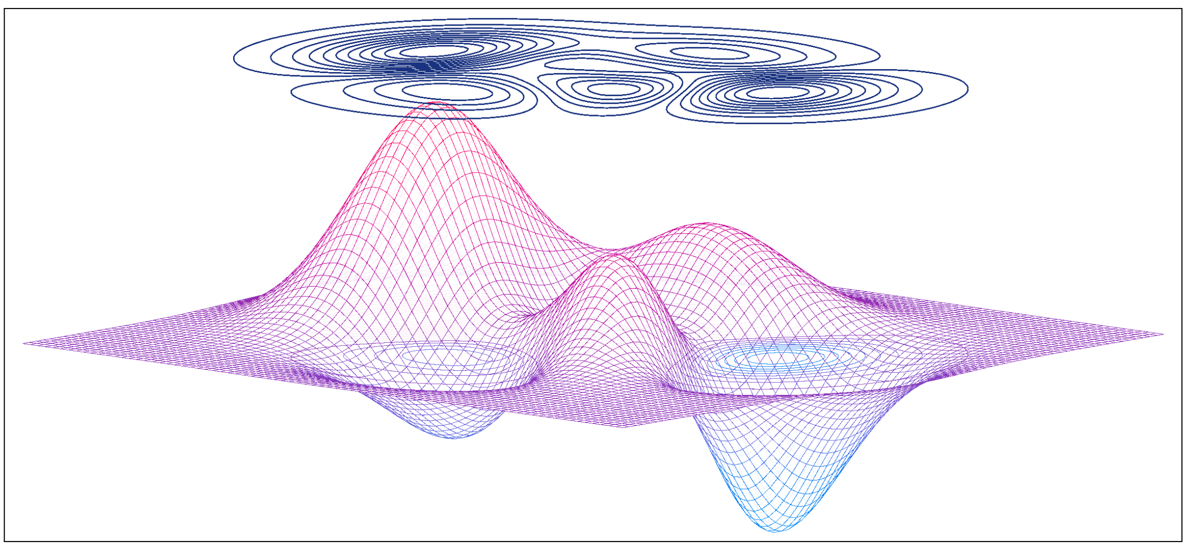
Multivariable Calculus

July 25, 2025



First Edition

Contents

Introduction

Welcome to the first edition of my guide on multivariable calculus, also known as Calculus III in the United States. We will cover vectors, multivariable differentiation, multiple integrals, vector fields, and the fundamental theorems of vector calculus. I don't intend for this to be as comprehensive as an entire semester or rigorous enough for a math major, but rather the most condensed version possible. Thus, I focus on clear concepts and getting you what you need to know for future learning. Material is divided up into three clear parts. You should work through every **EXAMPLE** on your own to test your knowledge.

There is indeed a lot of content, potentially quadruple what you learned in single-variable calculus (AP Calculus BC or Calculus I and Calculus II). That being said, you can do this. The material here is the backbone of your future math and science courses. Make sure to develop a foundation here so you can move on to more interesting things!

Recall that learning is a process, not a destination. It truly takes thousands of hours to fully learn many areas of mathematics. The best way to use this series is to combine it with other resources like textbooks and lectures. Read through a variety of explanations and keep revisiting problems. Remember to practice metacognition and create a routine that is conducive to metalearning. I highly recommend turning your work into a "guide" that you can look back on in the future.

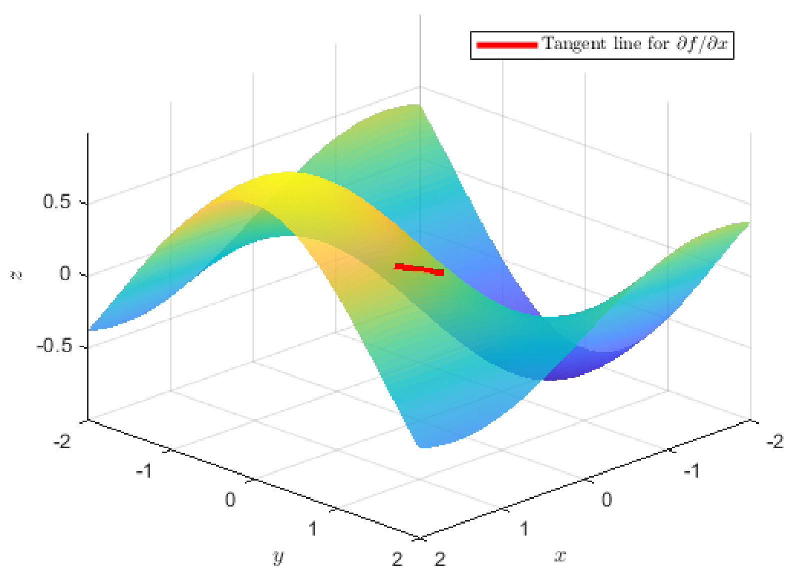
All sources used to create this guide are listed in **Bibliography**. I highly recommend checking out all of the resources there for more practice problems. Any images not attributed to someone else were created by and belong to me.

Part I

Vectors and Derivatives in $\mathbb{R}^3$ Space

In **Part I**, we will explore the concepts and tools that let us extend single-variable calculus into higher dimensions. We will cover vectors and how to operate on them, three-dimensional analytic geometry, calculus with vector-valued functions, multivariable limits, partial derivatives, and the classic applications of multivariable differentiation. By the end of this, you will have learned about

- Visualizing and computing vectors (i.e. dot products, cross products)
- Using vector-valued functions to describe curves and functions in space
- Geometry of lines, planes, and surfaces
- How concepts in single-variable calculus work with multiple variables
- Using partial and directional derivatives to analyze functions



Geometric interpretation of the partial derivative $\partial f/\partial x$ for the surface $f(x, y) = \sin(x)\cos(y)$ at the point $(1, 1)$

1 Review

1.1 Single-Variable Calculus

In single-variable calculus, you studied three main concepts: the limit, the derivative, and the integral.

The **limit** is perhaps the single most important definition in all of calculus, but it is also one of the most difficult to grasp. It describes how a function behaves as it gets infinitely close to a certain point.

> **Definition:** A function f approaches the limit L near a. If for every $\varepsilon > 0$ there exists a $\delta > 0$ where the following is true for all x
>
> $$0 < |x - a| < \delta,$$
>
> Then
>
> $$|f(x) - L| < \varepsilon.$$
>
> where x is the input variable, a is the point x approaches, ε represents how close we want $f(x)$ to be to L, and δ represents how close x must be to a to achieve that closeness.

In simpler terms, this means that no matter how close you want $f(x)$ to be to L, you can always find a range around a where this happens.

The general notation for a limit is

$$\lim_{x \to a} f(x) = L.$$

Recall that a function cannot approach two different limits near a. Near a, if f approaches L and f approaches another limit value m in the other direction, then $L = m$. That is, the right-hand and left-hand limits must agree:

$$\lim_{x \to a^-} f(x) = m = \lim_{x \to a^+} f(x) = L$$

Here are the most important properties of limits. These properties allow you to simplify expressions and compute limits quickly. In each case, we assume $\lim_{x\to a} f(x)$ and $\lim_{x\to a} g(x)$ exist. Let c be any constant.

Sum Rule:

$$\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$$

Difference Rule:

$$\lim_{x\to a}[f(x) - g(x)] = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$$

Constant Multiple Rule:

$$\lim_{x\to a}[c \cdot f(x)] = c \cdot \lim_{x\to a} f(x)$$

Product Rule:

$$\lim_{x\to a}[f(x) \cdot g(x)] = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$$

Quotient Rule:

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)} \quad \text{if } \lim_{x\to a} g(x) \neq 0$$

Power Rule:

$$\lim_{x\to a}[f(x)]^n = \left[\lim_{x\to a} f(x)\right]^n \quad \text{for integer } n$$

Root Rule:

$$\lim_{x\to a} \sqrt[n]{f(x)} = \sqrt[n]{\lim_{x\to a} f(x)} \quad \text{if the limit exists and the root is defined.}$$

The **derivative** is likely the first idea you came across in calculus that you found truly powerful.

For the points $(a, f(a))$ and $(a+h, f(a+h))$, the slope of the secant line between them is given by $\frac{f(a+h)-f(a)}{h}$ where $h \neq 0$. If you then think about the limit of the slope of a tangent line through $(a, f(a))$, you would have the expression

$$\lim_{h \to 0} \frac{f(a+h)-f(a)}{h}.$$

From this, we can then acquire the definition of the derivative:

Definition: A function f is differentiable at a if

$$\lim_{h \to 0} \frac{f(a+h)-f(a)}{h} \in \mathbb{R}.$$

This is of course known as the derivative of f at a and written as $f'(a)$.

The derivative measures the instantaneous rate of change of a function at a point. Geometrically, that would be the slope of the tangent line to the graph of f at that point. This makes derivatives an essential tool for analyzing functions.

- If $f'(a) > 0$, the tangent line has a positive slope, thus the function is increasing near a.
- If $f'(a) < 0$, the tangent line has a negative slope, thus the function is decreasing near a.
- If $f'(a) = 0$, the tangent line is flat, thus the function may have a local maximum, minimum, or a point of inflection at a.

The derivative also determines concavity through the second derivative. $f''(x) > 0$ means the graph is concave up and $f''(x) < 0$ means the graph is concave down.

In practice, there are rules we have for computing derivatives quickly:

- Power rule: $\frac{d}{dx}[x^n] = nx^{n-1}$
- Sum rule: $\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x)$
- Product rule: $\frac{d}{dx}[f(x)g(x)] = f'(x)g(x) + f(x)g'(x)$
- Quotient rule: $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$

The chain rule is one of the most important.

We start with two functions. The function $g(x)$ takes an input x and produces an output $g(x)$. The function $f(x)$ takes an input and produces an output $f(x)$. Suppose its input will be the result of $g(x)$.

When we form the corresponding composite function $f(g(x))$, we are first applying g to x, then applying f to the result. The behavior of $f(g(x))$ near $x = a$ depends on two things:

- How g behaves near a, since $g(x)$ determines the input to f.
- How f behaves near $g(a)$, since f takes $g(x)$ as its input.

It is therefore reasonable to require that f be differentiable at $g(a)$ in order for the derivative of $f(g(x))$ to exist at $x = a$. In other words, we need both functions to behave nicely at the right points: g must be differentiable at a to control how the input changes, and f must be differentiable at $g(a)$ to control how the output responds to those changes.

The derivative of $f(g(x))$ with respect to x is:

$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x)$$

L'Hôpital's Rule is used to evaluate limits that result in indeterminate forms such as $\frac{0}{0}$ or $\frac{\infty}{\infty}$. Suppose $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, or both limits are infinite. If f and g are differentiable near a (with $g'(x) \neq 0$ near a) and the limit $\lim_{x\to a} \frac{f'(x)}{g'(x)}$ exists (or is ∞ or $-\infty$). Then,

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}.$$

Derivatives tell us how functions change, whereas **integrals** tell us how much functions accumulate.

Suppose you divide the interval $[a, b]$ into four subintervals $[t_0, t_1], [t_1, t_2], [t_2, t_3], [t_3, t_4]$. where $a = t_0 < t_1 < t_2 < t_3 < t_4 = b$. On the ith interval $[t_{i-1}, t_i]$, the minimum value of f is m_i and the maximum value is M_i. Thus, the sum

$$s = m_1(t_1 - t_0) + m_2(t_2 - t_1) + m_3(t_3 - t_2) + m_4(t_4 - t_3)$$

represents the total area of rectangles inside the region $R(f, a, b)$. On the other hand, the sum

$$S = M_1(t_1 - t_0) + M_2(t_2 - t_1) + M_3(t_3 - t_2) + M_4(t_4 - t_3)$$

represents the total area of rectangles that make up the region $R(f, a, b)$. Based on this, it must be true for any division of subintervals that $s \leq A \leq S$.

A partition of the interval $[a, b]$ is a finite collection of points

$$P = \{t_0, t_1, \ldots, t_n\} \in [a, b]$$

where $a = t_0 < t_1 < \ldots < t_{n-1} < t_n = b$. For f bounded on $[a, b]$ with partition P,

The lower sum is defined as

$$L(f, P) = \sum_{i=1}^{n} m_i(t_i - t_{i-1}),$$

and the upper sum is defined as

$$U(f, P) = \sum_{i=1}^{n} M_i(t_i - t_{i-1}),$$

where m_i and M_i are the minimum and maximum values of f on the subinterval $[t_{i-1}, t_i]$, respectively.

These sums represent the total area of rectangles approximating the region under the curve. The lower sum uses the smallest value of the function on each subinterval, while the upper sum uses the largest value. As the partition becomes smaller, these sums get closer together, eventually leading to the exact integral.

Definition: If f is continuous on $[a, b]$, then f is integrable on $[a, b]$. In this case, the integral of f over $[a, b]$ is defined as:

$$\int_a^b f(x)\,dx$$

which represents the total area under the curve of f from a to b.

For any partition P, we always have:

$$L(f, P) \leq \int_a^b f(x)\,dx \leq U(f, P).$$

The **fundamental theorem of calculus** connects derivatives and integrals, showing that they are essentially inverse operations. It has two parts. Here is the first:

If f is continuous on $[a, b]$ and F is any antiderivative of f, meaning $F'(x) = f(x)$, then

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

This tells us that we can evaluate a definite integral by finding an antiderivative. In other words, you can add up all of the tiny changes of a quantity to get its total change. The definite integral represents the net change of the antiderivative.

Here is the second:

If f is continuous on $[a, b]$, then the function

$$F(x) = \int_a^x f(t)\,dt$$

is differentiable on $[a, b]$, and

$$F'(x) = f(x).$$

In other words, taking the derivative of an integral simply returns the original function.

To evaluate integrals, you used the following techniques:

Common Antiderivatives:

These are the most frequently used integrals:

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C \quad (\text{for } n \neq -1)$$
$$\int e^x \, dx = e^x + C$$
$$\frac{1}{x} \, dx = \ln|x| + C$$

u-Substitution:

This is used to simplify integrals by changing variables. It is the reverse of the chain rule:

$$\int f(g(x)) \, g'(x) \, dx = \int f(u) \, du$$

Integration by Parts:

This technique is based on the product rule for derivatives:

$$\int u \, dv = uv - \int v \, du$$

1.2 Parametric Equations and Polar Coordinates

Instead of describing curves with a single equation like $y = f(x)$, we can represent them using a parameter t:

$$x = f(t), \quad y = g(t)$$

These are called **parametric equations**. The parameter t often represents time or another quantity that controls the motion along the curve, which we call a parametric curve.

To find the slope of the curve at a point, we use:

$$\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}$$

This formula gives the rate of change of y with respect to x in terms of derivatives with respect to the parameter t.

If a curve C is described by the parametric equations

$$x = f(t), \quad y = g(t)$$

for $\alpha \leq t \leq \beta$, and if $f'(t)$ and $g'(t)$ are continuous on $[\alpha, \beta]$, then the length of the curve is given by:

$$L = \int_{\alpha}^{\beta} \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

or equivalently,

$$L = \int_{\alpha}^{\beta} \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

For the parametric curve where

$$x = f(t), \quad y = g(t), \quad \alpha \leq t \leq \beta,$$

and $f'(t)$ and $g'(t)$ are continuous, the differential arc length element is

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

Surface area is given by

$$S_y = \int_\alpha^\beta 2\pi y \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

for revolving around the x-axis or

$$S_x = \int_\alpha^\beta 2\pi x \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt$$

for revolving around the y-axis.

With this, we can show that the surface area of a sphere of radius r is $4\pi r^2$.

The sphere is obtained by rotating the semicircle

$$x = r\cos t, \quad y = r\sin t, \quad 0 \le t \le \pi$$

about the x-axis.

First, we compute the derivatives:

$$\frac{dx}{dt} = -r\sin t, \quad \frac{dy}{dt} = r\cos t$$

Using the formula for surface area,

$$S = \int_0^\pi 2\pi r\sin t\sqrt{(-r\sin t)^2 + (r\cos t)^2}\,dt = 2\pi\int_0^\pi r\sin t\sqrt{r^2(\sin^2 t + \cos^2 t)}\,dt$$

Simplifying yields

$$\int_0^\pi 2\pi r\sin t \cdot r\,dt.$$

Now we can evaluate:

$$2\pi r^2\int_0^\pi \sin t\,dt = 2\pi r^2\left[-\cos t\right]_0^\pi = 2\pi r^2(-\cos\pi + \cos 0) = 2\pi r^2(1+1) = 4\pi r^2$$

This confirms that the surface area of a sphere of radius r is $4\pi r^2$.

In addition to describing curves with Cartesian coordinates (x, y), we can also use **polar coordinates**, which are based on the distance from the origin and the angle from the positive x-axis.

In polar coordinates, each point is described by:

$$x = r\cos\theta, \quad y = r\sin\theta$$

where r is the distance from the origin and θ is the angle measured from the positive x-axis.

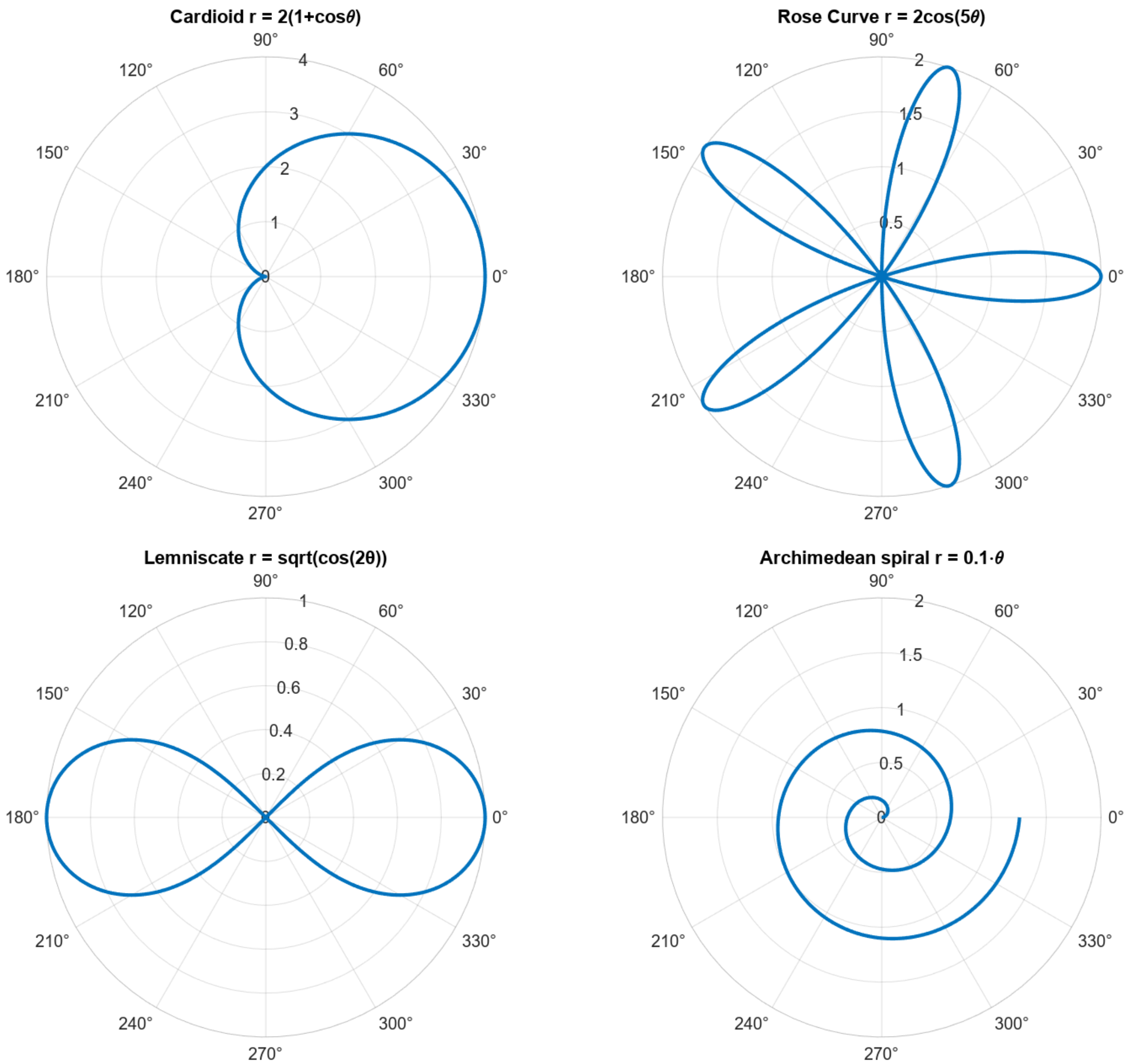
Polar coordinates are especially useful for describing curves with circular symmetry such as circles, spirals, and rose curves. To find a tangent line to a polar curve $r = f(\theta)$, θ is really a parameter with parametric equations

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

Then, the slope of the tangent line) is given by:

$$\frac{dy}{dx} = \frac{\frac{dy}{d\theta}}{\frac{dx}{d\theta}} = \frac{\frac{dr}{d\theta}\sin\theta + r\cos\theta}{\frac{dr}{d\theta}\cos\theta - r\sin\theta}.$$

Here is a gallery of some common polar curves:



In addition to graphing curves in polar coordinates, you can also compute areas and arc lengths.

Area Enclosed by a Polar Curve:

The area enclosed by the polar curve $r = r(\theta)$ from $\theta = \alpha$ to $\theta = \beta$ is given by

$$A = \frac{1}{2}\int_{\alpha}^{\beta} [r(\theta)]^2 \, d\theta.$$

Arc Length of a Polar Curve:

The length of a polar curve $r = r(\theta)$ from $\theta = \alpha$ to $\theta = \beta$ is

$$L = \int_{\alpha}^{\beta} \sqrt{[r(\theta)]^2 + \left[\frac{dr}{d\theta}\right]^2} \, d\theta.$$

1.3 Infinite Sequences and Series

A sequence is an ordered list of numbers:

$$a_1, a_2, a_3, a_4, ..., a_n...$$

Alternatively,

$$\{a_n\}_{n=1}^{\infty}$$

A sequence can also be expressed as a recurrence relation of the form $a_{n+1} = f(a_n)$ for $n \in \mathbb{N}$, where a_1 must be given or as an explicit formula of the form $a_n = f(n)$ for $n \in \mathbb{N}$.

Sequences are often analyzed by studying their limits.

Limit of a Sequence:

If $f(x)$ is a function such that $f(n) = a_n$ for all $n \in \mathbb{N}$, then the limit of the sequence $\{a_n\}_{n=1}^{\infty}$ is given by:

$$\lim_{n\to\infty} a_n = \lim_{x\to\infty} f(x).$$

That is, if a_n can be made arbitrarily close to L by taking n sufficiently large, then we say the limit of the sequence $\{a_n\}_{n=1}^{\infty}$ is L.

If $\lim_{n\to\infty} a_n$ exists, then we say that $\{a_n\}_{n=1}^{\infty}$ **converges** to L. Otherwise, we say the sequences diverges.

These are the terms used to describe the long-term behavior of a sequence. Let $n \in \mathbb{N}$. A sequence $\{a_n\}_{n=1}^{\infty}$ is:

- **Increasing** if $a_{n+1} > a_n$ for all n.
- **Nondecreasing** if $a_{n+1} \geq a_n$ for all n.
- **Decreasing** if $a_{n+1} < a_n$ for all n.
- **Nonincreasing** if $a_{n+1} \leq a_n$ for all n.
- **Monotonic** if it is either nondecreasing or nonincreasing.
- **Bounded** if there exists $M \in \mathbb{R}$ such that $|a_n| \leq M$ for all n.

A **sequence** $\{a_n\}_{n=1}^{\infty}$ is called a **geometric sequence** if each term is obtained by multiplying the previous term by a fixed constant r. That is,

$$a_{n+1} = ra_n.$$

All geometric sequences can be written as:

$$a_n = a \cdot r^{n-1}$$

for some constant $a \in \mathbb{R}$.

Limit of a Geometric Sequence:

For $r \in \mathbb{R}$, we have:

$$\lim_{n\to\infty} r^n = \begin{cases} 0, & \text{if } |r| < 1, \\ 1, & \text{if } r = 1, \\ \text{does not exist}, & \text{if } |r| \geq 1. \end{cases}$$

This is of course used to determine whether the associated geometric series converges or diverges. When $|r| < 1$, the terms will shrink rapidly and thus converge to a finite value. When $r = 1$, the terms are constant and grow without bounds, so the series diverges.

Infinite Series:

Given a sequence $\{a_n\}_{n=1}^{\infty}$, we define the **sequence of partial sums** $\{S_n\}_{n=1}^{\infty}$ by:

$$S_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^{n} a_k.$$

The infinite series $\sum_{k=1}^{\infty} a_k$ is defined as

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} S_n.$$

If this limit exists, we say the series **converges**. If the limit does not exist, we say the series **diverges**.

A **geometric series** has the general form

$$a + ar + ar^2 + ar^3 + \cdots = \sum_{n=1}^{\infty} ar^{n-1}.$$

The partial sum of the first n terms is denoted S_n:

$$S_n = a + ar + ar^2 + \cdots + ar^{n-1} = \sum_{k=1}^{n} ar^{k-1}.$$

By factoring and solving, we get the formula for the partial sum:

$$S_n = a \cdot \frac{1 - r^n}{1 - r}, \quad \text{where } r \neq 1$$

Geometric Series Test: The infinite geometric series $\sum_{k=1}^{n} ar^{k-1}$ converges if $|r| < 1$, in which case the sum is:

$$\lim_{n \to \infty} S_n = \frac{a}{1 - r}.$$

If $|r| \geq 1$, the series diverges.

If the infinite series $\sum a_n$ converges, then it must satisfy:

$$\lim_{n \to \infty} a_n = 0.$$

You can sometimes save yourself time by proving divergence:

Divergence Test: If $\lim_{n \to \infty} a_n \neq 0$, then the series $\sum a_n$ diverges. This test **only** checks for divergence and says nothing about convergence.

The famous **harmonic series** is given by:

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

Despite the fact that terms of the series approach zero, the harmonic series diverges.

A **telescoping series** is an infinite series where consecutive terms partially cancel when computing partial sums, leaving only some of the initial and final terms. To apply this technique, we express each term as a difference:

$$b_n = a_n - a_{n+1}$$

Then,

$$\sum_{n=1}^{\infty} b_n = \lim_{N\to\infty} \sum_{n=1}^{N} (a_n - a_{n+1}) = \lim_{N\to\infty} (a_1 - a_N) = a_1 - \lim_{N\to\infty} a_N.$$

The series converges if $\lim_{N\to\infty} a_N$ exists.

A general tool for testing series is the integral test:

Integral Test: If a function $f(x)$ is continuous, positive, and decreasing on $[1, \infty)$ and if $a_n = f(n)$, the series $\sum_{n=1}^{\infty} a_n$ is convergent if and only if the improper integral $\int_1^{\infty} f(x)\ dx$ is convergent. Otherwise, it is divergent.

One of the most widely used convergence tests for series in a certain form is as follows.

p-**Series Test:** A series in the form

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges if $p > 1$ and diverges if $p \leq 1$.

Many series, such as the harmonic series ($p = 1$), can be tested using this. If you cannot get a series written in this specific form, you will have to use another test.

Ratio Test:

Let $\sum_{n=1}^{\infty} a_n$ be an infinite series with positive terms $a_n > 0$ for all $n \in \mathbb{N}$.

This limit compares the size of successive terms in the series:

$$r = \lim_{k \to \infty} \frac{a_{k+1}}{a_k}$$

1. If $0 \leq r < 1$, the series converges.
2. If $r > 1$ (including $r = \infty$), the series diverges.
3. If $r = 1$, the ratio test is inconclusive.

The ratio test is especially useful for series involving factorials or exponentials.

Root Test: For a series $\sum_{n=1}^{\infty} a_n$ with nonnegative terms, compute the limit

$$L = \lim_{n \to \infty} \sqrt[n]{a_n}.$$

- If $0 \leq L < 1$, the series converges absolutely.
- If $L > 1$ or $L = \infty$, the series diverges.
- If $L = 1$, the test is inconclusive.

The root test is especially useful for series with terms raised to powers or involving exponentials.

Some series are difficult to compare directly. When this happens, we can compare them to simpler series.

Direct Comparison Test: Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be infinite series with positive terms. If the terms satisfy

$$0 < a_k \leq b_k \text{ and } 0 < b_k \leq a_k \text{ for all large enough } k \in \mathbb{N},$$

then

- If $\sum_{n=1}^{\infty} b_n$ converges, the series also converges.
- If $\sum_{n=1}^{\infty} b_n$ diverges, the series also diverges.

There is no clear set of guidelines for when to use the direct comparison test. If you can satisfy the inequalities quickly and get a fast answer, then this test is a solid choice. If you cannot, you should use the following test:

Limit Comparison Test: Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be infinite series with positive terms. Compute

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L.$$

Then

- If $0 < L < \infty$, then either both series either converge or diverge.
- If $L = 0$ and $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges.
- If $L = \infty$ and $\sum_{n=1}^{\infty} b_n$ diverges, then $\sum_{n=1}^{\infty} a_n$ also diverges.

Before we move on, let's summarize all the tests we've covered:

When testing whether a series with positive terms converges, here is a reasonable strategy to follow:

1. **Start with the Divergence Test:** If $\lim_{n\to\infty} a_n \neq 0$, the series diverges immediately.
2. **Check for Special Series:** Determine if your series matches or can be rewritten as one of these cases:
   - Geometric series
   - p-series
   - Telescoping series
   - Harmonic series
3. **Consider the Integral Test:** If the series terms look like an integrable function, the integral test might apply.
4. **Use the Ratio or Root Test for Factorials and Exponentials:** If the terms involve $n!$, n^n, or a^n, try the ratio test or possibly the root test.
5. **Use Comparison Tests for Rational Terms:** If the terms are rational functions of k or involve roots of rational functions, try the direct comparison test or the limit comparison test using known classic series from Step 2.

Until now, we've focused on series with positive terms. However, many important series have terms that aren't positive. These require different tools to analyze.

An alternating series has terms that switch signs, typically written as

$$a_n = \sum_{n=1}^{\infty} (-1)^{n-1} b_n \quad \text{or} \quad a_n = \sum_{n=1}^{\infty} (-1)^n b_n,$$

where $a_n > 0$. The factor $(-1)^n$ or $(-1)^{n-1}$ causes the terms to alternate between positive and negative.

Here is an example:

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

Here is another example:

$$\sum_{n=1}^{\infty} (-1)^n \frac{n}{n+1} = -\frac{1}{2} + \frac{2}{3} - \frac{3}{4} + \frac{4}{5} - \frac{5}{6} + \cdots$$

In general, the n-th term of an alternating series takes one of these forms:

- $a_n = (-1)^{n-1} b_n$
- $a_n = (-1)^n b_n$

where $b_n > 0$ and $b_n = |a_n|$. If the terms of an alternate series decrease towards 0 in magnitude, then the series converges:

Alternating Series Test: For the alternating series $\sum_{n=1}^{\infty} (-1)^{n-1} b_n$, evaluate the following:

1. $b_{n+1} \leq b_n$ for all n
2. $\lim_{n\to\infty} b_n = 0$,

If both are true, the series converges.

The alternating harmonic series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$$

converges even though the harmonic series itself diverges. This is because the terms can be rearranged in a way that make the alternating signs cancel out.

Let

$$\sum_{n=1}^{\infty} a_n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots$$

be the alternating harmonic series.

