Line Integrals

In this section we introduce a new type of integral. However, let's first revisit parametric equations, i.e. writing down parametric equations to describe a given curve. You should have seen some of this in a standard Calculus II course. Here are some of the more basic curves that we will need to parameterize, as well as the limits on the parameter if required.

1. The ellipse with equation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ has two parameterizations, one for each of the counter-clockwise and clockwise directions:
   - In the counter-clockwise direction, we have $x = a\cos(t)$, $y = b\sin(t)$, for $0 \le t \le 2\pi$.
   - In the clockwise direction, we have $x = a\cos(t)$, $y = -b\sin(t)$, for $0 \le t \le 2\pi$.
2. The circle with equation $x^2 + y^2 = r^2$ also has two parameterizations, one for the counter-clockwise and one for the clockwise directions:
   - In the counter-clockwise direction, we have $x = r\cos(t)$, $y = r\sin(t)$, for $0 \le t \le 2\pi$.
   - In the clockwise direction, we have $x = r\cos(t)$, $y = -r\sin(t)$, for $0 \le t \le 2\pi$.
3. For the curve $y = f(x)$, we let $x = t$ and $y = f(t)$.
4. For the curve $x = g(y)$, we let $x = g(t)$ and $y = t$.
5. For the line segment from (x_0, y_0, z_0) to (x_1, y_1, z_1), we have $\mathbf{r}(t) = (1-t)\langle x_0, y_0, z_0\rangle + t\langle x_1, y_1, z_1\rangle$, for $0 \le t \le 1$.

For the ellipse and the circle we've given two parameterizations, one tracing out the curve clockwise and the other counter-clockwise. As we'll eventually see the direction that the curve is traced out can, on occasion, change the answer. Also, both of these "start" on the positive x-axis at $t = 0$.

Now let's move on to line integrals. In Calculus I, we integrated a single-variable function $f(x)$ over an interval $[a, b]$. There, we were thinking of x as taking all the values in this interval starting at a and ending at b. With line integrals we will start with integrating the two-variable function $f(x, y)$, and the values of x and y that we're going to use will be the points (x, y) that lie on a curve C. Note that this is different from the double integrals that we were working with in the previous chapter, where the points came out of some two-dimensional region.

Consider a curve C. Assume it is *smooth* (defined later) and has parametric equations

$$x = h(t) \quad y = g(t) \quad a \le t \le b$$

Often we will want to write the parameterize the curve as a vector function. Thus

$$\mathbf{r}(t) = h(t)\,\mathbf{i} + g(t)\,\mathbf{j} \quad a \le t \le b$$

Definition 5.2.1. *A curve C is called **smooth** if the vector $\mathbf{r}'(t) \neq 0$ for all t and $\mathbf{r}'(t)$ is continuous.*

Definition 5.2.2. *The **line integral** of a function $f(x, y)$ along a curve C is equal to*

$$\iint_C f(x,y)\,ds$$

We use a ds here to acknowledge the fact that we are moving along a curve C instead of the x- or y- axes (dx and dy, respectively). Due to this choice, the above definition is sometimes called the **line integral of f with respect to arc length**.

We actually have seen ds before. If you recall from Calculus II the formula we used to find the arc length of a parametric curve with $x = x(t)$ and $y = y(t)$ to be

$$L = \int_a^b ds, \quad \text{where } ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt$$

So, in order to compute a line integral we convert everything into the subsequent parametric equations:

$$\boxed{\int_C f(x,y)\,ds = \int_a^b f(h(t), g(t))\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt}$$

Don't forget to plug in the parametric equations into the function. If we use the vector form of the parametrization, we can simplify the notation using

$$\sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} = |\mathbf{r}'(t)|$$

where $|\mathbf{r}'(t)|$ is the magnitude of the position vector $\mathbf{r}(t)$. With this notation, the line integral is

$$\boxed{\int_C f(x,y)\,ds = \int_a^b f(h(t), g(t))\,|\mathbf{r}'(t)|\,dt}$$

Problem 5.2.1. *Evaluate the line integral $\int_C xy^2\,ds$, where C is the right half of the unit circle, $x^2+y^2 = 1$, traced out in the counter-clockwise direction.*

Solution: We first parameterize the curve C.

$$x(t) = 4\cos t \quad \text{and} \quad y(t) = 4\sin t$$

and the right half of the circle will involve the range $-\frac{\pi}{2} \le t \le \frac{\pi}{2}$ if C is traced out in the counter-clockwise direction. Next, we take the derivatives of the parametric equations and determine the arc length element ds:

$$\frac{dx}{dt} = -4\sin t \quad \frac{dy}{dt} = 4\cos t$$
$$ds = \sqrt{16\sin^2 t + 16\cos^2 t}\,dt = 4\,dt$$

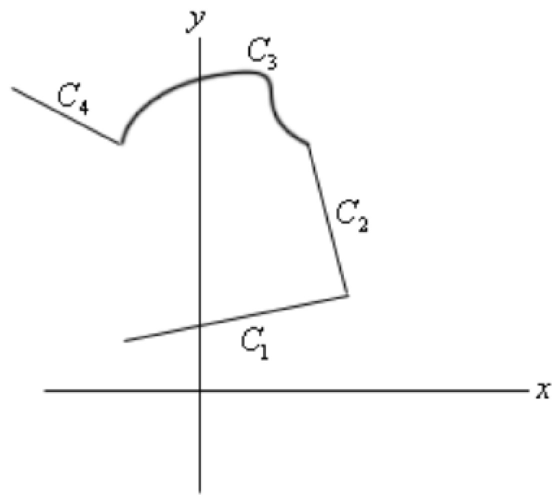
The value of the line integral is then

$$\begin{aligned}\int_C xy^2\,ds &= \int_{-\pi/2}^{\pi/2} 4\cos t(4\sin t)^2(4)\,dt \\ &= 256\int_{-\pi/2}^{\pi/2} \cos t\sin^2 t\,dt \\ &= \frac{256}{3}\sin^3 t\Big]_{-\pi/2}^{\pi/2} \\ &= \boxed{\frac{512}{3}}\end{aligned}$$

Line Integrals over Piecewise Smooth Curves We now discuss another type of line integral, which involves a little more nuance than just along smooth curves.

Definition 5.2.3. *A **piecewise-smooth curve** is any curve that can be written as the union of a finite number of smooth curves, $C_1, ..., C_n$ where the endpoint of C_i is the initial point of C_{i+1}.*

Below is an illustration of such a curve, with $n = 4$.

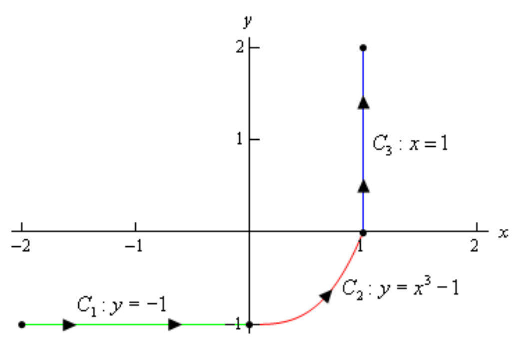


Evaluating line integrals over piecewise-smooth curves is relatively simple. All we do is evaluate the line integral over each individual piece and add them up. In this case, the line integral is

$$\int_C f(x,y)\,ds = \int_{C_1} f(x,y)\,ds + \int_{C_2} f(x,y)\,ds + \int_{C_3} f(x,y)\,ds + \int_{C_4} f(x,y)\,ds$$

Let's work out an example.

Problem 5.2.2. *Evaluate the line integral $\int_C 4x^3\,ds$ where C is the curve shown below.*



Solution: The first step is to parameterize each of the individual curves:

$$\begin{aligned} C_1 &: x = t, y = -1, && -2 \le t \le 0 \\ C_2 &: x = t, y = t^3 - 1, && 0 \le t \le 1 \\ C_3 &: x = 1, y = t, && 0 \le t \le 2 \end{aligned}$$

The next step is to compute the line integral over each of these curves

$$\int_{C_1} 4x^3\,ds = \int_{-2}^{0} 4t^3\sqrt{(1)^2+(0)^2}\,dt = \int_{-2}^{0} 4t^3\,dt = t^4\big]_{-2}^{0} = -16$$

$$\begin{aligned}\int_{C_2} 4x^3\,ds &= \int_0^1 4t^3\sqrt{(1)^2+(3t^2)^2}\,dt\\ &= \int_0^1 4t^3\sqrt{1+9t^4}\,dt\\ &= \frac{1}{9}\left(\frac{2}{3}\right)(1+9t^4)^{3/2}\big]_0^1\\ &= \frac{2}{27}\left(10^{3/2}-1\right)\end{aligned}$$

$$\int_{C_3} 4x^3\,ds = \int_0^2 4(1)^3\sqrt{(0)^2+(1)^2}\,dt = \int_0^2 4\,dt = 8$$

Finally, we take the sum of these three values to get the value of the line integral on C:

$$\begin{aligned}\int_C 4x^3\,ds &= \int_{C_1} 4x^3\,ds + \int_{C_2} 4x^3\,ds + \int_{C_3} 4x^3\,ds\\ &= -16 + \frac{2}{27}\left(10^{3/2}-1\right) + 8\\ &= \boxed{\frac{2}{27}\left(10^{3/2}-1\right) - 8}\end{aligned}$$

Remark. Notice how we put direction arrows on the curve. The direction of motion along a curve *may* change the value of the line integral as we will see very soon.

Line Integrals in Two Dimensions Let's consider what happens to the value of the line integral over a line segment when we change the direction of the path.

Problem 5.2.3. *Evaluate* $\int_C x^2\,ds$ *where* C *is:*

(a) the line segment from $(-2,-1)$ *to* $(1,2)$
(b) the line segment from $(1,2)$ *to* $(-2,-1)$.

Solution to part a: The parameterization of this line segment is

$$\begin{aligned}\mathbf{r}(t) &= (1-t)\langle -2,-1\rangle + t\langle 1,2\rangle\\ &= \langle -2+3t, -1+3t\rangle\end{aligned}$$

for $0 \le t \le 1$. That means the parametric equations are

$$x(t) = -2+3t \quad y = -1+3t$$

So using this path, the line integral is

$$\begin{aligned}\int_C x^2\,ds &= \int_0^1 (-2+3t)^2\sqrt{(3)^2+(3)^2}\,dt\\ &= 3\sqrt{2}\left(\frac{1}{9}\right)(-2+3t)^3\big]_0^1\\ &= \frac{\sqrt{2}}{3}(9) = \boxed{3\sqrt{2}}\end{aligned}$$

Solution to part b: Again, we parameterize the line segment.

$$\begin{aligned}\mathbf{r}(t) &= (1-t)\langle 1,2\rangle + t\langle -2,-1\rangle \\ &= \langle 1-3t, 2-3t\rangle\end{aligned}$$

for $0 \le t \le 1$. Be careful and note that we are switching the direction of the curve C, and this will also change the parameterization so we need to ensure that we start and end at the proper point. Here is the line integral.

$$\begin{aligned}\int_C x^2\,ds &= \int_0^1 (1-3t)^2\sqrt{(3)^2+(3)^2}\,dt \\ &= 3\sqrt{2}\left(-\frac{1}{3}\right)\cdot\frac{1}{3}(1-3t)^3\Big]_0^1 \\ &= -\frac{\sqrt{2}}{3}(-9) = \boxed{3\sqrt{2}}\end{aligned}$$

Remark. It appears that when we switch the direction of the curve, the value of the line integral (with respect to arc length) does not change. For such line integrals, this will always be the case. But there are other types of line integrals where this will not be the case.

Let's think about it more systematically. Suppose that a curve C has the parameterization $x = h(t)$, $y = g(t)$. Let A and B be the (respective) initial and final points on the curve. Then, the parameterization determines an *orientation* for C, in which the *positive* direction occurs as t increases.

Then let $-C$ be the curve with the same points as C, however in this case B and A are the (respective) initial and final points on the curve. As we traverse this curve, t is still increasing. Essentially, the curves C and $-C$ are identical, but the direction has been reversed.

As a result, line integrals with respect to arc length have the following property:

$$\boxed{\int_C f(x,y)\,ds = \int_{-C} f(x,y)\,ds}$$

So, for line integrals (with respect to arc length), we can change the direction of the curve without changing the value of the integral. This is useful to remember as some line integrals will be easier to compute in one direction than the other.

Problem 5.2.4. *Evaluate* $\int_C x\,ds$ *for the following curves.*

(a) $C_1 : y = x^2$, $-1 \le x \le 1$
(b) C_2*: the line segment from* $(-2,2)$ *to* $(2,2)$.
(c) C_3*: the line segment from* $(2,2)$ *to* $(-2,2)$.

Solution to part a: The parametric equations of C_1 are

$$x(t) = t \quad y(t) = t^2 \qquad -1 \le t \le 1$$

So the value of the line integral is

$$\int_{C_1} x\,ds = \int_{-1}^1 t\sqrt{1+4t^2}\,dt = \frac{1}{12}\left(1+4t^2\right)^{3/2}\Big|_{-1}^1 = \boxed{0}$$

Solution to part b: We don't have to use the generic formula for parameterizing a line segment here. Really, C_2 is just a portion of the graph of the line $y = 1$. Using that parameterization, we have

$$x(t) = t \quad y(t) = 1, \quad -2 \le t \le 2$$

The line integral over C_2 is

$$\int_{C_2} x\,ds = \int_{-2}^{2} t\sqrt{(1)^2+(0)^2}\,dt = \frac{1}{2}t^2|_{-2}^{2} = \boxed{0}$$

Solution to part c: We don't really need to do anything for this problem because we know $C_3 = -C_2$, and changing the direction of the curve does not affect the value of the line integral (with respect to arc length). Thus, the answer is $\boxed{0}$.

Line Integrals in Three Dimensions For three-dimensional curves, we have the following parameterization

$$x = x(t) \quad y = y(t) \quad z = z(t) \quad a \le t \le b$$

so the line integral over C is given by

$$\boxed{\int_C f(x,y,z)\,dz = \int_a^b f(x(t),y(t),z(t))\sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}\,dt}$$

As a vector function, the parameterization is

$$\mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$$

and the magnitude of this position vector is

$$|\mathbf{r}(t)| = \sqrt{\left(\frac{dx}{dt}\right)^2+\left(\frac{dy}{dt}\right)^2+\left(\frac{dz}{dt}\right)^2}$$

With this notation, the line integral can be written as

$$\boxed{\int_C f(x,y,z)\,ds = \int_a^b f(x(t),y(t),z(t))\,|\mathbf{r}(t)|\,dt}$$

Problem 5.2.5. *Evaluate $\int_C xyz\,ds$ where C is the helix given by $\boldsymbol{r}(t) = \langle\cos(t),\sin(t),2t\rangle$, $0 \le t \le 2\pi$.*

Solution: We are given the parameterization of C, so the line integral is

$$\begin{aligned}\int_C xyz\,ds &= \int_0^{2\pi} 2t\cos(t)\sin(t)\sqrt{\sin^2 t+\cos^2 t+4}\,dt\\ &= \int_0^{2\pi} 2t\left(\frac{1}{2}\sin(2t)\right)\sqrt{1+4}\,dt\\ &= \sqrt{5}\int_0^{2\pi} t\sin(2t)\,dt\\ &= \sqrt{5}\left(\frac{1}{4}\sin(2t)-\frac{t}{2}\cos(2t)\right)|_0^{2\pi}\\ &= \boxed{-\sqrt{5}\pi}\end{aligned}$$

***Remark*.** We used integration by parts to compute the integral of $t\sin(2t)$.

Line Integrals Without Respect to Arc Length In most cases, we won't be dealing with line integrals with respect to arc length; rather, we'll have line integrals with respect to a coordinate axis (x, y, or z). Let's start with two-dimensions. A curve C has parameterization

$$x = x(t) \quad y = y(t) \quad a \leq t \leq b$$

Definition 5.2.4. *The* ***line integral of f with respect to x*** *is*

$$\int_C f(x,y)\,dx = \int_a^b f(x(t),y(t))\,x'(t)\,dt$$

and the ***line integral of f with respect to y*** *is*

$$\int_C f(x,y)\,dy = \int_a^b f(x(t),y(t))\,y'(t)\,dt$$

Notice how similar these formulas are to the line integral with respect to arc length? The ONLY difference lies in the differential, i.e. using a dx or dy instead of a ds element. It's always a good idea to make sure what differential you are using so you don't confuse different types of line integrals. For the case of x and y, they often appear together, so we have the following shorthand notation:

$$\boxed{\int_C P\,dx + Q\,dy = \int_C P\,dx + \int_C Q\,dy}$$

Problem 5.2.6. *Evaluate the line integral* $\int_C \sin(\pi y)\,dy + yx^2\,dx$ *where:*

(a) C is the line segment from $(0,2)$ to $(1,4)$
(b) C is the line segment from $(1,4)$ to $(0,2)$

Solution to part a: The curve C is parameterized as

$$\mathbf{r}(t) = (1-t)\langle 0,2\rangle + t\langle 1,4\rangle = \langle t, 2+2t\rangle \quad 0 \leq t \leq 1$$

so the line integral is

$$\begin{aligned}\int_C \sin(\pi y)\,dy + yx^2\,dx &= \int_C \sin(\pi y)\,dy + \int_C yx^2\,dx \\ &= \int_0^1 \sin(\pi(2+2t))(2)\,dt + \int_0^1 (2+2t)^2(t)^2(1)\,dt \\ &= -\frac{1}{\pi}\cos(2\pi + 2\pi t)|_0^1 + \left(\frac{2}{3}t^3 + \frac{1}{2}t^4\right)|_0^1 \\ &= \boxed{\frac{7}{6}}\end{aligned}$$

Solution to part b: We simply opposed the direction of the curve as in part (a). Here is the new parameterization:

$$\mathbf{r}(t) = (1-t)\langle 1,4\rangle + t\langle 0,2\rangle = \langle 1-t, 4-2t\rangle \quad 0 \leq t \leq 1$$

So the line integral is

$$\begin{aligned}\int_C \sin(\pi y)\,dy + yx^2\,dx &= \int_C \sin(\pi y)\,dy + \int_C yx^2\,dx \\ &= \int_0^1 \sin(\pi(4-2t))(-2)\,dt + \int_0^1 (4-2t)(1-t)^2(-1)\,dt \\ &= -\frac{1}{\pi}\cos(4\pi - 2\pi t)|_0^1 - \left(-\frac{1}{2}t^4 + \frac{8}{3}t^3 - 5t^2 + 4t\right)|_0^1 \\ &= \boxed{-\frac{7}{6}}\end{aligned}$$

Remark. We observe that switching the direction of the curve gave us the opposite sign of the values of the line integral. It is a fact that this happens for all such line integrals:

$$\int_{-C} f(x,y)\,dx = -\int_C f(x,y)\,dx, \quad \int_{-C} f(x,y)\,dy = -\int_C f(x,y)\,dy$$

$$\therefore \int_{-C} P\,dx + Q\,dy = -\int_C P\,dx + Q\,dy$$

Extending to Three Dimensions We will simply pick up a third parameter, $z = z(t)$, and the line integral with respect to this parameter is simply

$$\int_C f(x,y,z)\,dz = \int_a^b f(x(t),y(t),z(t))\,z'(t)\,dt$$

and the parametric are, of course:

$$x = x(t) \quad y = y(t) \quad z = z(t) \quad a \le t \le b$$

For evaluating a line integral combining all three, the shorthand notation is

$$\boxed{\int_C P\,dx + Q\,dy + R\,dz = \int_C P(x,y,z)\,dx + \int_C Q(x,y,z)\,dy + \int_C R(x,y,z)\,dz}$$

Problem 5.2.7. *Evaluate the line integral* $\int_C y\,dx + x\,dy + z\,dz$ *where* C *is given by the parametric equations* $x = \cos t$, $y = \sin t$, $z = e^{2t}$, $0 \le t \le 2\pi$.