It is known that

$$\sum_{n=1}^{\infty} a_n = \ln 2.$$

Multiplying the series by $\frac{1}{2}$, we get

$$\sum_{n=1}^{\infty} \frac{1}{2} a_n = \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\right) = \frac{\ln 2}{2}.$$

Now, we define a new series $\sum_{n=1}^{\infty} b_n$ such that $b_{2n-1} = 0$ and $b_{2n} = \frac{a_n}{2}$ for all $n \geq 1$. It also converges:

$$\sum_{n=1}^{\infty} b_n = 0 + \frac{1}{2} \cdot \frac{1}{2} + 0 + \frac{1}{2} \cdot \frac{1}{4} + 0 + \frac{1}{2} \cdot \frac{1}{6} + \cdots = \frac{\ln 2}{2}$$

By the properties of convergent series:

$$\sum_{n=1}^{\infty} (a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n = \ln 2 + \frac{\ln 2}{2} = \frac{3 \ln 2}{2}.$$

Now, explicitly writing out the terms of $a_n + b_n$, we get

$$(1+0) + \left(-\frac{1}{2} + \frac{1}{2}\right) + \left(\frac{1}{3} + 0\right) + \left(-\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{5} + 0\right) + \left(-\frac{1}{6} + \frac{1}{6}\right) \cdots$$

This simplifies to

$$1+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\cdots$$

This is a rearrangement of the alternating harmonic series. Therefore,

$$1+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\cdots=\frac{3\ln 2}{2}.$$

Absolute and Conditional Convergence

- If $\sum_{n=1}^{\infty}|a_n|$ converges, we say that $\sum_{n=1}^{\infty}a_n$ **converges absolutely.**
- If $\sum_{n=1}^{\infty}|a_n|$ diverges and $\sum_{n=1}^{\infty}a_n$ converges, we say $\sum_{n=1}^{\infty}a_n$ **converges conditionally.** Otherwise, it diverges.

Power series, Taylor series, and Maclaurin series all follow the same basic structure; they are infinite sums involving powers of $(x-a)$.

- **Power Series:**

$$\sum_{n=0}^{\infty}c_n(x-a)^n$$

  The coefficients c_n can be any constants. The series is centered at a constant real number a, and its convergence depends on x.

- **Taylor Series:**

$$\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n$$

  This is a special type of power series where the coefficient terms are determined by $c_n=\frac{f^{(n)}(a)}{n!}$. The n-th order Taylor polynomial of f with its center at a denoted by $p_n(x)$ has the property that its value, slope, and all derivatives up to order n match those of f at $x=a$. For a Taylor series to be useful, you need to know the values of x for which the series converges and the values of x for which the output of the series representation equals f.

- **Maclaurin Series:**

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

  This is simply a Taylor series centered at $a = 0$.

To summarize, any Taylor series or Maclaurin series is a power series, but not every power series can be represented as a Taylor series.

Another special case of a Taylor series is the **binomial series**.

For $p \in \mathbb{R}$ and $k \in \mathbb{N}$, the **binomial coefficients** are defined by:

$$\binom{p}{k} = \frac{p(p-1)(p-2)\cdots(p-k+1)}{k!}, \quad \binom{p}{0} = 1.$$

The binomial series for $f(x) = (1+x)^p$ is given by:

$$\sum_{k=0}^{\infty} \binom{p}{k} x^k,$$

which explicitly expands to:

$$1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots.$$

This series converges for $|x| < 1$. In some cases, it may also converge at the endpoints depending on p. If $p \in \mathbb{Z}$, the series terminates after a finite number of terms and becomes a polynomial:

$$(x+a)^p = \sum_{k=0}^{p} \binom{p}{k} x^{p-k} a^k$$

Now, let's get into convergence:

Convergence of Power Series: A power series centered at a,

$$\sum_{n=0}^{\infty} c_n (x-a)^n,$$

converges in exactly one of three ways:

1. **Infinite Radius of Convergence:** The series converges for all $x \in \mathbb{R}$. In this case, the radius of convergence is $R = \infty$.
2. **Finite, Positive Radius of Convergence:** There exists a real number $R > 0$ such that the series converges for all $|x-a| < R$ and diverges for all $|x-a| > R$. The radius of convergence is $R \in \mathbb{N}$.
3. **Zero Radius of Convergence:** The series converges only at $x = a$, where the radius of convergence is $R = 0$.

Convergence of Taylor Series: For a function $f(x)$ with continuous derivatives of all orders on an interval I that contains constant a , the Taylor series centered at a is:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

The series converges if and only if

$$\lim_{n \to \infty} R_n(x) = 0.$$

Remember that a Maclaurin series is just a special Taylor series, so this convergence test applies for a Maclaurin series too.

These series are particularly useful in numerical methods. If you are interested, please check out my guide: *MATLAB Applications Part 1: Numerical Methods.*

2 Vectors and the Geometry of Space

Before you studied single-variable calculus, you had to first learn learn about numbers, arithmetic, symbolic manipulation of functions, and how functions behave given certain inputs. In multivariable calculus, we have functions that exist as surfaces in space. We will have to thus do that foundational work all over again.

2.1 Vectors in 2D Space

We begin by studying points in $\mathbb{R}^2$:

$$\mathbb{R}^2 = \{(x, y) \,:\, x, y \in \mathbb{R}\}$$

This represents the set of all ordered pairs (x, y) such that x and y are real numbers.

Points in $\mathbb{R}^2$ are defined as ordered pairs of real numbers:

$$P(x, y)$$

We graph point P on a 2D graph by moving x-units along the x-axis) and y-units along the y-axis).

A **vector** represents both *magnitude* and *direction*, written as:

$$\vec{x} = \langle x_2 - x_1,\, y_2 - y_1 \rangle$$

Vectors are denoted with an arrow above the symbol $\vec{x}$ or are lowercase and bolded $\mathbf{x}$.

If no starting point is specified, we assume the vector starts at the origin $(0, 0)$. The starting point of a vector is known as its *tail*.

For a vector in $\mathbb{R}^2$

$$\vec{x} = \langle x, y \rangle,$$

x is the first component and y is the second component.

EXAMPLE 2.1

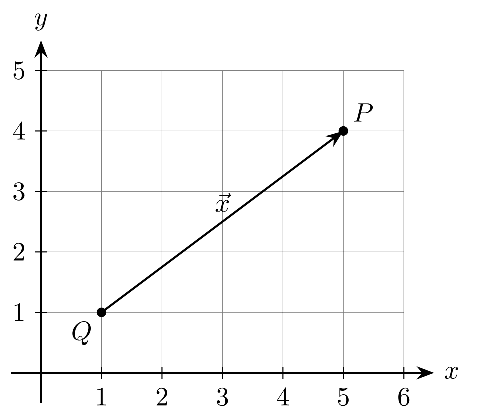
Draw vector $\vec{x} = \overrightarrow{QP}$ with $Q(1,1)$ and $P(5,4)$ and then interpret the result.

Solution:

We compute

$$\vec{x} = \langle 5-1,\, 4-1\rangle = \langle 4,3\rangle.$$

And then we graph:



The vector $\vec{x} = \langle 4,3\rangle$ represents the displacement from point Q to point P. It means:

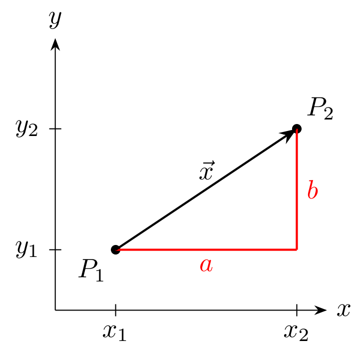
- Move 4 units in the x-direction (to the right).
- Move 3 units in the y-direction (upward).

Thus, $\vec{x}$ describes the exact movement required to go from Q to P in the plane.

There are infinitely many vectors that have the same components $\langle 4{,}3\rangle$, but they differ by their starting point. Our tail is Q and our head is P.

Let $\vec{x} = \overrightarrow{P_1P_2}$ where $P_1 = (x_1, y_1) \quad$ and $\quad P_2 = (x_2, y_2)$. This is computed as $\vec{x} = \langle x_2 - x_1,\, y_2 - y_1 \rangle$.

The graph looks as follows:



Thus, $\vec{x} = \langle a, b \rangle$ where $a = x_2 - x_1$ and $b = y_2 - y_1$. We go a units over and b units up. This should remind you of slope. We can then use the Pythagorean theorem to find the magnitude of $\vec{x}$:

By the Pythagorean theorem, we have:

$$c^2 = a^2 + b^2$$
$$c^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2$$
$$c = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

This gives the distance between points P_1 and P_2.

Next, we can say

$$\|\vec{x}\|_2 = \sqrt{a^2 + b^2}$$

where $\|\vec{x}\|_2$ denotes the **2-norm** (also called the *magnitude* or *length*) of vector $\vec{x}$.

EXAMPLE 2.2

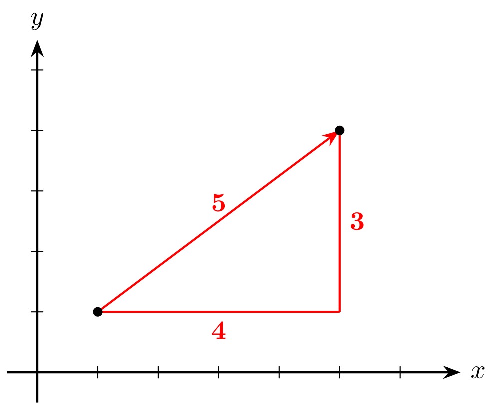
Find and graph $\|\vec{x}\|_2$ where $\vec{x} = \overrightarrow{QP}$ with $Q(1,1)$, $P(5,4)$.

Solution:

$$\begin{aligned} \vec{x} &= \langle 5-1,\, 4-1 \rangle \\ &= \langle 4, 3 \rangle \end{aligned}$$

$$\begin{aligned} \|\vec{x}\|_2^2 &= 4^2 + 3^2 \\ \|\vec{x}\|_2^2 &= \sqrt{25} = 5 \end{aligned}$$

And the graph looks like this:



The magnitude of the vector is 5 in the direction $\langle 4, 3 \rangle$.

Recall that the 2-norm operation maps vectors in $\mathbb{R}^2$ to real numbers. That is,

$$\| \cdot \|_2 : \mathbb{R}^2 \to \mathbb{R}.$$

This means it takes a vector in $\mathbb{R}^2$ and outputs a real number representing its magnitude.

We now define a new operation called **vector addition**:

$$+ : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$$

This map takes in two vectors in $\mathbb{R}^2$ and outputs a new vector in $\mathbb{R}^2$.

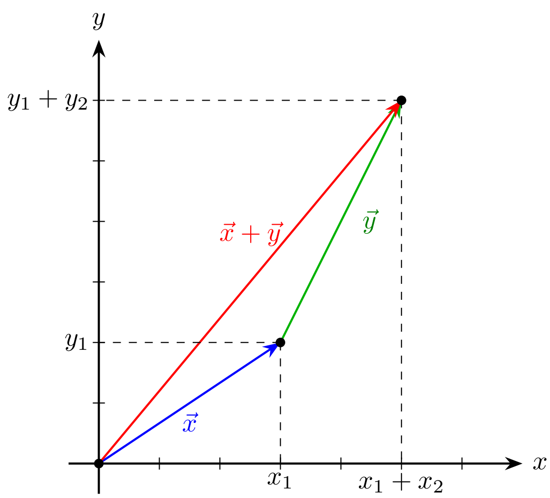
Let

$$\vec{x} = \langle x_1, y_1 \rangle \quad \text{and} \quad \vec{y} = \langle x_2, y_2 \rangle.$$

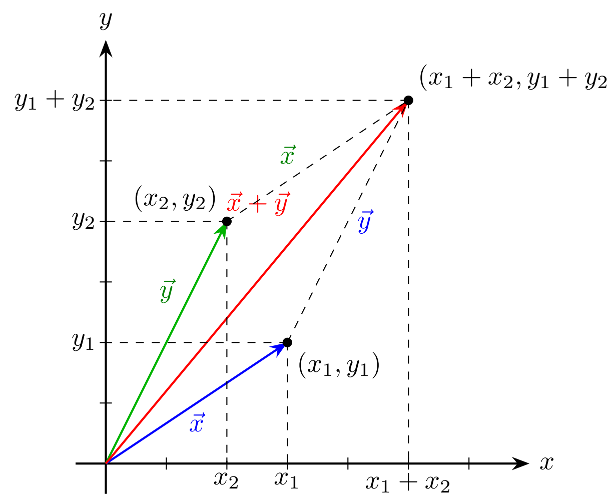
Then,

$$\vec{x} + \vec{y} = \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle = \langle x_1 + x_2,\ y_1 + y_2 \rangle.$$

We can visualize these using the triangle law and the parallelogram law:



Triangle Law of Vector Addition



Parallelogram Law of Vector Addition

In the triangle law, you place the vectors *tip-to-tail*. The sum is the line connecting the tip and head.

In the parallelogram method, both vectors start at the same point. The sum lies along the diagonal of the parallelogram.

The parallelogram law is also a geometric proof that tells us that vector addition is commutative. Both paths of the parallelogram lead to the same result. That is, $\vec{x} + \vec{y} = \vec{y} + \vec{x}$.

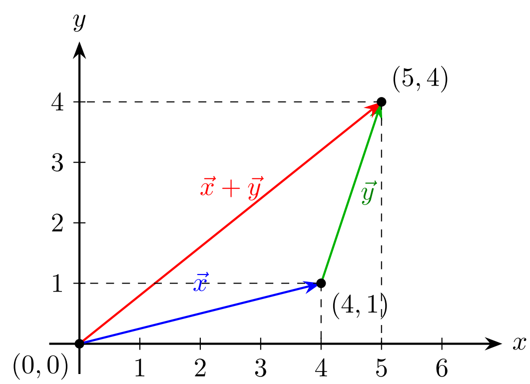
EXAMPLE 2.3

Let $\vec{x} = \langle 4, 1 \rangle$ and $\vec{y} = \langle 1, 3 \rangle$. Find $\vec{x} + \vec{y}$ and graph the vectors using the triangle law.

Solution:

$$\begin{aligned} \vec{x} + \vec{y} &= \langle 4, 1 \rangle + \langle 1, 3 \rangle \\ &= \langle 4 + 1, 1 + 3 \rangle \\ &= \langle 5, 4 \rangle \end{aligned}$$

Let's visualize using the triangle law:



We now introduce a third operation called **scalar-vector multiplication**. This operation takes a real number (called a *scalar*) and a vector in $\mathbb{R}^2$, and produces another vector in $\mathbb{R}^2$.

The operation is defined as follows:

$$\cdot : \mathbb{R} \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

That is, scalar-vector multiplication maps a scalar and a vector to a new vector.

Let $c \in \mathbb{R}$ and $\vec{x} \in \mathbb{R}^2$ with $\vec{x} = \langle x_1, y_1 \rangle$.

Then,

$$\begin{aligned} c \cdot \vec{x} &= c \cdot \langle x_1, y_1 \rangle \\ &= \langle c \cdot x_1,\, c \cdot y_1 \rangle = \langle cx_1,\, cy_1 \rangle \end{aligned}$$

EXAMPLE 2.4

Multiply $c = 2$ by $\vec{x} = \langle 1, 1 \rangle$. Then, reverse the direction of $\vec{x}$.

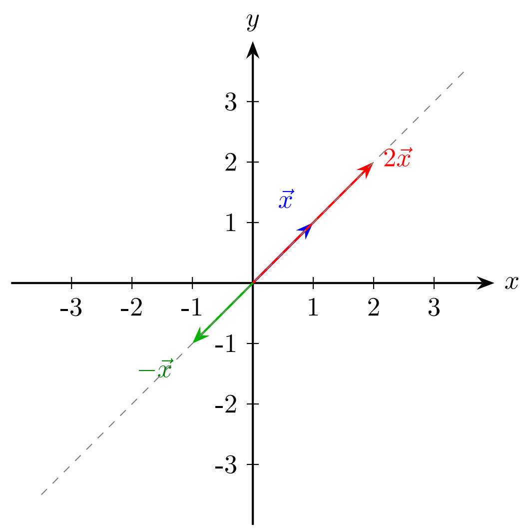
Solution:

$$
\begin{aligned}
2 \cdot \vec{x} &= 2 \cdot \langle 1, 1 \rangle \\
&= \langle 2 \cdot 1, 2 \cdot 1 \rangle \\
&= \langle 2, 2 \rangle
\end{aligned}
$$

We can reverse the direction multiplying by -1:

$$
\begin{aligned}
-1 \cdot \vec{x} &= -1 \cdot \langle 1, 1 \rangle \\
&= \langle -1 \cdot 1, -1 \cdot 1 \rangle \\
&= \langle -1, -1 \rangle
\end{aligned}
$$

Now, let's visualize it graphically:



Two vectors $\vec{x}$ and $\vec{y}$ with an initial point at the origin are in the same direction if and only if $\vec{x} = c \cdot \vec{y}$ for some scalar $c \in \mathbb{R}$.

The operation of **vector subtraction** is defined as:

$$- : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}^2$$

That is, vector subtraction maps two vectors in $\mathbb{R}^2$ to their difference, producing a new vector in $\mathbb{R}^2$.

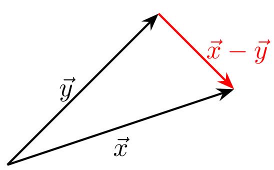
We have

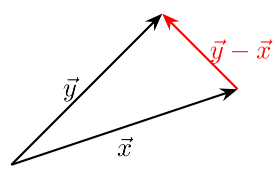
$$\vec{x} - \vec{y} = \langle x_1, y_1 \rangle - \langle x_2, y_2 \rangle$$

This can be rewritten using scalar multiplication:

$$\begin{aligned} \vec{x} + (-1) \cdot \vec{y} &= \langle x_1, y_1 \rangle + (-1) \cdot \langle x_2, y_2 \rangle \\ &= \langle x_1, y_1 \rangle + \langle -x_2, -y_2 \rangle \\ &= \langle x_1 - x_2,\, y_1 - y_2 \rangle \end{aligned}$$

Thus, subtracting two vectors simply means subtracts their components.





We say two vectors are equal if and only if they are equal in magnitude and direction. That is,

$$\begin{gathered} \vec{x} = \vec{y} \\ \vec{x} - \vec{y} = 0 = \langle 0, 0 \rangle \\ \|\vec{x} - \vec{y}\|_2 = 0 \end{gathered}$$

2.2 Vectors in 3D Space

In single-variable calculus, recall how we did this:

$$\frac{d}{dx}[F(x)] = f(x) = F'(x)$$

where we grouped the input and output of $F(x)$ into ordered pairs $(x, f(x))$.

We now introduce functions of two variables:

$$z = F(x, y)$$

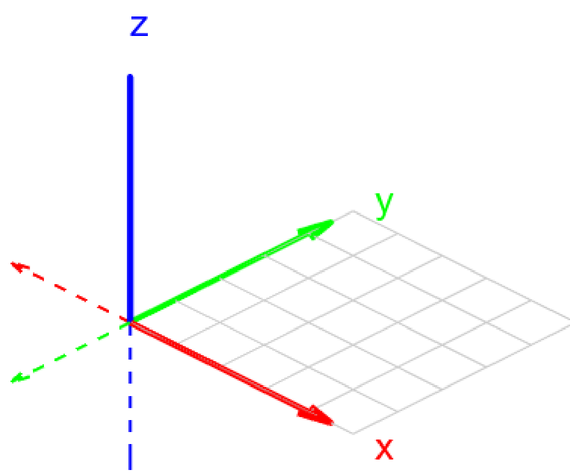
This function takes two inputs, x and y, and produces an output z.

If we group the input and output together, we get an ordered triple:

$$(x, y, z) = (x, y, F(x, y))$$

This represents a point in $\mathbb{R}^3$.

To graph an ordered triplet (x, y, z), we need to draw three **orthogonal** axes (perpendicular to each other) that intersect at the origin. This system is known as a three-dimensional Euclidean space $\mathbb{R}^3$:



A fundamental part of this is understanding the three main coordinate planes. These planes act as "boundaries" between positive and negative regions of space and are where one of the three coordinates is zero.

xy-plane:

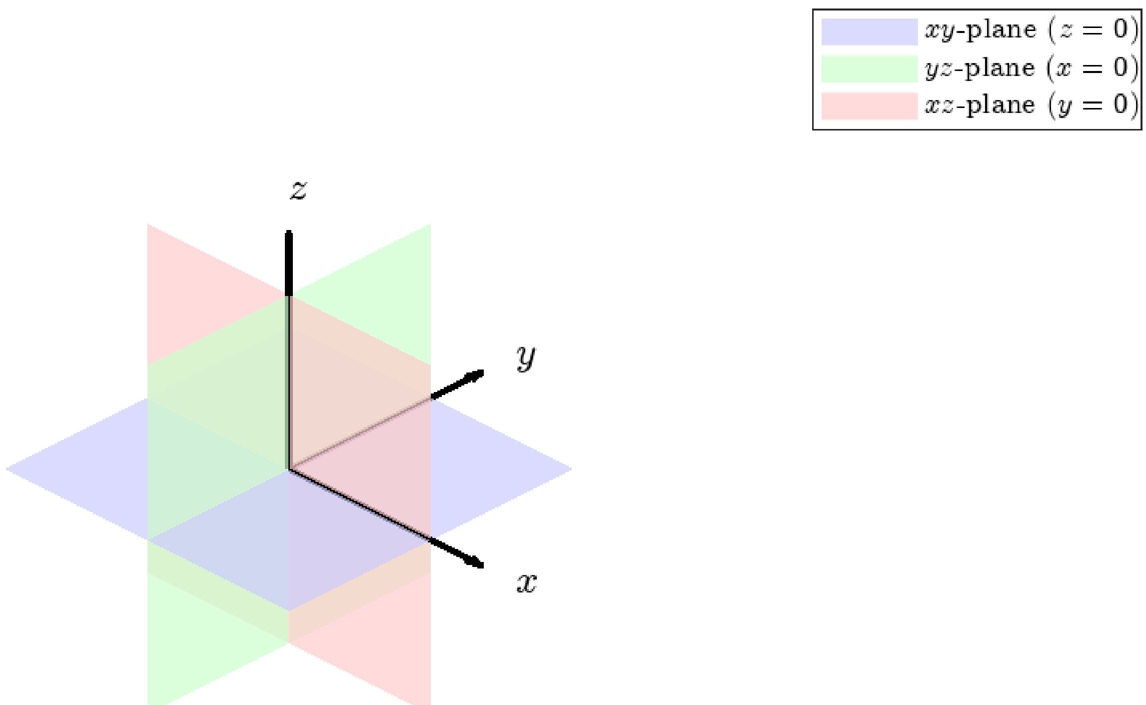
$$z = 0 \Rightarrow \{(x, y, z) \in \mathbb{R}^3 : z = 0\} \;=\; \{(x, y, 0) : x, y \in \mathbb{R}\}.$$

yz-plane:

$$x = 0 \Rightarrow \{(x, y, z) \in \mathbb{R}^3 : x = 0\} \;=\; \{(0, y, z) : y, z \in \mathbb{R}\}.$$

xz-plane:

$$y = 0 \Rightarrow \{(x, y, z) \in \mathbb{R}^3 : y = 0\} \;=\; \{(x, 0, z) : x, z \in \mathbb{R}\}.$$



The Three Coordinate Planes in $\mathbb{R}^3$

In the current orientation facing out of the page, you can think of the xz-plane like the right wall of a house, the yz-plane the left wall, and the xy-plane the floor.

These surfaces are hugely important when graphing because visualizing functions of multiple variables can often be a challenge.

EXAMPLE 2.5

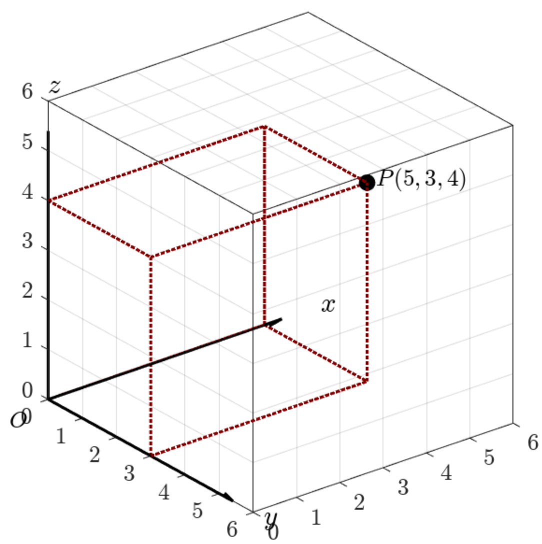
Graph the points $O(0,0,0)$ and $P(5,3,4)$.

Solution:

To locate point P, we move

- 5 units along the x-axis,
- 3 units along the y-axis, and
- 4 units upward along the z-axis.

We can then use dotted lines to connect P to the coordinate planes. This forms a rectangular prism that shows how P projects onto the three planes. Let's now graph:



EXAMPLE 2.5 (CONTINUED)

Please run the MATLAB code yourself and have a look!

```
figure
hold on
axis equal
grid on

xlim([0 6])
ylim([0 6])
zlim([0 6])

plot3(5, 3, 4, 'ko', 'MarkerFaceColor', 'k', 'MarkerSize', 8)

line([0 5], [0 0], [0 0], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([5 5], [0 3], [0 0], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([5 5], [3 3], [0 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 0], [0 3], [0 0], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 0], [0 0], [0 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 0], [3 3], [0 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([5 5], [0 0], [0 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 5], [3 3], [0 0], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 5], [0 0], [4 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 0], [0 3], [4 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([5 5], [0 3], [4 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)
line([0 5], [3 3], [4 4], 'LineStyle', ':', 'Color', [0.5 0 0], 'LineWidth',
    1.5)

quiver3(0, 0, 0, 6, 0, 0, 'k', 'LineWidth', 1.5)
quiver3(0, 0, 0, 0, 6, 0, 'k', 'LineWidth', 1.5)
quiver3(0, 0, 0, 0, 0, 6, 'k', 'LineWidth', 1.5)

set(gca, 'XTick', 0:1:6, 'YTick', 0:1:6, 'ZTick', 0:1:6, ...
    'FontSize', 12, 'TickLabelInterpreter', 'latex')

text(6.3, 0, 0, '$x$', 'Interpreter', 'latex', 'FontSize', 14)
text(0, 6.3, 0, '$y$', 'Interpreter', 'latex', 'FontSize', 14)
text(0, 0, 6.3, '$z$', 'Interpreter', 'latex', 'FontSize', 14)
text(5.2, 3, 4, '$P(5,3,4)$', 'Interpreter', 'latex', 'FontSize', 12)
text(-0.5, -0.5, -0.5, '$0$', 'Interpreter', 'latex', 'FontSize', 12)

view(45, 30)
box on
```

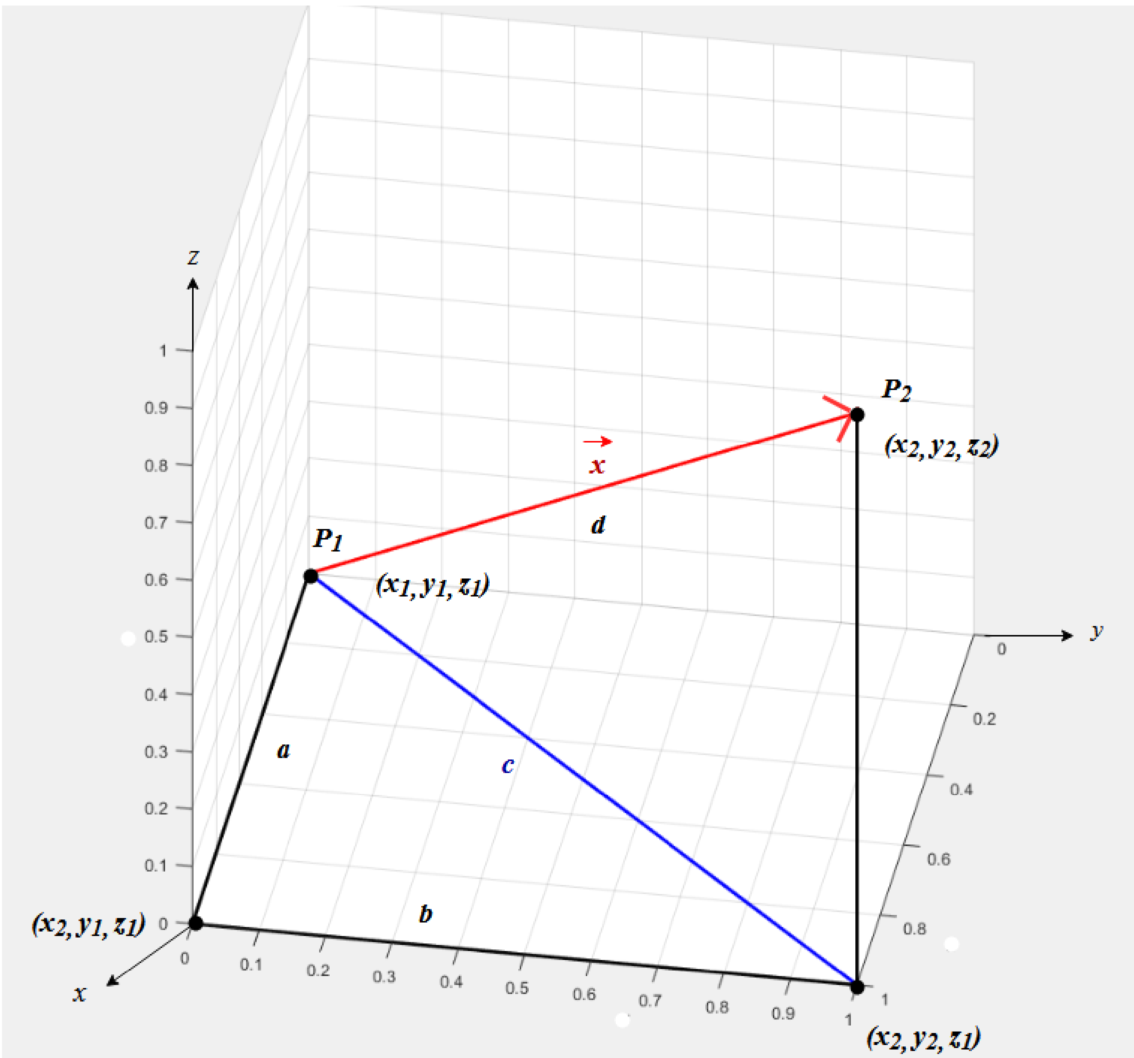
ex2point5plot.m

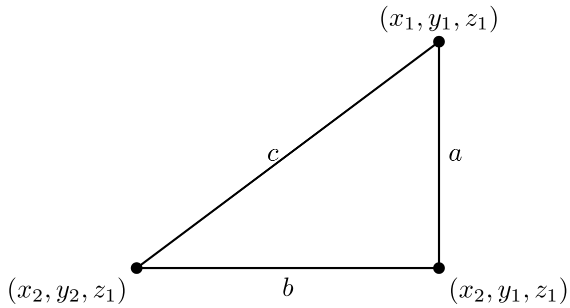
We now extend the Pythagorean theorem into three dimensions to define the 2-norm of a vector in $\mathbb{R}^3$.

Let $P_1 = (x_1, y_1, z_1)$ and $P_2 = (x_2, y_2, z_2)$ be two points in $\mathbb{R}^3$. We define the vector $\vec{x}$ from P_1 to P_2 as:

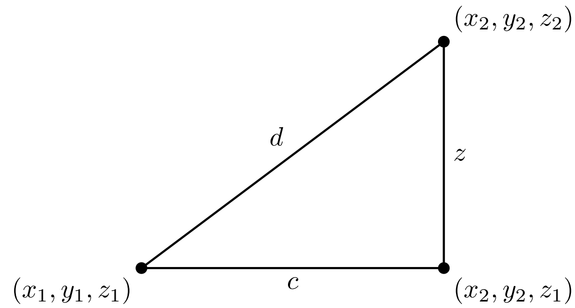
$$\vec{x} = \overrightarrow{P_1P_2} = \langle x_2 - x_1,\, y_2 - y_1,\, z_2 - z_1 \rangle$$

The 2-norm of $\vec{x}$ measures the straight-line distance between P_1 and P_2. We will analyze two right triangles as seen in this graph:





Triangle 2: Projection onto xy-plane to compute horizontal distance b



Triangle 1: Vertical slice showing the full distance

Let's begin solving for the 2-norm. We are trying to find $d = \|\vec{x}\|_2$.

By the Pythagorean theorem, we can see that $a = x_2 - x_1$ and $b = y_2 - y_1$ in Triangle 1. In Triangle 2, we can see that $d^2 = c^2 + z^2$ where $z = z_2 - z_1$. Thus $d^2 = a^2 + b^2 + z^2 = \|\vec{x}\|_2^2$. From that, we can substitute in our values for a, b, c and then simplify:

$$
\begin{aligned}
\|\vec{x}\|_2^2 &= (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2 \\
\|\vec{x}\|_2 &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}
\end{aligned}
$$

Similarly to last time, we assume that the initial point of a vector is at the origin $(0, 0, 0)$ unless otherwise stated.

EXAMPLE 2.6

Find $\|\vec{x}\|_2$ where $\vec{x} = \overrightarrow{PQ}$ and $P(2, -1, 7) \quad \text{and} \quad Q(1, -3, 5)$.

Solution:

$$
\begin{aligned}
\vec{x} &= \langle 1 - 2,\ -3 - (-1),\ 5 - 7 \rangle \\
&= \langle -1,\ -2,\ -2 \rangle
\end{aligned}
$$

$$
\begin{aligned}
\|\vec{x}\|_2 &= \sqrt{(-1)^2 + (-2)^2 + (-2)^2} \\
&= \sqrt{1 + 4 + 4} \\
&= \sqrt{9} = 3
\end{aligned}
$$

The 2-norm is defined as the following mapping:

$$\|\cdot\|_2 : \mathbb{R}^3 \to \mathbb{R},$$

which means it takes a vector from $\mathbb{R}^3$ and maps it to a real number representing its magnitude.

For a vector $\vec{x} = \langle x_1, y_1, z_1 \rangle$, we define the 2-norm by

$$\|\vec{x}\|_2 = \sqrt{x_1^2 + y_1^2 + z_1^2}$$

Then, we can say that $\|\vec{x}\|_2^2 = x_1^2 + y_1^2 + z_1^2$.

Let's move on to vector addition and scalar-vector multiplication in $\mathbb{R}^3$.

We define **vector addition** as a binary operation:

$$+ : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

This operation takes two vectors in $\mathbb{R}^3$ and outputs another vector in $\mathbb{R}^3$.

Explicitly, for vectors

$$\vec{x} = \langle x_1, y_1, z_1 \rangle \quad \text{and} \quad \vec{y} = \langle x_2, y_2, z_2 \rangle,$$

their sum is given by

$$\begin{aligned} \vec{x} \underbrace{+}_{\text{(a)}} \vec{y} &= \langle x_1, y_1, z_1 \rangle \underbrace{+}_{\text{(a)}} \langle x_2, y_2, z_2 \rangle \\ &= \langle x_1 \underbrace{+}_{\text{(b)}} x_2,\ y_1 \underbrace{+}_{\text{(b)}} y_2,\ z_1 \underbrace{+}_{\text{(b)}} z_2 \rangle. \end{aligned}$$

The addition denoted by (a) refers to adding entire vectors whereas (b) refers to scalar addition. One vector addition in $\mathbb{R}^3$ thus corresponds to three scalar additions in $\mathbb{R}$. This rule also scales to any number of dimensions; vector addition in $\mathbb{R}^n$ corresponds to n scalar additions in $\mathbb{R}$.

EXAMPLE 2.7

Find $\vec{x}+\vec{y}$ where $\vec{x}=\langle 1,0,1\rangle$ and $\vec{y}=\langle -1,1,3\rangle$.

Solution:

$$\begin{aligned}\vec{x}+\vec{y} &= \langle 1,0,1\rangle + \langle -1,1,3\rangle \\ &= \langle 1+(-1),\, 0+1,\, 1+3\rangle \\ &= \langle 0,1,4\rangle\end{aligned}$$

We will now define scalar-vector multiplication:

$$\cdot : \mathbb{R}\times\mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

Let $c\in\mathbb{R}$ and $\vec{x}\in\mathbb{R}^3$ with $\vec{x}=\langle x_1,y_1,z_1\rangle$. Then

$$c\cdot\vec{x} = c\cdot\langle x_1,y_1,z_1\rangle = \langle c\cdot x_1,\, c\cdot y_1,\, c\cdot z_1\rangle.$$

EXAMPLE 2.8

Let $\vec{z}=\langle 1,-1,3\rangle$. Find (a) $0\cdot\vec{z}$ and (b) $-1\cdot\vec{z}$.

Solution:

(a) $0\cdot\vec{z} = 0\cdot\langle 1,-1,3\rangle = \langle 0\cdot 1,\, 0\cdot(-1),\, 0\cdot 3\rangle = \langle 0,0,0\rangle$.

Multiplying any vector by zero produces the zero vector.

(b) $-1\cdot\vec{z} = -1\cdot\langle 1,-1,3\rangle = \langle (-1)\cdot 1,\, (-1)\cdot(-1),\, (-1)\cdot 3\rangle = \langle -1,1,-3\rangle$.

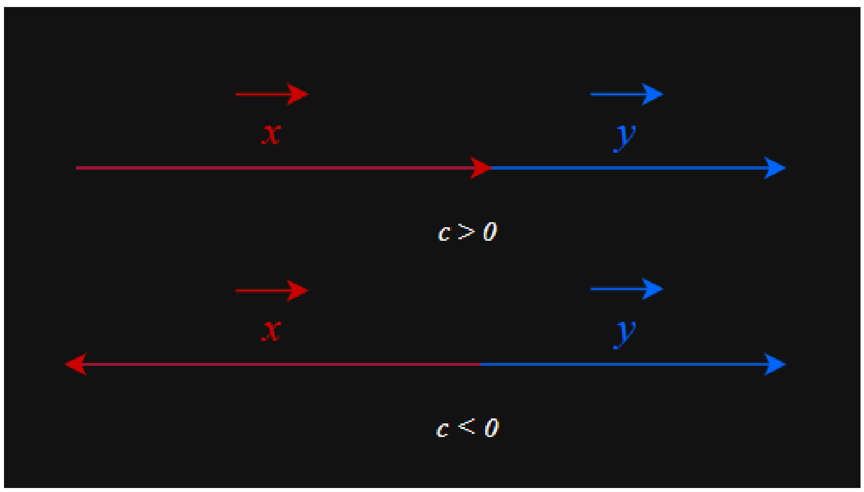
Multiplying a vector by -1 reverses its direction.

Two nonzero vectors $\vec{x}$ and $\vec{y}$ point in the same direction if there exists a scalar c such that:

$$\vec{x} = c\vec{y}$$

where if $c > 0$, $\vec{x}$ points in the same direction as $\vec{y}$, scaled by $|c|$. If $c < 0$, $\vec{x}$ points in the opposite direction, scaled by $|c|$. Furthermore, we can say that

- If $c > 0$, the vectors have the same direction and orientation.
- If $c < 0$, the vectors have the same direction but opposite orientation.



A **circle** in $\mathbb{R}^2$ is defined as the collection of points in the plane that lie at a fixed distance r from a given center (h, k).

Formally,

$$C = \left\{ (x, y) \; : \; \sqrt{(x-h)^2 + (y-k)^2} = r \right\}.$$

From this, we can get the standard equation of a circle

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

We can generalize circles in $\mathbb{R}^2$ as spheres in $\mathbb{R}^3$. We define a sphere as the set of all points (x, y, z) that are r units away from the center of our sphere (h, k, l). That is,

$$S = \left\{(x, y, z) \in \mathbb{R}^3 \,:\, \sqrt{(x-h)^2 + (y-k)^2 + (z-l)^2} = r\right\}.$$

We can express the sphere centered at (h, k, l) with radius r in two equivalent forms.

First, in normalized form:

$$S = \left\{(x, y, z) \in \mathbb{R}^3 \,:\, \frac{(x-h)^2}{r^2} + \frac{(y-k)^2}{r^2} + \frac{(z-l)^2}{r^2} = 1\right\}$$

This shows the sphere as a collection of points in $\mathbb{R}^3$.

Equivalently, we can write the sphere in its standard implicit form:

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

Let's stop to think about scalars and vectors in the real world. Recall that scalars only have magnitude while vectors have magnitude *and* direction. For instance, take mass. A block with a mass of 12 kg has a magnitude of 12 but kilograms do not have a direction. Thus, mass is a scalar. On the other hand, take weight. A block that weighs 5 N has a magnitude of 5 and also has a direction. Weight is gravitational force, so gravity is exerting a force on the block 5 N in the direction of the center of the Earth. Thus, weight is a vector. Length, area, volume, speed, mass, density, pressure, work, power, temperature, energy, entropy, electric current, and time are all examples of scalar quantities. Displacement, velocity, acceleration, momentum, force, drag, weight, and electric field are all examples of vector quantities.

Properties of Vectors

Let vectors $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ and scalars $c, d \in \mathbb{R}$. Then, the following properties hold:

1. **Commutative property:** $\vec{a} + \vec{b} = \vec{b} + \vec{a}$
2. **Associative property:** $(\vec{a} + \vec{b}) + \vec{c} = \vec{a} + (\vec{b} + \vec{c})$
3. **Additive identity property:** $\vec{a} + \vec{0} = \vec{a}$
4. **Additive inverse property:** $\vec{a} + (-\vec{a}) = \vec{0}$
5. **Distributive property for scalars:** $c(\vec{a} + \vec{b}) = c\vec{a} + c\vec{b}$
6. **Distributive property across scalars:** $(c + d)\vec{a} = c\vec{a} + d\vec{a}$
7. **Associative property for scalar multiplication:** $c(d\vec{a}) = (cd)\vec{a}$
8. **Multiplicative identity:** $1 \cdot \vec{a} = \vec{a}$

In three-dimensional space, the vectors

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

are called the **standard basis vectors**. These vectors each have length 1 and point in the directions of the positive x-axis, y-axis, and z-axis, respectively.

A vector $\vec{a} = \langle a_1, a_2, a_3 \rangle$ in $\mathbb{R}^3$ can be expressed as a linear combination of these basis vectors:

$$\vec{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$$

EXAMPLE 2.9

Let $\vec{a} = \mathbf{i} + 2\mathbf{j} - 3\mathbf{k}$ and $\vec{b} = 4\mathbf{i} + 7\mathbf{k}$. Express the vector $2\vec{a} + 3\vec{b}$ in terms of $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$.

Solution:

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

A **unit vector** is a vector of length 1. The standard basis vectors $\mathbf{i}, \mathbf{k}$, and $\mathbf{k}$ are all unit vectors. Given a nonzero vector $\vec{a}$, we can create a unit vector in its direction by dividing by its norm:

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

To verify this, let $c = \frac{1}{\|\vec{a}\|}$. Then $\mathbf{u} = c\vec{a}$ and c is a positive scalar, so $\mathbf{u}$ has the same direction as $\vec{a}$. We can also find unit length:

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

The process of finding unit vectors is known as **normalization**. To do this, you use scalar multiplication. Let's go through a few examples:

EXAMPLE 2.10

Find the unit vector $\mathbf{u}$ in the direction of $\vec{a} = 2\mathbf{i} - \mathbf{j} - 2\mathbf{k}$.

Solution:

First, let's find the length of the given vector:

$$\|\vec{a}\| = \sqrt{2^2 + (-1)^2 + (-2)^2} = \sqrt{4+1+4} = \sqrt{9} = 3$$

Thus, we divide by length to get

$$\mathbf{u} = \frac{1}{3}\langle 2\mathbf{i}, -1\mathbf{j}, -2\mathbf{k}\rangle = \left\langle \frac{2}{3}, -\frac{1}{3}, -\frac{2}{3} \right\rangle.$$

EXAMPLE 2.11

Let $\vec{v} = \langle 9, 2\rangle$. Find a vector with magnitude 5 that points in the opposite direction of $\vec{v}$.

Solution:

First, we compute the magnitude of $\vec{v}$:

$$\|\vec{v}\| = \sqrt{9^2 + 2^2} = \sqrt{85}$$

Then, the unit vector in the direction of $\vec{v}$ is:

$$\frac{1}{\sqrt{85}}\langle 9, 2\rangle = \left\langle \frac{9}{\sqrt{85}}, \frac{2}{\sqrt{85}} \right\rangle$$

To get a vector with magnitude 5 in the opposite direction, we multiply by -5:

$$-5 \cdot \left\langle \frac{9}{\sqrt{85}}, \frac{2}{\sqrt{85}} \right\rangle = \left\langle \frac{-45}{\sqrt{85}}, \frac{-10}{\sqrt{85}} \right\rangle$$

Thus, the desired vector is:

$$\left\langle \frac{-45}{\sqrt{85}}, \frac{-10}{\sqrt{85}} \right\rangle$$

EXAMPLE 2.12

Let $\vec{a} = \langle 16, -11 \rangle$ and let $\vec{b}$ be a unit vector that forms an angle of $225°$ with the positive x-axis. Express $\vec{a}$ and $\vec{b}$ in terms of the standard unit vectors $\mathbf{i}$ and $\mathbf{j}$.

Solution:

We can write $\vec{a}$ as $\vec{a} = 16\mathbf{i} - 11\mathbf{j}$. Next, since $\vec{b}$ is a unit vector at angle $225°$, we can express it as $\vec{b} = \cos 225° \, \mathbf{i} + \sin 225° \, \mathbf{j}$.

Evaluating, we get $\cos 225° = -\frac{\sqrt{2}}{2}$ and $\sin 225° = -\frac{\sqrt{2}}{2}$. Substituting,

$$\vec{b} = -\frac{\sqrt{2}}{2}\,\mathbf{i} - \frac{\sqrt{2}}{2}\,\mathbf{j}.$$

Finally, we have the following:

$$\vec{a} = 16\mathbf{i} - 11\mathbf{j}$$
$$\vec{b} = -\frac{\sqrt{2}}{2}\,\mathbf{i} - \frac{\sqrt{2}}{2}\,\mathbf{j}$$

2.3 The Dot Product

We now have a very important operation in $\mathbb{R}^2$ and $\mathbb{R}^3$ called the inner product, or more specifically the **dot product**. The dot product serves as a "multiplication" operation between vectors:

$$\cdot : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$$
$$\cdot : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}$$

Definition: If $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$, then the dot product of $\vec{u}$ and $\vec{v}$ is

$$\vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Properties of the Dot Product: For vectors $\vec{u}, \vec{v}, \vec{w}$ and scalar c, the following hold:

1. $\vec{u} \cdot \vec{u} = |\vec{u}|^2$
2. $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
3. $\vec{u} \cdot (\vec{v} + \vec{w}) = \vec{u} \cdot \vec{v} + \vec{u} \cdot \vec{w}$
4. $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$
5. $\vec{0} \cdot \vec{u} = 0$

EXAMPLE 2.13

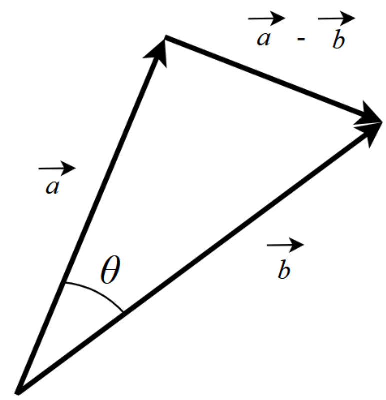
Compute $\vec{a} \cdot \vec{b}$, where $\vec{a} = \langle 2, 9, -1 \rangle$ and $\vec{b} = \langle -3, 1, -4 \rangle$.

Solution:

We apply the definition of the dot product:

$$\begin{aligned} \vec{u} \cdot \vec{v} &= (2)(-3) + (9)(1) + (-1)(-4) \\ &= -6 + 9 + 4 \\ &= 7 \end{aligned}$$

We can use the dot product to find the angle between two vectors. Let's say we have a triangle of vectors $\vec{a}$, $\vec{b}$, and $\vec{a} - \vec{b}$:



The norms of the values on each of the three sides would give you the side lengths of the triangle. Applying the law of cosines here gives

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

We can rewrite the left-hand side using the dot product:

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

Substituting into the law of cosines yields the following:

$$\begin{aligned} \|\vec{b}\|^2 - 2\vec{a} \cdot \vec{b} + \|\vec{a}\|^2 &= \|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\| \cos\theta \\ -2\vec{a} \cdot \vec{b} &= -2\|\vec{a}\|\|\vec{b}\| \cos\theta \\ \vec{a} \cdot \vec{b} &= \|\vec{a}\|\|\vec{b}\| \cos\theta \end{aligned}$$

Thus, for an angle θ between two vectors $\vec{a}$ and $\vec{b}$, we have

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

And to find the angle between two nonzero vectors,

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

EXAMPLE 2.14

Find angle in radians formed by the vectors $\vec{a} = \langle 1, 2, 0\rangle$ and $\vec{b} = \langle 2, 4, 1\rangle$.

Solution:

First, we compute the dot product

$$\vec{a}\cdot\vec{b} = (1)(2) + (2)(4) + (0)(1) = 2 + 8 + 0 = 10.$$

Next, we can compute the magnitudes:

$$\|\vec{a}\| = \sqrt{1^2 + 2^2 + 0^2} = \sqrt{5}$$

$$\|\vec{b}\| = \sqrt{2^2 + 4^2 + 1^2} = \sqrt{21}$$

Finally, we can compute θ.

$$\cos\theta = \frac{10}{\sqrt{105}} \Rightarrow \theta = \cos^{-1}\left(\frac{10}{\sqrt{105}}\right).$$

Thus, the angle is approximately 0.22 radians.

Two vectors $\vec{a}$ and $\vec{b}$ are orthogonal (perpendicular) if and only if

$$\vec{a} \cdot \vec{b} = 0$$

We will prove this. Let $\vec{a}$ and $\vec{b}$ be nonzero vectors with angle θ between them. First, assume $\vec{a} \cdot \vec{b} = 0$. Then

$$\|\vec{a}\|\|\vec{b}\| \cos\theta = 0.$$

For $\|\vec{a}\| \neq 0$ and $\|\vec{b}\| \neq 0$, we must have $\cos\theta = 0$. Thus, $\theta = 90°$, and the vectors are now orthogonal.

Now we can assume $\vec{a}$ and $\vec{b}$ are orthogonal. Plugging in $\theta = 90°$ gets us

$$\begin{aligned}\vec{a} \cdot \vec{b} &= \|\vec{a}\|\|\vec{b}\| \cos\theta \\ &= \|\vec{a}\|\|\vec{b}\| \cos 90° \\ &= \|\vec{a}\|\|\vec{b}\|(0) \\ &= 0.\end{aligned}$$

If two vectors $\vec{a}$ and $\vec{b}$ are orthogonal, we write $\vec{a} \perp \vec{b}$. Orthogonal vectors will always form a right angle when their initial points are aligned.

EXAMPLE 2.15

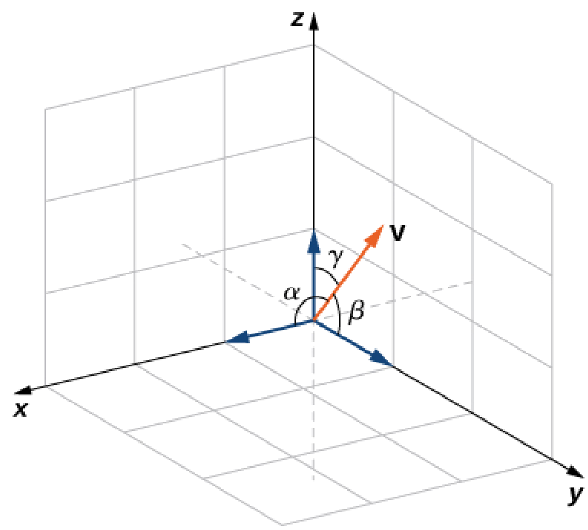
For which value of x is $\vec{a} = \langle 2, 8, -1\rangle$ orthogonal to $\vec{b} = \langle x, -1, 2\rangle$?

Solution:

Since $\vec{a}$ and $\vec{b}$ are orthogonal, we must have $\vec{a} \cdot \vec{b} = 0$:

$$\vec{a} \cdot \vec{b} = (2)(x) + (8)(-1) + (-1)(2) = 2x - 10$$
$$2x - 10 = 0 \Rightarrow x = 5$$

The angles a nonzero vector makes with each of the coordinate axes are called the **direction angles** α, β, and γ. These are very important in real-world applications. For example, in engineering, you can use direction angles to carefully calculate the orientation of a robot or the trajectory of a missile. The cosines of the direction angles are called **direction cosines**. In the following image, angle α is formed by vector $\vec{v}$ and unit vector $\mathbf{i}$, angle β is formed by vector $\vec{v}$, and unit vector $\mathbf{j}$, and angle γ is formed by vector $\vec{v}$ and unit vector $\mathbf{k}$:



Direction angles α, β, and γ. Image credit: Strang & Herman

Let's find the general form. For a nonzero vector $\vec{a}$, we have the following:

$$\cos\alpha = \frac{\vec{a}\cdot\mathbf{i}}{\|\vec{a}\|\|\mathbf{i}\|} = \frac{a_1}{\|\vec{a}\|}, \quad \cos\beta = \frac{\vec{a}\cdot\mathbf{j}}{\|\vec{a}\|\|\mathbf{j}\|} = \frac{a_2}{\|\vec{a}\|}, \quad \cos\gamma = \frac{\vec{a}\cdot\mathbf{k}}{\|\vec{a}\|\|\mathbf{k}\|} = \frac{a_3}{\|\vec{a}\|}$$

Squaring and adding these equations, we obtain

$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

We can then acquire

$$\begin{aligned}\vec{a} &= \langle a_1, a_2, a_3\rangle \\ &= \langle \|\vec{a}\|\cos\alpha,\ \|\vec{a}\|\cos\beta,\ \|\vec{a}\|\cos\gamma\rangle \\ &= \|\vec{a}\|\langle\cos\alpha, \cos\beta, \cos\gamma\rangle. \end{aligned} \tag{1}$$

Thus, the unit vector in the direction of $\vec{a}$ can be written as

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

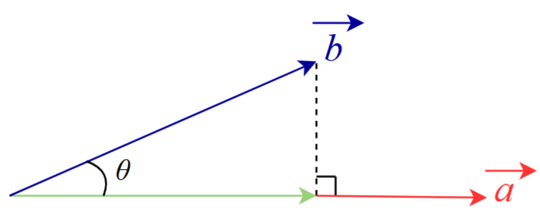
Adding two vectors together creates a resultant vector. But if we need to break a vector down into its components, we can use **vector projections**.

The **scalar projection** of $\vec{b}$ onto $\vec{a}$, also called the component of $\vec{b}$ along $\vec{a}$, measures the magnitude of the projection of $\vec{b}$ in the direction of $\vec{a}$. It is given by

$$\operatorname{comp}_{\vec{a}} \vec{b} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|}.$$

The **vector projection** of $\vec{b}$ onto $\vec{a}$ gives the actual vector in the direction of $\vec{a}$ whose length equals the scalar projection. That is, $\|\operatorname{proj}_{\vec{a}} \vec{b}\| = \operatorname{comp}_{\vec{a}} \vec{b}$. It is computed by multiplying the scalar projection by the unit vector in the direction of $\vec{a}$:

$$\operatorname{proj}_{\vec{a}} \vec{b} = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|^2}\vec{a}$$



The scalar projection and length of vector projection $\|\operatorname{proj}_{\vec{a}}\vec{b}\|$ are in green.

EXAMPLE 2.16

Find the scalar projection and vector projection of $\vec{b} = \langle 1, 1, 2\rangle$ onto $\vec{a} = \langle -2, 3, 1\rangle$.

Solution:

We first compute the magnitude of $\vec{a}$ which is $\|\vec{a}\| = \sqrt{(-2)^2 + 3^2 + 1^2} = \sqrt{14}$.

Now we compute the scalar projection:

$$\text{comp}_{\vec{a}}\,\vec{b} = \frac{(-2)(1) + 3(1) + 1(2)}{\sqrt{14}} = \frac{3}{\sqrt{14}}$$

We can now use this to find the vector projection

$$\text{proj}_{\vec{a}}\,\vec{b} = \frac{\text{comp}_{\vec{a}}\,\vec{b}}{\|\vec{a}\|}\vec{a} = \frac{3}{\sqrt{14}} \cdot \frac{\vec{a}}{\sqrt{14}} = \frac{3}{14}\vec{a} = \frac{3}{14}\langle -2, 3, 1\rangle = \langle -\frac{3}{7}, \frac{9}{14}, \frac{3}{14}\rangle.$$

Vectors can also be used to represent quantities of items. As a matter of fact, the idea of using vectors to store data is one of the most powerful in the modern day. You will revisit this in linear algebra.

EXAMPLE 2.17

A local market sells bread, milk, eggs, and apples. They pay \$0.25 per loaf of bread, \$0.25 per bottle of milk, \$0.50 per dozen eggs, and \$0.20 per apple. Bread sells for \$2.50 , milk for \$1.50, eggs for \$4.50, and apples for \$1.25. Last year, the market sold 1258 loaves of bread, 342 bottles of milk, 2426 dozens of eggs, and 1354 apples. Use vectors and dot products to compute last year's total sales and profit.

Solution:

The cost, price, and quantity vectors are:

$$\vec{c} = (0.25, 0.25, 0.50, 0.20),\ \vec{p} = (2.50, 1.50, 4.50, 1.25),\ \vec{q} = (1258, 342, 2426, 1354)$$

Total sales are given by

$$\begin{aligned}\vec{p} \cdot \vec{q} &= (2.50, 1.50, 4.50, 1.25) \cdot (1258, 342, 2426, 1354)\\ &= \$3145 + \$513 + \$10917 + \$1692.5 = \$16267.5\end{aligned}$$

Total cost is given by

$$\begin{aligned}\vec{c} \cdot \vec{q} &= (0.25, 0.25, 0.50, 0.20) \cdot (1258, 342, 2426, 1354)\\ &= \$314.5 + \$85.5 + \$1213 + \$270.8 = \$1883.8\end{aligned}$$

Profit:

$$\vec{p} \cdot \vec{q} - \vec{c} \cdot \vec{q} = 16267.5 - 1883.8 = 14383.7$$

Thus, the market made \$14,383.70 in profit last year.

EXAMPLE 2.18

Express $\vec{b} = \langle 8, -3, -3\rangle$ as a sum of orthogonal vectors such that one of the vectors has the same direction as $\vec{a} = \langle 2, 3, 2\rangle$.

Solution:

Let $\vec{p}$ represent the projection of $\vec{b}$ onto $\vec{a}$:

$$\vec{p} = \text{proj}_{\vec{a}}\,\vec{b} = \frac{\vec{a}\cdot\vec{b}}{\|\vec{a}\|^2}\vec{a} = \frac{\langle 2,3,2\rangle\cdot\langle 8,-3,-3\rangle}{\|\langle 2,3,2\rangle\|^2}\langle 2,3,2\rangle$$

$$= \frac{16-9-6}{2^2+3^2+2^2}\langle 2,3,2\rangle = \frac{1}{17}\langle 2,3,2\rangle = \left\langle \frac{2}{17}, \frac{3}{17}, \frac{2}{17}\right\rangle$$

Then,

$$\vec{q} = \vec{b} - \vec{p} = \langle 8,-3,-3\rangle - \left\langle \frac{2}{17}, \frac{3}{17}, \frac{2}{17}\right\rangle = \left\langle \frac{134}{17}, \frac{-54}{17}, \frac{-53}{17}\right\rangle.$$

To check our work, we can verify that $\vec{p}$ and $\vec{q}$ are orthogonal using the dot product:

$$\vec{p}\cdot\vec{q} = \left\langle \frac{2}{17}, \frac{3}{17}, \frac{2}{17}\right\rangle \cdot \left\langle \frac{134}{17}, \frac{-54}{17}, \frac{-53}{17}\right\rangle$$

$$= \frac{268}{289} + \frac{-162}{289} + \frac{-106}{289} = 0$$

Then,

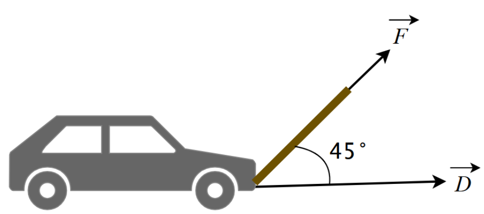
$$\vec{b} = \vec{p} + \vec{q} = \left\langle \frac{2}{17}, \frac{3}{17}, \frac{2}{17}\right\rangle + \left\langle \frac{134}{17}, \frac{-54}{17}, \frac{-53}{17}\right\rangle.$$

One of the most common applications of dot products is in physics when we want to calculate *work*. Work W is done by a force when it transfers energy to move an object. For a constant force F that moves an object over a distance d, the formula is $W = Fd$. However, this only works when the force acts in the same direction as the object being displaced. If the displacement vector $\vec{D} = \overrightarrow{PQ}$ is pointing in a different direction and gets the object from point P to point Q, the work done by the force $\vec{F}$ acts at an angle θ and is given by

$$W = \vec{F} \cdot \overrightarrow{PQ} = \|\vec{F}\| \|\overrightarrow{PQ}\| \cos\theta.$$

EXAMPLE 2.19

A car is being pulled a distance of 100 m along a horizontal path that goes from point P to point Q by a constant force of 5000 N. The rope is held at an angle of $\theta = 45°$.



Solution:

The work done by the force is:

$$W = \|\vec{F}\| \, \|\overrightarrow{PQ}\| \cos\theta = (5000\,\mathrm{N})(100\,\mathrm{m})(\cos 45°) = 353553.4\,\mathrm{J}$$

2.4 The Cross Product

The **cross product** mapping is as follows:

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \longrightarrow \mathbb{R}^3$$

It can be used to take two vectors in $\mathbb{R}^3$ and output a third vector in $\mathbb{R}^3$ that is orthogonal to the original two vectors.

Given two nonzero vectors $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$. Let $\vec{c} = \langle c_1, c_2, c_3 \rangle$. Then we have $\vec{a} \cdot \vec{c} = 0$ and $\vec{b} \cdot \vec{c} = 0$.

That is,

$$a_1c_1 + a_2c_2 + a_3c_3 = 0,$$
$$b_1c_1 + b_2c_2 + b_3c_3 = 0.$$

To eliminate c_3, we multiply the first equation by b_2 and the second by a_2, then subtract:

$$(a_1b_2 - a_2b_1)c_1 + (a_3b_2 - a_2b_3)c_3 = 0$$

This has the form $pc_1 + qc_3 = 0$. Solving, we get $c_1 = -q$ and $c_3 = p$. Thus,

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

We then substitute in those results to get

$$c_3 = a_1b_2 - a_2b_1.$$

The resulting vector is:

$$\vec{c} = \langle a_2b_3 - a_3b_2,\ a_3b_1 - a_1b_3,\ a_1b_2 - a_2b_1 \rangle.$$

This is called the **cross product** of $\vec{a}$ and $\vec{b}$, which we write as $\vec{a} \times \vec{b}$.

Definition: If $\vec{u} = \langle u_1, u_2, u_3 \rangle$ and $\vec{v} = \langle v_1, v_2, v_3 \rangle$, then the cross product of $\vec{u}$ and $\vec{v}$ is

$$\vec{u} \times \vec{v} = \langle u_2v_3 - u_3v_2,\, u_3v_1 - u_1v_3,\, u_1v_2 - u_2v_1 \rangle.$$

Properties of the Cross Product: For vectors $\vec{u}, \vec{v}, \vec{w}$, and scalar c, the following hold:

1. $\vec{u} \times \vec{v} = -(\vec{v} \times \vec{u})$
2. $\vec{u} \times (\vec{v} + \vec{w}) = \vec{u} \times \vec{v} + \vec{u} \times \vec{w}$
3. $c(\vec{u} \times \vec{v}) = (c\vec{u}) \times \vec{v} = \vec{u} \times (c\vec{v})$
4. $\vec{u} \times \vec{0} = \vec{0} \times \vec{u} = \vec{0}$
5. $\vec{u} \times \vec{u} = \vec{0}$
6. $\vec{u} \cdot (\vec{v} \times \vec{w}) = (\vec{u} \times \vec{v}) \cdot \vec{w}$
7. $\vec{u} \times (\vec{v} \times \vec{w}) = (\vec{u} \cdot \vec{w})\vec{v} - (\vec{u} \cdot \vec{v})\vec{w}$

Notice that the cross product produces a vector, unlike the dot product which produces a scalar. Thus, the cross product can only take vectors in $\mathbb{R}^3$.