Solution: The parameterization has been given, so the line integral is now trivial.

$$\begin{aligned}\int_C y\,dx + x\,dy + z\,dz &= \int_C y\,dx + \int_C x\,dy + \int_C z\,dz \\ &= \int_0^{2\pi} \sin t(-\sin t)\,dt + \int_0^{2\pi} \cos t(\cos t)\,dt + \int_0^{2\pi} e^{2t}\left(2e^{2t}\right)\,dt \\ &= -\int_0^{2\pi} \sin^2 t\,dt + \int_0^{2\pi} \cos^2 t\,dt + \int_0^{2\pi} 2e^{4t}\,dt \\ &= -\frac{1}{2}\int_0^{2\pi} (1-\cos(2t))\,dt + \frac{1}{2}\int_0^{2\pi} (1+\cos(2t))\,dt + \int_0^{2\pi} 2e^{4t}\,dt \\ &= \left(-\frac{1}{2}\left(t - \frac{1}{2}\sin(2t)\right) + \frac{1}{2}\left(t + \frac{1}{2}\sin(2t)\right) + \frac{1}{2}e^{4t}\right)|_0^{2\pi} \\ &= \boxed{\frac{1}{2}\left(e^{8\pi} - 1\right)}\end{aligned}$$

Line Integrals of Vector Fields The last topic of this section involves computing line integrals of vector fields. Let $\mathbf{F}$ be a vector field on $\mathbb{R}^3$ such that

$$\mathbf{F}(x, y, z) = P(x, y, z)\,\mathbf{i} + Q(x, y, z)\,\mathbf{j} + R(x, y, z)\,\mathbf{k}$$

and the three-dimensional smooth curve has vector equation

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k} \quad a \leq t \leq b$$

Definition 5.2.5. *The* ***line integral of*** $\mathbf{F}$ ***along*** C *is given by*

$$\int_C \boldsymbol{F} \cdot d\boldsymbol{r} = \int_a^b \boldsymbol{F}(\boldsymbol{r}(t)) \cdot \boldsymbol{r}'(t)\,dt$$

Consider the notation on the left side. This is actually the dot product of the vector field and the differential of $\mathbf{r}$, the position vector. Also, the $\mathbf{F}(\mathbf{r}(t))$ is a shorthand for

$$\mathbf{F}(\mathbf{r}(t)) = \mathbf{F}(x(t), y(t), z(t))$$

We can also relate line integrals of vector fields to line integrals with respect to arc length by recognizing

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \mathbf{F} \cdot \mathbf{T}\,ds$$

where $\mathbf{T}$ is the unit tangent vector:

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

Here's how it actually works:

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_C \mathbf{F} \cdot \mathbf{T}\,ds \\
&= \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}|\mathbf{r}'(t)|\,dt \\
&= \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,dt
\end{aligned}$$

Generally, we use the first form because it is much easier in computations. Let's try a couple of problems.

Problem 5.2.8. *Evaluate the line integral* $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$, *where* $\boldsymbol{F}(x, y, z) = 8x^2yz\,\boldsymbol{i} + 5z\,\boldsymbol{j} - 4xy\,\boldsymbol{k}$ *and* C *is the curve* $\boldsymbol{r}(t) = t\,\boldsymbol{i} + t^2\,\boldsymbol{j} + t^3\,\boldsymbol{k}$, *for* $0 \leq t \leq 1$.

Solution: The first step is to evaluate $\mathbf{F}(\mathbf{r}(t))$, or the value of the vector field along the curve:

$$\mathbf{F}(\mathbf{r}(t)) = 8t^2(t^2)(t^3)\,\mathbf{i} + 5t^3\,\mathbf{j} - 4t(t^2)\,\mathbf{k} = 8t^7\,\mathbf{i} + 5t^3\,\mathbf{j} - 4t^3\,\mathbf{k}$$

The derivative of the parameterization, $\mathbf{r}'(t)$, is equal to

$$\mathbf{r}'(t) = \mathbf{i} + 2t\,\mathbf{j} + 3t^2\,\mathbf{k}$$

so the dot product, $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$, is given by

$$\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = 8t^7 + 10t^4 - 12t^5$$

so the line integral is

$$\begin{aligned}
\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (8t^7 + 10t^4 - 12t^5)\,dt \\
&= (t^8 + 2t^5 - 2t^6)\big|_0^1 \\
&= \boxed{1}
\end{aligned}$$

Problem 5.2.9. *Evaluate the line integral $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$, where $\boldsymbol{F}(x, y, z) = xz\,\boldsymbol{j} + xy\,\boldsymbol{k}$ and C is the line segment from $(-1, 1, 0)$ to $(4, 0, 1)$.*

Solution: The first step here is to parameterize C.

$$\begin{aligned}\mathbf{r}(t) &= (1-t)\langle -1, 1, 0\rangle + t\langle 4, 0, 1\rangle \\ &= \langle -1 + 5t, 1 - t, t\rangle\end{aligned}$$

for $0 \le t \le 1$. The next step is to evaluate the vector field at the curve:

$$\mathbf{F}(\mathbf{r}(t)) = (-1+5t)(t)\,\mathbf{j} + (-1+5t)(1-t)\,\mathbf{k} = 0\,\mathbf{i} + (-t+5t^2)\,\mathbf{j} + (-1+5t-5t^2)\,\mathbf{k}$$

We then find the derivative of the parameterization, or $\mathbf{r}'(t)$:

$$\mathbf{r}'(t) = 5\,\mathbf{i} - \mathbf{j} + \mathbf{k}$$

The dot product $\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)$ is equal to

$$\mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t) = 6t - 10t^2 - 1$$

so the line integral is

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_0^1 (6t - 10t^2 - 1)\,dt \\ &= \left(3t^2 - \frac{10}{3}t^3 - t\right)\Big|_0^1 \\ &= \boxed{-\frac{4}{3}}\end{aligned}$$

Let's end this section with a nice result we get by generalizing the relationship between line integrals of vector fields to those of parameters x, y, and z. In particular, let $\mathbf{F}(x, y, z) = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ and the curve C have parameterization $\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$, for $a \le t \le b$. Then

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= \int_a^b (P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}) \cdot (x'\,\mathbf{i} + y'\,\mathbf{j} + z'\,\mathbf{k})\,dt \\ &= \int_a^b (Px' + Qy' + Rz')\,dt \\ &= \int_a^b Px'\,dt + \int_a^b Qy'\,dt + \int_a^b Rz'\,dt \\ &= \int_C P\,dx + \int_C Q\,dy + \int_C R\,dz \\ &= \int_C P\,dx + Q\,dy + R\,dz\end{aligned}$$

so in general, we can directly evaluate a line integral over a vector field by pairing its components with the components of the position vector:

$$\boxed{\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C P\,dx + Q\,dy + R\,dz}$$

And because vector fields can be defined in terms of line integrals with respect to x, y, and z, the idea that reversing the direction of path with line integrals changes the sign of the numerical value should apply to line integrals of vector fields.

$$\boxed{\int_{-C} \mathbf{F} \cdot d\mathbf{r} = -\int_C \mathbf{F} \cdot d\mathbf{r}}$$

5.3 Fundamental Theorem of Line Integrals

In single-variable calculus, we learned about the Fundamental Theorem of Calculus, which stated for a function $F(x)$ and its derivative $F'(x)$, we have the following property of definite integrals:

$$\int_a^b F'(x)\,dx = F(b) - F(a)$$

There's a version of this theorem for line integrals over certain vector fields.

Theorem 5.3.1. *Suppose C is a smooth curve given by $\boldsymbol{r}(t)$, $a \le t \le b$. Also, let f be a function whose gradient ∇f is continuous on the curve. Then*

$$\int_C \nabla f \cdot d\boldsymbol{r} = f(\boldsymbol{r}(b)) - f(\boldsymbol{r}(a))$$

Note that $\mathbf{r}(a)$ and $\mathbf{r}(b)$ are the initial and terminal points, respectively, on C. Also, we did not specify the number of variables because this theorem will hold true regardless.

Proof. For the purpose of this proof we will assume three dimensions, although it can be done in any number of dimensions.

$$\begin{aligned}
\int_C \nabla f \cdot d\mathbf{r} &= \int_a^b \nabla f(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,dt \\
&= \int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt}\right) dt \quad \text{Evaluate the line integral} \\
&= \int_a^b \frac{d}{dt}\left[f(\mathbf{r}(t))\right] dt \quad \text{Chain Rule} \\
&= f(\mathbf{r}(b)) - f(\mathbf{r}(a)) \quad \text{Fundamental Theorem of Calculus}
\end{aligned}$$

and the proof is complete. □

Problem 5.3.1. *Evaluate $\int_C \nabla f \cdot d\boldsymbol{r}$, where $f(x, y, z) = \cos(\pi x) + \sin(\pi y) - xyz$ and C is a path that starts at $(1, \frac{1}{2}, 2)$ and ends at $(2, 1, -1)$.*

Solution: Let $\mathbf{r}(t)$, $a \le t \le b$ be any path starting at $(1, \frac{1}{2}, 2)$ and ending at $(2, 1, -1)$. Then we have two position vectors below:

$$\mathbf{r}(a) = \left\langle 1, \frac{1}{2}, 2 \right\rangle \qquad \mathbf{r}(b) = \langle 2, 1, -1 \rangle$$

So the value of the line integral is

$$\begin{aligned}
\int_C \nabla f \cdot d\mathbf{r} &= f(2, 1, -1) - f\left(1, \frac{1}{2}, 2\right) \\
&= \cos(2\pi) + \sin\pi - 2(1)(-1) - \left(\cos\pi + \sin\left(\frac{\pi}{2}\right) - 1\left(\frac{1}{2}\right)(2)\right) \\
&= \boxed{4}
\end{aligned}$$

The main takeaway for this problem is not on how to compute the line integral, as it is pretty straightforward; all we did we plug in the terminal and initial points and subtract the values on f.

What really matters is the bigger idea behind this example (and the Fundamental Theorem of Calculus): for this type of line integral, we **don't** need to know the actual path taken between the two points. Regardless of whatever path we choose, the value of the line integral will be the same.

Earlier, when we first talked about line integrals (even before introducing vector fields along with it), we saw that changing the path usually *does* change the value of the integral. Now, we've found a special kind of line integral where changing the path *does not* affect the result.

To explore this further, we have got some definitions below. The first one is a review of something we've already seen (but it's important here, so it's worth repeating). The remaining definitions will be new.

Definition 5.3.1. *Suppose that $\boldsymbol{F}$ is a continuous vector field in some parameter domain D. The following hold true:*

1. *$\boldsymbol{F}$ is a* ***conservative*** *vector field if there is a function f such that $\boldsymbol{F} = \nabla f$. The function f is called a* ***potential function*** *for the vector field.*
2. *The line integral $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$ is called* ***path-independent*** *if $\int_{C_1} \boldsymbol{F} \cdot d\boldsymbol{r}$ and $\int_{C_2} \boldsymbol{F} \cdot d\boldsymbol{r}$ for any paths C_1 and C_2 in D with the same initial and final points.*
3. *A path C is said to be* ***closed*** *if its initial and final points are the same, e.g. circle.*
4. *A path C is said to be* ***simple*** *if it does not cross itself. A circle is a simple curve, a figure 8 curve is not.*
5. *A region D is called* ***open*** *if it does not contain any of its boundary points.*
6. *A region D is* ***connected*** *if we can connect any two points in the region with a path that lies completely in D.*
7. *A region D is* ***simply-connected*** *if it is connected and it contains no holes.*

From these definitions, the following statements are valid:

1. The line integral $\int_C \nabla f \cdot d\mathbf{r}$ is path-independent. Trivial proof; the Fundamental Theorem for Line Integrals states that in order to evaluate this line integral, all we need are the initial and final points. Therefore, the line integral must be independent of path.
2. If $\mathbf{F}$ is a conservative vector field, then $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path-independent. This is also easy to prove. If $\mathbf{F}$ is conservative then there exists a potential function such that $\mathbf{F} = \nabla f$. Then $\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \nabla f \cdot d\mathbf{r}$. Since the line integral on the right side of the equation is path-independent by definition, then $\int_C \mathbf{F} \cdot d\mathbf{r}$ must be path-independent.
3. If $\mathbf{F}$ is a continuous vector field on an open, connected region D and if $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path-independent for any path in D, then $\mathbf{F}$ is a conservative vector field on D. This is the converse of the previous statement
4. If $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path-independent, then $\int_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed path C.
5. If $\int_C \mathbf{F} \cdot d\mathbf{r} = 0$ for every closed path, then $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path-independent. This is the converse of the previous statement.

Problem 5.3.2. *Given that $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$ is independent of path, compute the line integral $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$, where C is the ellipse given by $\frac{(x-5)^2}{4} + \frac{y^2}{9} = 1$ with counterclockwise rotation.*

Solution: You're probably wondering, how can we solve this problem without the vector field? Well, it's actually quite simple because the vector field is unnecessary in this problem. We are told that C is an ellipse. More importantly, it is a *closed* curve. The value of a line integral of this form around a closed path will be zero given that it is path-independent. Therefore, $\int_C \mathbf{F} \cdot d\mathbf{r} = \boxed{0}$.

Problem 5.3.3. *Evaluate* $\int_C \nabla f \cdot d\boldsymbol{r}$ *where* $f(x, y) = x^3(3 - y^2) + 4y$ *and* C *is given by* $\boldsymbol{r}(t) = \langle 3 - t^2, 5 - t\rangle$ *with* $-2 \leq t \leq 3$.

Solution: We need to integrate over a gradient vector field, ∇f, and so the integral is set up to apply the Fundamental Theorem of Line Integrals. We need two "points," which are really position vectors, using the endpoints of our parameter interval:

$$\mathbf{r}(-2) = \langle -1, 7\rangle \qquad \mathbf{r}(3) = \langle -6, 2\rangle$$

The Fundamental Theorem of Line Integrals gives us

$$\begin{aligned}\int_C \nabla f \cdot d\mathbf{r} &= f(\mathbf{r}(3)) - f(\mathbf{r}(-2)) \\ &= f(-6, 2) - f(-1, 7) \\ &= 224 - 74 = \boxed{150}\end{aligned}$$

Conservative Vector Fields We have established the notion that if a vector field $\mathbf{F}$ is conservative then the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ is independent of path. These imply that we can easily evaluate the line integral using a potential function for $\mathbf{F}$. We now propose two questions:

1. How can we determine if a vector field $\mathbf{F}$ is conservative?
2. If $\mathbf{F}$ is a conservative vector field, how can we find a potential function for the vector field?

Theorem 5.3.2. *Let* $\boldsymbol{F} = P\,\boldsymbol{i} + Q\,\boldsymbol{j}$ *be a two-dimensional vector on an open and simply-connected region* D. *If* P *and* Q *have continuous first order partial derivatives in* D *and*

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

then $\boldsymbol{F}$ *is a conservative vector field.*

Problem 5.3.4. *Determine if the following vector field is conservative or not.*

$$\boldsymbol{F}(x, y) = (x^2 - yx)\,\boldsymbol{i} + (y^2 - xy)\,\boldsymbol{j}$$

Solution: In this case, $P = x^2 - yx$ and $Q = y^2 - xy$. The partial derivatives are $\frac{\partial P}{\partial y} = -x$ and $\frac{\partial Q}{\partial x} = -y$. Since these two partial derivatives are not equal, the vector field is $\boxed{\text{not conservative}}$.

Now we need to address how to find a potential function for a conservative vector field. Actually, this is quite easy to do. We know that for $\mathbf{F} = \nabla f$, we have

$$\nabla f = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} = P\,\mathbf{i} + Q\,\mathbf{j} = \mathbf{F}$$

Setting corresponding components equal to each other, we have

$$\frac{\partial f}{\partial x} = P \quad \text{and} \quad \frac{\partial f}{\partial y} = Q$$

Integrating each of these with respect to the appropriate independent variable, we have

$$f(x, y) = \int P(x, y)\,dx \quad \text{or} \quad f(x, y) = \int Q(x, y)\,dy$$

Let's reinforce this concept with a simple problem.