The *right-hand rule* gives the direction of the cross product:

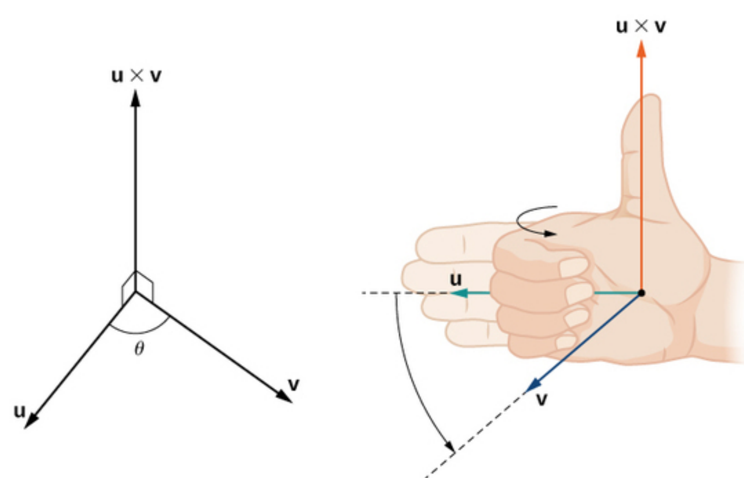


Image credit: Strang & Herman

Point your fingers in the direction of $\vec{u}$ and then curl your fingers in the direction of $\vec{v}$. Your thumb now points in the direction of the cross product.

The cross products of the standard unit vectors $\mathbf{i}, \mathbf{j}$, and $\mathbf{k}$ follow the rule $\mathbf{i} \times \mathbf{i} = \mathbf{j} \times \mathbf{j} = \mathbf{k} \times \mathbf{k} = \vec{0}$. Additionally, they have some important properties:

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

EXAMPLE 2.20

Use the properties of the cross product to compute $(\mathbf{i} \times \mathbf{k}) \times (\mathbf{k} \times \mathbf{j})$, and then multiply the result by $\mathbf{k}$.

Solution:

First, we compute

$$\mathbf{i} \times \mathbf{k} = -\mathbf{j}.$$

Next,

$$\mathbf{k} \times \mathbf{j} = -\mathbf{i}.$$

Now we compute the cross product

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

This simplifies to

$$\mathbf{j} \times \mathbf{i} = -\mathbf{k}.$$

Now the result multiply by $\mathbf{k}$:

$$-\mathbf{k} \times \mathbf{k} = -\vec{0} = \vec{0}$$

EXAMPLE 2.21

Find $\vec{p}\times\vec{q}$ for $\vec{p} = \langle 5, 1, 2\rangle$ and $\vec{q} = \langle -2, 0, 1\rangle$. Express the answer using standard unit vectors.

Solution:

We have

$$p_1 = 5,\ p_2 = 1,\ p_3 = 2,\ q_1 = -2,\ q_2 = 0, \text{and } q_3 = 1.$$

Now compute each component. The first component is

$$p_2q_3 - p_3q_2 = 1 \cdot 1 - 2 \cdot 0 = 1.$$

The second component is

$$p_3q_1 - p_1q_3 = 2 \cdot (-2) - 5 \cdot 1 = -4 - 5 = -9.$$

The third component is

$$p_1q_2 - p_2q_1 = 5 \cdot 0 - 1 \cdot (-2) = 0 + 2 = 2.$$

Thus,

$$\vec{p} \times \vec{q} = \langle 1, -9, 2\rangle.$$

And finally, expressed in terms of standard unit vectors, we have

$$\vec{p} \times \vec{q} = \mathbf{i} - 9\mathbf{j} + 2\mathbf{k}.$$

The cross product of two *standard unit vectors* is not only equal but also parallel to the third. They point in the same direction:

- $\mathbf{i} \times \mathbf{j} \parallel \mathbf{k}$
- $\mathbf{i} \times \mathbf{k} \parallel \mathbf{j}$
- $\mathbf{j} \times \mathbf{k} \parallel \mathbf{i}$

On the other hand, two nonzero vectors $\vec{a}$ and $\vec{b}$ are parallel if and only if $\vec{a} \times \vec{b} = \vec{0}$.

Let $\vec{a} = \langle a_1, a_2, a_3 \rangle$ and $\vec{b} = \langle b_1, b_2, b_3 \rangle$ be vectors with angle θ between them. Then,

$$\begin{gathered}\|\vec{a} \times \vec{b}\|^2 = (a_2b_3 - a_3b_2)^2 + (a_3b_1 - a_1b_3)^2 + (a_1b_2 - a_2b_1)^2. \\ = a_2^2b_3^2 - 2a_2a_3b_2b_3 + a_3^2b_2^2 + a_3^2b_1^2 - 2a_1a_3b_1b_3 + a_1^2b_3^2 + a_1^2b_2^2 - 2a_1a_2b_1b_2 + a_2^2b_1^2. \\ = a_1^2(b_2^2 + b_3^2) + a_2^2(b_1^2 + b_3^2) + a_3^2(b_1^2 + b_2^2) - 2a_1a_2b_1b_2 - 2a_1a_3b_1b_3 - 2a_2a_3b_2b_3.\end{gathered}$$

This simplifies to:

$$\|\vec{a} \times \vec{b}\|^2 = \|\vec{a}\|^2\|\vec{b}\|^2 - (\vec{a} \cdot \vec{b})^2.$$

Thus,

$$\begin{gathered}\|\vec{a} \times \vec{b}\|^2 = \|\vec{a}\|^2\|\vec{b}\|^2 - \|\vec{a}\|^2\|\vec{b}\|^2 \cos^2\theta, \\ = \|\vec{a}\|^2\|\vec{b}\|^2(1 - \cos^2\theta), \\ = \|\vec{a}\|^2\|\vec{b}\|^2 \sin^2\theta.\end{gathered}$$

Because $\sqrt{\sin^2\theta} = \sin\theta$ for $0 \leq \theta \leq 180°$, we obtain

$$\|\vec{a} \times \vec{b}\| = \|\vec{a}\|\,\|\vec{b}\| \sin\theta.$$

This gives the **magnitude of the cross product**.

We can express the cross product using the far easier determinant notation:

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

This determinant expands as:

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

A **determinant of order 2** is defined by

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

For example,

$$\begin{vmatrix} -2 & 1 \\ 4 & -6 \end{vmatrix} = (-2)(-6) - (1)(4) = 12 - 4 = 8.$$

A **determinant of order 3** can be defined in terms of second-order determinants as follows:

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

EXAMPLE 2.22

Use the determinant formula to show that $\vec{a} \times \vec{b}$ is orthogonal to $\vec{a}$.

Solution:

We compute the dot product:

$$(\vec{a} \times \vec{b}) \cdot \vec{a} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \cdot \langle a_1, a_2, a_3 \rangle$$

Expanding the determinant

$$= (a_2b_3 - a_3b_2)a_1 - (a_1b_3 - a_3b_1)a_2 + (a_1b_2 - a_2b_1)a_3$$

Simplifying

$$= a_1a_2b_3 - a_1a_3b_2 - a_1a_2b_3 + a_2a_3b_1 + a_1a_3b_2 - a_2a_3b_1 = 0$$

Thus, $\vec{a} \times \vec{b}$ is orthogonal to $\vec{a}$.

EXAMPLE 2.23

Use determinant notation to compute $\vec{a} \times \vec{b}$, where $\vec{a} = \langle 8, 2, 3 \rangle$ and $\vec{b} = \langle -1, 0, 4 \rangle$.

Solution:

We set up the determinant by placing the standard unit vectors in the first row, the components of $\vec{a}$ in the second row, and the components of $\vec{b}$ in the third row:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 8 & 2 & 3 \\ -1 & 0 & 4 \end{vmatrix}$$

We expand the determinant:

$$= \begin{vmatrix} 2 & 3 \\ 0 & 4 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 8 & 3 \\ -1 & 4 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 8 & 2 \\ -1 & 0 \end{vmatrix} \mathbf{k}$$

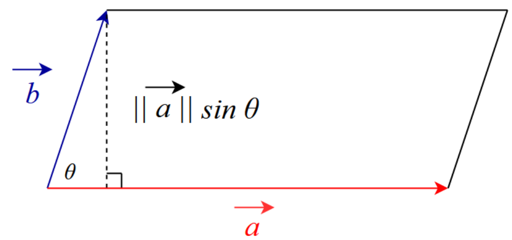
$$= (2 \cdot 4 - 3 \cdot 0)\mathbf{i} - (8 \cdot 4 - 3 \cdot (-1))\mathbf{j} + (8 \cdot 0 - 2 \cdot (-1))\mathbf{k}$$

$$= (8)\mathbf{i} - (32 + 3)\mathbf{j} + (2)\mathbf{k}$$

Thus,

$$\vec{a} \times \vec{b} = 8\mathbf{i} - 35\mathbf{j} + 2\mathbf{k}.$$

Vectors can also be used to find the area of a parallelogram:



Recall that the area A of a parallelogram is given by base $\times$ height. In this case, we have

$$A = \|\vec{a}\|(\|\vec{b}\| \sin\theta) = \|\vec{a} \times \vec{b}\|.$$

This shows that the magnitude of the cross product $\vec{a}\times\vec{b}$ is equivalent to the area of the parallelogram determined by $\vec{a}$ and $\vec{b}$. Furthermore, this means we can determine how "perpendicular" two vectors are from the area of a parallelogram. A parallelogram with large area (θ near $90°$) means vectors are nearly perpendicular and small area (θ near $0°$ or $180°$) means they are nearly parallel.

EXAMPLE 2.24

Find the area of the parallelogram $PQRS$ with vertices $P(1,1,0)$, $Q(7,1,0)$, $R(9,4,2)$, and $S(3,4,2)$.

Solution:

We first compute two adjacent vectors along the parallelogram:

$$\vec{PQ} = \langle 7-1,\ 1-1,\ 0-0\rangle = \langle 6,0,0\rangle$$

$$\vec{PS} = \langle 3-1,\ 4-1,\ 2-0\rangle = \langle 2,3,2\rangle$$

The area of the parallelogram is given by the magnitude of the cross product:

$$A = \|\vec{PQ} \times \vec{PS}\|$$

We compute the cross product:

$$\vec{PQ} \times \vec{PS} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 6 & 0 & 0 \\ 2 & 3 & 2 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 3 & 2 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 6 & 0 \\ 2 & 2 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 6 & 0 \\ 2 & 3 \end{vmatrix}$$

Calculating the minors

$$= (0\cdot 2 - 0\cdot 3)\mathbf{i} - (6\cdot 2 - 0\cdot 2)\mathbf{j} + (6\cdot 3 - 0\cdot 2)\mathbf{k} = 0\mathbf{i} - 12\mathbf{j} + 18\mathbf{k} = \langle 0,-12,18\rangle$$

The magnitude gives area:

$$A = \|\vec{PQ} \times \vec{PS}\| = \sqrt{0^2 + (-12)^2 + 18^2} = \sqrt{468} = 2\sqrt{117} = 6\sqrt{13} \text{ units squared}$$

The dot product of a vector with the cross product of two others vectors is called the triple scalar product. For vectors $\vec{a}, \vec{b}$, and $\vec{c}$, this would be $\vec{a} \cdot (\vec{b} \times \vec{c})$. As a determinant, the scalar triple product looks as follows:

$$\begin{aligned}
\vec{a} \cdot (\vec{b} \times \vec{c}) &= \langle a_1, a_2, a_3 \rangle \cdot \langle b_2c_3 - b_3c_2,\, b_3c_1 - b_1c_3,\, b_1c_2 - b_2c_1 \rangle \\
&= a_1(b_2c_3 - b_3c_2) + a_2(b_3c_1 - b_1c_3) + a_3(b_1c_2 - b_2c_1) \\
&= a_1b_2c_3 - a_1b_3c_2 - a_2b_1c_3 + a_2b_3c_1 + a_3b_1c_2 - a_3b_2c_1 \\
&= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}
\end{aligned}$$

EXAMPLE 2.25

Use the scalar triple product to show that the vectors $\vec{a} = \langle 1, 4, -7\rangle$, $\vec{b} = \langle 2, -1, 4\rangle$, and $\vec{c} = \langle 0, -9, 18\rangle$ are coplanar.

Solution:

We compute the scalar triple product as follows:

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

Let's expand the determinant:

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

And compute the minors:

$$= 1 \cdot ((-1)(18) - (4)(-9)) - 4 \cdot ((2)(18) - (4)(0)) + (-7) \cdot ((2)(-9) - (-1)(0)) = 0$$

The scalar triple product is zero, which means the vectors are coplanar.

The volume V of the parallelepiped (a three-dimensional prism with six faces that are parallelograms) determined by the vectors $\vec{a}, \vec{b}$, and $\vec{c}$ is given by $V = |\vec{a} \cdot (\vec{b} \times \vec{c})|$. The height h is given by the scalar projection of $\vec{a}$ onto $\vec{b} \times \vec{c}$:

$$h = \left\| \operatorname{proj}_{\vec{b} \times \vec{c}} \vec{a} \right\| = \left| \frac{\vec{a} \cdot (\vec{b} \times \vec{c})}{\|\vec{b} \times \vec{c}\|} \right|$$

If you want to check, multiplying this by height by the base $\|\vec{b} \times \vec{c}\|$ would get us the volume of the parallelepiped.

EXAMPLE 2.26

Find a vector orthogonal to the plane containing the points $P = (5, 2, -1)$, $Q = (-2, 4, 3)$, and $R = (1, -1, 2)$.

Solution:

The plane must contain the vectors $\overrightarrow{PQ}$ and $\overrightarrow{QR}$:

$$\overrightarrow{PQ} = \langle -2-5,\ 4-2,\ 3-(-1)\rangle = \langle -7, 2, 4\rangle$$

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

The cross product $\overrightarrow{PQ} \times \overrightarrow{QR}$ outputs a vector orthogonal to both $\overrightarrow{PQ}$ and $\overrightarrow{QR}$:

$$\overrightarrow{PQ} \times \overrightarrow{QR} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -7 & 2 & 4 \\ 3 & -5 & -1 \end{vmatrix}$$

$$= \mathbf{i}\begin{vmatrix} 2 & 4 \\ -5 & -1 \end{vmatrix} - \mathbf{j}\begin{vmatrix} -7 & 4 \\ 3 & -1 \end{vmatrix} + \mathbf{k}\begin{vmatrix} -7 & 2 \\ 3 & -5 \end{vmatrix}$$

$$= \mathbf{i}\cdot(2\cdot(-1) - 4\cdot(-5)) - \mathbf{j}\cdot((-7)\cdot(-1) - 4\cdot 3) + \mathbf{k}\cdot((-7)\cdot(-5) - 2\cdot 3)$$

$$= \mathbf{i}\cdot(-2+20) - \mathbf{j}\cdot(7-12) + \mathbf{k}\cdot(35-6)$$

$$= 18\mathbf{i} + 5\mathbf{j} + 29\mathbf{k}$$

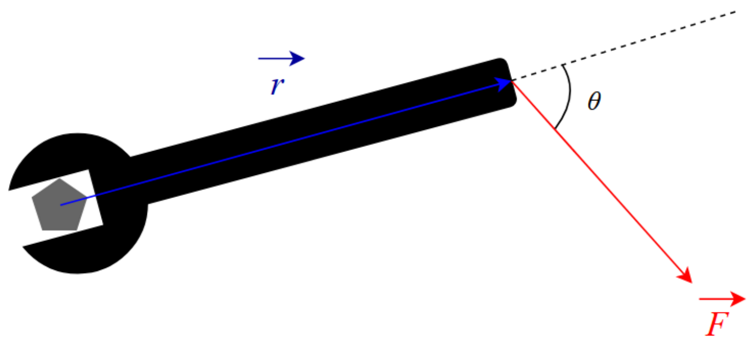
Thus, the vector $\langle 18, 5, 29\rangle$ is orthogonal to the plane containing the points P, Q, and R.

Torque $\vec{\tau}$ is the moment of force that causes rotation around an axis of rotation. Turning a screwdriver to tighten a screw or moving a door both create a torque. For a position vector $\vec{r}$ that starts on the axis of rotation and has a terminal point where the force is applied and an applied force vector $\vec{T}$, we have

$$\vec{\tau} = \vec{r} \times \vec{F}.$$

EXAMPLE 2.27

A bolt is being tightened by a force of $\|\vec{F}\| = 50$ N using a wrench with $\|\vec{r}\| = 0.25$ m. The angle between the wrench and the force vector $\vec{F}$ is $\theta = 75°$ as shown. Find the magnitude of the torque about the center of the bolt.



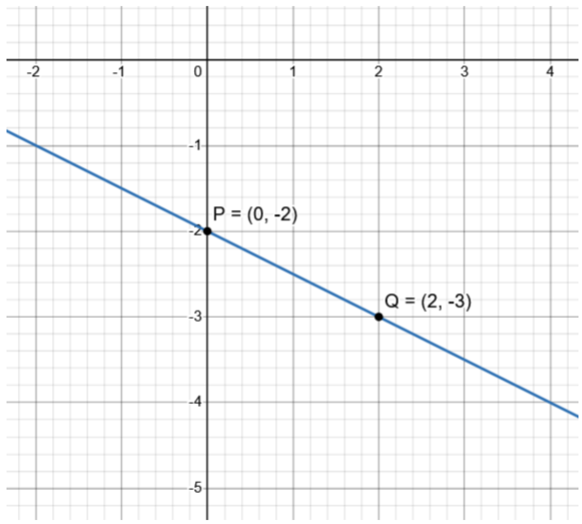
Solution:

We substitute in the givens:

$$\|\vec{\tau}\| = (0.25 \text{ m})(50 \text{ N}) \sin 75° = 12.07 \text{ Nm}$$

2.5 Lines and Planes

Let's say we wanted to find the equation of the following line:



We would simply find the slope $m = \frac{-3-(-2)}{2-0} = -\frac{1}{2}$. Then with a y-intercept of -2, we would have $y - (-2) = -\frac{1}{2}(x - 0) \Rightarrow y = -\frac{1}{2}x - 2$. Let's now write it as a *vector-valued function*. That is, in the form

$$\vec{r}(t) = \langle x, y \rangle.$$

Let's now substitute in our values:

$$\begin{aligned} &\langle x, -\frac{1}{2}x - 2 \rangle \\ &= \langle x, -\frac{1}{2}x \rangle + \langle 0, -2 \rangle \\ &= x\langle 1, -\frac{1}{2} \rangle + \langle 0, -2 \rangle \end{aligned}$$

More commonly, we would actually write this in the form $\vec{r}(t) = \langle x(t), y(t) \rangle$. We will let $x = 2t$ and we now have

$$t \underbrace{\langle 2, -1 \rangle}_{\vec{v}} + \underbrace{\langle 0, -2 \rangle}_{\vec{r}_0}.$$

This is in the form $\vec{r}(t) = t\vec{v} + \vec{r}_0$. We call this the **vector equation of a line**. We can rewrite this further:

$$\begin{aligned} \vec{r}(t) &= \langle x_0, y_0 \rangle + t\langle a, b \rangle \\ &= \langle x_0, y_0 \rangle + t\langle a, b \rangle \\ &= \langle x_0 + at, y_0 + bt \rangle \end{aligned}$$

Combining these gets us the **parametric equations of a line** in 2D:

$$\begin{aligned} x(t) &= x_0 + at \\ y(t) &= y_0 + bt \end{aligned}$$

The scalar equation for an ellipse in $\mathbb{R}^2$ centered at (h, k) with x-semiaxis length a and y-semiaxis length b is given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} = 1.$$

To convert this into the vector-valued equation for an ellipse in $\mathbb{R}^2$, we would simply rewrite it in the form $\vec{r}(t) = \langle x(t), y(t) \rangle$. We would have $\vec{r}(t) = \langle h + a\cos t, k + b\cos t \rangle$ with $0 \le t \le 2\pi$.

We can now extend this to lines in $\mathbb{R}^3$. Let's say we have a point on our line $\vec{r}_0 = \langle x_0, y_0, z_0 \rangle$ and a direction for the line $\vec{v} = \langle a, b, c \rangle$:

$$\begin{aligned} \vec{r}(t) &= \vec{r}_0 + t\vec{v} \\ &= \langle x_0, y_0, z_0 \rangle + t\langle a, b, c \rangle \\ &= \langle x_0 + at, y_0 + bt, z_0 + ct \rangle \\ &= \langle x(t), y(t), z(t) \rangle \end{aligned}$$

From this, we can now get the parametric equations of a line in $\mathbb{R}^3$:

$$\begin{aligned} x &= x_0 + at \\ y &= y_0 + bt \\ z &= z_0 + ct \end{aligned}$$

For nonzero a, b, and c, we can solve for t to get the *symmetric equations of a line*:

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

EXAMPLE 2.28

Find the parametric and symmetric equations for the line in $\mathbb{R}^3$ passing through the points $P_0(2, 4, -3)$ and $P_1(3, -1, 1)$.

Solution:

We first find the direction vector $\vec{v}$ by subtracting P_0 from P_1:

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

Thus, the vector equation of the line is

$$\vec{r}(t) = \langle 2, 4, -3\rangle + t\langle 1, -5, 4\rangle = \langle 2 + t, 4 - 5t, -3 + 4t\rangle.$$

This gives the following parametric equations:

$$x(t) = 2 + t,\ \ y(t) = 4 - 5t,\ \ z(t) = -3 + 4t$$

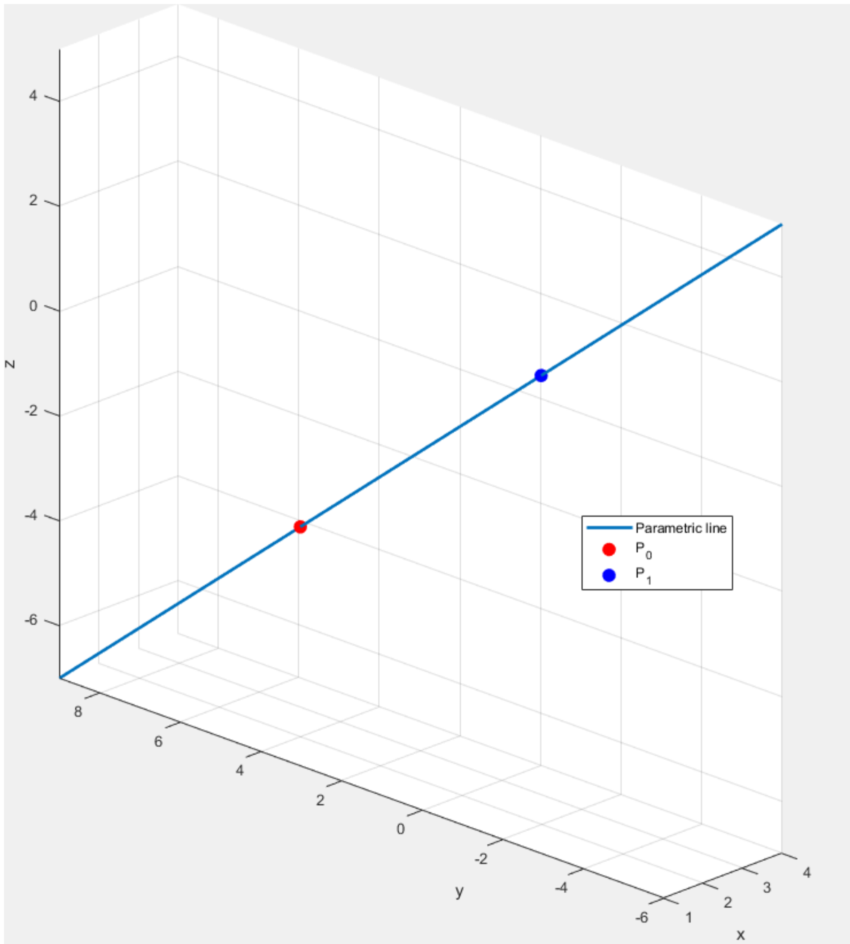
Now, we solve for t in each equation to get the symmetric forms:

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

Therefore, the symmetric equation of the line is

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

We can also graph **EXAMPLE** 2.28:



EXAMPLE 2.29

Find parametric and symmetric equations of the line passing through the points $P_0 = (1, 4, -2)$ and $P_1 = (-3, 5, 0)$.

Solution:

First, find the direction vector by subtracting the position vectors of the points. In other words, find a vector parallel to the line:

$$\vec{v} = \overrightarrow{P_0P_1} = \langle -3 - 1,\, 5 - 4,\, 0 - (-2) \rangle = \langle -4, 1, 2 \rangle$$

We now write the parametric equations of the line using $P_0 = (1, 4, -2)$ as the initial point:

$$x = 1 - 4t,\;\; y = 4 + t,\;\; z = -2 + 2t$$

Solving each equation for t, we obtain the symmetric equation of the line:

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

Sometimes, we don't want the equation of an entire line. The solution to this is to use the equation of only a **line segment**. To do this, we simply restrict the parameter t.

Let points $P = (x_0, y_0, z_0)$ and $Q = (x_1, y_1, z_1)$ lie on a line. Their position vectors are

$$\vec{p} = \langle x_0, y_0, z_0 \rangle,\;\; \vec{q} = \langle x_1, y_1, z_1 \rangle.$$

The vector equation of the line passing through P and Q is

$$\begin{aligned}
\vec{r}(t) &= \vec{p} + t\,\overrightarrow{PQ}, \\
&= \langle x_0, y_0, z_0 \rangle + t\langle x_1 - x_0,\, y_1 - y_0,\, z_1 - z_0 \rangle \\
&= \langle x_0, y_0, z_0 \rangle + t\left(\langle x_1, y_1, z_1 \rangle - \langle x_0, y_0, z_0 \rangle\right) \\
&= \langle x_0, y_0, z_0 \rangle + t\langle x_1, y_1, z_1 \rangle - t\langle x_0, y_0, z_0 \rangle \\
&= (1 - t)\langle x_0, y_0, z_0 \rangle + t\langle x_1, y_1, z_1 \rangle \\
&= (1 - t)\vec{p} + t\vec{q}.
\end{aligned}$$

Thus, the vector equation of the line segment from P to Q is:

$$\vec{r}(t) = (1-t)\vec{p} + t\vec{q},\ 0 \le t \le 1.$$

When $t = 0$, we are at $P = \vec{r}(0) = \vec{p}$. When $t = 1$, we are at $Q = \vec{r}(1) = \vec{q}$.

If our domain were $(-\infty, \infty)$, we would have an infinitely long line. With the domain restriction, the equation smoothly traces the line segment from P to Q as t moves from 0 to 1.

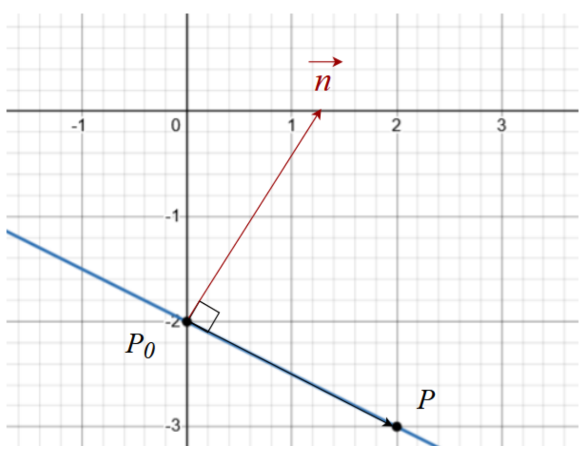
We can also find the parametric equations of a line segment:

$$\begin{aligned} \vec{r} &= \vec{p} + t\,\overrightarrow{PQ} \\ \langle x, y, z\rangle &= \langle x_0, y_0, z_0\rangle + t\langle x_1 - x_0,\ y_1 - y_0,\ z_1 - z_0\rangle \\ &= \langle x_0 + t(x_1 - x_0),\ y_0 + t(y_1 - y_0),\ z_0 + t(z_1 - z_0)\rangle \end{aligned}$$

Then, we have

$$\begin{aligned} x &= x_0 + t(x_1 - x_0); \\ y &= y_0 + t(y_1 - y_0); \\ z &= z_0 + t(z_1 - z_0), \quad 0 \le t \le 1. \end{aligned}$$

Previously, we started with the explicit equation for a 2D line $y = -\frac{1}{2}x - 2$ and then found the vector equation for the line $\vec{r}(t) = t\langle 2, -1\rangle + \langle 0, -2\rangle = \langle -2-t, 2t\rangle$. Here, our slope $\vec{v} = \langle 2, -1\rangle$ corresponds to a direction vector that defines the direction orthogonal to the vector between P_0 and P.



Let's define the line as all points $P(x, y)$ such that vector $\vec{x} = \overrightarrow{P_0P}$ from $P_0 = (x_0, y_0)$ to $P(x, y)$ is orthogonal to $\vec{n} = \langle b, -a\rangle$. That is,

$$\vec{n} \cdot \vec{x} = 0$$

And now we substitute:

$$\vec{n} \cdot \langle x - x_0, y - y_0\rangle = 0$$

Because $\vec{v} = \langle a, b\rangle$, we know that $\vec{n} = \langle b, -a\rangle$. Continuing on,

$$\begin{aligned}
b(x - x_0) - a(y - y_0) &= 0, \\
bx - bx_0 - ay + ay_0 &= 0, \\
bx - ay &= bx_0 - ay_0.
\end{aligned}$$

This can be written as $Ax + By = -C$ where $A = b$, $B = -a$, and $C = bx_0 - ay_0$.

Let's now extrapolate and find the equations of planes. We know that a line in space is determined by a point and a direction, but a plane is determined by a point $P_0(x_0, y_0, z_0)$ and a normal vector $\vec{n} = \langle a, b, c\rangle$ that is orthogonal to the plane. More specifically, the plane is the collection of all points $P_0(x, y, z) \in \mathbb{R}^3$ so that $\vec{x} = \overrightarrow{P_0P}$ is orthogonal to $\vec{n}$. That is, $\vec{n} \cdot \vec{x} = 0$. Substituting $\vec{x} = \overrightarrow{P_0P} = \langle x - x_0, y - y_0, z - z_0\rangle$, we can get the *dot product equation for a plane*:

$$\vec{n} \cdot \overrightarrow{P_0P} = 0$$

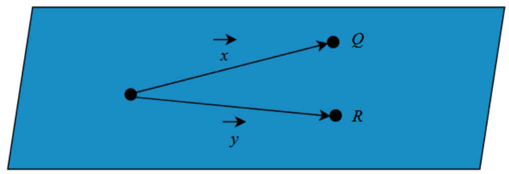
From this, we substitute in to get $\langle a, b, c\rangle \cdot \langle x - x_0, y - y_0, z - z_0\rangle = 0$. Continuing on,

$$\begin{gathered}
a \cdot (x - x_0) + b \cdot (y - y_0) + c \cdot (z - z_0) = 0 \\
ax + by + cz = d, \text{ where } d = ax_0 + by_0 + cz_0
\end{gathered}$$

Both of these are known as the scalar equation of the plane. The second equation is linear in x, y, and z.

EXAMPLE 2.30

Find the equation of the plane passing through the points $P_0(2,-1,3), Q(1,4,0)$, and $R(0,-1,5)$.



Solution:

First, we compute two vectors lying on the plane:

$$\overrightarrow{P_0Q} = \langle 1-2,\ 4-(-1),\ 0-3\rangle = \langle -1,5,-3\rangle$$
$$\overrightarrow{P_0R} = \langle 0-2,\ -1-(-1),\ 5-3\rangle = \langle -2,0,2\rangle$$

We take the cross product:

$$\vec{n} = \overrightarrow{P_0Q} \times \overrightarrow{P_0R} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -1 & 5 & -3 \\ -2 & 0 & 2 \end{vmatrix} = \mathbf{i}\begin{vmatrix} 5 & -3 \\ 0 & 2 \end{vmatrix} - \mathbf{j}\begin{vmatrix} -1 & -3 \\ -2 & 2 \end{vmatrix} + \mathbf{k}\begin{vmatrix} -1 & 5 \\ -2 & 0 \end{vmatrix}.$$

Let's now compute:

$$\vec{n} = \mathbf{i}(5\cdot 2 - (-3)\cdot 0) - \mathbf{j}((-1)(2) - (-3)(-2)) + \mathbf{k}((-1)(0) - 5\cdot(-2))$$
$$\vec{n} = \mathbf{i}(10) - \mathbf{j}(-2-6) + \mathbf{k}(0+10) = \langle 10,8,10\rangle.$$

Thus, the normal vector is $\vec{n} = \langle 10,8,10\rangle$. We can substitute point $P_0(2,-1,3)$ into the equation of the plane:

$$10(x-2) + 8(y-(-1)) + 10(z-3) = 10x + 8y + 10z = 42 = 0$$
$$5x + 4y + 5z = 21$$

Two planes are parallel if their normal vectors are parallel. For instance, the planes $x + 2y - 3z = 4$ and $2x + 4y - 6z = 3$ have normal vectors $\vec{n}_1 = \langle 1, 2, -3\rangle$ and $\vec{n}_2 = \langle 2, 4, -6\rangle$. Since $\vec{n}_2 = 2\vec{n}_1$, one is a scalar multiple of the other and they both point in the same direction. Thus, the planes are parallel. If two planes are not parallel, then they intersect in a straight line:

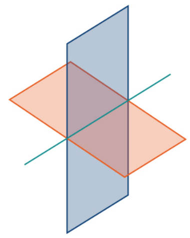


Image credit: Strang & Herman

The angle between the two planes is the same as the acute angle between their normal vectors:

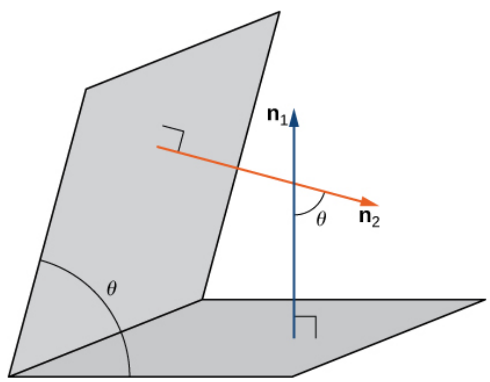


Image credit: Strang & Herman

We can find this angle θ using the following equation:

$$\cos\theta = \frac{|\vec{n}_1 \cdot \vec{n}_2|}{\|\vec{n}_1\|\|\vec{n}_2\|}$$

EXAMPLE 2.31

Find the angle between the planes $x + y + z = 1$ and $x - 2y + 3z = 1$. Then, find symmetric equations for the line of intersection between them.

Solution:

The angle between two planes is equivalent to the angle between their normal vectors. Let $\vec{n}_1 = \langle 1, 1, 1 \rangle$ and $\vec{n}_2 = \langle 1, -2, 3 \rangle$. The angle θ between the planes is given by:

$$\cos\theta = \frac{\vec{n}_1 \cdot \vec{n}_2}{\|\vec{n}_1\|\|\vec{n}_2\|} = \frac{1(1) + 1(-2) + 1(3)}{\sqrt{1^2 + 1^2 + 1^2} \cdot \sqrt{1^2 + (-2)^2 + 3^2}} = \frac{1 - 2 + 3}{\sqrt{3} \cdot \sqrt{14}} = \frac{2}{\sqrt{42}}$$

So the angle is:

$$\theta = \cos^{-1}\left(\frac{2}{\sqrt{42}}\right)$$

To find the line of intersection, we need a direction vector $\vec{d}$ that lies on both planes. This is given by the cross product of their normals:

$$\vec{d} = \vec{n}_1 \times \vec{n}_2 = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 1 & 1 \\ 1 & -2 & 3 \end{vmatrix} = \mathbf{i}(1 \cdot 3 - 1 \cdot (-2)) - \mathbf{j}(1 \cdot 3 - 1 \cdot 1) + \mathbf{k}(1 \cdot (-2) - 1 \cdot 1)$$

$$= \langle 5, -2, -3 \rangle$$

To find a point on the line, set $z = 0$ and solve the system

$$x + y = 1, \; x - 2y = 1.$$

We get $y = 0$ and $x = 1$. Thus, a point on the line is $(1, 0, 0)$ and the symmetric equations are

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

If we need to find the distance between two parallel planes, we simply locate a point on one and find the distance between it and the other plane. Suppose a plane with normal vector $\vec{n}$ passes through point Q. The distance d between the plane and another point P is given by

$$d = \left\|\text{proj}_{\vec{n}}\, \overrightarrow{QP}\right\| = \left|\text{comp}_{\vec{n}}\, \overrightarrow{QP}\right| = \frac{|\overrightarrow{QP} \cdot \vec{n}|}{\|\vec{n}\|}$$

The distance from the point (x_0, y_0, z_0) to the plane $ax + by + cz + k = 0$, where $\vec{n} = \langle a, b, c\rangle$ is the normal vector and $Q = (x_1, y_1, z_1)$ is any point on the plane. Substituting into the formula yields

$$d = \frac{|a(x_0 - x_1) + b(y_0 - y_1) + c(z_0 - z_1)|}{\sqrt{a^2 + b^2 + c^2}} = \frac{|ax_0 + by_0 + cz_0 + k|}{\sqrt{a^2 + b^2 + c^2}}.$$

EXAMPLE 2.32

Find the distance between the parallel planes $5x - 2y + z = 6$ and $5x - 2y + z = -3$.

Solution:

Since the planes are parallel, we know their normal vector is shared: $\vec{n} = \langle 5, -2, 1\rangle$. We pick a point on one of the planes. I will go with a point on the second plane. Then we set $x = 0$, $y = 0$, and solve for z:

$$5(0) - 2(0) + z = -3 \Rightarrow z = -3$$

So we can now use point $P(0, 0, -3)$ and compute the distance to the other plane:

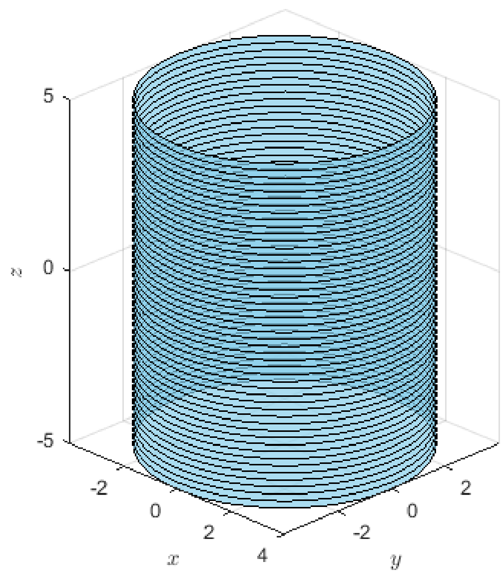
$$d = \frac{|5(0) - 2(0) + (-3) - 6|}{\sqrt{5^2 + (-2)^2 + 1^2}} = \frac{|-9|}{\sqrt{25 + 4 + 1}} = \frac{9}{\sqrt{30}} = \frac{3\sqrt{30}}{10} \text{ units}$$

2.6 Quadric Surfaces

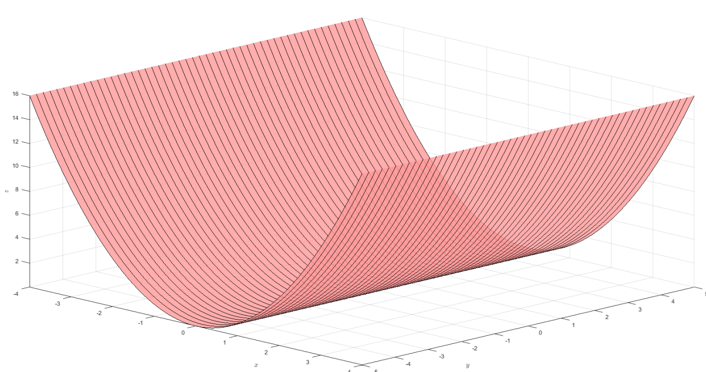
Planes and spheres are specific cases of three-dimensional figures called *surfaces*. We will continue our exploration of surfaces with **cylinders**. We can define a cylinder as a surface that contains all lines parallel to a given line that pass through a given plane curve. These parallel lines are called *rulings*.

This means that any curve can form a cylinder. Cylindrical surfaces do not have to be circular.

The equation $x^2 + y^2 = 16$ describes a circle in $\mathbb{R}^2$ centered at the origin with radius 4. In $\mathbb{R}^3$, this would represent a surface. If we stack many of them on top of each other, we'd have a cylinder. In other words, we can extend a curve along an axis (or a straight line) to form a cylinder:



In similar fashion, *parabolic cylinders* are created by extending many parabolas along a straight line. Here is the graph of $z = x^2$:



Quadric surfaces are described using quadratic equations and are generalized in the form

$$Ax^2 + By^2 + Cz^2 + Dxy + Exz + Fyz + Gx + Hy + Jz + K = 0,$$

where $A, B, C, ..., J, K$ are nonlinear real number coefficients. Through translation and rotation, we can come up with two standard forms:

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

When surfaces intersect a plane parallel to one of the coordinate planes, the cross sections created are known as *traces*. When a quadric surface intersects a coordinate plane, its trace is a conic section.

We begin our study of quadric surfaces with **ellipsoids**. All of the traces of an ellipsoid are ellipses. Ellipsoids are given by

$$\frac{(x-h)^2}{a^2} + \frac{(y-k)^2}{b^2} + \frac{(z-m)^2}{c^2} = 1.$$

If we assume that the center (h, k, m) is $(0, 0, 0)$, we get

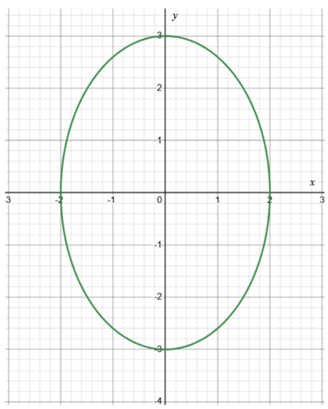
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z}{c^2} = 1.$$

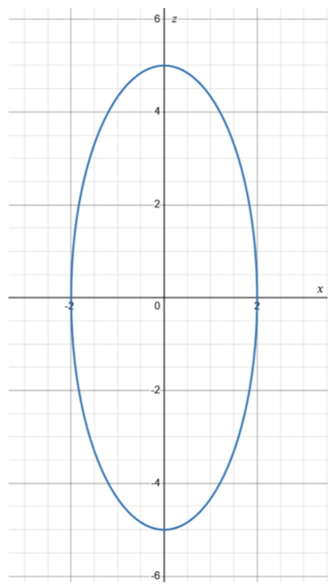
EXAMPLE 2.33

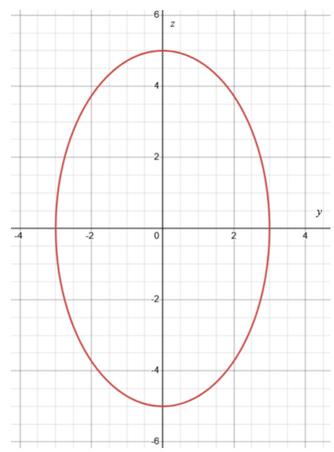
Graph the ellipsoid $\frac{x^2}{2^2} + \frac{y^2}{3^2} + \frac{z^2}{5^2} = 1$ by first finding the traces. Assume $k = 0$.

Solution:

We can find the traces by setting $z = 0$ for the trace in the xy-plane, $y = 0$ for the trace in the xz-plane, and $x = 0$ for the trace in the yz-plane. The results are shown below, respectively:

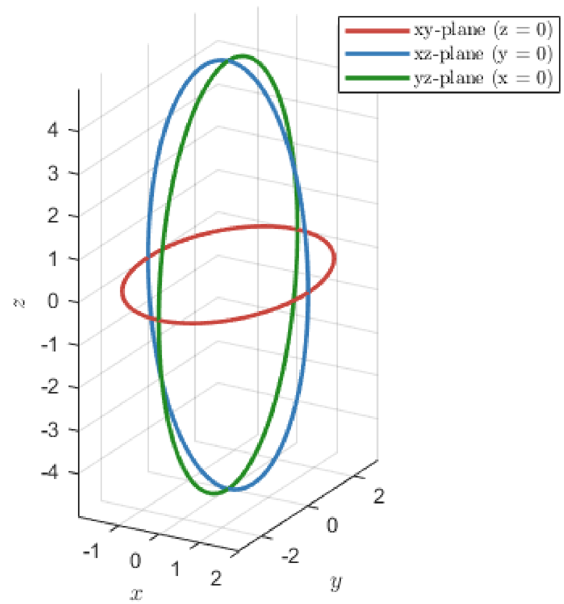


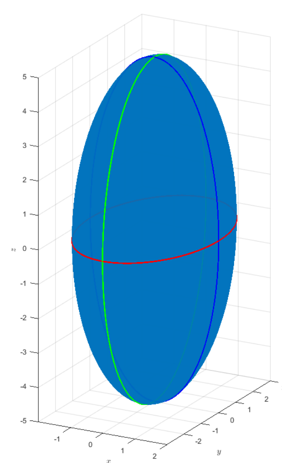




EXAMPLE 2.33 (CONTINUED)

Let's now sketch these traces in 3D:





If a quadric surface has elliptical traces in the xy-plane, but parabolic traces in the xz-plane and yz-plane, it is called an **elliptic paraboloid**. The equation is of the form

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{x}{c}$$

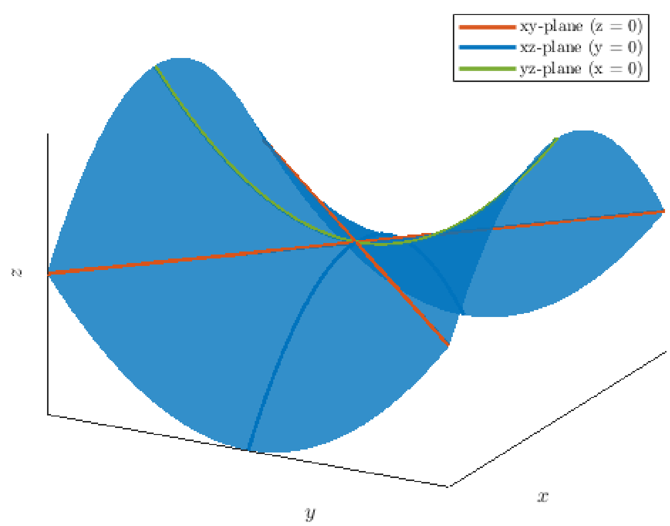
EXAMPLE 2.34

Determine what surface $z = y^2 - x^2$ is by first finding its traces.

Solution:

Generally, the traces in the xy-plane are found by setting $z = k$ which gives us $k = y^2 - x^2$. This is a family of hyperbolas. The traces in the xz-plane are found by setting $y = k$ which gives us $z = k^2 - x^2$. These are parabolas that open downward. The traces in the yz-plane are found by setting $x = k$ which gives us $z = y^2 - k^2$. These are parabolas that open upward.

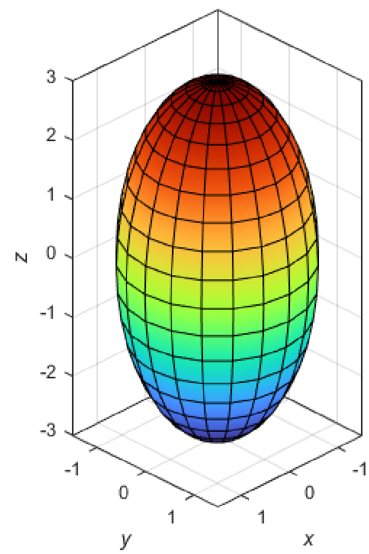
We can now set $k = 0$ in each case to get $y^2 = x^2$ in the xy-plane, $z = -x^2$ in the xz-plane, and $z = y^2$ in the yz-plane. Here is the graph:



This saddle-like shape means it is a **hyperbolic paraboloid**.

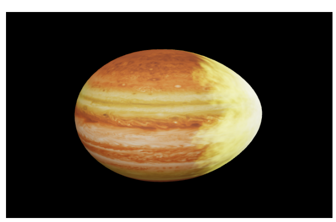
Seventeen quadric surfaces can be derived from the general equation, but you only need to know the most common six. Let's go through a summary of each.

Ellipsoid:



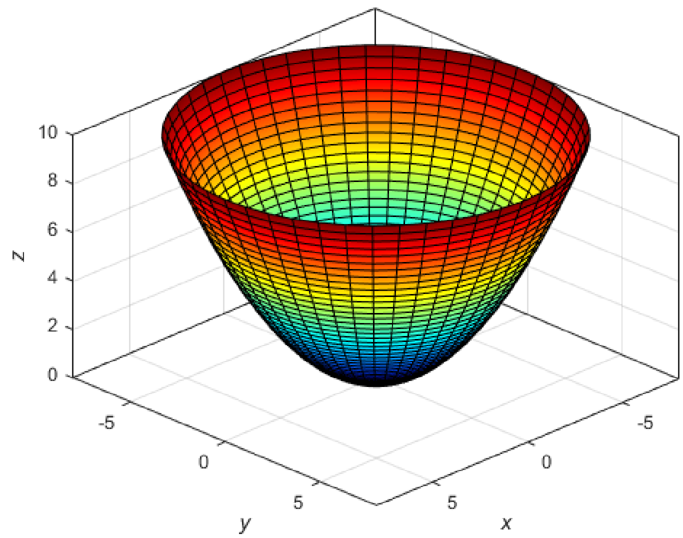
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

All traces are ellipses.
If $a = b = c$, the ellipsoid is a sphere.
Real-world example: Certain planets



Exoplanet WASP-12 b. Image credit: NASA

Elliptic Paraboloid:

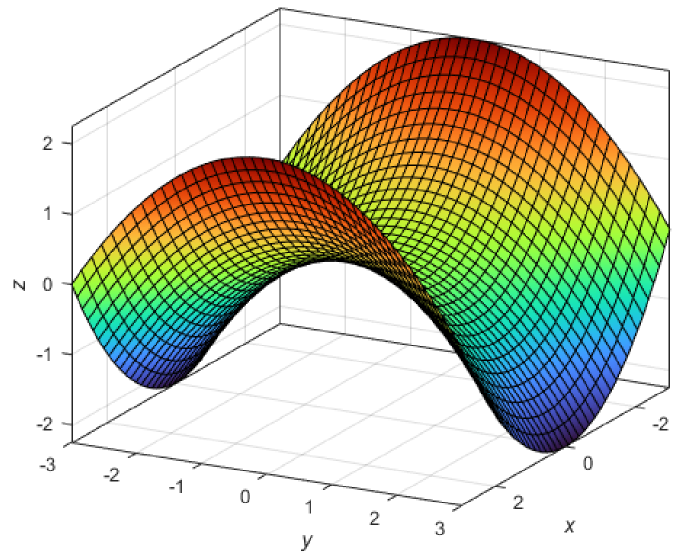


$$\frac{z}{c} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

Horizontal traces are ellipses. Vertical traces are parabolas.
The variable raised to the first power indicates the axis of the paraboloid.
Real-world example: Satellite dishes

Deep Space Station 53 Antenna in Madrid. Image credit: NASA

Hyperbolic Paraboloid:

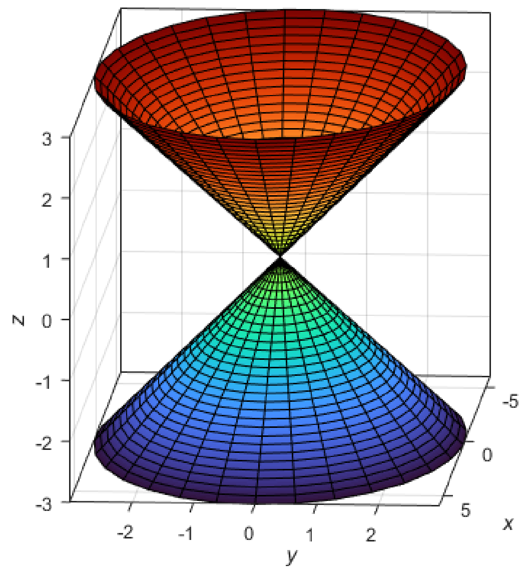


$$\frac{z}{c} = \frac{x^2}{a^2} - \frac{y^2}{b^2}$$

Horizontal traces are hyperbolas. Vertical traces are parabolas.
The axis of the surface corresponds to the linear variable.
Real-world example: Certain roofs in modern architecture

Olympic Stadium in Munich. Image credit: Olympiapark München

Cone (elliptic cone):

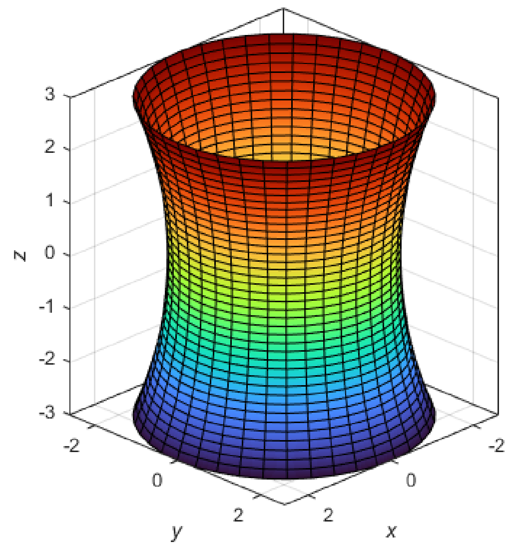


$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 0$$

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 $k = 0$. The traces in the coordinate planes parallel to the axis are intersecting lines.
The axis of the surface corresponds to the variable with a negative coefficient.
Real world-example: Volcanoes

Mount Shishaldin in Alaska. Image credit: National Geographic

Hyperboloid of One Sheet:

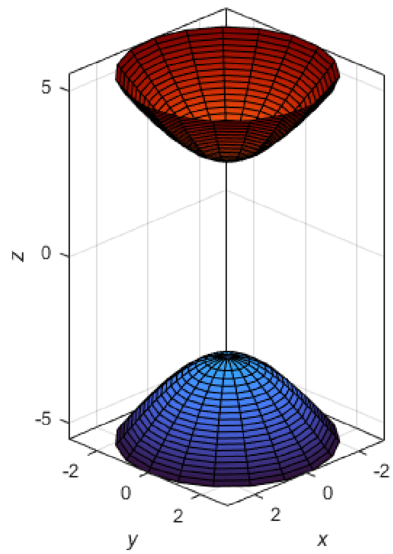


$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

Horizontal traces are ellipses. Vertical traces are hyperbolas.
The axis of symmetry corresponds to the variable whose coefficient is negative.
Real-world example: Nuclear cooling towers

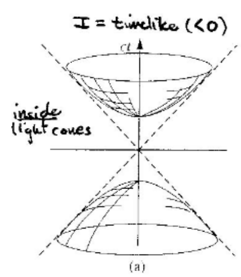
Nine Mile Point in New York. Image credit: Constellation Energy

Hyperboloid of Two Sheets:



$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Horizontal traces in $z = k$ are ellipses if $k > c$ or $k < -c$. Vertical traces are hyperbolas.
The axis of the surface corresponds to the variable with a positive coefficient.
The two negative terms indicate two sheets.
Real-world example: Can sometimes appear in the geometry of spacetime



Minkowski space diagram showing timelike separation inside a light cone. Image credit: Errerde, UIUC.

EXAMPLE 2.35

Identify the surface represented by the equation $9x^2 + y^2 - z^2 + 2z - 10 = 0$.

Solution:

We begin by completing the square on the z-terms and rewriting the expression in standard form:

$$9x^2 + y^2 - z^2 + 2z = 10$$

$$9x^2 + y^2 - \left(z^2 - 2z\right) = 10$$

$$9x^2 + y^2 - \left[(z-1)^2 - 1\right] = 10$$

$$9x^2 + y^2 - (z-1)^2 + 1 = 10$$

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

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

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

This is the equation of a hyperboloid of one sheet centered at $(0, 0, 1)$ with an axis of symmetry along the z-axis.

3 Vector-Valued Functions

Now that we have a foundation in the properties of vectors and their geometric interpretations, we will extend our focus to studying curves in planes and three-dimensional space.

3.1 Limits of Vector-Valued Functions

A **vector-valued function**, or **vector function**, is simply a function $\vec{r}(t)$ whose domain is a set of real numbers and whose range is a set of vectors. We will focus on vector functions whose values are in $\mathbb{R}^3$.

Let $f(t)$, $g(t)$, and $h(t)$ be the real-valued component functions of $\vec{r}(t)$. Then, the general form of a vector-valued function is

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

EXAMPLE 3.1

For the vector-valued function $\vec{r}(t) = (t^2 - 3t)\,\mathbf{i} + (4t + 1)\,\mathbf{j}$, evaluate $\vec{r}(0)$, $\vec{r}(1)$, and $\vec{r}(-4)$. Determine if the function has any domain restrictions.

Solution:

We substitute each value of t into the expression for $\vec{r}(t)$:

$$\vec{r}(0) = (0^2 - 3 \cdot 0)\,\mathbf{i} + (4 \cdot 0 + 1)\,\mathbf{j} = 0\,\mathbf{i} + 1\,\mathbf{j} = \mathbf{j}$$

$$\vec{r}(1) = (1^2 - 3 \cdot 1)\,\mathbf{i} + (4 \cdot 1 + 1)\,\mathbf{j} = (-2)\,\mathbf{i} + 5\,\mathbf{j}$$

$$\vec{r}(-4) = ((-4)^2 - 3 \cdot (-4))\,\mathbf{i} + (4 \cdot (-4) + 1)\,\mathbf{j} = (16 + 12)\,\mathbf{i} + (-16 + 1)\,\mathbf{j} = 28\,\mathbf{i} - 15\,\mathbf{j}$$

There are no domain restrictions because both components of $\vec{r}(t)$ are polynomial functions, which are defined for all real numbers.

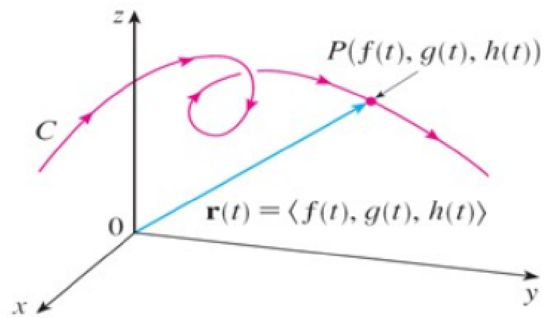
A vector is considered to be in *standard condition* if the initial point is located at the origin. We generally graph vectors in the domain of the function in standard position to guarantee the uniqueness of the graph. The graph of the vector function

$$\vec{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$$

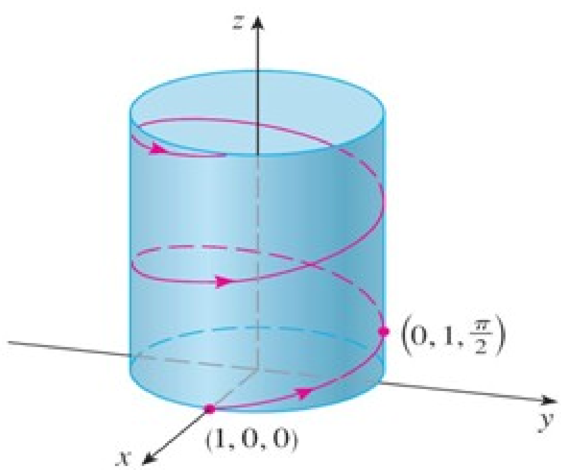
consists of the set of all points $(f(t), g(t))$ and is called a **plane curve**. The graph of the vector function

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

consists of the set of all points $(f(t), g(t), h(t))$ and is called a **space curve**. We refer to the vector function representation of plane curves and space curves as *vector parameterization* of a curve.



A curve being traced out by the moving position vector $\vec{r}(t)$. Image credit: Stewart



The projection of the curve onto the xy-plane is given by $\vec{r}(t) = \langle \cos t, \sin t, 0 \rangle$. The curve spirals upward around a cylinder. Image credit: Stewart

EXAMPLE 3.2

Graph the plane curve represented by $\vec{r}(t) = 4\cos(t^3)\,\mathbf{i} + 3\sin(t^3)\,\mathbf{j},\ 0 \le t \le \sqrt{2\pi}$.

Solution:

First, complete your table of values:

t	$\vec{r}(t)$	t	$\vec{r}(t)$
0	$4\,\mathbf{i}$	$\sqrt[3]{\pi}$	$-4\,\mathbf{i}$
$\sqrt[3]{\frac{\pi}{4}}$	$2\sqrt{2}\,\mathbf{i} + \frac{3\sqrt{2}}{2}\,\mathbf{j}$	$\sqrt[3]{\frac{5\pi}{4}}$	$-2\sqrt{2}\,\mathbf{i} - \frac{3\sqrt{2}}{2}\,\mathbf{j}$
$\sqrt[3]{\frac{\pi}{2}}$	$3\,\mathbf{j}$	$\sqrt[3]{\frac{3\pi}{2}}$	$-3\,\mathbf{j}$
$\sqrt[3]{\frac{3\pi}{4}}$	$-2\sqrt{2}\,\mathbf{i} + \frac{3\sqrt{2}}{2}\,\mathbf{j}$	$\sqrt[3]{\frac{7\pi}{4}}$	$2\sqrt{2}\,\mathbf{i} - \frac{3\sqrt{2}}{2}\,\mathbf{j}$
$\sqrt[3]{2\pi}$	$4\,\mathbf{i}$		

We can now graph:

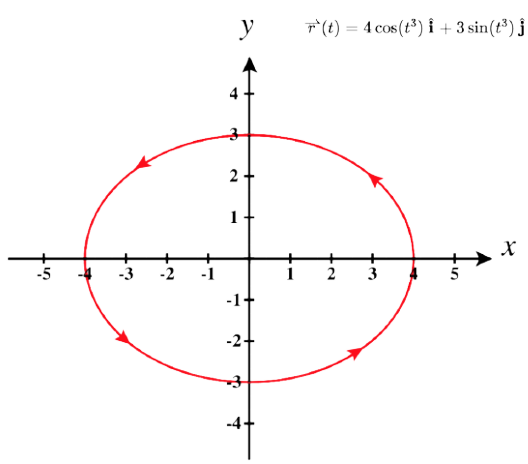


Image credit: Strang & Herman

EXAMPLE 3.3

Graph the space curve represented by $\vec{r}(t) = 4\cos(t)\,\mathbf{i} + 4\sin(t)\,\mathbf{j} + t\mathbf{k},\ 0 \le t \le 4\pi$.

Solution:

First, complete your table of values:

t	$\vec{r}(t)$	t	$\vec{r}(t)$
0	$4\,\mathbf{i}$	π	$-4\,\mathbf{i} + \pi\,\mathbf{k}$
$\frac{\pi}{4}$	$2\sqrt{2}\,\mathbf{i} + 2\sqrt{2}\,\mathbf{j} + \frac{\pi}{4}\,\mathbf{k}$	$\frac{5\pi}{4}$	$-2\sqrt{2}\,\mathbf{i} - 2\sqrt{2}\,\mathbf{j} + \frac{5\pi}{4}\,\mathbf{k}$
$\frac{\pi}{2}$	$4\,\mathbf{j} + \frac{\pi}{2}\,\mathbf{k}$	$\frac{3\pi}{2}$	$-4\,\mathbf{j} + \frac{3\pi}{2}\,\mathbf{k}$
$\frac{3\pi}{4}$	$-2\sqrt{2}\,\mathbf{i} + 2\sqrt{2}\,\mathbf{j} + \frac{3\pi}{4}\,\mathbf{k}$	$\frac{7\pi}{4}$	$2\sqrt{2}\,\mathbf{i} - 2\sqrt{2}\,\mathbf{j} + \frac{7\pi}{4}\,\mathbf{k}$
2π	$4\,\mathbf{i} + 2\pi\,\mathbf{k}$		

We can now graph:

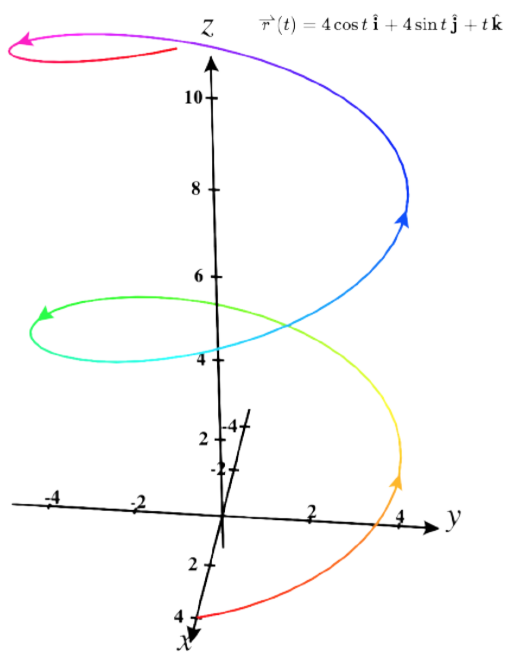


Image credit: Strang & Herman

Given a vector-valued function $\vec{r}(t) = f(t)\mathbf{i}, g(t)\mathbf{j}$, we can define $x = f(t)$, $y = g(t)$, and $z = h(t)$, which are parametric equations. In other words, a vector-valued function is very similar to a parametric equation. For a vector-valued function, you are often only interested in specific domains of the parameter. If you restrict the domain, you can essentially make a parametric curve and a vector-valued function trace out the same path. Points on the vector-valued graph simply represent the head of the vector that originates from the origin whereas parametric curves treat each point as a location. Thus, since we can parameterize a curve defined by a given function, we can also represent any plane curve as a vector-valued function.