Problem 5.3.5. *Determine if the following vector field is conservative. If it is, find a potential function.*

$$\boldsymbol{F}(x,y) = (2x^3y^4 + x)\,\boldsymbol{i} + (2x^4y^3 + y)\,\boldsymbol{j}$$

Solution: First, we identify P and Q and check the corresponding partial derivatives.

$$P = 2x^3y^4 + x \quad \text{and} \quad Q = 2x^4y^3 + y$$

We observe that

$$\frac{\partial P}{\partial y} = 8x^3y^3 = \frac{\partial Q}{\partial x}$$

We can show that $\mathbf{F}$ is $\boxed{\text{conservative}}$. Now let's find the potential function. Since

$$\frac{\partial f}{\partial x} = 2x^3y^4 + x \quad \text{and} \quad \frac{\partial f}{\partial y} = 2x^4y^3 + y$$

then we can integrate either equation to solve for f:

$$f(x,y) = \int (2x^3y^4 + x)\,dx \quad \text{or} \quad f(x,y) = \int (2x^4y^3 + y)\,dy$$

If we take the first integral, then

$$\begin{aligned} f(x,y) &= \int (2x^3y^4 + x)\,dx \\ &= \frac{1}{2}x^4y^4 + \frac{1}{2}x^2 + h(y) \end{aligned}$$

where $h(y)$ is treated as our "constant of integration." It is not too hard to find $h(y)$. Until this point, we used the P component of $\mathbf{F}$, but we also need to use the Q component to finish the problem. Recall that Q is the partial derivative of f with respect to y. So let's differentiate the above result with respect to y to get

$$\frac{\partial f}{\partial y} = 2x^4y^3 + h'(y) = 2x^4y^3 + y = Q$$

By inspection, we know

$$h'(y) = y$$

so $h(y) = \int y\,dy = \frac{1}{2}y^2 + C$. Putting everything together, we see a potential function for the vector field as

$$\boxed{f(x,y) = \frac{1}{2}x^4y^4 + \frac{1}{2}x^2 + \frac{1}{2}y^2 + C}$$

Remark. We can always check our work by verifying $\nabla f = \mathbf{F}$. Also, C can be any real number so there are an infinite number of potential functions of the form we obtained.

Now, we (currently) do not have a way to determine whether a three-dimensional vector field is conservative, however, if we are given the assumption that it is, then we can also find a potential function. The only difference is that it will take another step. Here:

$$\nabla f = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} + \frac{\partial f}{\partial z}\,\mathbf{k} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k} = \mathbf{F}$$

Problem 5.3.6. *Given that the following vector field* $\boldsymbol{F}$ *is conservative, find a potential function for* $\boldsymbol{F}$.

$$\boldsymbol{F} = 2xy^3z^4\,\boldsymbol{i} + 3x^2y^2z^4\,\boldsymbol{j} + 4x^2y^3z^3\,\boldsymbol{k}$$

Solution: We have three equalities:

$$\frac{\partial f}{\partial x} = 2xy^3z^4 \quad \frac{\partial f}{\partial y} = 3x^2y^2z^4 \quad \frac{\partial f}{\partial z} = 4x^2y^3z^3$$

Let's start by integrating the first with respect to x:

$$f(x, y, z) = \int 2xy^3z^4\,dx = x^2y^3z^4 + g(y, z)$$

The "constant of integration" is a function of both y and z because differentiating anything of that form with respect to x will result in 0.

Now we differentiate with respect to y and set it equal to the Q component:

$$\frac{\partial f}{\partial y} = 3x^2y^2z^4 + g_y(y, z) = 3x^2y^2z^4 = Q$$

We can observe that $g_y(y, z) = 0$. Since differentiating $g(y, z)$ with respect to y results in 0, $g(y, z)$ can at most be a function of z, so we write $g(y, z) = h(z)$ and

$$f(x, y, z) = x^2y^3z^4 + h(z)$$

Differentiating with respect to z and setting it equal to the R component gives

$$\frac{\partial f}{\partial z} = 4x^2y^3z^3 + h'(z) = 4x^2y^3z^3 = R$$

We see that $h'(z) = 0$, and $h(z) = \int h'(z)\,dz = C$, where C is any real number. A potential function for $\mathbf{F}$ is then

$$\boxed{f(x, y, z) = x^2y^3z^4 + C}$$

Problem 5.3.7. *Evaluate $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$, where $\boldsymbol{F}$ is the vector field from problem* 5.3.5 *and C is the curve given by $\boldsymbol{r}(t) = (t\cos(\pi t) - 1)\,\boldsymbol{i} + \sin\left(\frac{\pi t}{2}\right)\,\boldsymbol{j}$, with $0 \le t \le 1$.*

Solution: We know that the vector field in problem 5.3.5 is conservative, and the potential function f for which $\mathbf{F} = \nabla f$ is

$$f(x, y) = \frac{1}{2}x^4y^4 + \frac{1}{2}x^2 + \frac{1}{2}y^2 + C$$

Since $\mathbf{F}$ is independent, then the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ must be independent of path, so all we need to do is use the Fundamental Theorem of Line Integrals:

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C \nabla f \cdot d\mathbf{r} = f(\mathbf{r}(1)) - f(\mathbf{r}(0))$$

where $\mathbf{r}(1)$ and $\mathbf{r}(0)$ are the respective position vectors

$$\mathbf{r}(1) = \langle -2, 1\rangle \quad \text{and} \quad \mathbf{r}(0) = \langle -1, 0\rangle$$

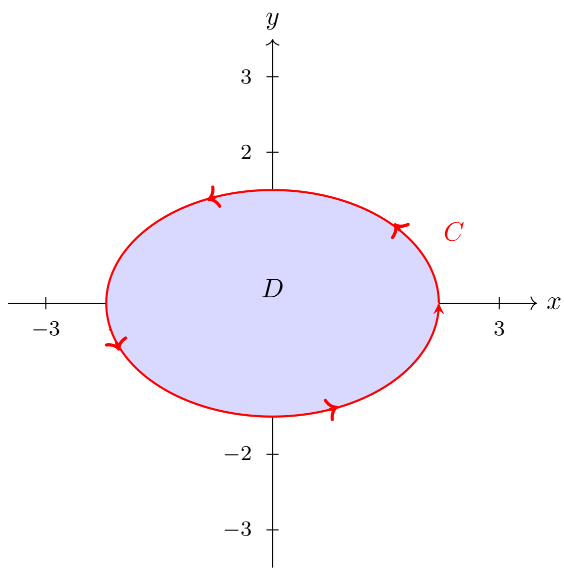
So the value of the line integral is

$$\begin{aligned}\int_C \mathbf{F} \cdot d\mathbf{r} &= f(-2, 1) - f(-1, 0) \\ &= \left(\frac{21}{2} + C\right) - \left(\frac{1}{2} + C\right) \\ &= \boxed{10}\end{aligned}$$

5.4 Green's Theorem

In this section, we will talk about a powerful theorem that relates a line integral around a simple closed curve to a double integral of the plane region enclosed by that curve.

Core Idea Let C be a simple closed curve and let D be the region bounded by C, as shown below. Assume that D consists all points inside and on the boundary of C. In Green's Theorem we understand that the curve is traversed counterclockwise, giving C *positive* orientation.



Theorem 5.4.1. *Let C be a positively oriented, piecewise-smooth, simple closed curve in the plane and let D be the region bounded by C. If P and Q have continuous partial derivatives on an open region containing D, then*

$$\int_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$$

Notice that $\int_C P\,dx + Q\,dy$ is another way of expressing $\int_C \mathbf{F}\cdot d\mathbf{r}$, where $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j}$.

The proof of Green's Theorem is non-trivial, but let us consider the special case where the region is both type I and type II (see section 4.2 in this book). We define such regions as **simple regions**.

We can prove Green's Theorem if we can establish the following:

$$\int_C P\,dx = -\iint_D \frac{\partial P}{\partial y}\,dA$$

and

$$\int_C Q\,dy = \iint_D \frac{\partial Q}{\partial x}\,dA$$

Let $P(x, y)$ and $Q(x, y)$ have continuous partial derivatives on a set containing a closed, positively oriented, piecewise-smooth curve C and its interior plane region D.
Since D is both type I and II, we can decompose the boundary C in two ways:
For **type I**, we slice vertically, yielding $D = \{(x, y)\,|\,a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$. Meanwhile, for **type II**, we slice horizontally, yielding $D = \{(x, y)\,|\,c \le y \le d,\ h_1(y) \le x \le h_2(y)\}$.
We know that

$$\oint_C P\,dx + Q\,dy = \oint_C P\,dx + \oint_C Q\,dy$$

Using the type I description, we get

$$\oint_C Q\,dy = \int_a^b Q(x, g_2(x))g_2'(x)\,dx - \int_a^b Q(x, g_1(x))g_1'(x)\,dx + \int_a^b \int_{g_1(x)}^{g_2(x)} \frac{\partial Q(x,y)}{\partial x}\,dy\,dx$$

It's important to realize that on the top and bottom pieces, $dy = g'(x)\,dx$, but all the horizontal segments cancel out via symmetry, except for the difference between top and bottom, so it's easiest to apply the Fundamental Theorem of Calculus on each vertical slice like this:

$$\int_{g_1(x)}^{g_2(x)} \frac{\partial Q}{\partial x}(x,y)\,dy = Q(x, g_2(x))g_2'(x) - Q(x, g_1(x))g_1'(x)$$

which effectively reduces to $\oint Q\,dy = \iint_D \frac{\partial Q}{\partial x}\,dA$.
Similarly, we apply horizontal slicing on the $P\,dx$ component:

$$-\oint_C P\,dx = \iint_D \frac{\partial P}{\partial y}\,dA \therefore \oint_C P\,dx = -\iint_D \frac{\partial P}{\partial y}\,dA$$

and finally, combine the two results to yield

$$\oint_C P\,dx + Q\,dy = -\iint_D \frac{\partial P}{\partial y}\,dA + \iint_D \frac{\partial Q}{\partial x}\,dA = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dA$$

and the proof is complete. □

Problem 5.4.1. *Evaluate $\int_C x^4\,dx + xy\,dy$, where C is the triangular curve consisting of the line segments from $(0,0)$ to $(1,0)$, from $(1,0)$ to $(0,1)$, and from $(0,1)$ to $(0,0)$.*

Solution: Note that the plane region D enclosed by C is a simple region. Moreover, C has positive orientation, because this order of vertex traversal is counterclockwise. Therefore, we can use Green's Theorem, much faster than otherwise having to evaluate three line integrals.

$$\begin{aligned}\int_C x^4\,dx + xy\,dy &= \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)dA \\ &= \int_0^1 \int_0^{1-x} (y-0)\,dy\,dx \\ &= \int_0^1 \left[\frac{1}{2}y^2\right]_{y=0}^{y=1-x} dx = \frac{1}{2}\int_0^1 (1-x)^2\,dx \\ &= -\frac{1}{6}(1-x)^3\,\Big|_0^1 = \boxed{\frac{1}{6}}\end{aligned}$$

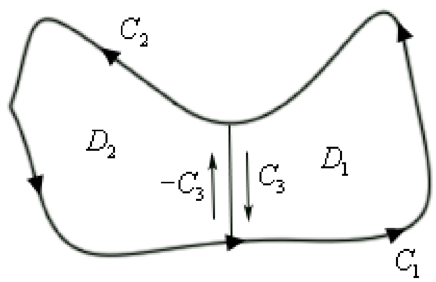
Problem 5.4.2. *Evaluate $\oint_C \left(3y - e^{\sin x}\right)\,dx + \left(7x + \sqrt{y^4+1}\right)\,dy$, where C is the circular disk $x^2+y^2 = 9$.*

Solution: Assume that C is positively oriented, i.e. the circle is traversed counterclockwise. Additionally, the region D bounded by the curve is the disk $x^2 + y^2 \le 9$, so it makes sense to use polar coordinates immediately after setting up the integral via Green's Theorem.

$$\begin{aligned}\oint \left(3y - e^{\sin x}\right)\,dx + \left(7x + \sqrt{y^4+1}\right)\,dy &= \iint_D \left[\frac{\partial}{\partial x}\left(7x + \sqrt{y^4+1}\right) - \frac{\partial}{\partial y}\left(3y - e^{\sin x}\right)\right]dA \\ &= \int_0^{2\pi}\int_0^3 (7-3)\,r\,dr\,d\theta = 4\int_0^{2\pi} d\theta \int_0^3 r\,dr \\ &= \boxed{36\pi}\end{aligned}$$

Handling Regions That Are Not Simple We know that Green's Theorem only applies to regions bounded by simple closed curves, i.e. the regions do not have holes. However, many regions in fact have holes in them, so let's see how we can deal with them.

Consider the region below. Even though it does not technically have holes, the arguments we will use are essentially the same for regions that do have holes in them.



The region D is the union of D_1 and D_2, or $D = D_1 \cup D_2$. The region D_1 is bounded by $C_1 \cup C_3$ while the region D_2 is bounded by $C_2 \cup (-C_3)$. Both of these boundaries are also positively oriented: in other words, as we traverse each boundary the corresponding region is always located to the *left*. The overall boundary curve is:

$$C = (C_1 \cup C_3) \cup (C_2 \cup (-C_3)) = C_1 \cup C_2$$

Because the paths C_3 and $-C_3$ "cancel" each other. Let's demonstrate this, first using properties of double integrals.

$$\iint_D (Q_x - P_y)\, dA = \iint_{D_1 \cup D_2} (Q_x - P_y)\, dA = \iint_{D_1} (Q_x - P_y)\, dA + \iint_{D_2} (Q_x - P_y)\, dA$$

According to Green's Theorem, we can convert these double integrals into line integrals, and also break up the line integrals into separate line integrals for each portion of the boundary.

$$\begin{aligned}\iint_D (Q_x - P_y)\, dA &= \iint_{D_1} (Q_x - P_y)\, dA + \iint_{D_2} (Q_x - P_y)\, dA \\ &= \oint_{C_1 \cup C_3} P\,dx + Q\,dy + \oint_{C_2 \cup (-C_3)} P\,dx + Q\,dy \\ &= \oint_{C_1} P\,dx + Q\,dy + \oint_{C_3} P\,dx + Q\,dy + \oint_{C_2} P\,dx + Q\,dy + \oint_{-C_3} P\,dx + Q\,dy\end{aligned}$$

Since $\oint_{-C_3} P\,dx + Q\,dy = -\oint_{C_3} P\,dx + Q\,dy$, we get

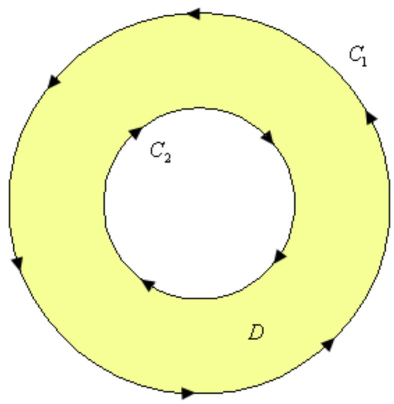
$$\begin{aligned}\oint_{C_1} P\,dx + Q\,dy + \oint_{C_3} P\,dx + Q\,dy + \oint_{C_2} P\,dx + Q\,dy + \oint_{-C_3} P\,dx + Q\,dy \\ = \oint_{C_1} P\,dx + Q\,dy + \oint_{C_2} P\,dx + Q\,dy\end{aligned}$$

Finally, put all the line integrals together to get

$$\begin{aligned}\iint_D (Q_x - P_y)\,dA &= \oint_{C_1} P\,dx + Q\,dy + \oint_{C_2} P\,dx + Q\,dy \\ &= \oint_{C_1 \cup C_2} P\,dx + Q\,dy \\ &= \oint_C P\,dx + Q\,dy\end{aligned}$$

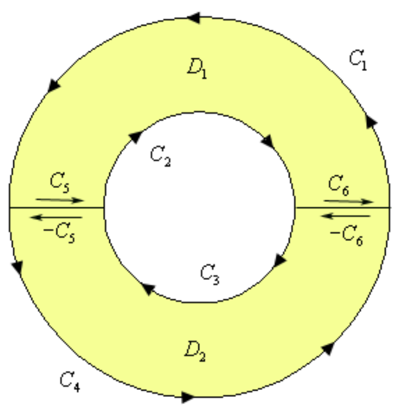
Alright, so what's the takeaway from this? At first, it might feel like we just did a lot of work to get a result we already knew from Green's Theorem. But here's the cool part: when we break up the region like we did, the parts of the line integral on the curve in the middle of the region actually cancel each other out because they point in opposite directions. This is super useful for when we're working with regions that have holes in them.

For instance, look at the ring below:



Notice this: both curves are positively oriented because the region D is on the left as we move along the curve in the direction shown. But here's something interesting: curve C_2 seems to break the original rule about positive orientation. We usually say a curve is positively oriented if you go counter-clockwise, right? But that rule only works for regions without holes, i.e. simple regions. When there's a hole, we need to use a different definition for orientation−one we talked about earlier.

Since this region has a hole, we apparently can't use Green's Theorem or any line integral with a curve $C = C_1 \cup C_2$. However, we can cut up the disk and rename all the various portions of the curves to get the following sketch:



The upper half boundary of D_1 is given by $C_1 \cup C_2 \cup C_5 \cup C_6$ and the lower half boundary of D_2 is $C_3 \cup C_4 \cup (-C_5) \cup (-C_6)$. What's special about these two halves? They don't have any holes, i.e. they are simple regions, so we can use Green's Theorem!

$$\begin{aligned}\iint_D (Q_x - P_y)\, dA &= \iint_{D_1} (Q_x - P_y)\, dA + \iint_{D_2} (Q_x - P_y)\, dA \\ &= \oint_{C_1 \cup C_2 \cup C_5 \cup C_6} P\, dx + Q\, dy + \oint_{C_3 \cup C_4 \cup (-C_5) \cup (-C_6)} P\, dx + Q\, dy\end{aligned}$$

We can now break up these line integrals into line integrals on each piece of the boundary. Recall that boundaries with the same curve but opposite direction "cancel" each other.

$$\begin{aligned}\iint_D (Q_x - P_y)\, dA &= \iint_{D_1} (Q_x - P_y)\, dA + \iint_{D_2} (Q_x - P_y)\, dA \\ &= \oint_{C_1} P\, dx + Q\, dy + \oint_{C_2} P\, dx + Q\, dy + \oint_{C_3} P\, dx + Q\, dy + \oint_{C_4} P\, dx + Q\, dy\end{aligned}$$

Finally, we can add up the line integrals to obtain

$$\begin{aligned}\iint_D (Q_x - P_y)\, dA &= \oint_{C_1 \cup C_2 \cup C_3 \cup C_4} P\, dx + Q\, dy \\ &= \oint_C P\, dx + Q\, dy\end{aligned}$$

The key takeaway from this case is that we could have just used Green's Theorem on the ring/disk from the start. This will be true in general for regions that contain holes. Let's look at a problem.

Problem 5.4.3. *Evaluate $\oint_C y^3\, dx - x^3\, dy$ where C are the two circles of radius 2 and 1 centered at the origin with positive orientation.*

Solution: Based on the description of C, it is best to use polar coordinates after converting the line integral into a double integral.

$$\begin{aligned}\oint_C y^3\, dx - x^3\, dy &= \iint_D (-3x^2 - 3y^2)\, dA \\ &= -3 \iint_D (x^2 + y^2)\, dA \\ &= -3 \int_0^{2\pi} \int_1^2 r^3\, dr\, d\theta \\ &= -3 \int_0^{2\pi} \frac{1}{4} r^4 \Big|_1^2\, d\theta \\ &= -3 \cdot \frac{15}{4} \int_0^{2\pi} d\theta \\ &= \boxed{-\frac{45\pi}{2}}\end{aligned}$$

Using Green's Theorem to Find Area We can actually use the reverse of Green's Theorem to calculate the area A of some region D. Since this measurement is defined as $A = \iint_D 1\, dA$, we need a vector field with components P and Q such that

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1$$

We identify three possible sets $\{P, Q\}$:

$$\begin{array}{lll} P(x,y) = 0 & P(x,y) = -y & P(x,y) = -\frac{1}{2}y \\ Q(x,y) = x & Q(x,y) = 0 & Q(x,y) = \frac{1}{2}x \end{array}$$

Green's Theorem expresses the area of the plane region D in three equations:

$$\boxed{A = \oint_C x\,dy = -\oint_C y\,dx = \frac{1}{2}\oint_C x\,dy - y\,dx}$$

Problem 5.4.4. *Show, using Green's Theorem, that the area of an ellipse with standard-form equation* $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ *is given by* $A = \pi ab$*.*

Solution: The most important thing to do here is to parametrize the ellipse. The relevant equations are $x = a\cos t$ and $y = b\sin t$, for $0 \le t \le 2\pi$. Therefore, Green's Theorem gives us

$$\begin{aligned} A &= \frac{1}{2}\int_C x\,dy - y\,dx \\ &= \frac{1}{2}\int_0^{2\pi} (a\cos t)(b\cos t)\,dt - (b\sin t)(-a\sin t)\,dt \\ &= \frac{ab}{2}\int_0^{2\pi} dt = \boxed{\pi ab} \end{aligned}$$

5.5 Curl and Divergence

When dealing with vector fields, there are two special operations that play a fundamental role in applications to fluid flow, electricity, magnetism, etc. One operation produces a scalar field while the other produces a vector field.

Curl If $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ is a vector field on $\mathbb{R}^3$ and the partial derivatives of the component functions P, Q, and R all exist, then the curl of $\mathbf{F}$ is the vector field on $\mathbb{R}^3$ defined by

$$\boxed{\operatorname{curl}\mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}}$$

Obviously, this is pretty dense and hard to just rote memorize. We introduce the vector differential operator:

$$\nabla = \mathbf{i}\,\frac{\partial}{\partial x} + \mathbf{j}\,\frac{\partial}{\partial y} + \mathbf{k}\,\frac{\partial}{\partial z}$$

If it operates on a scalar function f, it produces the gradient of f:

$$\nabla f = \mathbf{i}\,\frac{\partial f}{\partial x} + \mathbf{j}\,\frac{\partial f}{\partial y} + \mathbf{k}\,\frac{\partial f}{\partial z} = \frac{\partial f}{\partial x}\,\mathbf{i} + \frac{\partial f}{\partial y}\,\mathbf{j} + \frac{\partial f}{\partial z}\,\mathbf{k}$$

If you think of ∇ as a vector with components $\partial/\partial x$, $\partial/\partial y$, and $\partial/\partial z$, the formal cross product of ∇ with the vector field $\mathbf{F}$ are as follows:

$$\begin{aligned} \nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix} \\ &= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} \\ &= \operatorname{curl}\mathbf{F} \end{aligned}$$

Hence the easiest way to remember the formula for the curl of a vector field is to understand it as a vector cross product:

$$\boxed{\operatorname{curl}\mathbf{F} = \nabla \times \mathbf{F}}$$

Problem 5.5.1. *If* $\boldsymbol{F}(x, y, z) = xz\,\boldsymbol{i} + xyz\,\boldsymbol{j} - y^2\,\boldsymbol{k}$, *find curl* $\boldsymbol{F}$.

Solution: Using the cross product of the differential operator and the vector field, we get

$$\begin{aligned}
\text{curl } \mathbf{F} = \nabla \times \mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz & xyz & -y^2 \end{vmatrix} \\
&= \left[\frac{\partial}{\partial y}(-y^2) - \frac{\partial}{\partial z}(xyz)\right]\mathbf{i} - \left[\frac{\partial}{\partial x}(-y^2) - \frac{\partial}{\partial z}(xz)\right]\mathbf{j} \\
&\quad + \left[\frac{\partial}{\partial x}(xyz) - \frac{\partial}{\partial y}(xz)\right]\mathbf{k} \\
&= (-2y - xy)\mathbf{i} - (0 - x)\mathbf{j} + (yz - 0)\mathbf{k} \\
&= \boxed{-y(2 + x)\mathbf{i} + x\mathbf{j} + yz\mathbf{k}}
\end{aligned}$$

Remember that ∇f is also a vector field on $\mathbb{R}^3$ and so we can also compute its curl. The following theorem states the curl of a gradient vector field is $\mathbf{0}$. Note that this is the zero **vector**, not the number 0.

Theorem 5.5.1. *If f is a function of three variables with continuous second order partial derivatives, then the curl of its gradient vector field is the zero vector.*

$$\textit{curl } (\nabla f) = \boldsymbol{0}$$

We can prove this theorem pretty easily.

$$\begin{aligned}
\text{curl } (\nabla f) = \nabla \times (\nabla f) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix} \\
&= \left(\frac{\partial^2 f}{\partial y\,\partial z} - \frac{\partial^2 f}{\partial z\,\partial y}\right)\mathbf{i} + \left(\frac{\partial^2 f}{\partial z\,\partial x} - \frac{\partial^2 f}{\partial x\,\partial z}\right)\mathbf{j} + \left(\frac{\partial^2 f}{\partial x\,\partial y} - \frac{\partial^2 f}{\partial y\,\partial x}\right)\mathbf{k} \\
&= 0\,\mathbf{i} + 0\,\mathbf{j} + 0\,\mathbf{k} = \mathbf{0}
\end{aligned}$$

and the proof is complete via Clairaut's Theorem. □

Since a conservative vector field $\mathbf{F}$ has the property $\mathbf{F} = \nabla f$, the above theorem can be rephrased:

$$\text{If } \mathbf{F} \text{ is conservative, then curl } \mathbf{F} = \mathbf{0}$$

This rephrasing actually helps us demonstrate when a vector field is *not* conservative.

Problem 5.5.2. *Show that the vector field* $\boldsymbol{F}$ *in problem* 5.5.1 *is not conservative.*

Solution: From the solution to the stated problem, we know that

$$\text{curl } \mathbf{F} = -y(2 + x)\,\mathbf{i} + x\,\mathbf{j} + yz\,\mathbf{k}$$

which is not the zero vector $\mathbf{0}$. Therefore, by the remarks prior to this problem, we know that $\mathbf{F}$ is not conservative.

Note: The converse of Theorem 5.5.1 is not generally true, but the theorem that follows states that the converse is only true if $\mathbf{F}$ is defined everywhere. In other words, if the domain is simply-connected (i.e. contains no "holes"). However, its proof requires Stokes' Theorem and will be demonstrated at the end of section 5.8 of this book.

Theorem 5.5.2. *If* $\boldsymbol{F}$ *is a vector field on all of* $\mathbb{R}^3$ *whose component functions all have continuous partial derivatives and curl* $\boldsymbol{F} = \boldsymbol{0}$*, then* $\boldsymbol{F}$ *is a conservative vector field.*

Problem 5.5.3. *Show that the vector field*

$$\boldsymbol{F}(x, y, z) = y^2z^3\,\boldsymbol{i} + 2xyz^3\,\boldsymbol{j} + 3xy^2z^2\,\boldsymbol{k}$$

is conservative. Additionally, find a function f *such that* $\boldsymbol{F} = \nabla f$*.*

Solution: The curl of $\mathbf{F}$ is found by computing the cross product $\nabla \times \mathbf{F}$:

$$\begin{aligned} \text{curl}\,\mathbf{F} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y^2z^3 & 2xyz^3 & 3xy^2z^2 \end{vmatrix} \\ &= (6xyz^2 - 6xyz^2)\,\mathbf{i} - (3y^2z^2 - 3y^2z^2)\,\mathbf{j} + (2yz^3 - 2yz^3)\,\mathbf{k} \\ &= \mathbf{0} \end{aligned}$$

Since curl $\mathbf{F} = \mathbf{0}$ and the domain of $\mathbf{F}$ ($\mathbb{R}^3$) is simply-connected, $\mathbf{F}$ is a conservative vector field according to Theorem 5.5.2.

Now to find f, we use the technique that was provided in section 5.3, by use of the potential function: we want a function f such that $f_x = y^2z^3$, $f_y = 2xyz^3$, and $f_z = 3xy^2z^2$. If we integrate the first equation in the set with respect to x, we get

$$f(x, y, z) = xy^2z^3 + g(y, z)$$

If we differentiate this result with respect to y, we get

$$f_y(x, y, z) = 2xyz^2 + g_y(y, z)$$

If we compare this with the second equation in the set ($f_Y = 2xyz^3$), we find that $g_y(y, z) = 0$. Thus, $g(y, z)$ only depends on z and can be written as $h(z)$. Finally, differentiate with respect to z and get

$$f_z(x, y, z) = 3xy^2z^2 + h'(z)$$

and comparison with $f_z = 3xy^2z^2$ we can easily see that $h'(z) = 0$. Thus, $h(z) = C$ where C is any constant. Therefore, we find $\boxed{f(x, y, z) = xy^2z^3 + C}$.

Divergence If $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ is a vector field on $\mathbb{R}^3$ and the partial derivatives P_x, Q_y, and R_z exist, then the **divergence of F** is the function of three variables defined by

$$\boxed{\text{div}\,\mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}}$$

Notice something here. curl $\mathbf{F}$ is a vector field, while div $\mathbf{F}$ is a scalar field. In terms of the gradient operator $\nabla = (\partial/\partial x)\,\mathbf{i} + (\partial/\partial y)\,\mathbf{j} + (\partial/\partial z)\,\mathbf{k}$, so the divergence of $\mathbf{F}$ can be symbolically described as the dot product of ∇ with $\mathbf{F}$:

$$\boxed{\text{div}\,\mathbf{F} = \nabla \cdot \mathbf{F}}$$

Problem 5.5.4. *A vector field on* $\mathbb{R}^3$ *is given by*

$$\boldsymbol{F}(x, y, z) = (2xy^2 - z\sin(y))\,\boldsymbol{i} + (e^x + 3y)\,\boldsymbol{j} + (x - z^2)\,\boldsymbol{k}$$

Find div $\boldsymbol{F}$*.*

Solution: The divergence is the dot product of the gradient operator and the vector field.

$$\begin{aligned}\text{div }\mathbf{F} &= \nabla \cdot \mathbf{F}\\ &= \frac{\partial}{\partial x}(2xy^2 - zsin(y)) + \frac{\partial}{\partial y}(e^x + 3y) + \frac{\partial}{\partial z}(x - z^2)\\ &= 2y^2 + 3 + (-2z)\\ &= \boxed{2y^2 - 2z + 3}\end{aligned}$$

We already know that if $\mathbf{F}$ is a vector field on $\mathbb{R}^3$, then its curl is also a vector field on $\mathbb{R}^3$. Then we can also find its divergence. The following theorem demonstrates that the result is 0.

Theorem 5.5.3. *If $\boldsymbol{F} = P\,\boldsymbol{i} + Q\,\boldsymbol{j} + R\,\boldsymbol{k}$ is a field on $\mathbb{R}^3$ and P, Q, and R all have continuous second partial derivatives, then*

$$\textit{div curl } \boldsymbol{F} = 0$$

Note that this result is the scalar quantity, 0, not the zero vector, $\mathbf{0}$.
Let's prove the theorem using the definitions of curl and divergence.

$$\begin{aligned}\text{div curl }\mathbf{F} &= \nabla \cdot (\nabla \times \mathbf{F})\\ &= \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\\ &= \frac{\partial^2 R}{\partial x\,\partial y} - \frac{\partial^2 Q}{\partial x\,\partial z} + \frac{\partial^2 P}{\partial y, \partial z} - \frac{\partial^2 R}{\partial y\,\partial x} + \frac{\partial^2 Q}{\partial z\,\partial x} - \frac{\partial^2 P}{\partial z\,\partial y}\\ &= 0\end{aligned}$$

as pairwise terms cancel out via Clairaut's Theorem, and the proof is complete. □

Problem 5.5.5. *Prove that the vector field $\boldsymbol{F}(x, y, z) = xz\,\boldsymbol{i} + xyz\,\boldsymbol{j} - y^2\,\boldsymbol{k}$ cannot be written as the curl of another vector field, i.e. $\boldsymbol{F} \neq$ curl $\boldsymbol{G}$ for any vector field $\boldsymbol{G}$.*

Solution: The divergence of $\mathbf{F}$ is given by

$$\begin{aligned}\text{div }\mathbf{F} &= \nabla \cdot \mathbf{F}\\ &= \frac{\partial}{\partial x}(xz) + \frac{\partial}{\partial y}(xyz) + \frac{\partial}{\partial z}(-y^2)\\ &= z + xz + 0\\ &= z(1 + x)\end{aligned}$$

and is therefore nonzero. If it were true that $\mathbf{F} = \text{curl }\mathbf{G}$, then Theorem 5.5.3 would mandate

$$\text{div }\mathbf{F} = \text{div curl }\mathbf{G} = 0$$

but this is a contradiction, since div $\mathbf{F} = z(1 + x) \neq 0$. Therefore, $\mathbf{F}$ cannot be the curl of another vector field, and the proof is complete. □

Vector Forms of Green's Theorem The final topic of this section is to cover the two different vector representations of Green's Theorem, which we studied in the previous section. As you probably guessed, they both involve the gradient operator, ∇.

The first one involves the curl of a vector field.

$$\boxed{\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_D (\text{curl }\mathbf{F}) \cdot \mathbf{k}\, dA}$$

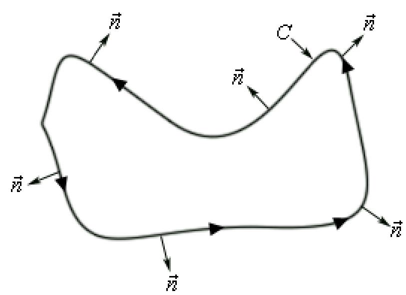
where $\mathbf{k}$ is the standard unit vector pointing in the positive z direction.

The second form uses the divergence of a vector field. This one is a little different because we need the outward unit normal vector to the curve C. If we parameterize

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$$

then the outward normal is given by:

$$\mathbf{n} = \frac{y'(t)}{|\mathbf{r}'(t)|}\,\mathbf{i} - \frac{x'(t)}{|\mathbf{r}'(t)|}\,\mathbf{j}$$



The image below shows how the unit normal vectors function, as they point perpendicularly outward from a curve C at various points. The vector form of Green's Theorem using the divergence is

$$\boxed{\oint_C \mathbf{F}\cdot\mathbf{n}\,ds = \iint_D (\text{div }\mathbf{F})\,dA}$$

Both these vector forms will be very useful when we see higher-dimensional versions of the ideas that we have about vector fields and line integrals so far. For example, the curl of a vector field will become significant when studying *Stokes' Theorem* (covered in section 5.8), and the divergence will be critical to understanding the *Divergence Theorem* (covered in section 5.9).

5.6 Parametric Surfaces

Before, we parameterized a curve by taking some values of t, our parameter, from an interval $[a, b]$ and then plugging into the parametric equation

$$\mathbf{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$$

so that the resulting set of vectors would serve as position vectors for the points on the curve.

Introduction to Parametric Surfaces For surfaces, we'll do something similar. We choose points (u, v) from a two-dimensional region D and plug them into the parametric equation

$$\mathbf{r}(u, v) = x(u, v)\,\mathbf{i} + y(u, v)\,\mathbf{j} + z(u, v)\,\mathbf{k}$$

and the resulting set of vectors will be the position vectors for the points on the surface S that we wish to parameterize. This is often called the parametric representation of the **parametric surface** S.

Many times, we will need to write down parametric equations for a surface S.

$$x = x(u, v) \quad y = y(u, v) \quad z = z(u, v)$$

Problem 5.6.1. *Determine the surface given by the parametric representation below.*

$$\boldsymbol{r}(u,v) = u\,\boldsymbol{i} + u\cos v\,\boldsymbol{j} + u\sin v\,\boldsymbol{k}$$

Solution: The parametric equations for S are

$$x = u \quad y = u\cos v \quad z = u\sin v$$

Notice that if you sum the squares of y and z, you get

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

So let's eliminate the parameters u and v. We are then left with

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

so the surface is a $\boxed{\text{cone}}$ that opens along the x-axis.

In most cases, however, we will need to be able to write down the parametric equations of a surface rather than identify the surface from the parametric representation.