The **limit** $\vec{L}$ of a vector-valued function $\vec{r}$ as it approaches a is written as

$$\lim_{t \to a} \vec{r}(t) = \vec{L},$$

provided

$$\lim_{t \to a} \left\| \vec{r}(t) - \vec{L} \right\| = 0.$$

More practically, for $\vec{r}(t) = \langle f(t), g(t), h(t) \rangle$, the limit is given by

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

provided that the limits of the component functions exist.

Similarly, the componentwise definition can be written as follows. The limit of the vector-valued function $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j} + h(t)\,\mathbf{k}$ as t approaches a is given by

$$\lim_{t \to a} \vec{r}(t) = \left[\lim_{t \to a} f(t)\right] \mathbf{i} + \left[\lim_{t \to a} g(t)\right] \mathbf{j} + \left[\lim_{t \to a} h(t)\right] \mathbf{k},$$

provided that all three limits exist.

EXAMPLE 3.4

Find $\lim\limits_{t\to 0} \vec{r}(t)$, where $\vec{r}(t) = (1+t^3)\,\mathbf{i} + te^{-t}\,\mathbf{j} + \frac{\sin t}{t}\,\mathbf{k}$.

Solution:

We evaluate each coordinate limit separately:

$$\lim_{t\to 0} \vec{r}(t) = \left[\lim_{t\to 0}(1+t^3)\right]\mathbf{i} + \left[\lim_{t\to 0}\left(te^{-t}\right)\right]\mathbf{j} + \left[\lim_{t\to 0}\left(\frac{\sin t}{t}\right)\right]\mathbf{k}$$

$$= 1\,\mathbf{i} + 0\,\mathbf{j} + 1\,\mathbf{k} = \mathbf{i} + \mathbf{k}$$

EXAMPLE 3.5

Calculate $\lim\limits_{t\to 2} \vec{r}(t)$, where $\vec{r}(t) = \sqrt{t^2+3t-1}\,\mathbf{i} - (4t-3)\,\mathbf{j} - \sin\left(\frac{(t+1)\pi}{2}\right)\mathbf{k}$.

Solution:

We evaluate each coordinate limit separately:

$$\lim_{t\to 2} \vec{r}(t) = \left[\lim_{t\to 2}\sqrt{t^2+3t-1}\right]\mathbf{i} + \left[\lim_{t\to 2}(-4t+3)\right]\mathbf{j} + \left[\lim_{t\to 2}\left(-\sin\left(\frac{(t+1)\pi}{2}\right)\right)\right]\mathbf{k}$$

$$= \sqrt{4+6-1}\,\mathbf{i} + (-8+3)\,\mathbf{j} + \left(-\sin\left(\frac{3\pi}{2}\right)\right)\mathbf{k} = 3\,\mathbf{i} - 5\,\mathbf{j} + 1\,\mathbf{k}$$

Let f, g, and h be functions of t. Then, the vector-valued function

$$\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j}$$

is **continuous** at the point $t = a$ if the following three conditions hold:

1. $\vec{r}(a)$ exists,
2. $\lim_{t \to a} \vec{r}(t)$ exists,
3. $\lim_{t \to a} \vec{r}(t) = \vec{r}(a)$.

Similarly, the vector-valued function

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

is **continuous** at the point $t = a$ if the following three conditions hold:

1. $\vec{r}(a)$ exists,
2. $\lim_{t \to a} \vec{r}(t)$ exists,
3. $\lim_{t \to a} \vec{r}(t) = \vec{r}(a)$.

3.2 Derivatives and Integrals of Vector-Valued Functions

Let's say we have the vector-valued function $\vec{r}(t)$. The average rate of change between t_1 and t_2 is given by

$$\frac{\vec{r}(t_2) - \vec{r}(t_1)}{t_2 - t_1}.$$

This is known as the *secant vector*:

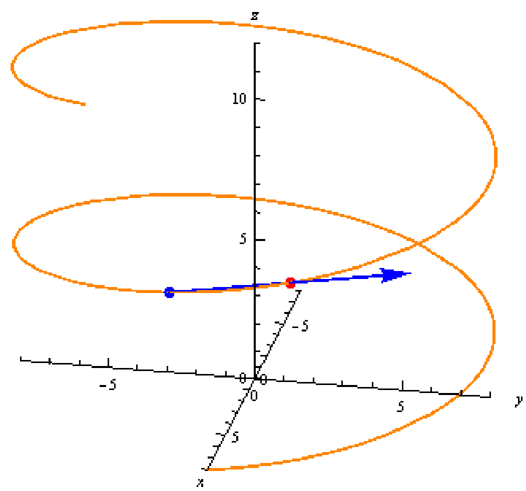


Image credit: UMich

The average rate of change will approach the derivative as t_2 gets closer to t_1 and thus becomes a better estimate at that point. The derivative is a vector that is always tangent to the curve at t_1. It gives us the instantaneous rate of change.

The derivative of a vector-valued function $\vec{r}(t)$ is defined as:

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

provided the limit exists.

If $\vec{r}\,'(t)$ exists for all t in an open interval (a, b), then $\vec{r}(t)$ is differentiable over (a, b). If $\vec{r}\,'(t)$ exists for all t in an open interval (a, b), then $\vec{r}(t)$ is differentiable over (a, b).

For $\vec{r}(t)$ to be differentiable on the closed interval $[a, b]$, the following one-sided limits must also exist:

$$\vec{r}\,'(a) = \lim_{h\to 0^+} \frac{\vec{r}(a+h) - \vec{r}(a)}{h} \quad \text{and} \quad \vec{r}\,'(b) = \lim_{h\to 0^-} \frac{\vec{r}(b+h) - \vec{r}(b)}{h}$$

EXAMPLE 3.6

Use the definition to calculate the derivative of the vector-valued function $\vec{r}(t) = (3t+4)\,\mathbf{i} + (t^2 - 4t + 3)\,\mathbf{j}$.

Solution:

$$\lim_{h\to 0} \frac{[3(t+h)+4]\;\mathbf{i} + \left[(t+h)^2 - 4(t+h) + 3\right]\;\mathbf{j} - \left[(3t+4)\,\mathbf{i} + (t^2 - 4t + 3)\,\mathbf{j}\right]}{h}$$

$$= \lim_{h\to 0} \frac{(3t+3h+4)\,\mathbf{i} - (3t+4)\,\mathbf{i} + \left(t^2 + 2th + h^2 - 4t - 4h + 3\right)\;\mathbf{j} - (t^2 - 4t + 3)\,\mathbf{j}}{h}$$

$$= \lim_{h\to 0} \frac{(3h)\,\mathbf{i} + \left(2th + h^2 - 4h\right)\;\mathbf{j}}{h}$$

$$= \lim_{h\to 0} \left(3\,\mathbf{i} + (2t + h - 4)\,\mathbf{j}\right)$$

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

EXAMPLE 3.7

Find an equation for the line tangent to the curve $\vec{r}(t) = \langle t^2,\ 3 - t^2,\ t^3 \rangle$ at $t = 1$.

Solution:

The tangent line will pass through $\vec{r}(1)$ and point in the direction of $\vec{r}\,'(1)$.
First, compute the derivative componentwise:

$$\vec{r}\,'(t) = \left\langle 2t,\ -2t,\ 3t^2 \right\rangle$$

We evaluate at $t = 1$ to get $\vec{r}(1) = \langle 1,\ 2,\ 1 \rangle$ and $\vec{r}\,'(1) = \langle 2,\ -2,\ 3 \rangle$.
Thus, the parametric equations for the tangent line are

$$\vec{L}(t) = \vec{r}(1) + t\,\vec{r}\,'(1) = \langle 1,\ 2,\ 1 \rangle + t\,\langle 2,\ -2,\ 3 \rangle = \langle 1 + 2t,\ 2 - 2t,\ 1 + 3t \rangle.$$

The tangent line (shown in blue) is a continuation of the derivative vector in both directions:

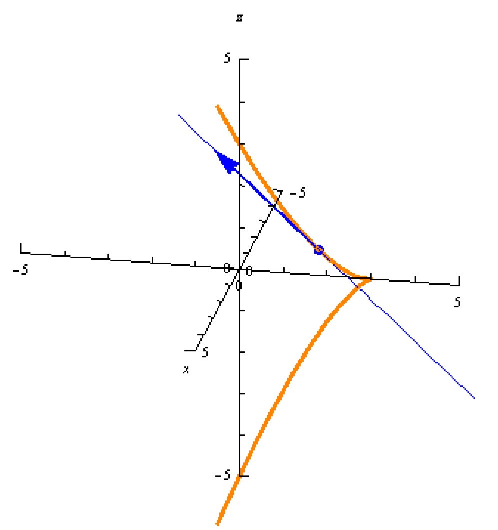


Image credit: UMich

Geometrically, $\vec{r}'(t)$ is the *tangent vector* to the curve at time t. If $\vec{r}(t)$ traces the path of a particle, then $\vec{r}'(t)$ points in the direction of motion at that instant.

We also define the **unit tangent vector** as:

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

The unit tangent vector tells you the direction in which a curve is heading at a specific point. It is essentially the normalized version of the derivative vector.

And now for an essential theorem on the differentiation of vector-valued functions:

Let $f(t), g(t), h(t)$ be differentiable functions of t.

1. If $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j}$, then

$$\vec{r}'(t) = f'(t)\,\mathbf{i} + g'(t)\,\mathbf{j}.$$

2. If $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j} + h(t)\,\mathbf{k}$, then

$$\vec{r}'(t) = f'(t)\,\mathbf{i} + g'(t)\,\mathbf{j} + h'(t)\,\mathbf{k}.$$

EXAMPLE 3.9

Calculate the derivative of the vector-valued function $\vec{r}(t) = (t \ln t)\,\mathbf{i} + (5e^t)\,\mathbf{j} + (\cos t - \sin t)\,\mathbf{k}$.

Solution:

Differentiate each component:

$$\frac{d}{dt}(t \ln t) = \ln t + 1$$

$$\frac{d}{dt}(5e^t) = 5e^t$$

$$\frac{d}{dt}(\cos t - \sin t) = -\sin t - \cos t$$

After gathering we have our answer:

$$\vec{r}\,'(t) = (\ln t + 1)\,\mathbf{i} + 5e^t\,\mathbf{j} + (-\sin t - \cos t)\,\mathbf{k}$$

EXAMPLE 3.10

(a) Find the derivative of $\vec{r}(t) = (1 + t^3)\,\mathbf{i} + te^{-t}\,\mathbf{j} + \sin(2t)\,\mathbf{k}$.
(b) Find the unit tangent vector at the point where $t = 0$.

Solution:

(a) Differentiate each component to get $\vec{r}\,'(t) = 3t^2\,\mathbf{i} + (1-t)e^{-t}\,\mathbf{j} + 2\cos(2t)\,\mathbf{k}$.

(b) Then, evaluate at $t = 0$ to get $\vec{r}(0) = \mathbf{i}$ and $\vec{r}\,'(0) = 0\,\mathbf{i} + 1\,\mathbf{j} + 2\,\mathbf{k}$. Now we can plug these in to find the unit tangent vector at $(1, 0, 0)$:

$$\mathbf{T}(0) = \frac{\vec{r}\,'(0)}{\|\vec{r}\,'(0)\|} = \frac{\mathbf{j} + 2\mathbf{k}}{\sqrt{1+4}} = \frac{1}{\sqrt{5}}\,\mathbf{j} + \frac{2}{\sqrt{5}}\,\mathbf{k}$$

EXAMPLE 3.11

Find parametric equations for the tangent line to the helix $\vec{r}(t) = \langle 2\cos t, \sin t, t\rangle$ at the point $(0, 1, \frac{\pi}{2})$. Then graph the helix and tangent line.

Solution:

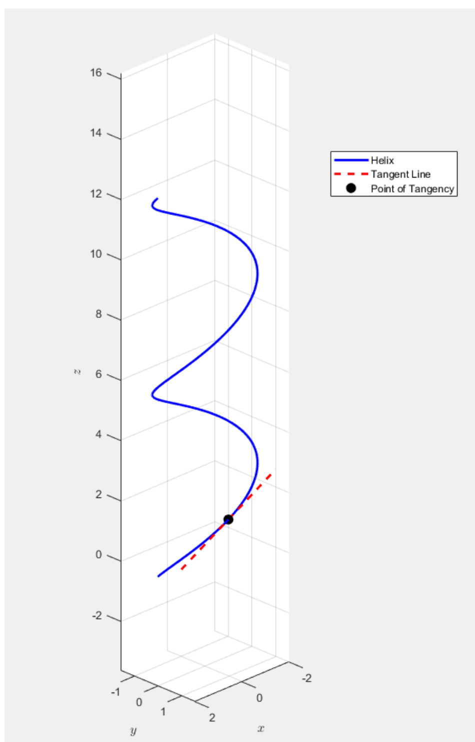
Differentiate to get $\vec{r}\,'(t) = \langle -2\sin t, \cos t, 1\rangle$. Then, evaluate at $t = \frac{\pi}{2}$:

$$\vec{r}\,'\left(\frac{\pi}{2}\right) = \langle -2,\ 0,\ 1\rangle$$

Thus, the tangent line at the point $(0, 1, \frac{\pi}{2})$ has direction vector $\langle -2, 0, 1\rangle$, and the parametric equations are

$$x = -2t,\ y = 1,\ z = \frac{\pi}{2} + t.$$

Let's now graph:



EXAMPLE 3.11 (CONTINUED)

Please run the MATLAB code yourself and have a look!

```
% Helix parameters
t = linspace(0, 4*pi, 1000);
x = 2 * cos(t);
y = sin(t);
z = t;

t0 = pi/2; % Tangent point
p0 = [2*cos(t0), sin(t0), t0];
v = [-2, 0, 1];

% Tangent line parameter
s = linspace(-1, 1, 200);
x_tan = p0(1) + v(1)*s;
y_tan = p0(2) + v(2)*s;
z_tan = p0(3) + v(3)*s;

figure;
hold on;

% Helix
plot3(x, y, z, 'b', 'LineWidth', 2);
% Tangent line
plot3(x_tan, y_tan, z_tan, 'r--', 'LineWidth', 2);
% Point of tangency
plot3(p0(1), p0(2), p0(3), 'ko', 'MarkerSize', 8, 'MarkerFaceColor', 'k');

xlabel('$x$', 'Interpreter', 'latex');
ylabel('$y$', 'Interpreter', 'latex');
zlabel('$z$', 'Interpreter', 'latex');
title('Helix and Tangent Line at t = \pi/2');
legend('Helix', 'Tangent Line', 'Point of Tangency');
grid on;
axis equal;
view(135, 25);
```

ex3point11plot.m

Properties of Differentiation for Vector-Valued Functions

Let $\vec{u}(t)$ and $\vec{v}(t)$ be differentiable vector-valued functions of t, let $f(t)$ be a differentiable scalar-valued function of t that takes only real numbers, and let c be a constant. Then the following derivative rules hold:

1. $\dfrac{d}{dt}\left[c\vec{u}(t)\right] = c\vec{u}\,'(t)$ (scalar multiple)
2. $\dfrac{d}{dt}\left[\vec{u}(t) \pm \vec{v}(t)\right] = \vec{u}\,'(t) \pm \vec{v}\,'(t)$ (sum and difference)
3. $\dfrac{d}{dt}\left[f(t)\vec{u}(t)\right] = f'(t)\vec{u}(t) + f(t)\vec{u}\,'(t)$ (scalar product)
4. $\dfrac{d}{dt}\left[\vec{u}(t) \cdot \vec{v}(t)\right] = \vec{u}\,'(t) \cdot \vec{v}(t) + \vec{u}(t) \cdot \vec{v}\,'(t)$ (dot product)
5. $\dfrac{d}{dt}\left[\vec{u}(t) \times \vec{v}(t)\right] = \vec{u}\,'(t) \times \vec{v}(t) + \vec{u}(t) \times \vec{v}\,'(t)$ (cross product)
6. $\dfrac{d}{dt}\left[\vec{u}(f(t))\right] = \vec{u}\,'(f(t)) \cdot f'(t)$ (chain rule)
7. If $\vec{u}(t) \cdot \vec{u}(t) = c$, then $\vec{u}(t) \cdot \vec{u}\,'(t) = 0$. (orthogonality condition)

EXAMPLE 3.12

Given the vector-valued functions $\vec{r}(t) = (6t+8)\,\mathbf{i} + (4t^2+2t-3)\,\mathbf{j} + 5t\,\mathbf{k}$ and $\vec{u}(t) = (t^2-3)\,\mathbf{i} + (2t+4)\,\mathbf{j} + (t^3-3t)\,\mathbf{k}$, calculate each of the following derivatives:

(a) $\dfrac{d}{dt}\,[\vec{r}(t)\cdot\vec{u}(t)]$

(b) $\dfrac{d}{dt}\,[\vec{u}(t)\times\vec{u}(t)]$

Solution:

(a) We compute the derivatives componentwise:

$$\vec{r}\,'(t) = 6\,\mathbf{i} + (8t+2)\,\mathbf{j} + 5\,\mathbf{k},\ \ \vec{u}\,'(t) = 2t\,\mathbf{i} + 2\,\mathbf{j} + (3t^2-3)\,\mathbf{k}$$

By the dot product rule, $\frac{d}{dt}\,[\vec{r}(t)\cdot\vec{u}(t)] = \vec{r}\,'(t)\cdot\vec{u}(t) + \vec{r}(t)\cdot\vec{u}\,'(t)$. Thus we have

$$\begin{aligned}
&(6\,\mathbf{i} + (8t+2)\,\mathbf{j} + 5\,\mathbf{k})\cdot\big((t^2-3)\,\mathbf{i} + (2t+4)\,\mathbf{j} + (t^3-3t)\,\mathbf{k}\big)\\
&+\big((6t+8)\,\mathbf{i} + (4t^2+2t-3)\,\mathbf{j} + 5t\,\mathbf{k}\big)\cdot\big(2t\,\mathbf{i} + 2\,\mathbf{j} + (3t^2-3)\,\mathbf{k}\big)
\end{aligned}$$

$$= 6(t^2-3) + (8t+2)(2t+4) + 5(t^3-3t) + 2t(6t+8) + 2(4t^2+2t-3) + 5t(3t^2-3)$$

$$= 6t^2 - 18 + 16t^2 + 32t + 4t + 8 + 5t^3 - 15t + 12t^2 + 16t + 8t^2 + 4t - 6 + 15t^3 - 15t$$

$$= 20t^3 + 42t^2 + 26t - 16.$$

EXAMPLE 3.12 (CONTINUED)

(b) We use the product rule for cross products which says $\frac{d}{dt}\left[\vec{u}(t)\times\vec{u}(t)\right] = \vec{u}\,'(t)\times\vec{u}(t) + \vec{u}(t)\times\vec{u}\,'(t)$.

Since the cross product of a vector with itself is zero ($\vec{u}(t)\times\vec{u}(t) = \vec{0}$), we know that $\vec{u}\,'(t)\times\vec{u}(t)+\vec{u}(t)\times\vec{u}\,'(t) = \vec{0}$.

We are trying to find $\frac{d}{dt}\left[\vec{u}(t)\times\vec{u}(t)\right] = \vec{u}(t)\times\vec{u}\,''(t)$. For $\vec{u}(t) = (t^2-3)\,\mathbf{i}+(2t+4)\,\mathbf{j}+(t^3-3t)\,\mathbf{k}$, we compute

$$\vec{u}\,''(t) = \frac{d}{dt}\left[2t\,\mathbf{i}+2\,\mathbf{j}+(3t^2-3)\,\mathbf{k}\right] = 2\,\mathbf{i}+6t\,\mathbf{k}.$$

Now we compute the cross product:

$$\frac{d}{dt}\left[\vec{u}(t)\times\vec{u}(t)\right] = \vec{u}(t)\times\vec{u}\,''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t^2-3 & 2t+4 & t^3-3t \\ 2 & 0 & 6t \end{vmatrix}$$

$$= \mathbf{i}\left[(2t+4)(6t)-0\right] - \mathbf{j}\left[(t^2-3)(6t)-2(t^3-3t)\right] + \mathbf{k}\left[(t^2-3)(0)-2(2t+4)\right]$$

$$= \mathbf{i}\cdot 6t(2t+4) - \mathbf{j}\left[6t(t^2-3)-2(t^3-3t)\right] - \mathbf{k}\cdot 4(2t+4)$$

$$= (12t^2+24t)\,\mathbf{i} + (12t^2-4t^3)\,\mathbf{j} - (4t+8)\,\mathbf{k}$$

We express the integral of a continuous vector-valued function $\vec{r}(t)$ in terms of the integrals of its component functions:

Integrals of Vector-Valued Functions

Let $f(t), g(t)$, and $h(t)$ be integrable real-valued functions over the interval $[a, b]$.

1. The **indefinite integral** of a vector-valued function $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j}$ is given by

$$\int \vec{r}(t)\,dt = \left(\int f(t)\,dt\right)\mathbf{i} + \left(\int g(t)\,dt\right)\mathbf{j}.$$

The **definite integral** is:

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

2. The **indefinite integral** of a vector-valued function $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j} + h(t)\,\mathbf{k}$ is given by

$$\int \vec{r}(t)\,dt = \left(\int f(t)\,dt\right)\mathbf{i} + \left(\int g(t)\,dt\right)\mathbf{j} + \left(\int h(t)\,dt\right)\mathbf{k}.$$

The **definite integral** is

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

Just like with real-valued functions, each of the component integrals contains an integration constant.

Say we have the following component integrals in two dimensions:

$$\int f(t)\,dt = F(t) + C_1 \quad \text{and} \quad \int g(t)\,dt = G(t) + C_2,$$

where F and G are antiderivatives of f and g, respectively.

Then, for the vector-valued function

$$\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j},$$

its indefinite integral is computed component-by-component:

$$\begin{gathered}\int \left(f(t)\,\mathbf{i} + g(t)\,\mathbf{j}\right)\,dt = \left(\int f(t)\,dt\right)\,\mathbf{i} + \left(\int g(t)\,dt\right)\,\mathbf{j} \\ = \left(F(t) + C_1\right)\,\mathbf{i} + \left(G(t) + C_2\right)\,\mathbf{j} = F(t)\,\mathbf{i} + G(t)\,\mathbf{j} + C_1\,\mathbf{i} + C_2\,\mathbf{j} \\ = F(t)\,\mathbf{i} + G(t)\,\mathbf{j} + \vec{C},\end{gathered}$$

where $\vec{C} = C_1\,\mathbf{i} + C_2\,\mathbf{j}$ is a constant vector.

The integration constants became a constant vector.

The fundamental theorem of calculus can be extended to continuous vector functions:

Let $\vec{r}(t) = f(t)\,\mathbf{i} + g(t)\,\mathbf{j} + h(t)\,\mathbf{k}$ be a continuous vector-valued function on the interval $[a, b]$, and suppose that $\vec{R}(t)$ is an antiderivative of $\vec{r}(t)$; that is, $\vec{R}\,'(t) = \vec{r}(t)$. Then:

$$\int_a^b \vec{r}(t)\,dt = \vec{R}(t)\Big|_a^b = \vec{R}(b) - \vec{R}(a)$$

In component form, this becomes

$$\int_a^b f(t)\,dt\,\mathbf{i} + \int_a^b g(t)\,dt\,\mathbf{j} + \int_a^b h(t)\,dt\,\mathbf{k} = [F(b) - F(a)]\,\mathbf{i} + [G(b) - G(a)]\,\mathbf{j} + [H(b) - H(a)]\,\mathbf{k},$$

where F, G, and H are antiderivatives of f, g, and h, respectively.

EXAMPLE 3.13

Evaluate the following definite integrals:

(a) $\int \big((3t^2 + 2t)\,\mathbf{i} + (3t - 6)\,\mathbf{j} + (6t^3 + 5t^2 - 4)\,\mathbf{k}\big)\,dt$
(b) $\int \langle t,\ t^2,\ t^3\rangle \times \langle t^3,\ t^2,\ t\rangle\,dt$

Solution:

(a) Break the integral into componentwise integrals:

$$= \left[\int (3t^2 + 2t)\,dt\right]\mathbf{i} + \left[\int (3t - 6)\,dt\right]\mathbf{j} + \left[\int (6t^3 + 5t^2 - 4)\,dt\right]\mathbf{k}$$
$$= (t^3 + t^2)\,\mathbf{i} + \left(\frac{3}{2}t^2 - 6t\right)\mathbf{j} + \left(\frac{3}{2}t^4 + \frac{5}{3}t^3 - 4t\right)\mathbf{k} + \vec{C}$$

(b) First, compute the cross product $\langle t,\ t^2,\ t^3\rangle \times \langle t^3,\ t^2,\ t\rangle$:

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ t & t^2 & t^3 \\ t^3 & t^2 & t \end{vmatrix} = \big(t^2(t) - t^3(t^2)\big)\,\mathbf{i} - \big(t(t) - t^3(t^3)\big)\,\mathbf{j} + \big(t(t^2) - t^2(t^3)\big)\,\mathbf{k}$$
$$= (t^3 - t^5)\,\mathbf{i} - (t^2 - t^6)\,\mathbf{j} + (t^3 - t^5)\,\mathbf{k}$$

Now we can integrate:

$$\int (t^3 - t^5)\,\mathbf{i} - (t^2 - t^6)\,\mathbf{j} + (t^3 - t^5)\,\mathbf{k}\ \ dt$$
$$= \left(\frac{t^4}{4} - \frac{t^6}{6}\right)\mathbf{i} - \left(\frac{t^3}{3} - \frac{t^7}{7}\right)\mathbf{j} + \left(\frac{t^4}{4} - \frac{t^6}{6}\right)\mathbf{k} + \vec{C}$$

EXAMPLE 3.13

Do the following for the vector-valued function $\vec{r}(t) = 2\cos t\,\mathbf{i} + \sin t\,\mathbf{j} + 2t\,\mathbf{k}$:

(a) Find the indefinite integral $\displaystyle\int \vec{r}(t)\,dt$

(b) Evaluate the definite integral $\displaystyle\int_0^{\pi/2} \vec{r}(t)\,dt$

Solution:

(a)

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

(b)

$$\int_0^{\pi/2} \vec{r}(t)\,dt = \left[2\sin t\,\mathbf{i} - \cos t\,\mathbf{j} + t^2\,\mathbf{k}\right]_0^{\pi/2} = 2\,\mathbf{i} + \mathbf{j} + \frac{\pi^2}{4}\,\mathbf{k}$$

3.3 Arc Length and Curvature

Recall that the formula for the arc length s of a curve defined by the parametric functions $x = x(t)$ and $y = y(t)$ for $t_1 \le t \le t_2$ is

$$s = \int_{t_1}^{t_2} \sqrt{(x'(t))^2 + (y'(t))^2}\,dt.$$

Arc length for a smooth curve defined by the vector-valued function $\vec{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$ for $a \le t \le b$ is given by

$$s = \int_a^b \sqrt{(f'(t))^2 + (g'(t))^2}\,dt.$$

Similarly, arc length for the vector-valued function $\vec{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$ in three dimensions for the same interval is given by

$$s = \int_a^b \sqrt{(f'(t))^2 + (g'(t))^2 + (h'(t))^2}\, dt.$$

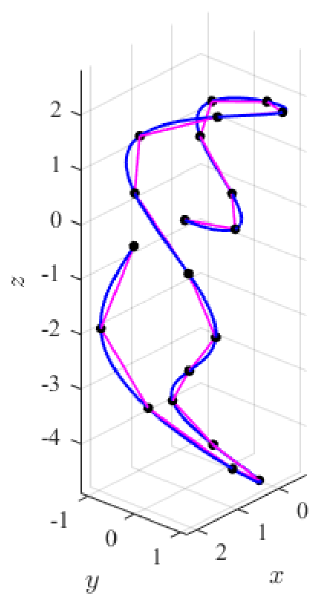
Suppose a smooth curve has the vector equation $\vec{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j}$ where $a \leq t \leq b$. Equivalently, you could say it has the parametric equations $x = f(t)$ and $y = g(t)$ where f', g', and h' are continuous. Then, the arc length of the *plane curve* is given by

$$s = \int_a^b \|\vec{r}\,'(t)\|\, dt = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2}\, dt.$$

Suppose a smooth curve has the vector equation $\vec{r}(t) = f(t)\mathbf{i} + g(t)\mathbf{j} + h(t)\mathbf{k}$ where $a \leq t \leq b$. Equivalently, you could say it has the parametric equations $x = f(t)$, $y = g(t)$, and $z = h(t)$ where f', g', and h' are continuous. Then, the arc length of the *space curve* is given by

$$s = \int_a^b \|\vec{r}\,'(t)\|\, dt = \int_a^b \sqrt{[f'(t)]^2 + [g'(t)]^2 + [h'(t)]^2}\, dt.$$

Recall how we can approximate area under a real-valued curve using Riemann sums. This plot shows how the length of a space curve (blue) is the limit of lengths of inscribed polygons (pink).



In other words, we can approximate the curve using line segments.

EXAMPLE 3.14

Find the length of the arc of the circular helix defined by the vector-valued function $\vec{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$ from the point $(1, 0, 0)$ to the point $(1, 0, 2\pi)$. Then, graph the result.

Solution:

We begin by computing the derivative of $\vec{r}(t)$:

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

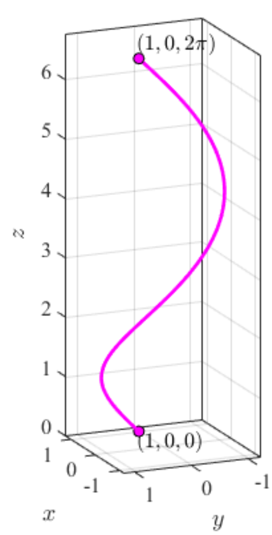
Now compute the magnitude:

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

The arc begins at $t = 0$ and ends at $t = 2\pi$, so the arc length is given by

$$s = \int_0^{2\pi} \|\vec{r}\,'(t)\|\,dt = \int_0^{2\pi} \sqrt{2}\,dt = \sqrt{2}\left(\int_0^{2\pi} dt\right) = \sqrt{2}(2\pi) = 2\sqrt{2}\pi$$

Let's graph:



A vector-valued function that describes a helix can be written in the form

$$\vec{r}(t) = R\cos\left(\frac{2\pi N t}{h}\right)\mathbf{i} + R\sin\left(\frac{2\pi N t}{h}\right)\mathbf{j} + t\,\mathbf{k},\ 0 \le t \le h,$$

where R represents the radius of the helix, h represents the height (distance between two consecutive turns), and the helix completes N turns. Let's now derive a formula for the arc length of this helix. First we have

$$\vec{r}\,'(t) = -\frac{2\pi NR}{h}\sin\left(\frac{2\pi Nt}{h}\right)\mathbf{i} + \frac{2\pi NR}{h}\cos\left(\frac{2\pi Nt}{h}\right)\mathbf{j} + \mathbf{k}.$$

Then,

$$s = \int_a^b \|\vec{r}\,'(t)\|\,dt$$

$$= \int_0^h \sqrt{\left(-\frac{2\pi NR}{h}\sin\left(\frac{2\pi Nt}{h}\right)\right)^2 + \left(\frac{2\pi NR}{h}\cos\left(\frac{2\pi Nt}{h}\right)\right)^2 + 1^2}\,dt$$

$$= \int_0^h \sqrt{\frac{4\pi^2N^2R^2}{h^2}\left(\sin^2\left(\frac{2\pi Nt}{h}\right) + \cos^2\left(\frac{2\pi Nt}{h}\right)\right) + 1}\,dt$$

$$= \int_0^h \sqrt{\frac{4\pi^2N^2R^2}{h^2} + 1}\,dt$$

$$= \left[t\sqrt{\frac{4\pi^2N^2R^2}{h^2} + 1}\right]_0^h$$

$$= h\sqrt{\frac{4\pi^2N^2R^2}{h^2} + 1}$$

$$= \sqrt{4\pi^2N^2R^2 + h^2}$$

This gives the formula for the length of a wire needed to form a helix with N turn, radius R, and height h.