Problem 5.6.2. *Give parametric representations for each of the following surfaces.*

(a) The elliptic paraboloid $x = 5y^2 + 2z^2 - 10$.
(b) The elliptic paraboloid $x = 5y^2 + 2z^2 - 10$ *in front of the* yz*-plane.*
(c) The sphere $x^2 + y^2 + z^2 = 30$.
(d) The cylinder $y^2 + z^2 = 25$.

Solution to part a: This is very easy to parameterize. We can choose y and z to be anything, and x will always have a value of $5y^2 + 2z^2 - 10$. Thus, the parametric equations are

$$x = 5y^2 + 2z^2 - 10 \quad y = y \quad z = z$$

Since the surface is in the form $x = f(y,z)$, the parametric representation is

$$\boxed{\mathbf{r}(y,z) = (5y^2 + 2z^2 - 10)\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}}$$

Solution to part b: The parametric representation is the same as in (a):

$$\mathbf{r}(y,z) = (5y^2 + 2z^2 - 10)\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$$

However, we need $x \geq 0$ since the surface lies in front of the yz-plane, we need

$$5y^2 + 2z^2 - 10 \geq 0 \therefore 5y^2 + 2z^2 \geq 0$$

Thus, the final parametric representation is

$$\boxed{\mathbf{r}(y,z) = (5y^2 + 2z^2 - 10)\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}, \quad 5y^2 + 2z^2 \geq 0}$$

Solution to part c: This one can be a little tricky, but we need to use our knowledge of spherical coordinates. A sphere centered at the origin with radius a is written as

$$\rho = a$$

and so the equation of this sphere is $\rho = \sqrt{30}$. Also, we have three formulas to convert from Cartesian to spherical coordinates:

$$x = \rho\sin\phi\cos\theta \quad y = \rho\sin\phi\sin\theta \quad z = \rho\cos\phi$$

We know that ρ is for our sphere so if we plug $\rho = \sqrt{30}$ into these conversions we will arrive at the parametric representation

$$\mathbf{r}(\theta, \phi) = \sqrt{30}\sin\phi\cos\theta\,\mathbf{i} + \sqrt{30}\sin\phi\sin\theta\,\mathbf{j} + \sqrt{30}\cos\phi\,\mathbf{k}$$

Since we are dealing with a full sphere, $0 \le \phi \le \pi$, and because we are not retracing the surface, $0 \le \theta \le 2\pi$. The final parametric representation is then

$$\boxed{\mathbf{r}(\theta, \phi) = \sqrt{30}\sin\phi\cos\theta\,\mathbf{i} + \sqrt{30}\sin\phi\sin\theta\,\mathbf{j} + \sqrt{30}\cos\phi\,\mathbf{k}, \quad 0 \le \phi \le \pi,\, 0 \le \theta \le 2\pi}$$

Solution to part d: As with (c), this can be a little tricky, but we will use cylindrical coordinates to simplify the problem-solving process. In cylindrical coordinates, a cylinder with radius a is described by the equation

$$r = a$$

and so the equation of this cylinder is $r = 5$. Now, since the cylinder is centered on the x-axis, we can set x to be anything (since we haven't restricted the height of the cylinder), while y and z will be written in polar form.

$$x = x \quad y = r\sin\theta \quad z = r\cos\theta$$

Since we know what r is, we can plug in to get the parametric representation in terms of x and θ. Additionally, we are only tracing the cylinder once, so $0 \le \theta \le 2\pi$, and our final answer is

$$\boxed{\mathbf{r}(x, \theta) = x\,\mathbf{i} + 5\sin\theta\,\mathbf{j} + 5\cos\theta\,\mathbf{k}, \quad 0 \le \theta \le 2\pi}$$

Generalized Parametric Representations Note that part (a) for problem 5.6.2 involved a function in the form $x = f(y, z)$ that we could quickly parameterize. Actually, this can always be done for functions in the basic forms:

$$\begin{aligned}
z = f(x, y) &\Rightarrow \mathbf{r}(x, y) = x\,\mathbf{i} + y\,\mathbf{j} + f(x, y)\,\mathbf{k} \\
x = f(y, z) &\Rightarrow \mathbf{r}(y, z) = f(y, z)\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k} \\
y = f(x, z) &\Rightarrow \mathbf{r}(x, z) = x\,\mathbf{i} + f(x, z)\,\mathbf{j} + z\,\mathbf{k}
\end{aligned}$$

Applications of Parametric Surfaces: Tangent Planes Consider a surface S traced out by the vector function

$$\mathbf{r}(u, v) = x(u, v)\,\mathbf{i} + y(u, v)\,\mathbf{j} + z(u, v)\,\mathbf{k}$$

at an arbitrary point P_0 with position vector $\mathbf{r}(u_0, v_0)$. Let u be fixed at $u = u_0$. Then $\mathbf{r}(u_0, v)$ is only dependent upon the v parameter and defines a grid curve C_1 lying on the surface. The tangent vector to C_1 at P_0 is obtained by taking the partial derivative of $\mathbf{r}$ with respect to v:

$$\mathbf{r}_v = \frac{\partial x}{\partial v}(u_0, v_0)\,\mathbf{i} + \frac{\partial y}{\partial v}(u_0, v_0)\,\mathbf{j} + \frac{\partial z}{\partial v}(u_0, v_0)\,\mathbf{k}$$

Similarly, we can fix $v = v_0$, so we get a grid curve C_2 given by $\mathbf{r}(u, v_0)$ lying on S, with tangent vector being

$$\mathbf{r}_u = \frac{\partial x}{\partial u}(u_0, v_0)\,\mathbf{i} + \frac{\partial y}{\partial u}(u_0, v_0)\,\mathbf{j} + \frac{\partial z}{\partial u}(u_0, v_0)\,\mathbf{k}$$

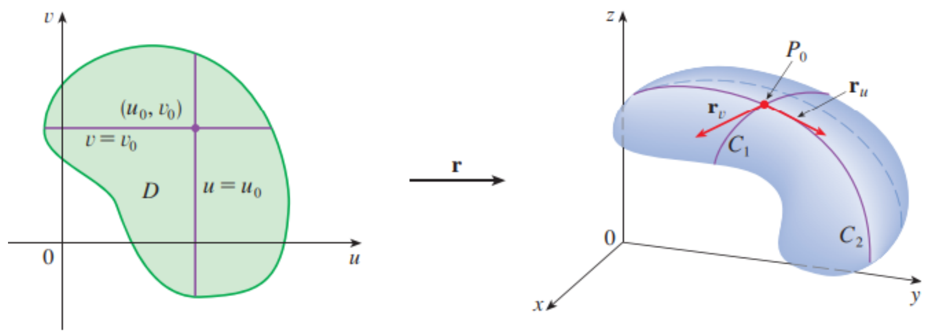


Image Credit: Calculus, Early Transcendentals

If the cross product $\mathbf{r}_u \times \mathbf{r}_v$ is never $\mathbf{0}$, then S is called a **smooth** surface (i.e. it has no "corners").

Definition 5.6.1. *For a smooth surface, the **tangent plane** is the plane that contains the two tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, and the vector $\mathbf{r}_u \times \mathbf{r}_v$ is the **normal vector** to the tangent plane.*

Problem 5.6.3. *Find the tangent plane to the surface with parametric equations $x = u^2$, $y = v^2$, and $z = u + 2v$ at the point $(1, 1, 3)$.*

Solution: The first step is to compute the tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$:

$$\mathbf{r}_u = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j} + \frac{\partial z}{\partial u}\,\mathbf{k} = 2u\,\mathbf{i} + \mathbf{k}$$
$$\mathbf{r}_v = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j} + \frac{\partial z}{\partial v}\,\mathbf{k} = 2v\,\mathbf{j} + 2\,\mathbf{k}$$

A normal vector $\mathbf{n}$ to the plane is

$$\begin{aligned}\mathbf{n} &= \mathbf{r}_u \times \mathbf{r}_v \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2u & 0 & 1 \\ 0 & 2v & 2 \end{vmatrix} \\ &= -2v\,\mathbf{i} - 4u\,\mathbf{j} + 4uv\,\mathbf{k}\end{aligned}$$

The point $(1, 1, 3)$ indicates that

$$1 = u^2 \quad 1 = v^2 \quad 3 = u + 2v$$

so the parameter values are $u = 1$, $v = 1$. If we substitute into the cross product, we get the normal vector

$$-2\,\mathbf{i} - 4\,\mathbf{j} + 4\,\mathbf{k}$$

Thus, the equation of the tangent plane is

$$-2(x-1) - 4(y-1) + 4(z-3) = 0 \therefore \boxed{x + 2y - 2z + 3 = 0}$$

Applications of Parametric Surfaces: Surface Area Consider a surface S whose parameter domain D is a rectangle. If we divide it into smaller sub-rectangles R_{ij}. Choose (u_i^*, v_j^*) to be the lower left corner of R_{ij}.

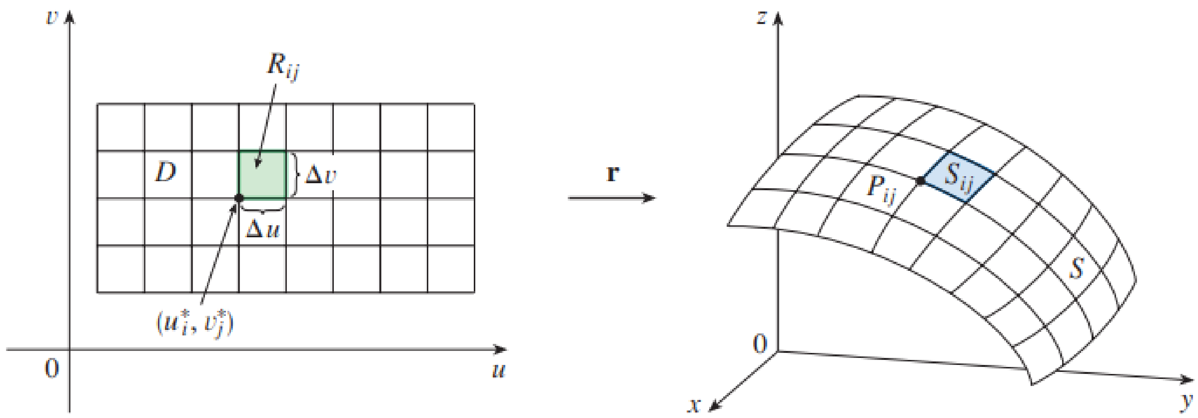


Image Credit: Calculus, Early Transcendentals

The part S_{ij} of the surface S corresponding to R_{ij} is called a *patch* and has the point P_{ij} with position vector $\mathbf{r}(u_i^*, v_j^*)$ as one of its corners. The tangent vectors at the point P_{ij} are

$$\mathbf{r}_u^* = \mathbf{r}_u(u_i^*, v_j^*) \quad \text{and} \quad \mathbf{r}_v^* = \mathbf{r}_v(u_i^*, v_j^*)$$

The two "edges" of the patch intersect at this special point. We can approximate them as vectors, which can also be approximated by the vectors $\Delta u\, \mathbf{r}_u^*$ and $\Delta v\, \mathbf{r}_v^*$ because partial derivatives can be approximated using different quotients. So S_{ij} can be approximated as a parallelogram determined by those two vectors. The area of this parallelogram is equal to the magnitude of the cross product, or

$$|(\Delta u\, \mathbf{r}_u^*) \times (\Delta v\, \mathbf{r}_v^*)| = |\mathbf{r}_u^* \times \mathbf{r}_v^*|\, \Delta u\, \Delta v$$

and so it can be approximated by

$$\sum_{i=1}^{m} \sum_{j=1}^{n} |\mathbf{r}_u^* \times \mathbf{r}_v^*|\, \Delta u\, \Delta v$$

and our intuition says that our approximation becomes more accurate as we continue increasing the number of sub-rectangles, and this double summation eventually becomes a double integral, $\iint_D |\mathbf{r}_u \times \mathbf{r}_v|\, du\, dv$.

Definition 5.6.2. *If a smooth parametric surface S is given by the equation*

$$\boldsymbol{r}(u, v) = x(u, v)\, \boldsymbol{i} + y(u, v)\, \boldsymbol{j} + z(u, v)\, \boldsymbol{k} \quad (u, v) \in D$$

and S is covered exactly once as (u, v) ranges throughout the parameter domain D, then the ***surface area*** *of S is*

$$A(S) = \iint_D |\boldsymbol{r}_u \times \boldsymbol{r}_v|\, dA$$

Problem 5.6.4. *Find the surface area of a sphere of radius a.*

Solution: The parametric equations describing this sphere are

$$x = a \sin\phi \cos\theta \quad y = a \sin\phi \sin\theta \quad z = a \cos\phi$$

where the parameter domain is

$$D = \{(\phi, \theta) \,|\, 0 \le \phi \le \pi,\, 0 \le \theta \le 2\pi\}$$

The two tangent vectors are $\mathbf{r}_\phi = a\cos\phi\cos\theta\,\mathbf{i} + a\cos\phi\sin\theta\,\mathbf{j} - a\sin\phi\,\mathbf{k}$ and $\mathbf{r}_\theta = -a\sin\phi\sin\theta\,\mathbf{i} + a\sin\phi\cos\theta\,\mathbf{j}$. Now we compute the cross product:

$$\begin{aligned}\mathbf{r}_\phi \times \mathbf{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} & \frac{\partial z}{\partial \phi} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \end{vmatrix} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a\cos\phi\cos\theta & a\cos\phi\sin\theta & -a\sin\phi \\ -a\sin\phi\sin\theta & a\sin\phi\cos\theta & 0 \end{vmatrix} \\ &= a^2\sin^2\phi\cos\theta\,\mathbf{i} + a^2\sin^2\phi\sin\theta\,\mathbf{j} + a^2\sin\phi\cos\phi\,\mathbf{k}\end{aligned}$$

The magnitude of the cross product, $|\mathbf{r}_\phi \times \mathbf{r}_\theta|$ is

$$\begin{aligned}|\mathbf{r}_\phi \times \mathbf{r}_\theta| &= \sqrt{a^4\sin^4\phi\cos^2\theta + a^4\sin^4\theta\sin^2\theta + a^4\sin^2\phi\cos^2\phi} \\ &= \sqrt{a^4\sin^4\phi + a^4\sin^2\phi\cos^2\phi} \\ &= \sqrt{a^4\sin^2\phi\,(\sin^2\phi + \cos^2\phi)} \\ &= \sqrt{a^4\sin^2\phi} = a^2\sin\phi\end{aligned}$$

as $\sin\phi \geq 0$ for $0 \leq \phi \leq \pi$. The area of the sphere is

$$\begin{aligned}A = \iint_D |\mathbf{r}_\phi \times \mathbf{r}_\theta|\,dA &= \int_0^{2\pi}\int_0^{\pi} a^2\sin\phi\,d\phi\,d\theta \\ &= a^2\int_0^{2\pi} d\theta \int_0^{\pi}\sin\phi\,d\phi = a^2(2\pi)(2) \\ &= \boxed{4\pi a^2}\end{aligned}$$

Finally, let's consider the special case of a surface S with equation $z = f(x, y)$, where (x, y) lies in D and f has continuous partial derivatives. Taking x and y to be parameters, the parametric equations are

$$x = x \quad y = y \quad z = f(x, y)$$

so the tangent vectors are

$$\mathbf{r}_x = \mathbf{i} + \left(\frac{\partial f}{\partial x}\right)\mathbf{k} \qquad \mathbf{r}_y = \mathbf{j} + \left(\frac{\partial f}{\partial y}\right)\mathbf{k}$$

and the cross product is

$$\mathbf{r}_x \times \mathbf{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{vmatrix} = -\frac{\partial f}{\partial x}\mathbf{i} - \frac{\partial f}{\partial y}\mathbf{j} + \mathbf{k}$$

Finally,

$$|\mathbf{r}_x \times \mathbf{r}_y| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}$$

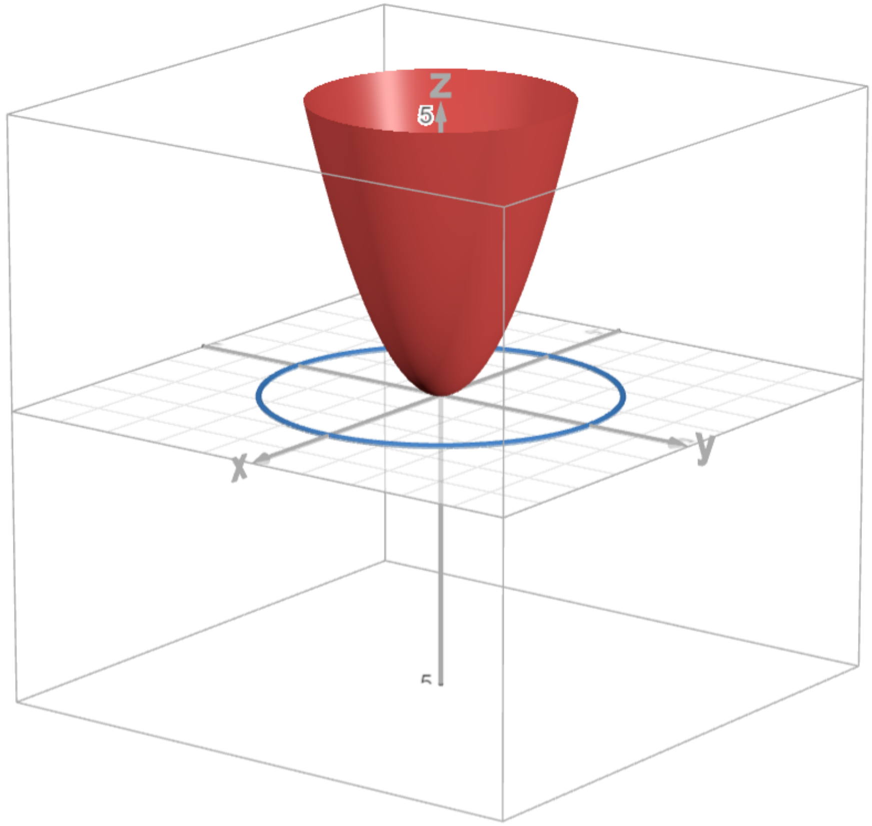
and the surface area becomes

$$\boxed{A(S) = \iint_D \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}\,dA}$$

which is the same result we derived in section 4.5!

Problem 5.6.5. *Calculate the surface area of the part of the paraboloid $z = x^2 + y^2$ that lies under the plane $z = 9$.*

Solution: As shown below, the plane intersects the paraboloid in the circle $x^2 + y^2 = 9$, $z = 9$.



Therefore, the surface lies above the disk D centered at the origin with radius 3, or $x^2 + y^2 \leq 9$. Applying the formula, we have

$$\begin{aligned} A &= \iint_D \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1} \\ &= \iint_D \sqrt{(2x)^2 + (2y)^2 + 1}\, dA = \iint_D \sqrt{4(x^2 + y^2) + 1} \end{aligned}$$

Because we are integrating over the disk $x^2 + y^2 \leq 9$, we will use polar coordinates to obtain

$$\begin{aligned} A &= \int_0^{2\pi} \int_0^3 \sqrt{4r^2 + 1}\, r\, dr\, d\theta = \int_0^{2\pi} d\theta \int_0^3 r\sqrt{4r^2 + 1}\, dr \\ &= 2\pi \left(\frac{1}{8}\right)\left(\frac{2}{3}\right)(1 + 4r^2)^{3/2}]_0^3 = \boxed{\frac{\pi}{6}\left(37\sqrt{37} - 1\right)} \end{aligned}$$

5.7 An Understanding of Surface Integrals

In section 5.2, we explored the relationship between line integrals and arc length. This is very similar to the relationship between surface integrals and surface area, which we will uncover in this section. Suppose f is a function of three variables whose domain includes a surface S. The surface integral of f over S is defined in a way, such that, when $f(x, y, z) = 1$, the value of the integral is equal to the surface area of S.

Parametric Surfaces Suppose a surface S is described by the parameterization:

$$\mathbf{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k} \quad (u,v) \in D$$

We assume that the parameter domain D is a rectangle, which can be divided into sub-rectangles R_{ij} with dimensions of Δu and Δv. Then S is divided into patches S_{ij}, as shown in the previous section.

Let's evaluate f at a point P_{ij}^* in each patch, multiply by the area ΔS_{ij} of the patch, and form the Riemann sum

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{ij}^*)\,\Delta S_{ij}$$

The accuracy of this approximation increases as we increase the number of patches, eventually giving us the **surface integral of f over the surface S**:

$$\boxed{\iint_S f(x,y,z)\,dS = \lim_{m,n\to\infty}\sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{ij}^*)\,\Delta S_{ij}}$$

To evaluate this surface integral we approximate ΔS_{ij} by using a parallelogram in the tangent plane. In the previous section, we demonstrated it to be

$$\Delta S_{ij} \approx |\mathbf{r}_u \times \mathbf{r}_v|\,\Delta u\,\Delta v$$

where $\mathbf{r}_u$ and $\mathbf{r}_v$ are the tangent vectors at a corner of S_{ij}. If the components are continuous and $\mathbf{r}_u$ and $\mathbf{r}_v$ are nonzero and not parallel in the interior of D, then

$$\boxed{\iint_S f(x,y,z)\,dS = \iint_D f(\mathbf{r}(u,v))|\mathbf{r}_u \times \mathbf{r}_v|\,dA}$$

Observe that when $f(x,y,z) = 1$, then we have

$$\iint_S 1\,dS = \iint_D |\mathbf{r}_u \times \mathbf{r}_v|\,dA = A(S)$$

Remember that $f(\mathbf{r}(u,v))$ is evaluated by writing $x = x(u,v)$, $y = y(u,v)$, and $z = z(u,v)$ in the formula for $f(x,y,z)$.

Problem 5.7.1. *Evaluate $\iint_S z\,dS$ where S is the upper half of a sphere of radius 2.*

Solution: The parameterization of this sphere is

$$\mathbf{r}(\phi,\theta) = 2\sin\phi\cos\theta\,\mathbf{i} + 2\sin\phi\sin\theta\,\mathbf{j} + 2\cos\phi\,\mathbf{k}$$

with $0 \le \theta \le 2\pi$ and $0 \le \phi \le \frac{\pi}{2}$, since we are only tracing the upper half of the sphere once. Now, we need to determine the tangent vectors, $\mathbf{r}_\phi$ and $\mathbf{r}_\theta$.

$$\begin{aligned}\mathbf{r}_\phi &= 2\cos\phi\cos\theta\,\mathbf{i} + 2\cos\phi\sin\theta\,\mathbf{j} - 2\sin\phi\,\mathbf{k}\\ \mathbf{r}_\theta &= -2\sin\phi\sin\theta\,\mathbf{i} + 2\sin\phi\cos\theta\,\mathbf{j}\end{aligned}$$

Now we take the cross product:

$$\begin{aligned}\mathbf{r}_\phi \times \mathbf{r}_\theta &= \begin{vmatrix}\mathbf{i} & \mathbf{j} & \mathbf{k}\\ 2\cos\phi\cos\theta & 2\cos\phi\sin\theta & -2\sin\phi\\ -2\sin\phi\sin\theta & 2\sin\phi\cos\theta & 0\end{vmatrix}\\ &= 4\sin^2\phi\cos\theta\,\mathbf{i} + 4\sin^2\phi\sin\theta\,\mathbf{j} + 4\sin\phi\cos\phi\,\mathbf{k}\end{aligned}$$

The magnitude of the cross product is

$$\begin{aligned}
|\mathbf{r}_\phi \times \mathbf{r}_\theta| &= \sqrt{16\sin^4\phi\cos^2\theta + 16\sin^4\phi\sin^2\theta + 16\sin^2\phi\cos^2\phi} \\
&= \sqrt{16\sin^4\phi(\cos^2\theta + \sin^2\theta) + 16\sin^2\phi\cos^2\phi} \\
&= \sqrt{16\sin^2\phi(\sin^2\phi + \cos^2\phi)} \\
&= 4\sqrt{\sin^2\phi} = 4\sin\phi
\end{aligned}$$

since $\sin\phi \geq 0$ when $0 \leq \phi \leq \pi/2$. The surface integral is then

$$\begin{aligned}
\iint_S z\,dS &= \iint_D 2\cos\phi\,(4\sin\phi)\;dA \\
&= \int_0^{2\pi}\int_0^{\pi/2} 4\sin(2\phi)\,d\phi\,d\theta \\
&= \int_0^{2\pi} (-2\cos(2\phi))\big]_0^{\pi/2}\,d\theta \\
&= \int_0^{2\pi} 4\,d\theta = \boxed{8\pi}
\end{aligned}$$

Graphs of Functions Any surface S with equation $z = g(x, y)$ can be regarded as a parametric surface, where

$$x = x \quad y = y \quad z = g(x, y)$$

which gives tangent vectors

$$\mathbf{r}_x = \mathbf{i} + \left(\frac{\partial g}{\partial x}\right)\mathbf{k} \qquad \mathbf{r}_y = \mathbf{j} + \left(\frac{\partial g}{\partial y}\right)\mathbf{k}$$

Also,

$$\mathbf{r}_x \times \mathbf{r}_y = -\frac{\partial g}{\partial x}\mathbf{i} - \frac{\partial g}{\partial y}\mathbf{j} + \mathbf{k}$$

and

$$|\mathbf{r}_x \times \mathbf{r}_y| = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}$$

so the surface integral becomes

$$\boxed{\iint_S f(x, y, z)\,dS = \iint_D f(x, y, g(x, y))\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}\,dA}$$

Note that similar, slightly modified formulas are used when we can project S into the xz- and yz-planes as well.

Problem 5.7.2. *Evaluate $\iint_S y\,dS$, where S is the surface $z = x + y^2$, $0 \leq x \leq 1$, $0 \leq y \leq 2$.*

Solution: We first find the partial derivatives:

$$\frac{\partial f}{\partial x} = 1 \quad \text{and} \quad \frac{\partial z}{\partial y} = 2y$$

Now we can easily evaluate the surface integral by applying the formula:

$$\begin{aligned}
\iint_S y\,dS &= \iint_D y\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\,dA \\
&= \int_0^1\int_0^2 y\sqrt{1+1+4y^2}\,dy\,dx \\
&= \int_0^1 dx\,\sqrt{2}\int_0^2 y\sqrt{1+2y^2}\,dy \\
&= \sqrt{2}\left(\frac{1}{2}\right)\left(\frac{2}{3}\right)(1+2y^2)^{3/2}]_0^2 \\
&= \boxed{\frac{13\sqrt{2}}{3}}
\end{aligned}$$

If S is a piecewise-smooth surface, that is, a finite union of smooth surfaces $S_1, S_2, \ldots, S_n$ that only intersect at the boundaries (i.e. do not overlap), then the surface integral of f over S is defined by

$$\boxed{\iint_S f(x,y,z)\,dS = \iint_{S_1} f(x,y,z)\,dS + \cdots + \iint_{S_n} f(x,y,z)\,dS}$$

Problem 5.7.3. *Evaluate $\iint_S z\,dS$, where S is the surface whose sides S_1 are given by the cylinder $x^2+y^2=1$, whose bottom S_2 is the disk $x^2+y^2\le 1$ in the plane $z=0$, and whose top S_3 is the part of the plane $z=1+x$ lying above S_2.*

Solution: For S_1 we can parameterize with θ and z, writing the subsequent equations:

$$x=\cos\theta \quad y=\sin\theta \quad z=z$$

where

$$0\le\theta\le 2\pi \quad \text{and} \quad 0\le z\le 1+x=1+\cos\theta$$

Therefore

$$\mathbf{r}_\theta\times\mathbf{r}_z = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \cos\theta\,\mathbf{i}+\sin\theta\,\mathbf{j}$$

and it is trivial to show

$$|\mathbf{r}_\theta\times\mathbf{r}_z| = \sqrt{\cos^2\theta+\sin^2\theta}=1$$

The surface integral over S_1 is

$$\begin{aligned}
\iint_{S_1} z\,dS &= \iint_D z\,|\mathbf{r}_\theta\times\mathbf{r}_z|\,dA \\
&= \int_0^{2\pi}\int_0^{1+\cos\theta} z\,dz\,d\theta = \int_0^{2\pi}\frac{1}{2}(1+\cos\theta)^2\,d\theta \\
&= \frac{1}{2}\int_0^{2\pi}\left[1+2\cos\theta+\frac{1}{2}(1+\cos 2\theta)\right]d\theta \\
&= \frac{1}{2}\left[\frac{3}{2}\theta+2\sin\theta+\frac{1}{4}\sin 2\theta\right]_0^{2\pi} = \frac{3\pi}{2}
\end{aligned}$$

The surface integral over S_2 is even more trivial. S_2 lies in the xy-plane, so in other words, the z-coordinate is zero and therefore

$$\iint_{S_2} z\,dS = \iint_{S_2} 0\,dS = 0$$

The top-most surface S_3 is located above the unit disk D, $x^2+y^2 \le 1$, and is part of the plane $z = 1+x$. If we take $g(x,y) = 1+x$ and convert to polar coordinates to simplify the double integral, we have

$$\begin{aligned}
\iint_{S_1} z\,dS &= \iint_D (1+x)\sqrt{1+\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2}\,dA \\
&= \int_0^{2\pi}\int_0^1 (1+r\cos\theta)\sqrt{1+1+0}\,r\,dr\,d\theta \\
&= \sqrt{2}\int_0^{2\pi}\int_0^1 (r+r^2\cos\theta)\,dr\,d\theta = \sqrt{2}\int_0^{2\pi}\left(\frac{1}{2}+\frac{1}{3}\cos\theta\right)d\theta \\
&= \sqrt{2}\left[\frac{\theta}{2}+\frac{\sin\theta}{3}\right]_0^{2\pi} = \sqrt{2}\,\pi
\end{aligned}$$

Thus, our final answer is

$$\begin{aligned}
\iint_S z\,dS &= \iint_{S_1} z\,dS + \iint_{S_2} z\,dS + \iint_{S_3} z\,dS \\
&= \frac{3\pi}{2} + 0 + \sqrt{2}\,\pi = \boxed{\left(\frac{3}{2}+\sqrt{2}\right)\pi}
\end{aligned}$$

Oriented Surfaces Let's start with a surface S with a tangent plane at every point (x,y,z) except for boundary points. There are exactly two normal vectors: $\mathbf{n}_1$ and $\mathbf{n}_2 = -\mathbf{n}_1$, representing the positive and negative orientations on S, respectively.

Definition 5.7.1. *A surface S is called an **oriented surface** if it is possible to choose a unit normal vector $\mathbf{n}$ at every point (x,y,z) so that $\mathbf{n}$ varies continuously over the surface.*

The given choice of $\mathbf{n}$ provides S with a certain orientation. For a surface $z = g(x,y)$, we take the cross product of the tangent vectors $\mathbf{r}_x \times \mathbf{r}_y$ and divide by the magnitude, $|\mathbf{r}_x \times \mathbf{r}_y|$ to get the unit normal vector

$$\mathbf{n} = \frac{-\frac{\partial g}{\partial x}\,\mathbf{i} - \frac{\partial g}{\partial y}\,\mathbf{j} + \mathbf{k}}{\sqrt{1+\left(\frac{\partial g}{\partial x}\right)^2+\left(\frac{\partial g}{\partial y}\right)^2}}$$

Since the $\mathbf{k}$-component is positive, this gives the *upward* orientation on S. Also, if S is smooth and parameterized as the vector function $\mathbf{r}(u,v)$, then we automatically assign the unit normal vector as

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}$$

and the opposite direction is $-\mathbf{n}$. For example, we can parameterize a sphere with radius a using spherical coordinates:

$$\mathbf{r}(\phi,\theta) = a\sin\phi\cos\theta\,\mathbf{i} + a\sin\phi\sin\theta\,\mathbf{j} + a\cos\phi\,\mathbf{k}$$

We also know that

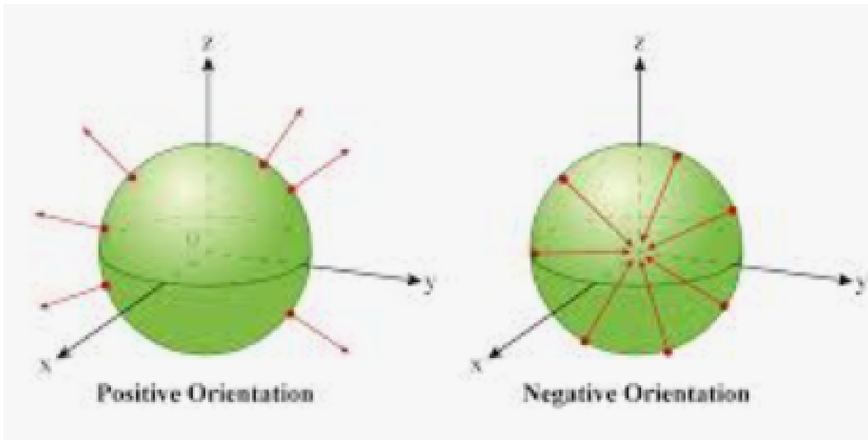
$$\mathbf{r}_\phi \times \mathbf{r}_\theta = a^2\sin^2\phi\cos\theta\,\mathbf{i} + a^2\sin^2\phi\sin\theta\,\mathbf{j} + a^2\sin\phi\cos\phi\,\mathbf{k}$$

and the magnitude of this cross product is $a^2 \sin\phi$. So, the orientation induced by the *position vector*, $\mathbf{r}(\phi, \theta)$ is defined by the unit normal

$$\mathbf{n} = \frac{\mathbf{r}_\phi \times \mathbf{r}_\theta}{|\mathbf{r}_\phi \times \mathbf{r}_\theta|} = \sin\phi\cos\theta\,\mathbf{i} + \sin\phi\sin\theta\,\mathbf{j} + \cos\phi\,\mathbf{k} = \frac{1}{a}\,\mathbf{r}(\phi, \theta)$$

We can infer from the above equation that $\mathbf{n}$ always points in the same direction as the position vector, $\mathbf{r}(\phi, \theta)$, that is, outward from the sphere. The opposite (inward) direction would be obtained by reversing the order of the parameters, because:

$$\mathbf{r}_\theta \times \mathbf{r}_\phi = -\mathbf{r}_\phi \times \mathbf{r}_\theta$$



Surface Integrals of Vector Fields and Flux Let S be an oriented surface with unit normal $\mathbf{n}$. Imagine a fluid with density $\rho(x, y, z)$ and velocity field $\mathbf{v}(x, y, z)$ flowing through the surface. The *rate of flow*, in mass per unit time, per unit area, is given by the vector field $\mathbf{F} = \rho\mathbf{v}$.

After dividing S into small patches S_{ij} (notice how we keep using this intuition?), these patches are nearly planar and so the mass of fluid per unit time crossing S_{ij} in the direction of $\mathbf{n}$ is approximately

$$(\rho\mathbf{v} \cdot \mathbf{n})\, A(S_{ij})$$

Why the dot product? Because this is the component of $\rho\mathbf{v}$ in the direction of $\mathbf{n}$. By summing all these quantities as the number of patches approaches infinity, the surface integral of the function $\rho\mathbf{v} \cdot \mathbf{n}$ over S is

$$\iint_S \rho\mathbf{v} \cdot \mathbf{n}\, dS = \iint_S \rho(x, y, z)\mathbf{v}(x, y, z) \cdot \mathbf{n}(x, y, z)\, dS$$

which is interpreted as the rate of flow through S. If we substitute $\mathbf{F} = \rho\mathbf{v}$, where $\mathbf{F}$ is a vector field on $\mathbb{R}^3$ and the integral given is

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, dS$$

Physicists often refer to this surface integral as the *flux integral* of $\mathbf{F}$ over S.