The arc length of a curve often arises naturally from the shape of the curve rather than a specific coordinate system. This is why *parameterization* of a curve with respect to arc length is useful. It lets us efficiently describe the motion and shape of a curve as a geometric object.

To do this, we need the arc length function:

Let $\vec{r}(t)$ describe a smooth curve for $t \geq a$. Then the arc length function is given by

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

Additionally,

$$\frac{ds}{dt} = \|\vec{r}\,'(t)\| > 0.$$

If $\|\vec{r}\,'(t)\| = 1$ for all $t \geq a$, then the parameter t represents the arc length measured from $t = a$.

If a curve $\vec{r}(t)$ is already given in terms of a parameter t, and you have the arc length function $s(t)$, then you can try to solve for t in terms of s. That is, find $t = t(s)$. Once you do that, you can reparametrize the curve in terms of s by substituting it in for t:

$$\vec{r}(s) = \vec{r}(t(s))$$

This now means you're describing the curve in terms of how far you've traveled along it from its starting point. For example, if $s = 3$, then $\vec{r}(t(3))$ is the position vector of the point that's 3 units of length along the curve from where you started.

EXAMPLE 3.15

Reparameterize the helix $\vec{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + t\,\mathbf{k}$ with respect to arc length measured from the point $(1, 0, 0)$ in the direction of increasing t.

Solution:
The initial point $(1, 0, 0)$ corresponds to $t = 0$. The derivative is $\vec{r}\,'(t) = -\sin t\,\mathbf{i} + \cos t\,\mathbf{j} + \mathbf{k}$. Then, we compute the magnitude:

$$\|\vec{r}\,'(t)\| = \sqrt{(-\sin t)^2 + (\cos t)^2 + (1)^2} = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{1+1} = \sqrt{2} = \frac{ds}{dt}$$

The arc length function is

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

Solving for t, we get $t = \frac{s}{\sqrt{2}}$. Substituting this into $\vec{r}(t)$, the reparameterization with respect to arc length is

$$\vec{r}(t(s)) = \cos\left(\frac{s}{\sqrt{2}}\right)\mathbf{i} + \sin\left(\frac{s}{\sqrt{2}}\right)\mathbf{j} + \left(\frac{s}{\sqrt{2}}\right)\mathbf{k}$$

You know that *smoothness* is a measure of the number of continuous derivatives a function has; a smooth curve has no corners or cusps. We can extend this idea to our current work. We call a parameterization $\vec{r}(t)$ smooth on an interval if its derivative is continuous and nonzero on that interval. A curve is called smooth if it has a smooth parameterization. The measurement of how sharply a smooth curve turns, or how quickly it changes direction is called **curvature**. For instance, a circle has constant curvature that is proportional to the size of its radius. We define curvature κ as the magnitude of the rate of change of the unit tangent vector with respect to arc length. The unit tangent vector has constant length, so only changes in direction contribute to the rate of change.

The curvature κ of a smooth curve in a plane or in space given by $\vec{r}(s)$ is given by

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

where $\mathbf{T}$ is the unit tangent vector and s is the arc length parameter.

For a smooth curve $\vec{r}(t)$, we can acquire a more useful form using the chain rule:

$$\frac{d\mathbf{T}}{dt} = \frac{d\mathbf{T}}{ds}\frac{ds}{dt} \Rightarrow \kappa(t) = \left\| \frac{d\mathbf{T}}{ds} \right\| = \frac{\|\mathbf{T}'(t)\|}{\|\vec{r}'(t)\|}$$

This form expresses curvature in terms of the parameter t instead of s.

For a three-dimensional curve, curvature can be given by the formula*

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

For a plane curve represented by function $y = f(x)$ where both y' and y'' exist, the curvature at the point (x, y) is given by

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

*If you are interested in the proof, here it is:

We begin with the definition of the unit tangent vector:

$$\mathbf{T}(t) = \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|} \quad \text{and} \quad \|\vec{r}'(t)\| = \frac{ds}{dt}$$

Differentiating $\vec{r}(t) = \|\vec{r}'(t)\|\, \mathbf{T}(t)$ using the product rule gives us

$$\vec{r}''(t) = \frac{d^2s}{dt^2}\mathbf{T}(t) + \frac{ds}{dt}\mathbf{T}'(t).$$

Now take the cross product of $\vec{r}'(t)$ and $\vec{r}''(t)$:

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

Since $\mathbf{T}(t) \perp \mathbf{T}'(t)$, their cross product is orthogonal, and its magnitude is:

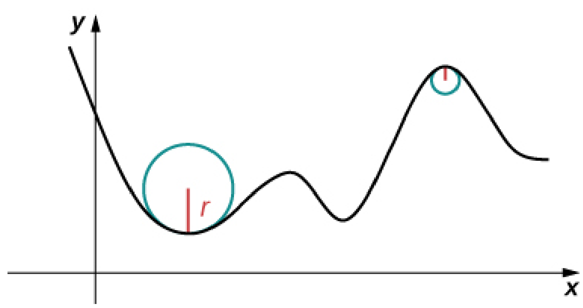
$$\|\vec{r}'(t) \times \vec{r}''(t)\| = \left(\frac{ds}{dt}\right)^2 \|\mathbf{T}(t) \times \mathbf{T}'(t)\| = \left(\frac{ds}{dt}\right)^2 \|\mathbf{T}'(t)\|$$

Now solve for $\|\mathbf{T}'(t)\|$:

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

Finally, recall that curvature is

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



This graph represents the curvature of a function $y = f(x)$ where curvature is inversely proportional to the radius of the inscribed circle. Image credit: Strang & Herman

EXAMPLE 3.16

Find the curvature of a circle of radius a. Assume the circle is centered at the origin.

Solution:

We will use this as our parameterization:

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

Let's differentiate with respect to t:

$$\vec{r}\,'(t) = -a\sin t\,\mathbf{i} + a\cos t\,\mathbf{j} \Rightarrow \|\vec{r}\,'(t)\| = a$$

The unit tangent vector is

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

Now we will differentiate $\mathbf{T}(t)$:

$$\mathbf{T}\,'(t) = -\cos t\,\mathbf{i} - \sin t\,\mathbf{j} \Rightarrow \|\mathbf{T}\,'(t)\| = 1$$

Plugging this into the formula for curvature yields

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

EXAMPLE 3.17

Find the curvature of the twisted cubic $\vec{r}(t) = \langle t,\ t^2,\ t^3\rangle$ at a general point and at $(0,\ 0,\ 0)$.

Solution:

We compute the derivatives and get $\vec{r}\,'(t) = \langle 1,\ 2t,\ 3t^2\rangle$ and $\vec{r}\,''(t) = \langle 0,\ 2,\ 6t\rangle$.

Next, compute the cross product:

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

Now compute the magnitudes:

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

Using the curvature formula yields the curvature at a general point:

$$\kappa(t) = \frac{\|\vec{r}\,'(t) \times \vec{r}\,''(t)\|}{\|\vec{r}\,'(t)\|^3} = \frac{\sqrt{4 + 36t^2 + 36t^4}}{(1 + 4t^2 + 9t^4)^{3/2}}$$

At the origin, where $t = 0$, we have

$$\kappa(0) = \frac{\sqrt{4}}{(1)^{3/2}} = \frac{2}{1} = 2.$$

At a given point on a smooth space curve $\vec{r}(t)$, there are many vectors that are orthogonal to the unit tangent vector $\mathbf{T}(t)$. Because $\mathbf{T}(t)$ is a unit vector, its magnitude is always 1 for all t. In other words, $\|\mathbf{T}(t)\| = 1$. And this implies that $\mathbf{T}(t) \cdot \mathbf{T}(t) = 1$. Then, we differentiate both sides with respect to t:

$$\frac{d}{dt}\left(\mathbf{T}(t) \cdot \mathbf{T}(t)\right) = 2\mathbf{T}(t) \cdot \vec{T}'(t) = 0$$

Thus $\mathbf{T}(t) \cdot \vec{T}'(t) = 0$, which tells us that $\vec{T}'(t)$ is always orthogonal to $\mathbf{T}(t)$. Therefore, it points in the direction that the tangent vector is turning. If we normalize this, we get a unit vecor that points in the direction of curvature which is known as the **principle unit normal vector $\mathbf{N}(t)$**.

For a three-dimensional smooth curve represented by $\vec{r}$ over an open interval where $\vec{T}'(t) \neq \vec{0}$, the principal unit normal vector at t is defined as

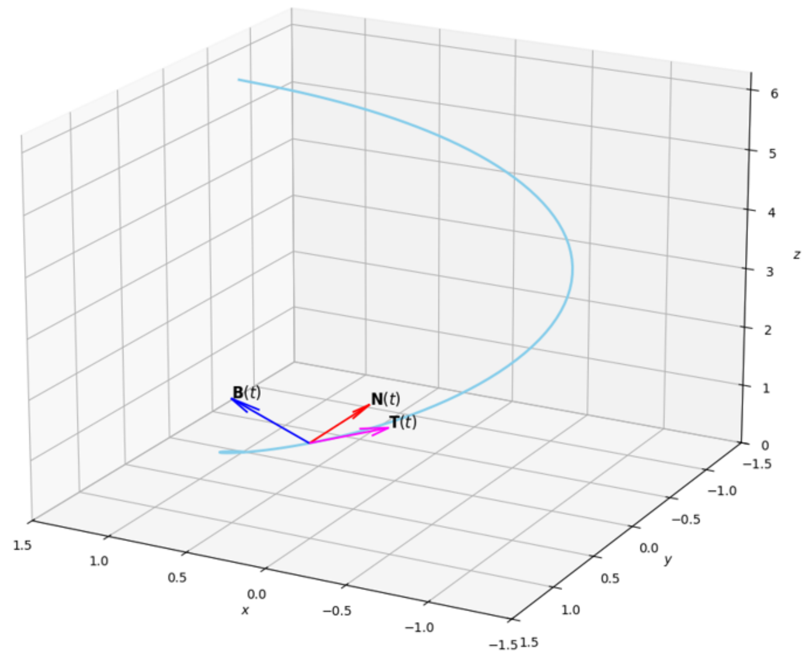
$$\mathbf{N}(t) = \frac{\mathbf{T}(t)}{\left\|\vec{T}'(t)\right\|}.$$

Then the binormal vector is given by

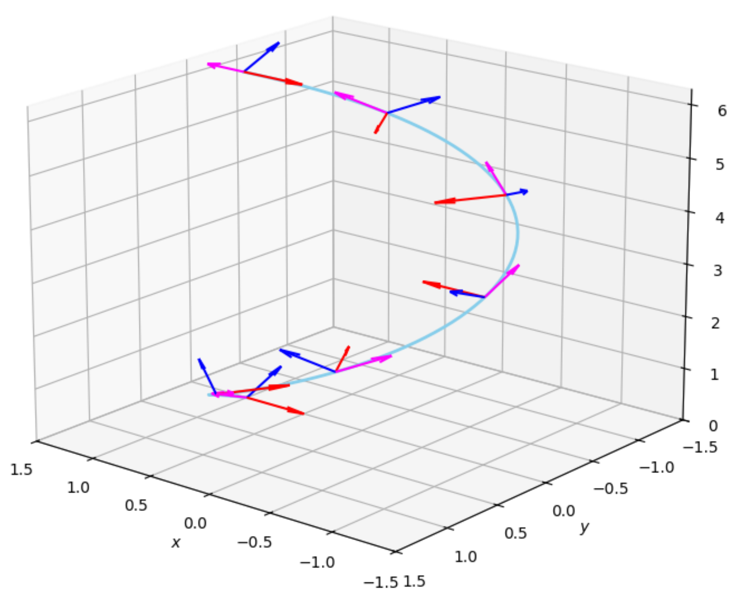
$$\mathbf{B}(t) = \mathbf{T}(t) \times \mathbf{N}(t),$$

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

The binormal vector is orthogonal to both the unit tangent vector and the normal vector. With that, we have completed the basics of what is known as the **Frenet-Serret frame**. You can think of it as the frame of reference for describing a curve's geometry as you move along it. This is important in the discipline of differential geometry.



Vectors **T**, **N**, and **B** on a helix



Here are those same vectors at multiple points along the helix.

EXAMPLE 3.18

Find the unit normal vector for the vector-valued function $\vec{r}(t) = 4\cos t\,\mathbf{i} - 4\sin t\,\mathbf{j}$.

Solution:

We first compute the unit tangent vector $\mathbf{T}(t) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)\|}$:

$$\begin{aligned}
&\frac{-4\sin t\,\mathbf{i} - 4\cos t\,\mathbf{j}}{\sqrt{(-4\sin t)^2 + (-4\cos t)^2}} \\
&= \frac{-4\sin t\,\mathbf{i} - 4\cos t\,\mathbf{j}}{\sqrt{16\sin^2 t + 16\cos^2 t}} = \frac{-4\sin t\,\mathbf{i} - 4\cos t\,\mathbf{j}}{\sqrt{16(\sin^2 t + \cos^2 t)}} \\
&= \frac{-4\sin t\,\mathbf{i} - 4\cos t\,\mathbf{j}}{\sqrt{16}} = \frac{-4\sin t\,\mathbf{i} - 4\cos t\,\mathbf{j}}{4} = -\sin t\,\mathbf{i} - \cos t\,\mathbf{j}
\end{aligned}$$

Now differentiate $\mathbf{T}(t)$ to find the unit normal vector $\mathbf{N}(t) = \frac{\vec{T}'(t)}{\|\vec{T}'(t)\|}$:

$$\vec{T}'(t) = -\cos t\,\mathbf{i} + \sin t\,\mathbf{j} \Rightarrow \|\vec{T}'(t)\| = \sqrt{(-\cos t)^2 + (\sin t)^2} = \sqrt{\cos^2 t + \sin^2 t} = \sqrt{1} = 1$$

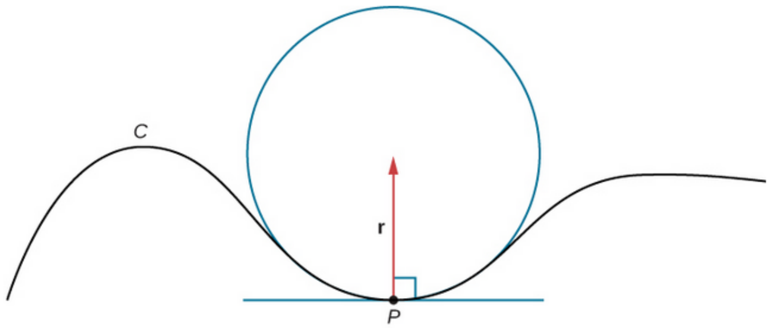
$$\mathbf{N}(t) = \frac{\frac{d}{dt}\left(-\sin t\,\mathbf{i} - \cos t\,\mathbf{j}\right)}{\left\|\frac{d}{dt}\left(-\sin t\,\mathbf{i} - \cos t\,\mathbf{j}\right)\right\|} = \frac{-\cos t\,\mathbf{i} + \sin t\,\mathbf{j}}{\sqrt{(-\cos t)^2 + (\sin t)^2}} = -\cos t\,\mathbf{i} + \sin t\,\mathbf{j}$$

We have a three-dimensional system that follows the Frenet-Serret frame. The **normal plane** at point P is the plane perpendicular to $\mathbf{T}(t)$, spanned by the vectors $\mathbf{N}(t)$ and $\mathbf{B}(t)$. It contains all of the lines orthogonal to $\mathbf{T}(t)$. The plane spanned by $\mathbf{T}(t)$ and $\mathbf{N}(t)$ is known as the **osculating plane** at P, which is the plane that best approximates the curve's local behavior at that point. The word "osculating" comes from the Latin word *osculum*, meaning "kiss." It is the plane that *kisses* the curve most closely at that point.

Now imagine a circle that lies in the osculating plane and shares the same position, tangent, and curvature as the curve at point P. This circle is called the **osculating circle** (or **circle of curvature**). It lies on the concave side of the curve (toward which $\mathbf{N}(t)$ points) and has radius of curvature

$$R = \frac{1}{\kappa},$$

where κ is the curvature at P.



This is the osculating circle at point P on the curve C. The circle is tangent to the curve at point P and lies in the osculating plane. Image credit: Strang & Herman

EXAMPLE 3.19

Find the equation of the osculating circle of the curve defined by the function $y = x^3 - 3x + 1$ at the point $x = 1$.

Solution:

First, we compute the curvature κ of the graph of a function $y = f(x)$ using the formula

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

For $f(x) = x^3 - 3x + 1$, we have $f'(x) = 3x^2 - 3$ and $f''(x) = 6x$. Plugging these in, we get

$$\kappa = \frac{|6x|}{[1 + (3x^2 - 3)^2]^{3/2}}.$$

At $x = 1$, this becomes

$$\kappa = \frac{6}{(1 + 0^2)^{3/2}} = \frac{6}{1} = 6.$$

So the radius of curvature is

$$R = \frac{1}{\kappa} = \frac{1}{6}.$$

EXAMPLE 3.19 (CONTINUED)

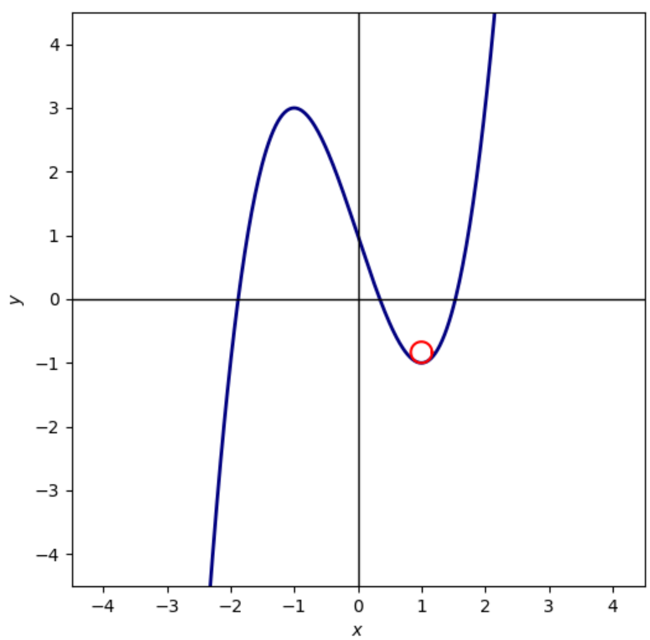
Now we need to find the coordinates of the center of the circle. When $x = 1$, the slope of the tangent line is $f'(1) = 0$. This tells us that the tangent line is horizontal here, so the normal vector which is always perpendicular to the tangent must be vertical. Thus, the center of the osculating circle is directly above $P = (1, f(1)) = (1, -1)$, in the direction of the normal. Thus, the center is

$$C = \left(1, -1 + \frac{1}{6}\right) = \left(1, -\frac{5}{6}\right).$$

Now we write the equation of the circle with radius $r = \frac{1}{6}$ and center $(h, k) = \left(1, -\frac{5}{6}\right)$ in the form $(x-h)^2 + (y-k)^2 = r^2$:

$$(x-1)^2 + \left(y + \frac{5}{6}\right)^2 = \left(\frac{1}{6}\right)^2.$$

The osculating circle (red) is drawn on the curve (blue):



3.4 Motion in Space

Previously, you studied motion along a straight line using scalar functions for position $x(t)$, velocity $v(t) = x'(t)$, and acceleration $a(t) = v'(t)$. You likely used these to describe the motion of something very simple such as a car, or more generally particle motion along the x-axis. In this section, we extend this idea to motion in two or three dimensions. We will use vector-valued functions to describe the motion of an object in space.

Suppose a particle is moving through space such that it can be described by a position vector $\vec{r}(t)$ at time t. For small values of h, the vector

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

approximates the direction of the particle as it moves along the curve $\vec{r}(t)$. Furthermore, its magnitude measures the absolute value of the displacement vector per unit time. Thus, the vector gives the average velocity over a time interval of length h. If you take its limit, you will get the velocity vector:

$$\vec{v}(t) = \lim_{h\to 0} \frac{\vec{r}(t+h) - \vec{r}(t)}{h} = \vec{r}\,'(t)$$

The velocity vector is the tangent vector and it points in the direction of the tangent line. If you simply wanted speed, you would compute $\|\vec{v}(t)\|$. Now, let's look at acceleration which is given by

$$\vec{a}(t) = \vec{v}\,'(t) = \vec{r}\,''(t).$$

We can summarize the formulas for motion as such:

Quantity	**Two Dimensions**	**Three Dimensions**
Position	$\vec{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j}$	$\vec{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j} + z(t)\,\mathbf{k}$
Velocity	$\vec{v}(t) = x'(t)\,\mathbf{i} + y'(t)\,\mathbf{j}$	$\vec{v}(t) = x'(t)\,\mathbf{i} + y'(t)\,\mathbf{j} + z'(t)\,\mathbf{k}$
Acceleration	$\vec{a}(t) = x''(t)\,\mathbf{i} + y''(t)\,\mathbf{j}$	$\vec{a}(t) = x''(t)\,\mathbf{i} + y''(t)\,\mathbf{j} + z''(t)\,\mathbf{k}$
Speed	$\|\vec{v}(t)\| = \sqrt{(x'(t))^2 + (y'(t))^2}$	$\|\vec{v}(t)\| = \sqrt{(x'(t))^2 + (y'(t))^2 + (z'(t))^2}$

Let's say you are driving on a curvy road. If you were to simply drive in a straight line, you would go off the road. The velocity at which you are traveling can be described by velocity vectors, which are tangent to the path traveled by your car:

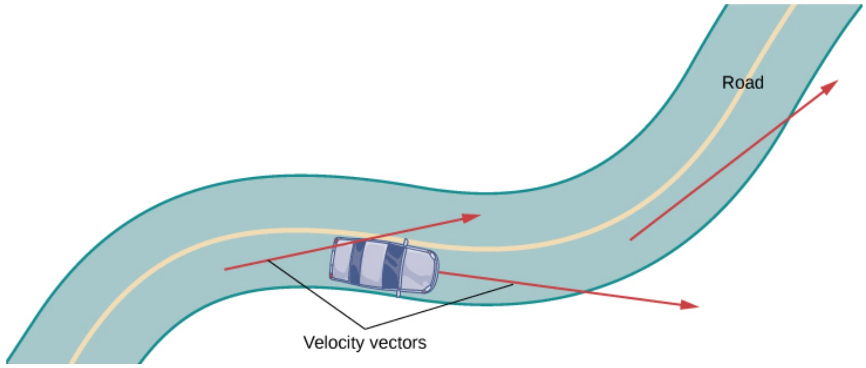


Image credit: Strang & Herman

We obviously don't want to crash into the barrier on the side of the road, so you have to turn your steering wheel to stay on the road. Despite the fact that the magnitude of your velocity (speed) is not changing, your *direction* is constantly going to change to keep you on the road. Your acceleration vector points to whichever direction you turn towards; if you turn right, your acceleration vector also points to the right. And when you turn left to go along the next segment of the road, your acceleration vector will point to the left. Even if you are keeping your foot's position on the gas pedal constant (constant speed), your velocity and acceleration vectors are constantly changing:

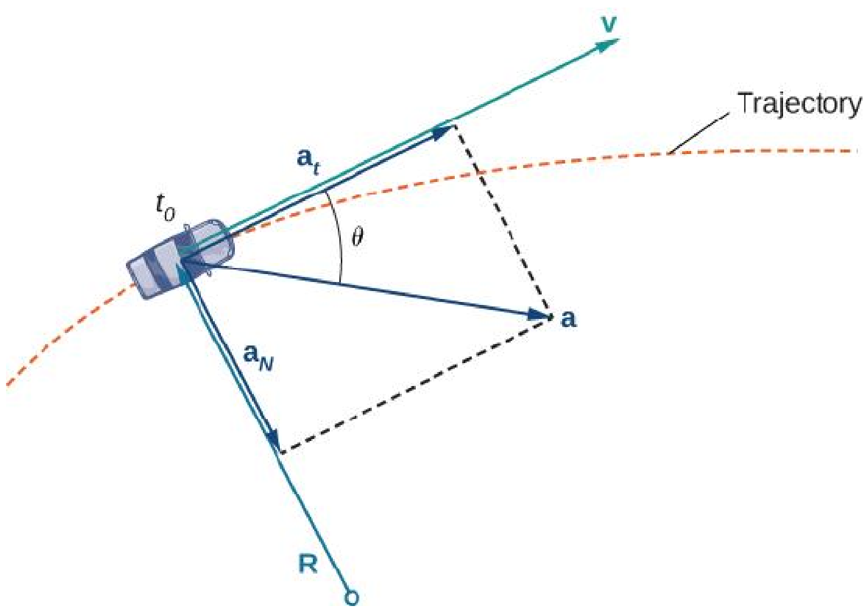


Image credit: Strang & Herman

EXAMPLE 3.20

The position vector of an object moving in a plane is given by $\vec{r}(t) = t^3\,\mathbf{i} + t^2\,\mathbf{j}$. Find the velocity, speed, and acceleration, $t = 1$. Then, graph the results.

Solution:

The velocity vector is the first derivative:

$$\vec{v}(t) = \vec{r}\,'(t) = 3t^2\,\mathbf{i} + 2t\,\mathbf{j}$$

The speed is the magnitude of the velocity:

$$\|\vec{v}(t)\| = \sqrt{(3t^2)^2 + (2t)^2} = \sqrt{9t^4 + 4t^2}$$

The acceleration vector is the second derivative:

$$\vec{a}(t) = \vec{r}\,''(t) = 6t\,\mathbf{i} + 2\,\mathbf{j}$$

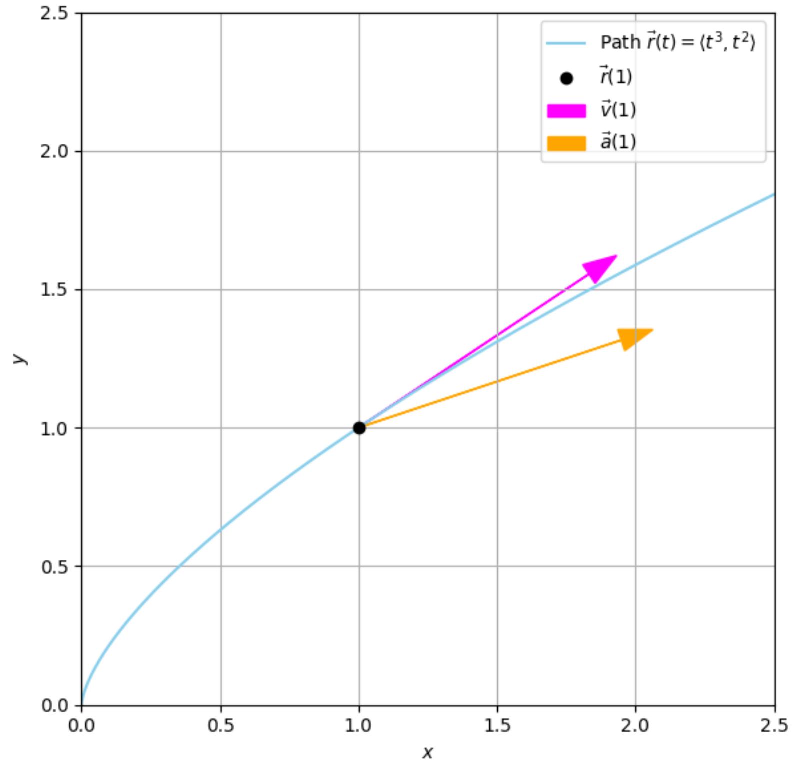
And now we evaluate each at $t = 1$:

$$\vec{v}(1) = 3\,\mathbf{i} + 2\,\mathbf{j} \Rightarrow \|\vec{v}(1)\| = \sqrt{13}$$

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

EXAMPLE 3.20 (CONTINUED)

Here is the graph:



EXAMPLE 3.21

A moving particle starts at an initial position $\vec{r}(0) = \langle 1, 0, 0 \rangle$ with initial velocity $\vec{v}(0) = \mathbf{i} - \mathbf{j} + \mathbf{k}$ and acceleration $\vec{a}(t) = 4t\,\mathbf{i} + 6t\,\mathbf{j} + \mathbf{k}$. Find the velocity and position vectors at time t.

Solution:

Since $\vec{a}(t) = \vec{v}\,'(t)$, we integrate to find velocity:

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

To determine $\vec{C}$, use the initial condition $\vec{v}(0) = \mathbf{i} - \mathbf{j} + \mathbf{k}$. Plug in to get

$$\vec{v}(0) = \vec{C} = \mathbf{i} - \mathbf{j} + \mathbf{k}.$$

Thus, the velocity vector becomes

$$\vec{v}(t) = (2t^2 + 1)\,\mathbf{i} + (3t^2 - 1)\,\mathbf{j} + (t + 1)\,\mathbf{k}.$$

Now integrate $\vec{v}(t)$ to find position

$$\vec{r}(t) = \int \vec{v}(t)\,dt = \int \big[(2t^2 + 1)\,\mathbf{i} + (3t^2 - 1)\,\mathbf{j} + (t + 1)\,\mathbf{k}\big]\,dt.$$

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

To find the constant of integration $\vec{D}$, plug in $\vec{r}(0) = \langle 1, 0, 0 \rangle$ to get $\vec{D} = \mathbf{i}$. Finally,

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

If $\vec{a}(t)$ is a known acceleration vector, then velocity is given by

$$\vec{v}(t) = \vec{v}(t_0) + \int_{t_0}^{t} \vec{a}(u)\, du.$$

Similarly, if velocity is known, then position is given by:

$$\vec{r}(t) = \vec{r}(t_0) + \int_{t_0}^{t} \vec{v}(u)\, du.$$

If the force acting on a particle is known, the acceleration can be determined using **Newton's second law**:

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

EXAMPLE 3.22

Suppose an object of mass m moves in a circular path at constant angular speed ω, with position vector $\vec{r}(t) = a\cos(\omega t)\,\mathbf{i} + a\sin(\omega t)\,\mathbf{j}$. Find the force vector.

Solution:

We differentiate to obtain velocity and acceleration:

$$\begin{aligned}\vec{v}(t) &= \vec{r}\,'(t) = -a\omega\sin(\omega t)\,\mathbf{i} + a\omega\cos(\omega t)\,\mathbf{j},\\ \vec{a}(t) &= \vec{v}\,'(t) = -a\omega^2\cos(\omega t)\,\mathbf{i} - a\omega^2\sin(\omega t)\,\mathbf{j}.\end{aligned}$$

Then, using Newton's second law yields

$$\vec{F}(t) = m\vec{a}(t) = -m\omega^2\left(a\cos(\omega t)\,\mathbf{i} + a\sin(\omega t)\,\mathbf{j}\right) = -m\omega^2\vec{r}(t).$$

This result shows that the force vector points opposite to the position vector and therefore always toward the origin. Such a force is called a centripetal (center-seeking) force.