Definition 5.7.2. *If $\boldsymbol{F}$ is a continuous vector field defined on an oriented suface S with unit normal vector $\boldsymbol{n}$, then the* ***surface integral of $\boldsymbol{F}$ over*** *S is*

$$\iint_S \boldsymbol{F} \cdot d\boldsymbol{S} = \iint_S \boldsymbol{F} \cdot \boldsymbol{n}\, dS$$

The value of this integral is equal to the ***flux*** *of $\boldsymbol{F}$ across the surface.*

Since we often define surfaces as vector functions, $\mathbf{r}(u,v)$, then $\mathbf{n}$ is replaced by the cross product of the tangent vectors $\mathbf{r}_u$ and $\mathbf{r}_v$, and then normalized by dividing by the magnitude of the cross product, $|\mathbf{r}_u \times \mathbf{r}_v|$:

$$\begin{aligned}\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_S \mathbf{F}\cdot\mathbf{n}\,dS = \iint_S \mathbf{F}\cdot\frac{\mathbf{r}_u\times\mathbf{r}_v}{|\mathbf{r}_u\times\mathbf{r}_v|}\,dS \\ = \iint_D \left[\mathbf{F}(\mathbf{r}(u,v))\cdot\frac{\mathbf{r}_u\times\mathbf{r}_v}{|\mathbf{r}_u\times\mathbf{r}_v|}\right]|\mathbf{r}_u\times\mathbf{r}_v|\,dA\end{aligned}$$

where D is the parameter domain. Thus, we can verify

$$\boxed{\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_D \mathbf{F}\cdot(\mathbf{r}_u\times\mathbf{r}_v)\,dA}$$

Note: this formula assumes that S has positive orientation induced by the normal vector $\mathbf{r}_u \times \mathbf{r}_v$. If the orientation turns out to be negative, we multiply the result by -1.

Problem 5.7.4. *Calculate the flux of the vector field $\boldsymbol{F}(x,y,z) = z\,\boldsymbol{i} + y\,\boldsymbol{j} + x\,\boldsymbol{k}$ across the unit sphere $\rho = 1$.*

Solution: Parameterize the sphere, setting $r = 1$ and taking ϕ and θ as the parameters:

$$\mathbf{r}(\phi,\theta) = \sin\phi\cos\theta\,\mathbf{i} + \sin\phi\sin\theta\,\mathbf{j} + \cos\phi\,\mathbf{k} \quad 0\le\phi\le\pi \quad 0\le\theta\le 2\pi$$

We also transform the vector field from $\mathbf{F}(x,y,z) \to \mathbf{F}(\mathbf{r}(\phi,\theta))$:

$$\mathbf{F}(\mathbf{r}(\phi,\theta)) = \cos\phi\,\mathbf{i} + \sin\phi\sin\theta\,\mathbf{j} + \sin\phi\cos\theta\,\mathbf{k}$$

Then, we take the tangent vectors $\mathbf{r}_\phi$ and $\mathbf{r}_\theta$.

$$\begin{aligned}\mathbf{r}_\phi &= \cos\phi\cos\theta\,\mathbf{i} + \cos\phi\sin\theta\,\mathbf{j} - \sin\phi\,\mathbf{k} \\ \mathbf{r}_\theta &= -\sin\phi\sin\theta\,\mathbf{i} + \sin\phi\cos\theta\,\mathbf{j}\end{aligned}$$

The cross product of these two vectors is

$$\mathbf{r}_\phi\times\mathbf{r}_\theta = \sin^2\phi\cos\theta\,\mathbf{i} + \sin^2\phi\sin\theta\,\mathbf{j} + \sin\phi\cos\theta\,\mathbf{k}$$

(Can you verify this yourself?)

Our last step before evaluating the flux integral is to determine the dot product, given by $\mathbf{F}(\mathbf{r}(\phi,\theta))\cdot(\mathbf{r}_\phi\times\mathbf{r}_\theta)$.

$$\mathbf{F}(\mathbf{r}(\phi,\theta))\cdot(\mathbf{r}_\phi\times\mathbf{r}_\theta) = \cos\phi\sin^2\phi\cos\theta + \sin^3\phi\sin^2\theta + \sin^2\phi\cos\phi\cos\theta$$

The flux of $\mathbf{F}$ across the sphere is then

$$\begin{aligned}\iint_S \mathbf{F}\cdot d\mathbf{S} &= \iint_D \mathbf{F}\cdot(\mathbf{r}_\phi\times\mathbf{r}_\theta)\,dA \\ &= \int_0^{2\pi}\int_0^{\pi}(2\sin^2\phi\cos\phi\cos\theta + \sin^3\phi\sin^2\theta)\,d\phi\,d\theta \\ &= 2\int_0^{\pi}\sin^2\phi\cos\phi\,d\phi\int_0^{2\pi}\cos\theta\,d\theta + \int_0^{\pi}\sin^3\phi\,d\phi + \int_0^{2\pi}\sin^2\theta\,d\theta \\ &= 0 + \int_0^{\pi}\sin^3\phi\,d\phi\int_0^{2\pi}\sin^2\theta\,d\theta \\ &= \boxed{\frac{4\pi}{3}}\end{aligned}$$

Remark. We used the fact that $\int_0^{2\pi} \cos\theta \, d\theta = 0$ to simplify our calculations.

Consider the special case of a surface defined by the function $z = g(x, y)$. Taking x and y as parameters, we have

$$\mathbf{F} \cdot (\mathbf{r}_x \times \mathbf{r}_y) = (P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}) \cdot \left(-\frac{\partial g}{\partial x}\,\mathbf{i} - \frac{\partial g}{\partial y}\,\mathbf{j} + \mathbf{k}\right)$$

so the surface integral becomes

$$\boxed{\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_D \left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right) dA}$$

Remark #1. This formula assumes the upward orientation on S (because of the positive $\mathbf{k}$-component); for a downward orientation, we multiply by -1.

Remark #2. We can easily modify the above equation for surfaces defined by functions $y = h(x, z)$ or $x = k(y, z)$.

Problem 5.7.5. *Evaluate $\iint_S \boldsymbol{F} \cdot d\boldsymbol{S}$, where $\boldsymbol{F}(x, y, z) = y\,\boldsymbol{i} + x\,\boldsymbol{j} + z\,\boldsymbol{k}$ and S is the boundary of the solid region E enclosed by the paraboloid $z = 1 - x^2 - y^2$ and the plane $z = 0$.*

Solution: According to the definition of a paraboloid, S consists of a parabolic top surface S_1 and a circular bottom surface S_2. Also, S is closed, so we use the convention of positive (outward) orientation. In other words, S_1 is oriented upward and D is the projection of S in onto the xy-plane, namely, the disk $x^2 + y^2 \leq 1$. On S_1, we have

$$P(x, y, z) = y \quad Q(x, y, z) = x \quad R(x, y, z) = z = 1 - x^2 - y^2$$

and the partial derivatives are

$$\frac{\partial g}{\partial x} = -2x \qquad \frac{\partial g}{\partial y} = -2y$$

So we apply the formula for surface integrals over functions of the form $z = g(x, y)$:

$$\begin{aligned}
\iint_{S_1} \mathbf{F} \cdot d\mathbf{S} &= \iint_D \left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right) dA \\
&= \iint_D [-y(-2x) - x(-2y) + 1 - x^2 - y^2]\, dA \\
&= \iint_D (1 + 4xy - x^2 - y^2)\, dA \\
&= \int_0^{2\pi} \int_0^1 (1 + 4r^2\cos\theta\sin\theta - r^2)\, r\, dr\, d\theta \\
&= \int_0^{2\pi} \int_0^1 (r - r^3 + 4r^3\cos\theta\sin\theta)\, dr\, d\theta \\
&= \int_0^{2\pi} \left(\frac{1}{4} + \cos\theta\sin\theta\right) d\theta \\
&= \frac{1}{4}(2\pi) + 0 = \frac{\pi}{2}
\end{aligned}$$

The disk S_2 is oriented downward, hence its unit normal is $\mathbf{n} = -\mathbf{k}$ and thus

$$\iint_{S_2} \mathbf{F} \cdot d\mathbf{S} = \iint_S \mathbf{F} \cdot (-\mathbf{k})\, dS = \iint_D (-z)\, dA = \iint_D 0\, dA = 0$$

since S_2 lies in the plane $z = 0$. Finally, we compute $\iint_S \mathbf{F} \cdot d\mathbf{S}$ as the sum of the surface integrals of S over the pieces S_1 and S_2:

$$\begin{aligned}\iint_S \mathbf{F} \cdot d\mathbf{S} &= \iint_{S_1} \mathbf{F} \cdot d\mathbf{S} + \iint_{S_2} \mathbf{F} \cdot d\mathbf{S} \\ &= \frac{\pi}{2} + 0 = \boxed{\frac{\pi}{2}}\end{aligned}$$

Application: Heat Flow Suppose the temperature at a point (x, y, z) in an object is $u(x, y, z)$.

Definition 5.7.3. *The **heat flow** is defined as the vector field*

$$\boldsymbol{F} = -K\,\nabla u$$

*where K is the experimental value of the **conductivity** of the substance.*

The rate of heat flow across the surface S is given by the integral

$$\iint_S \mathbf{F} \cdot d\mathbf{S} = -K \iint_S \nabla u \cdot d\mathbf{S}$$

Problem 5.7.6. *The temperature u in a metal ball is proportional to the square of the distance from the center of the ball. Find the rate of heat flow across a sphere S of radius a with center at the center of the ball.*

Solution: Without loss of generality, let's place the center of the ball at the origin. Therefore

$$u(x, y, z) = C(x^2 + y^2 + z^2)$$

where C is a proportionality constant. The heat flow is represented by the vector field

$$\mathbf{F}(x, y, z) = -K\,\nabla u = -KC(2x\,\mathbf{i} + 2y\,\mathbf{j} + 2z\,\mathbf{k})$$

where K is the metal's conductivity. Instead of parameterizing the sphere, we can recognize that the outward unit normal to the sphere $x^2 + y^2 + z^2 = a^2$ at the point (x, y, z) is

$$\mathbf{n} = \frac{1}{a}(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k})$$

and so the dot product $\mathbf{F} \cdot \mathbf{n}$ is

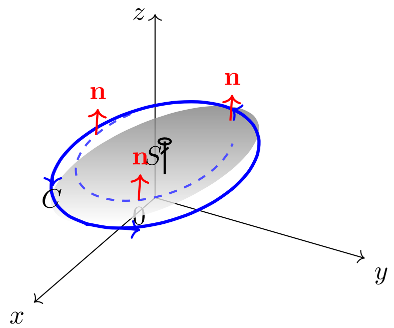
$$\mathbf{F} \cdot \mathbf{n} = -\frac{2KC}{a}(x^2 + y^2 + z^2)$$

Since S is a sphere, we can replace $x^2 + y^2 + z^2$ with a^2, so $\mathbf{F} \cdot \mathbf{n} = -2aKC$. The rate of heat flow across S is

$$\begin{aligned}\iint_S \mathbf{F} \cdot d\mathbf{S} &= \iint_S \mathbf{F} \cdot \mathbf{n}\,dS = -2aKC \iint_S dS \\ &= -2aKC \cdot A(S) = -2aKC \cdot 4\pi a^2 = \boxed{-8KC\pi a^3}\end{aligned}$$

5.8 Stokes' Theorem

In section 5.4, we discussed Green's Theorem, which relates a double integral over a plane region D to a line integral over the plane boundary curve C. In this section, we propose a concept that essentially functions as a higher-dimensional version of Green's Theorem. It relates a surface integral over a surface S to a line integral around the boundary curve of S (also called a space curve). The image below shows an oriented surface with a unit normal vector, $\mathbf{n}$. The surface's orientation induces a **positive orientation** of C. In other words, if you walk in the positive direction around the curve with your head pointing in the direction of the normal vector, the surface will *always* be to your left.



This idea is represented by Stokes' Theorem.

Definition 5.8.1. *Let S be an oriented piecewise-smooth surface that is bounded by a simple, closed, piecewise-smooth boundary curve C with positive orientation. Let $\boldsymbol{F}$ be a vector field whose components have continuous partial derivatives on an open region in $\mathbb{R}^3$. Then **Stokes' Theorem** states*

$$\int_C \boldsymbol{F}\cdot d\boldsymbol{r} = \iint_S \textit{curl } \boldsymbol{F}\cdot d\boldsymbol{S}$$

Since

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_C \mathbf{F}\cdot\mathbf{T}\,ds \quad \text{and} \quad \iint_S \text{curl } \mathbf{F}\cdot d\mathbf{S} = \iint_S \text{curl } \mathbf{F}\cdot\mathbf{n}\,dS$$

Stokes' Theorem essentially says that the line integral around C of the tangential component of $\mathbf{F}$ is equal to the surface integral of S of the normal component of the curl of $\mathbf{F}$.

Proof. A complete and rigorous proof of the general Stokes' Theorem is typically beyond the scope of a standard multivariable calculus course, but a proof for a special case where the surface is the graph of a function $z = f(x, y)$ over a region $D \subset \mathbb{R}^2$ can be demonstrated with the help of Green's Theorem.

For a vector field $\mathbf{F} = \langle P, Q, R\rangle$ with continuous partials, we wish to demonstrate

$$\oint_C \mathbf{F}\cdot d\mathbf{r} = \iint_S (\nabla\times\mathbf{F})\cdot d\mathbf{S}$$

Let's parameterize the surface S as

$$\mathbf{r}(x, y) = \langle x, y, f(x, y)\rangle$$

Then $\mathbf{r}_x \times \mathbf{r}_y = \langle -f_x, -f_y, 1\rangle$, $d\mathbf{S} = \langle -f_x, -f_y, 1\rangle\,dx\,dy$. (You can verify this result by computing the cross product of $\mathbf{r}_x$ with $\mathbf{r}_y$). The curl of $\mathbf{F}$ is equal to

$$\nabla\times\mathbf{F} = \langle R_y - Q_z, P_z - R_x, Q_x - P_y\rangle$$

so the dot product

$$(\nabla\times\mathbf{F})\cdot d\mathbf{S} = [Q_x - P_y + Q_z f_x - P_z f_y + R_x f_y - R_y f_x]\,dx\,dy$$

On the curve C, since $z = f(x, y)$, we have

$$d\mathbf{r} = \langle dx, dy, f_x\,dx + f_y\,dy\rangle$$

so ultimately

$$\mathbf{F} \cdot d\mathbf{r} = (P + Rf_x)\,dx + (Q + Rf_y)\,dy$$

Applying Green's Theorem yields

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_D \left(\frac{\partial}{\partial x}(Q + Rf_y) - \frac{\partial}{\partial y}(P + Rf_x) \right) dx\,dy$$

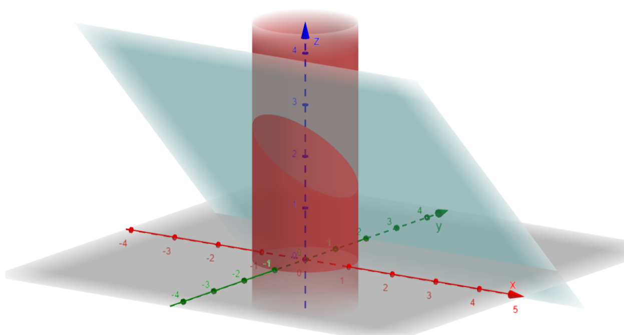
which simplifies to the same integrand above. Hence

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_S (\nabla \times \mathbf{F}) \cdot d\mathbf{S}$$

and the proof is complete. □

Problem 5.8.1. *Evaluate the line integral $\int_C \boldsymbol{F} \cdot d\boldsymbol{r}$, where $\boldsymbol{F}(x, y, z) = -y^2\,\boldsymbol{i} + x\,\boldsymbol{j} + z^2\,\boldsymbol{k}$ and C is the curve of intersection of the plane $y + z = 2$ and the cylinder $x^2 + y^2 = 1$. Assume that C is in the counterclockwise direction when viewed from above.*

Solution: The curve of intersection C is the slanted ellipse, as shown below.



We could parameterize the ellipse and evaluate $\int_C \mathbf{F} \cdot d\mathbf{r}$ directly, but it's easier to use Stokes' Theorem. The first step is to compute the curl.

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -y^2 & x & z^2 \end{vmatrix} = (1 + 2y)\,\mathbf{k}$$

Since we give S upward orientation, then C has positive orientation. The projection D of S in the xy-plane is the disk $x^2 + y^2 \leq 1$. Since the equation of the intersecting plane $y + z = 2$ in terms of $z = f(x, y)$ is $z = 2 - y$, we proceed with the following:

Let $\mathbf{r}(x, y)$ be the parametric description of S:

$$\mathbf{r}(x, y) = \langle x, y, f(x, y) \rangle = \langle x, y, 2 - y \rangle$$

so $\mathbf{r}_x = \langle 1, 0, 0 \rangle$ and $\mathbf{r}_y = \langle 0, 1, -1 \rangle$. The cross product is

$$\begin{aligned} \mathbf{r}_x \times \mathbf{r}_y &= \langle 1, 0, 0 \rangle \times \langle 0, 1, -1 \rangle \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{vmatrix} \\ &= 0\,\mathbf{i} + 1\,\mathbf{j} + 1\,\mathbf{k} \\ &= \langle 0, 1, 1 \rangle \end{aligned}$$

Now we will proceed with Stokes' Theorem:

$$\begin{aligned}\int_C \mathbf{F}\cdot d\mathbf{r} &= \iint_S (\nabla\times\mathbf{F})\cdot d\mathbf{S} = \iint_S (\nabla\times\mathbf{F})\cdot\mathbf{n}\,dS\\ &= \iint_S \langle 0,0,1+2y\rangle\cdot\langle 0,1,1\rangle\,dA = \iint_D (1+2y)\,dA\\ &= \int_0^{2\pi}\int_0^1 (1+2r\sin\theta)\,r\,dr\,d\theta\\ &= \int_0^{2\pi}\left[\frac{r^2}{2}+\frac{2r^3}{3}\sin\theta\right]_0^1 d\theta = \int_0^{2\pi}\left(\frac{1}{2}+\frac{2}{3}\sin\theta\right)d\theta\\ &= \frac{1}{2}(2\pi)+0 = \boxed{2\pi}\end{aligned}$$

Note: We converted the double integral into polar coordinates because D is a circular disk $x^2+y^2\le 1$. Also, $\mathbf{r}_x\times\mathbf{r}_y$ is the normal vector $\mathbf{n}$ perpendicular to the surface.

Remark. Stokes' Theorem allows us to compute a surface integral by simply knowing the values of the vector field $\mathbf{F}$ on the boundary curve C. In other words, if we had a second oriented surface with the same boundary curve C, then we get the same value for the surface integral:

$$\boxed{\iint_{S_1}\text{curl }\mathbf{F}\cdot d\mathbf{S} = \int_C \mathbf{F}\cdot d\mathbf{r} = \iint_{S_2}\text{curl }\mathbf{F}\cdot d\mathbf{S}}$$

Why is this important? Well, sometimes it may be difficult to integrate over one surface of a solid, but very easy to integrate over another. As long as the boundary curve for both surfaces remains the same, we can use this fact to our advantage.

Problem 5.8.2. *Use Stokes' Theorem to compute the surface integral $\iint_S(\nabla\times\boldsymbol{F})\cdot d\boldsymbol{S}$, where $\boldsymbol{F}(x,y,z) = xz\,\boldsymbol{i}+yz\,\boldsymbol{j}+xy\,\boldsymbol{k}$ and S is the portion of the sphere $x^2+y^2+z^2=4$ that lies inside the cylinder $x^2+y^2=1$ and above the xy-plane.*

Solution 1: To determine the boundary curve C, we solve the system of equations with $x^2+y^2+z^2=4$ and $x^2+y^2=1$. If we subtract the second equation from the first, we are left with $z^2=3$ which yields $z=\sqrt{3}$ or $z=-\sqrt{3}$. Because S lies above the xy-plane, $z>0 \therefore z=\sqrt{3}$. Thus, C is the circle given by the equations $x^2+y^2=1$, $z=\sqrt{3}$. If we parameterize C, we have

$$\mathbf{r}(t)=\cos t\,\mathbf{i}+\sin t\,\mathbf{j}+\sqrt{3}\,\mathbf{k}\quad 0\le t\le 2\pi$$

Differentiating this results in

$$\mathbf{r}'(t) = -\sin t\,\mathbf{i}+\cos t\,\mathbf{j}$$

Also, $\mathbf{F}=\mathbf{F}(\mathbf{r}(t))=\sqrt{3}\cos t\,\mathbf{i}+\sqrt{3}\sin t\,\mathbf{j}+\cos t\sin t\,\mathbf{k}$. Stokes' Theorem gives

$$\begin{aligned}\iint_S \text{curl }\mathbf{F}\cdot d\mathbf{S} = \int_C \mathbf{F}\cdot d\mathbf{r} &= \int_0^{2\pi}\mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt\\ &= \int_0^{2\pi}(-\sqrt{3}\cos t\sin t+\sqrt{3}\sin t\cos t)\,dt\\ &= \sqrt{3}\int_0^{2\pi}0\,dt = \boxed{0}\end{aligned}$$

Solution 2: We can use the fact that surface integrals over different surfaces with the same boundary curve C have the same value. Let S_1 be the disk inside the plane $z=\sqrt{3}$ inside the cylinder $x^2+y^2=1$. S_1 and S have the same boundary curve C (the disk $x^2+y^2=1$), so

$$\iint_S \text{curl }\mathbf{F}\cdot d\mathbf{S} = \iint_{S_1}\text{curl }\mathbf{F}\cdot d\mathbf{S}$$

Because S_1 is part of the horizontal plane, the upward normal vector is $\mathbf{k}$. First, we find curl $\mathbf{F}$.

$$\begin{aligned}\text{curl } \mathbf{F} &= \nabla \times \mathbf{F} \\ &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz & yz & xy \end{vmatrix} \\ &= (x-y)\,\mathbf{i} + (x-y)\,\mathbf{j}\end{aligned}$$

Therefore, the surface integral evaluates to

$$\begin{aligned}\iint_S \text{curl } \mathbf{F} \cdot d\mathbf{S} = \iint_{S_1} \text{curl } \mathbf{F} \cdot d\mathbf{S} &= \iint_{S_1} \text{curl } \mathbf{F} \cdot \mathbf{n}\, dS \\ &= \iint_{S_1} [(x-y)\,\mathbf{i} + (x-y)\,\mathbf{j}] \cdot \mathbf{k}\, dS \\ &= \iint_{S_1} 0\, dS = \boxed{0}\end{aligned}$$

5.9 The Divergence Theorem

In section 5.5, we saw that one of the vector versions for Green's Theorem is

$$\int_C \mathbf{F} \cdot \mathbf{n}\, ds = \iint_D \text{div } \mathbf{F}\, dA$$

where C is the positively-oriented boundary curve of the plane region D. If we wanted to extend this theorem to a vector field on $\mathbb{R}^3$, we might intuitively guess:

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, dS = \iiint_E \text{div } \mathbf{F}\, dV$$

where S is the boundary *surface* of the *solid* region E. It turns out this equation is actually true, and under appropriate hypotheses, it is called the Divergence Theorem.

Before we formally state the theorem and provide proof, it might be a good idea to review the various types of regions in which we performed triple integrals in section 4.6. These are the subsequent type 1, 2, and 3 regions that we refer to as **simple solid regions**, and we will state the Divergence Theorem in all cases. The boundary of E is a closed surface, with positive orientation; the outward normal vector $\mathbf{n}$ is directed outward from E.

Theorem 5.9.1. *Let E be a simple solid region and let S be the boundary surface of E, given positive (outward) orientation. Let $\boldsymbol{F}$ be a vector field whose component functions have continuous partial derivatives on an open region containing E. Then the* ***Divergence Theorem*** *states*

$$\iint_S \boldsymbol{F} \cdot d\boldsymbol{S} = \iiint_E \textit{div } \boldsymbol{F}\, dV$$

The Divergence Theorem essentially states that the flux of a vector field $\mathbf{F}$ across a boundary surface S is equal to the triple integral of the divergence of $\mathbf{F}$ over the solid region E.

Proof. Let $\mathbf{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ be a continuous vector field on $\mathbb{R}^3$. The divergence is

$$\text{div } \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

so we have

$$\iiint_E \text{div}\,\mathbf{F}\,dV = \iiint_E \frac{\partial P}{\partial x}\,dV + \iiint_E \frac{\partial Q}{\partial y}\,dV + \iiint_E \frac{\partial R}{\partial z}\,dV$$

If $\mathbf{n}$ is the outward unit normal of S, the surface integral on the left side of the theorem is

$$\begin{aligned}\iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_S \mathbf{F}\cdot\mathbf{n}\,dS &= \iint_S (P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k})\cdot\mathbf{n}\,dS \\ &= \iint_S P\,\mathbf{i}\cdot\mathbf{n}\,dS + \iint_S Q\,\mathbf{j}\cdot\mathbf{n}\,dS + \iint_S R\,\mathbf{k}\cdot\mathbf{n}\,dS\end{aligned}$$

In order to prove the Divergence Theorem, we need to prove the following three equalities:

$$\iint_S P\,\mathbf{i}\cdot\mathbf{n}\,dS = \iiint_E \frac{\partial P}{\partial x}\,dV$$

$$\iint_S Q\,\mathbf{j}\cdot\mathbf{n}\,dS = \iiint_E \frac{\partial Q}{\partial y}\,dV$$

$$\iint_S R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iiint_E \frac{\partial R}{\partial z}\,dV$$

Let's choose the last of these three equalities. In this case, E is a type 1 region, with

$$E = \{(x,y,z)\,|\,(x,y)\in D,\ u_1(x,y)\le z\le u_2(x,y)\}$$

where D is the projection of E onto the xy-plane. We can describe the triple integral as

$$\iiint_E \frac{\partial R}{\partial z}\,dV = \iint_D \left[\int_{u_1(x,y)}^{u_2(x,y)} \frac{\partial R}{\partial z}\,(x,y,z)\ dz\right]\,dA$$

and the Fundamental Theorem of Calculus gives

$$\iiint_E \frac{\partial R}{\partial z}\,dV = \iint_D \left[R(x,y,u_2(x,y)) - R(x,y,u_1(x,y))\right]\,dA \qquad (\star)$$

Let's now consider the structure of S, the boundary surface of E. It consists of at least two pieces: a bottom surface S_1 and a top surface S_2. It may contain a third vertical surface S_3, which lies above the boundary curve of D. Note that $\mathbf{k}$ and $\mathbf{n}$ are perpendicular, so the dot product $\mathbf{k}\cdot\mathbf{n}$ is 0 so

$$\iint_{S_3} R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iint_{S_3} 0\,dS = 0$$

So regardless of whether there is a third vertical surface, we surely know

$$\iint_S R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iint_{S_1} R\,\mathbf{k}\cdot\mathbf{n}\,dS + \iint_{S_2} R\,\mathbf{k}\cdot\mathbf{n}\,dS$$

The equation of the top surface S_2 is $z = u_2(x,y)$ and the outward unit normal $\mathbf{n}$ points upward, so we have

$$\iint_{S_2} R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iint_D R(x,y,u_2(x,y))\,dA$$

For the bottom surface S_1, its equation is the same, but the outward unit normal $\mathbf{n}$ points downward, so we multiply by -1 to get

$$\iint_{S_1} R\,\mathbf{k}\cdot\mathbf{n}\,dS = -\iint_D R(x,y,u_1(x,y))\,dA$$

So the value of the surface integral of $\mathbf{F} = R\,\mathbf{k}$ over S is

$$\iint_S R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iint_D \left[R(x,y,u_2(x,y)) - R(x,y,u_1(x,y))\right]\,dA$$

and comparison with the equation in $(\star)$ demonstrates

$$\iint_S R\,\mathbf{k}\cdot\mathbf{n}\,dS = \iiint_E \frac{\partial R}{\partial z}\,dV$$

Of the three equalities resulting from the theorem statement, the other two can be easily proved in a similar manner using a type 2 or type 3 region. Hence, the proof is complete. $\square$

Problem 5.9.1. *Calculate the flux of the vector field* $\boldsymbol{F} = z\,\boldsymbol{i} + y\,\boldsymbol{j} + x\,\boldsymbol{k}$ *across the sphere with equation* $x^2 + y^2 + z^2 = 9$.

Solution: Since the boundary surface S is a sphere of radius 3, it is a closed surface so we can apply the Divergence Theorem directly. The divergence of $\mathbf{F}$ is

$$\begin{aligned}\text{div }\mathbf{F} &= \nabla\cdot\mathbf{F}\\ &= \frac{\partial}{\partial x}(z) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(x)\\ &= 0 + 1 + 0\\ &= 1\end{aligned}$$

The flux of $\mathbf{F}$ across $S-$which is the surface of the region $E-$is then

$$\begin{aligned}\iint_S \mathbf{F}\cdot\mathbf{n}\,dS &= \iiint_E \text{div }\mathbf{F}\,dV\\ &= \iiint_E 1\,dV = V(E)\\ &= \frac{4}{3}\pi(3)^3 = \boxed{36\pi}\end{aligned}$$

Problem 5.9.2. *Evaluate* $\iint_S \boldsymbol{F}\cdot d\boldsymbol{S}$, *where* $\boldsymbol{F}(x,y,z) = xy\,\boldsymbol{i} + (3y^2 + e^{xz^2})\,\boldsymbol{j} + \sin(\pi xy)\,\boldsymbol{k}$, *and* S *is the surface of the region* E *bounded by the parabolic cylinder* $z = 1 - x^2$ *and the planes* $z = 0$, $y = 0$, *and* $y + z = 2$.

Solution: It is clearly too difficult to evaluate the surface integral directly. However, we quickly notice that S is a closed surface, and the divergence of $\mathbf{F}$ is

$$\begin{aligned}\text{div }\mathbf{F} &= \nabla\cdot\mathbf{F}\\ &= \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(3y^2 + e^{xz^2}) + \frac{\partial}{\partial z}(\sin(\pi xy))\\ &= y + 6y + 0\\ &= 7y\end{aligned}$$

which is much nicer than $\mathbf{F}$ itself. Therefore, we use the Divergence Theorem to transform the surface integral into a triple integral. The easiest approach is to express E as a type 3 region:

$$E = \{(x,y,z)\,|\,-1\le x\le 1,\ 0\le z\le 1,\ 0\le y\le 2-z\}$$

Then we have

$$\begin{aligned}\iint_S \mathbf{F}\cdot d\mathbf{S} &= \iiint_E \text{div } \mathbf{F}\, dV = \iiint_E 7y\, dV \\ &= 7\int_{-1}^{1}\int_0^{1-x^2}\int_0^{2-z} y\, dy\, dz\, dx \\ &= 7\int_{-1}^{1}\int_0^{1-x^2} \frac{(2-z)^2}{2}\, dz\, dx \\ &= \frac{7}{2}\int_{-1}^{1}\left[-\frac{(2-z)^3}{3}\right]_0^{1-x^2} dx \\ &= -\frac{7}{6}\int_{-1}^{1}[(x^2+1)^3-8]\, dx \\ &= -\frac{7}{3}\int_0^1 (x^6+3x^4+3x^2-7)\, dx \\ &= \boxed{\frac{184}{15}}\end{aligned}$$

Remark. In the second to last step of the solution, we performed the substitution $\int_{-a}^{a} f(x)\, dx = 2\int_0^a f(x)\, dx$, where $f(x)$ is an *even* function, i.e. $f(-x) = f(x)$, because the integrand, $(x^2+1)^3 - 8$ was even.

Solids Regions As Finite Unions of Simple Solid Regions Consider a region E lying between closed surfaces S_1 and S_2, with $S_1 \subset S_2$. Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be outward unit normals of S_1 and S_2, respectively. Then the boundary surface of E is $S = S_1 \cup S_2$ and the outward unit normal is given by $\mathbf{n} = -\mathbf{n}_1$ on S_1 and $\mathbf{n} = \mathbf{n}_2$ on S_2.

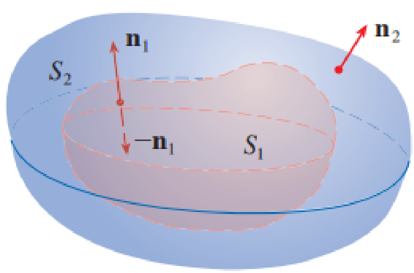


Image Credit: Calculus, Early Transcendentals

Applying the Divergence Theorem to S gives us

$$\begin{aligned}\iiint_E \text{div } \mathbf{F}\, dv &= \iint_S \mathbf{F}\cdot d\mathbf{S} = \iint_S \mathbf{F}\cdot\mathbf{n}\, dS \\ &= \iint_{S_1} \mathbf{F}\cdot(-\mathbf{n}_1)\, dS + \iint_{S_2} \mathbf{F}\cdot\mathbf{n}_2\, dS \\ &= -\iint_{S_1} \mathbf{F}\cdot d\mathbf{S} + \iint_{S_2} \mathbf{F}\cdot d\mathbf{S}\end{aligned}$$

Problem 5.9.3. *Recall from section* 5.1 *that the electric field is given by*

$$\boldsymbol{E}(\boldsymbol{x}) = \frac{\varepsilon Q}{|\boldsymbol{x}|^3}\,\boldsymbol{x}$$

where the electric charge Q is located at the origin and $\boldsymbol{x} = \langle x, y, z\rangle$ is a position vector in $\mathbb{R}^3$. Use the Divergence Theorem to prove that the electric flux of $\boldsymbol{E}$ through ANY closed surface S that encloses the origin is equal to

$$\iint_S \boldsymbol{E} \cdot d\boldsymbol{S} = 4\pi\varepsilon Q$$

Solution: This problem is particularly difficult because we do not have an explicit equation for the boundary surface S; it is *any* surface that encloses the origin. So let's use some wishful thinking and let S_1 be a sphere (chosen for its symmetry) centered at the origin with radius a, which lies between S_1 and a general boundary surface S. Applying the Divergence Theorem to this union of simple solid regions, we have

$$\iiint_E \text{div } \mathbf{E}\, dV = -\iint_{S_1} \mathbf{E} \cdot d\mathbf{S} + \iint_S \mathbf{E} \cdot d\mathbf{S}$$

We observe div $\mathbf{E} = \nabla \cdot \mathbf{E} = 0$ (you can verify this yourself!). Therefore

$$\iint_{S_1} \mathbf{E} \cdot d\mathbf{S} = \iint_S \mathbf{E} \cdot d\mathbf{S}$$

Why is this important? This calculation implies that we can compute the flux of $\mathbf{E}$ across S_1 due to its spherical structure. Additionally, the unit normal at a point (x, y, z) is given by $\mathbf{n} = \frac{\mathbf{x}}{|\mathbf{x}|}$. The dot product $\mathbf{E} \cdot \mathbf{n}$ is given by

$$\mathbf{E} \cdot \mathbf{n} = \frac{\varepsilon Q}{|\mathbf{x}|^3}\,\mathbf{x} \cdot \left(\frac{\mathbf{x}}{|\mathbf{x}|}\right) = \frac{\varepsilon Q}{|\mathbf{x}|^4}\,\mathbf{x} \cdot \mathbf{x} = \frac{\varepsilon Q}{|\mathbf{x}|^2} = \frac{\varepsilon Q}{a^2}$$

since the magnitude of the position vector is equal to the radius of S_1. Finally, the net electric flux of $\mathbf{E}$ across S is

$$\begin{aligned}\iint_S \mathbf{E} \cdot d\mathbf{S} &= \iint_{S_1} \mathbf{E} \cdot \mathbf{n}\, dS \\ &= \frac{\varepsilon Q}{a^2} \iint_{S_1} dS = \frac{\varepsilon Q}{a^2} \cdot A(S_1) \\ &= \frac{\varepsilon Q}{a^2} \cdot 4\pi a^2 = 4\pi\varepsilon Q\end{aligned}$$

and the proof is complete. □

Thank you so much for reading this book! I am honored to have contributed to your academic journey in some way!

Thanks,

Aditya Baisakh

CEO, TMAS Academy