We will now extend to **projectile motion**. Projectile motion is the movement of an object as it is launched through the air, only subject to gravitational acceleration which acts downward:

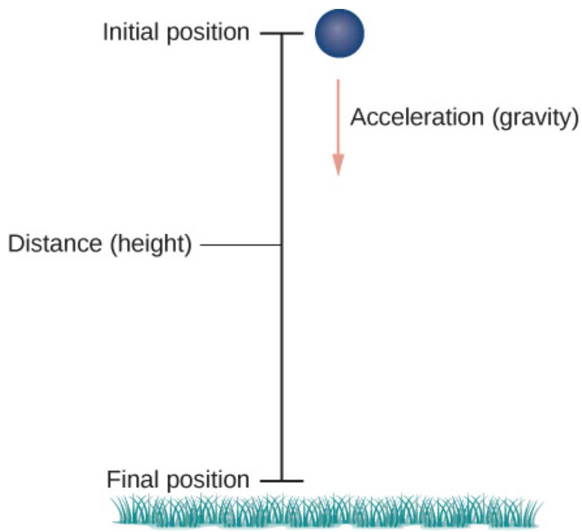


Image credit: Strang & Herman

Let the vertical axis be aligned with $\mathbf{j}$. Ignore air resistance. Newton's second law gives

$$\vec{F}_g = -m\vec{a} = -mg\mathbf{j} \Rightarrow \vec{a}(t) = -g\mathbf{j}.$$
$$\vec{v}\,'(t) = -g\,\mathbf{j}.$$

Let's now integrate. We have

$$\vec{v}(t) = \int -g\,\mathbf{j}\,dt = -gt\,\mathbf{j} + \vec{C}_1,$$

for some constant vector $\vec{C}_1$.

To determine $\vec{C}_1$, we use the initial condition $\vec{v}(0) = \vec{v}_0$.

Substituting this into our velocity equation yields the following:

$$\vec{v}(0) = -g(0)\,\mathbf{j} + \vec{C}_1 = \vec{v}_0 \Rightarrow \vec{C}_1 = \vec{v}_0$$

So the velocity vector becomes

$$\vec{v}(t) = -gt\,\mathbf{j} + \vec{v}_0.$$

and position becomes

$$\vec{s}(t) = -\frac{1}{2}gt^2\mathbf{j} + \vec{v}_0 t + \vec{s}_0.$$

Let's now factor in what happens if we launch from an angle.

We assume an object is launched from the origin with initial speed v_0 at an angle θ above the horizontal. The motion occurs in a vertical plane under the influence of gravity g, with no air resistance.

The position of the object at time t is given by these *scalar* parametric equations:

$$x(t) = (v_0 \cos\theta)\,t,\;\; y(t) = (v_0 \sin\theta)\,t - \frac{1}{2}gt^2$$

To find the horizontal range d, we solve for the time when the object returns to the ground, or when $y = 0$. Solving $0 = (v_0 \sin\theta)\,t - \frac{1}{2}gt^2 = t\left(v_0 \sin\theta - \frac{1}{2}gt\right)$ yields $t = 0$ or $t = \frac{2v_0 \sin\theta}{g}$.

The second root gives the total time of flight. Plug this into $x(t)$ to find the horizontal range:

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

Thus, the range is maximized when $\sin(2\theta) = 1$, which occurs when

$$\theta = \frac{\pi}{4} = 45^\circ.$$

Let's now transition to using *vectors*.

Suppose an object is launched from the origin at time $t = 0$, with initial speed v_0 at an angle θ above the horizontal. Then, we decompose the initial velocity vector $\vec{v}_0$ into horizontal and vertical components:

$$\vec{v}_0 = v_0 \cos\theta\,\mathbf{i} + v_0 \sin\theta\,\mathbf{j}.$$

We can use the fact that all acceleration is due to gravity and say $\vec{a}(t) = -g\,\mathbf{j}$. Then we can integrate with respect to time to obtain the position function $\vec{s}(t)$. The velocity function is:

$$\vec{v}(t) = \vec{v}_0 - gt\,\mathbf{j} = v_0\cos\theta\,\mathbf{i} + (v_0\sin\theta - gt)\,\mathbf{j}.$$

Integrating this velocity vector yields the position vector:

$$\begin{aligned}\vec{s}(t) &= \int \vec{v}(t)\,dt = \int (v_0\cos\theta\,\mathbf{i} + (v_0\sin\theta - gt)\,\mathbf{j})\,dt.\\ &= v_0t\cos\theta\,\mathbf{i} + \left(v_0t\sin\theta - \frac{1}{2}gt^2\right)\mathbf{j}.\end{aligned}$$

The coefficient of $\mathbf{i}$, $v_0t\cos\theta$, gives the horizontal displacement at time t. The coefficient of $\mathbf{j}$, $v_0t\sin\theta - \frac{1}{2}gt^2$, gives the vertical displacement at time t. Maximum height is when $v_y(t) = 0$. In other words, when

$$t = \frac{v_0\sin\theta}{g}.$$

Total flight time is

$$t = \frac{2v_0\sin\theta}{g}.$$

Substituting $t = \dfrac{2v_0\sin\theta}{g}$ into $\vec{s}(t)$:

$$\vec{s}\left(\frac{2v_0\sin\theta}{g}\right) = v_0\left(\frac{2v_0\sin\theta}{g}\right)\cos\theta\,\mathbf{i} + \left[v_0\left(\frac{2v_0\sin\theta}{g}\right)\sin\theta - \frac{1}{2}g\left(\frac{2v_0\sin\theta}{g}\right)^2\right]\mathbf{j}$$

The vertical term simplifies to zero, so we are left with:

$$\vec{s}\left(\frac{2v_0\sin\theta}{g}\right) = \left(\frac{2v_0^2\sin\theta\cos\theta}{g}\right)\mathbf{i} = \left(\frac{v_0^2\sin 2\theta}{g}\right)\mathbf{i}$$

So, the maximum horizontal distance, or range, is

$$R = \frac{v_0^2\sin 2\theta}{g}\mathbf{i}.$$

To maximize the range, we differentiate to find an angle that maximizes the range of the projectile:

$$\frac{d}{d\theta}\left(\frac{v_0^2 \sin 2\theta}{g}\right) = \frac{2v_0^2 \cos 2\theta}{g} = 0$$

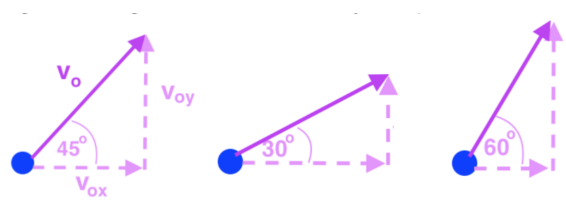
$$\cos 2\theta = 0 \Rightarrow \theta = 45^\circ$$

Therefore, the maximum range occurs when:

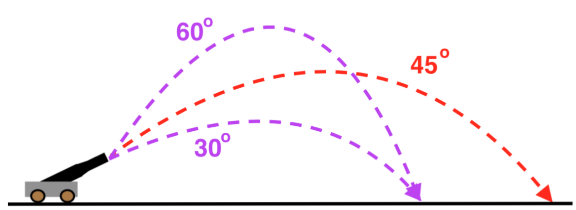
$$R = \frac{v_0^2 \sin 90^\circ}{g} = \frac{v_0^2}{g}$$

$$\vec{s}_{\text{max}} = \left(\frac{v_0^2}{g}\right)\mathbf{i}$$

We just analyzed the same physical motion using both a scalar and vector approach, and the result ($\theta = 45^\circ$) was the same! We looked at it from two different perspectives, but make no mistake: the physics behind it all do not change.



This shows projectile motion broken down into velocity components. In particular, this shows how changing the launch angle changes the velocity components. Image credit: Khan Academy



Higher launch angles have higher maximum height, but 45° is the maximized range. Image credit: Khan Academy

EXAMPLE 3.23

An archer fires an arrow at an angle of 40° above the horizontal with an initial speed of 98 m/s. The height of the archer is 1.715 m. Find the horizontal distance the arrow travels before it hits the ground.

Solution:

We have $s_0 = 1.715\,\text{m}$. We start by decomposing the initial velocity:

$$\vec{v}_0 = 98\cos(40°)\,\mathbf{i} + 98\sin(40°)\,\mathbf{j}$$

Using projectile motion with initial height, the position function is

$$\vec{s}(t) = v_0 t\cos\theta\,\mathbf{i} + \left(s_0 + v_0 t\sin\theta - \frac{1}{2}gt^2\right)\mathbf{j}.$$

We find when the projectile hits the ground by setting the vertical component to zero:

$$s_0 + v_0 t\sin\theta - \frac{1}{2}gt^2 = 0$$

We plug in the known values to get $1.715 + 98t\sin(40°) - \frac{1}{2}(9.8)t^2 = 0$. Solving this yields $t = 12.8829$ s. We can then find the horizontal distance the arrow travels before hitting the ground by plugging our values into the horizontal component of the position function:

$$\begin{aligned} x(t) &= 98t\cos 40° \\ &= (98\ \text{m/s})(12.8829\ \text{s})\cos 40° \\ &= 967.15\ \text{m} \end{aligned}$$

When it comes to studying particle motion, we can **decompose acceleration into two components**, one in the direction of the *tangent* and the other in the direction of the *normal.*

Let $v = \|\vec{v}(t)\|$ be the speed of the particle. Then the unit tangent vector is

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

Thus, $\vec{v} = v\,\mathbf{T}$.

Now differentiate both sides with respect to t:

$$\vec{a} = \vec{v}\,' = \frac{dv}{dt}\,\mathbf{T} + v\,\frac{d\mathbf{T}}{dt}$$

We now express $\frac{d\mathbf{T}}{dt}$ in terms of the unit normal vector $\mathbf{N}$. From the curvature formula:

$$\kappa = \frac{\|\mathbf{T}\,'\|}{\|\vec{r}\,'\|} = \frac{\|\mathbf{T}\,'\|}{v} \quad \Rightarrow \quad \|\mathbf{T}\,'\| = \kappa v$$

Since $\mathbf{T}\,' = \|\mathbf{T}\,'\|\,\mathbf{N}$, we have

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

Substitute back into the acceleration formula to get

$$\vec{a} = \frac{dv}{dt}\,\mathbf{T} + v\,(\kappa v\,\mathbf{N}) = \frac{dv}{dt}\,\mathbf{T} + \kappa v^2\,\mathbf{N}$$

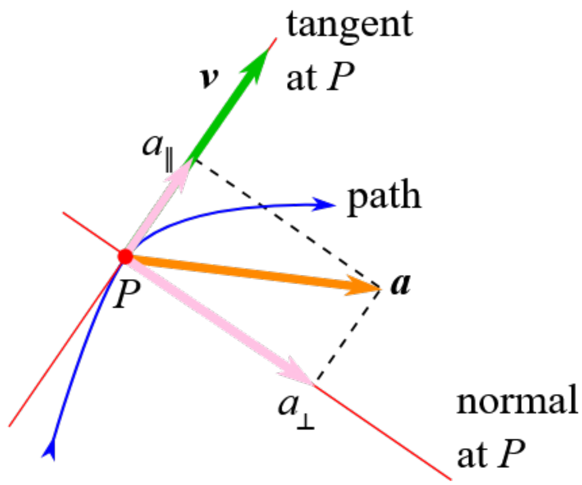
So the acceleration vector decomposes as:

$$\vec{a} = \frac{dv}{dt}\,\mathbf{T} + \kappa v^2\,\mathbf{N}$$

And now we can get our final formula. If an object is moving along a smooth curve $\vec{r}(t)$, we can express its acceleration as the sum of:

$$\vec{a}(t) = a_T\,\mathbf{T}(t) + a_N\,\mathbf{N}(t)$$

where $\mathbf{T}(t)$ is the unit tangent vector (direction of motion), $\mathbf{N}(t)$ is the unit normal vector (direction of curvature), $a_T = \frac{dv}{dt}$ is the tangential component (change in speed), and $a_N = \kappa v^2$ is the normal component (change in direction). You can also think of the tangential and normal components of acceleration as the parallel and perpendicular components, respectively:



The parallel component is aligned with velocity and is what changes speed. The perpendicular component is orthogonal to velocity and is what changes direction. Image credit: University of Manchester

Acceleration only lies in the osculating plane defined by $\mathbf{T}$ and $\mathbf{N}$. The binormal $\mathbf{B}$ is absent. This is because $\mathbf{T}$ represents direction and $\mathbf{N}$ represents direction; these two are all we need to describe acceleration. $\mathbf{B}$ represents *torsion*, or essentially rotation out of the curve's plane, which is not needed. Notice that the tangential component of acceleration is $\frac{dv}{dt}$, the rate of change of speed, whereas the normal component of acceleration is κv^2. κv^2 is curvature times the square of the speed, which is responsible for changing direction. This makes sense in the real world because if you take a sharp turn in a car, κ is large, so the component of the acceleration perpendicular to the motion is also large. This might result in you getting slammed against the car door. Going at a very high speed has an even more meaningful result because it is squared. Think about why the normal component of acceleration is sometimes called the centripetal component of acceleration.

Let's say we want expressions for a_T and a_N that depend only on $\vec{r}(t)$, $\vec{r}\,'(t)$, and $\vec{r}\,''(t)$. We start with the following:

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

Dot both sides with $\vec{v} = v\,\mathbf{T}$. Keep in mind that $\mathbf{T}\cdot\mathbf{T} = 1 \Rightarrow \mathbf{T}\cdot\mathbf{N} = 0$:

$$\vec{v}\cdot\vec{a} = v\,\mathbf{T}\cdot\left(v'\,\mathbf{T} + \kappa v^2\,\mathbf{N}\right) = vv'\,(\mathbf{T}\cdot\mathbf{T}) + \kappa v^3\,(\mathbf{T}\cdot\mathbf{N}) = vv'$$

Therefore

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

To get the normal component, we use the curvature formula:

$$\kappa = \frac{\left\|\vec{r}'(t) \times \vec{r}''(t)\right\|}{\|\vec{r}'(t)\|^3} \Rightarrow a_N = \kappa v^2 = \kappa \left\|\vec{r}'(t)\right\|^2$$

Finally, we have

$$a_N = \frac{\left\|\vec{r}'(t) \times \vec{r}''(t)\right\|}{\|\vec{r}'(t)\|}.$$

These give a_T and a_N without needing unit vectors, which can often be easier to compute.

EXAMPLE 3.24

A particle moves along a space curve with position function $\vec{r}(t) = \langle t^2, t^2, t^3 \rangle$. Find the tangential and normal components of the acceleration vector.

Solution:

We begin by computing the velocity, speed, and acceleration:

$$\vec{r}\,'(t) = \langle 2t, 2t, 3t^2 \rangle \Rightarrow \|\vec{r}\,'(t)\| = \sqrt{(2t)^2 + (2t)^2 + (3t^2)^2} = \sqrt{8t^2 + 9t^4},\ \vec{r}\,''(t) = \langle 2, 2, 6t \rangle$$

Now compute the dot product:

$$\vec{r}\,'(t) \cdot \vec{r}\,''(t) = (2t)(2) + (2t)(2) + (3t^2)(6t) = 4t + 4t + 18t^3 = 8t + 18t^3$$

This gets substituted into the formula for the tangential component:

$$a_T = \frac{\vec{r}\,'(t) \cdot \vec{r}\,''(t)}{\|\vec{r}\,'(t)\|} = \frac{8t + 18t^3}{\sqrt{8t^2 + 9t^4}}$$

Next, compute the cross product $\vec{r}\,'(t) \times \vec{r}\,''(t)$:

$$\vec{r}\,'(t) \times \vec{r}\,''(t) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2t & 2t & 3t^2 \\ 2 & 2 & 6t \end{vmatrix} = (2t)(6t) - (3t^2)(2)\,\mathbf{i} - \left[(2t)(6t) - (3t^2)(2)\right]\mathbf{j} = \langle 6t^2, -6t^2, 0 \rangle$$

$$\left\|\vec{r}\,'(t) \times \vec{r}\,''(t)\right\| = \sqrt{(6t^2)^2 + (-6t^2)^2} = \sqrt{72t^4} = 6\sqrt{2}\,t^2$$

So the normal component is

$$a_N = \frac{\left\|\vec{r}\,'(t) \times \vec{r}\,''(t)\right\|}{\|\vec{r}\,'(t)\|} = \frac{6\sqrt{2}\,t^2}{\sqrt{8t^2 + 9t^4}}.$$

Astronomer Johannes Kepler formulated three laws that describe the motion of planets. His work was published in 1609, which was 78 years before Newton published *Principia Mathematica* in 1687. Kepler's first law, also known as the law of ellipses, says that the path of any planet around the Sun follows an elliptical orbit with the Sun at one focus of the ellipse. Kepler's second law, also known as the law of equal areas, says that a line drawn from the center of the Sun to the center of a planet sweeps out equal areas in equal times. This means that a planet moves faster when it's closer to the Sun (at perihelion) and slower when it's farther from the Sun (at aphelion). Finally, Kepler's third law, also known as the law of harmonies, says that the ratio of the squares of the periods of any two planets is equal to the ratio of the cubes of the lengths of their semi-major orbital axes.

Kepler's laws are among the most important in all of physics, but Newton provided the rigorous mathematical foundation for all of it. Indeed, Kepler's laws are consequences of Newton's second law, law of universal gravitation, and work on calculus.

Assuming that the Sun is at the origin and a planet of mass m has position vector $\vec{r}(t)$. Newton's second law gives $\vec{F} = m\vec{a}$ and his law of gravitation gives

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

where $\vec{F}$ is the gravitational force on the planet, m is the mass of the planet, M is the mass of the sun, G is the gravitational constant, $r = \|\vec{r}\|$, and $\mathbf{u} = (\frac{1}{r})\vec{r}$ is the unit vector in the direction of $\vec{r}$.

From this, we find that

$$\vec{a} = -\frac{GM}{r^3}\vec{r}.$$

Since acceleration $\vec{a}$ is always parallel to $\vec{r}$, it follows that $\vec{r} \times \vec{a} = 0$, and so

$$\frac{d}{dt}\left(\vec{r} \times \vec{v}\right) = \vec{r}\,' \times \vec{v} + \vec{r} \times \vec{v}\,' = \vec{v} \times \vec{v} + \vec{r} \times \vec{a} = \vec{0}.$$

Therefore $\vec{r} \times \vec{v} = \vec{h}$ where $\vec{h}$ is a constant vector. We may also assume that $\vec{h}$ is nonzero and that $\vec{r}$ and $\vec{v}$ are not parallel. $\vec{h}$ is orthogonal to the plane of motion, meaning the planet moves in a fixed plane. The orbit is a plane curve.

We will now rewrite $\vec{h}$:

$$\vec{h} = \vec{r} \times \vec{v} = \vec{r} \times \vec{r}\,' = r\,\mathbf{u} \times (r\,\mathbf{u})'$$

Differentiate using the product rule:

$$= r\,\mathbf{u} \times (r'\,\mathbf{u} + r\,\mathbf{u}') = r\,r'\,\mathbf{u} \times \mathbf{u} + r^2\,\mathbf{u} \times \mathbf{u}'$$

Keep in mind that $\mathbf{u} \times \mathbf{u} = \vec{0}$:

$$\vec{h} = r^2\,\mathbf{u} \times \mathbf{u}'$$

We now have

$$\vec{a} \times \vec{h} = -\frac{GM}{r^2}\,\mathbf{u} \times \vec{h}$$

Keeping in mind that $\mathbf{u} \cdot \mathbf{u} = 1$ and $\|\mathbf{u}(t)\| = 1$ means that $\mathbf{u} \cdot \mathbf{u}' = 0$, we have

$$\vec{a} \times \vec{h} = GM\,\mathbf{u}.$$

Integrating both sides yields

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

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

At this point, it would be convenient to choose the coordinate axes so that the standard basis vector $\mathbf{k}$ points in the direction of the vector $\vec{h}$. This means that the planet moves in the xy-plane. Since $\vec{v} \times \vec{h}$ and $\mathbf{u}$ are perpendicular to $\vec{h}$, it follows that c lies in the xy-plane. We must then choose the x- and y-axes so that the vector $\mathbf{i}$ lies in the direction of $\vec{c}$. This simplifies the expression because now $\vec{c} = c\,\mathbf{i}$, and the angle θ becomes the polar angle between $\vec{r}$ and the x-axis.

Take the dot product of both sides with $\vec{r}$:

$$\vec{r} \cdot (\vec{v} \times \vec{h}) = GM\,\vec{r} \cdot \mathbf{u} + \vec{r} \cdot \vec{c} = GMr + r\,\|\vec{c}\|\cos\theta$$

This becomes

$$\vec{r} \cdot (\vec{v} \times \vec{h}) = r \cdot (GM + c\cos\theta) \Rightarrow r = \frac{\vec{r} \cdot (\vec{v} \times \vec{h})}{GM + c\cos\theta}.$$

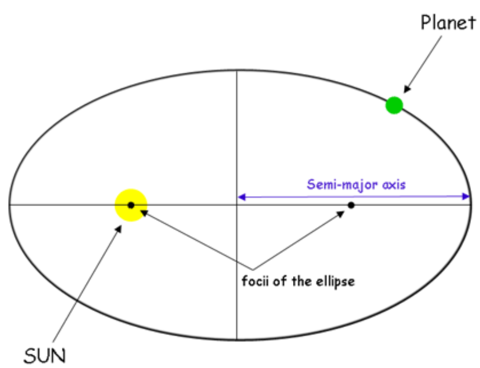
Letting eccentricity $e = \dfrac{c}{GM}$, we get

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

Finally, if we let $d = \dfrac{h^2}{c}$, then

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

This is the polar equation of a conic section with focus at the origin, semi-latus rectum d which controls the scale of the ellipse, and eccentricity e. A planet that stays in orbit must have an orbit that is a closed curve. Thus, this cannot be a parabola or a hyperbola; it must be an ellipse. And with that, we have proved Kepler's first law.



A visualization of Kepler's first law. Image credit: ESO

Let's move on to Kepler's second law. We express the position vector of the planet in polar coordinates as

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

To compute angular momentum, we evaluate the cross product $r \times v$. First, we differentiate:

$$v = \frac{d}{dt}\left(r\cos\theta\,\mathbf{i} + r\sin\theta\,\mathbf{j}\right) = r'\cos\theta\,\mathbf{i} - r\frac{d\theta}{dt}\sin\theta\,\mathbf{i} + r'\sin\theta\,\mathbf{j} + r\frac{d\theta}{dt}\cos\theta\,\mathbf{j}$$

Group terms as such:

$$v = r'(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) + r\frac{d\theta}{dt}(-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j})$$

Now the position vector is

$$r = r(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}).$$

Then the angular momentum is

$$\vec{h} = r \times v = r(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) \times \left[r'(\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) + r\frac{d\theta}{dt}(-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j})\right].$$

The cross product of any vector with itself is zero, so the first term vanishes. The only nonzero part is

$$\vec{h} = r \cdot r\frac{d\theta}{dt}\left((\cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}) \times (-\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j})\right).$$

This simplifies to

$$\vec{h} = r^2\frac{d\theta}{dt}\,\mathbf{k}.$$

So the magnitude of angular momentum is

$$\|\vec{h}\| = r^2\frac{d\theta}{dt}.$$

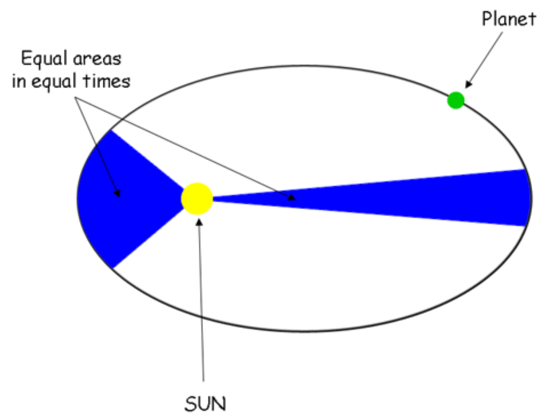
Now consider the area $A(t)$ swept out by the radius vector. In polar coordinates, the infinitesimal area swept is

$$dA = \frac{1}{2}r^2 d\theta \Rightarrow \frac{dA}{dt} = \frac{1}{2}r^2\frac{d\theta}{dt}.$$

Substitute the expression for $r^2\dfrac{d\theta}{dt}$:

$$\frac{dA}{dt} = \frac{1}{2}h$$

Since h is constant (from conservation of angular momentum), the rate of area sweep is constant. This proves Kepler's second law.



A visualization of Kepler's second law. Image credit: ESO

Let T be the period of a planet orbiting the Sun, the time it takes to complete one full revolution around its elliptical orbit. Suppose the ellipse has semi-major axis a and semi-minor axis b, so the total area enclosed by the orbit is πab.

From Kepler's second law, we know that the rate at which area is swept out is constant:

$$\frac{dA}{dt} = \frac{1}{2}h$$

Over one full revolution, the planet sweeps out the total area $A = \pi ab$. So we integrate over one full period:

$$\int_0^T \frac{dA}{dt}\,dt = A = \pi ab = \frac{1}{2}hT$$

Solving for T, we find:

$$T = \frac{2\pi ab}{h}$$

Next, recall that the polar equation of the orbit is:

$$r = \frac{ed}{1 + e\cos\theta} \quad \text{where} \quad d = \frac{h^2}{GM}$$

From conic geometry, the semi-latus rectum d of an ellipse is related to the axes by

$$d = \frac{b^2}{a}.$$

So we substitute to get

$$\frac{h^2}{GM} = \frac{b^2}{a} \Rightarrow h^2 = \frac{GMb^2}{a}.$$

Now plug this into our earlier formula for T:

$$T = \frac{2\pi ab}{h} \Rightarrow T^2 = \frac{4\pi^2 a^2 b^2}{h^2}$$

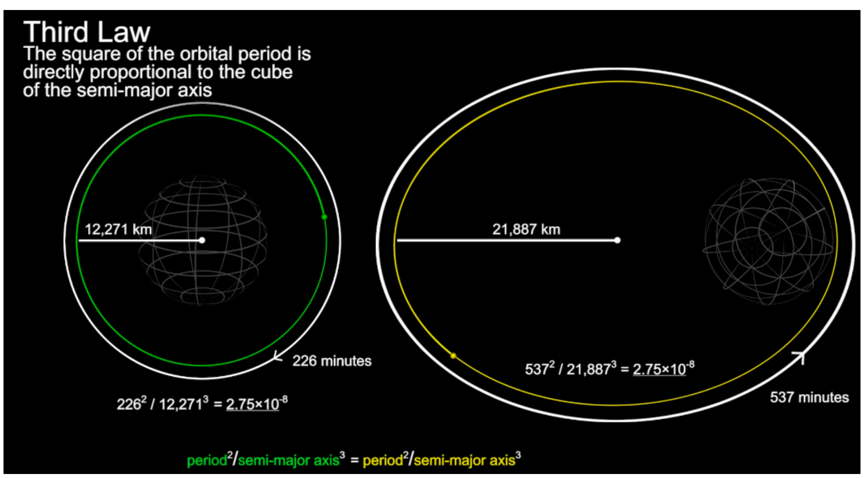
Substitute for h^2 to get

$$T^2 = \frac{4\pi^2 a^2 b^2}{\frac{GMb^2}{a}} = \frac{4\pi^2 a^3}{GM}.$$

This gives the final form:

$$T^2 = \frac{4\pi^2}{GM} a^3$$

This proves Kepler's third law.



A visualization of Kepler's third law. Image credit: NASA

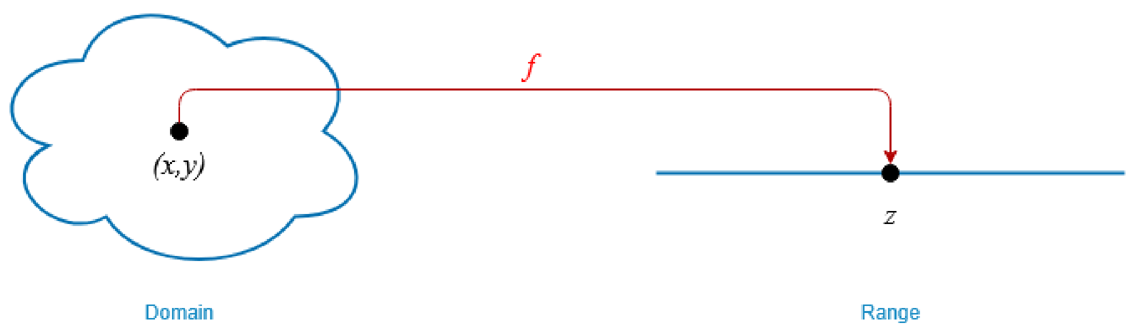
4 Multivariable Differentiation

Most of, if not all, the rigorous problems in your learning journey up to this point have conveniently been dependent on only one variable. However, the reality is that much of the real world's quantities depend on more than one variable. In this chapter, we will learn how to use and apply multivariable functions.

4.1 Multivariable Functions

For functions of a single variable, we map values of one variable to values of another variable. For functions of multiple variables, we map *multiple variables* to another variable.

A function of two variables $z = f(x, y)$ maps each ordered pair (x, y) in a subset D of the real plane $\mathbb{R}^2$ to a unique real number z. The set D is the domain of the function and its range is the subset of all real numbers $\mathbb{R}$ that has at least one ordered pair $(x, y) \in D$ such that $f(x, y) = z$.



For instance, wind chill refers to how cold it truly feels when it's windy. The measurement is known as the wind chill index W, and is dependent on the actual air temperature T and wind speed v. This would be written as $W = f(T, v)$. The following table contains the values of W:

Wind Chill Chart

Wind (mph)

Temperature (°F)

Calm	5	10	15	20	25	30	35	40	45	50	55	60
40	36	34	32	30	29	28	28	27	26	26	25	25
35	31	27	25	24	23	22	21	20	19	19	18	17
30	25	21	19	17	16	15	14	13	12	12	11	10
25	19	15	13	11	9	8	7	6	5	4	4	3
20	13	9	6	4	3	1	0	-1	-2	-3	-3	-4
15	7	3	0	-2	-4	-5	-7	-8	-9	-10	-11	-11
10	1	-4	-7	-9	-11	-12	-14	-15	-16	-17	-18	-19
5	-5	-10	-13	-15	-17	-19	-21	-22	-23	-24	-25	-26
0	-11	-16	-19	-22	-24	-26	-27	-29	-30	-31	-32	-33
-5	-16	-22	-26	-29	-31	-33	-34	-36	-37	-38	-39	-40
-10	-22	-28	-32	-35	-37	-39	-41	-43	-44	-45	-46	-48
-15	-28	-35	-39	-42	-44	-46	-48	-50	-51	-52	-54	-55
-20	-34	-41	-45	-48	-51	-53	-55	-57	-58	-60	-61	-62
-25	-40	-47	-51	-55	-58	-60	-62	-64	-65	-67	-68	-69
-30	-46	-53	-58	-61	-64	-67	-69	-71	-72	-74	-75	-76
-35	-52	-59	-64	-68	-71	-73	-76	-78	-79	-81	-82	-84
-40	-57	-66	-71	-74	-78	-80	-82	-84	-86	-88	-89	-91
-45	-63	-72	-77	-81	-84	-87	-89	-91	-93	-95	-97	-98

Frostbite Times

30 minutes
10 minutes
5 minutes

Wind Chill Chart. Image credit: NWS

When the temperature is 5°F and the wind speed is 40 mph, your body would feel as if it were −22°F. This would be written as

$$f(5, 40) = -22.$$

And this would mean that if you don't get to warmth within 30 minutes, you would get frostbite!

EXAMPLE 4.1

Find the domain and range of each of the following functions:

(a) $f(x, y) = 3x + 5y + 2$

(b) $g(x, y) = \sqrt{9 - x^2 - y^2}$

Solution:

(a) This is a linear function in two variables. There are no values that could cause either variable to be undefined, so the function is defined for all real inputs. Therefore, the domain is $\mathbb{R}^2$. To determine the range, note that for any real number z, we can solve the equation

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

We can set $y = 0$ to get a solution $(\frac{z-2}{3}, 0)$. This shows that every real z has at least one corresponding $(x, y) \in \mathbb{R}^2$ such that $f(x, y) = z$. Thus, the range is $\mathbb{R}$.

EXAMPLE 4.1 (CONTINUED)

(b) The function $g(x, y)$ contains a square root, so the expression inside must be nonnegative:

$$9 - x^2 - y^2 \geq 0 \Rightarrow x^2 + y^2 \leq 9$$

This inequality describes a solid disk of radius 3 centered at the origin. So the domain is:

$$\left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 9\right\}.$$

The maximum value of $g(x, y)$ occurs at the origin, where $x = y = 0$. This gives:

$$g(0, 0) = \sqrt{9 - 0 - 0} = 3$$

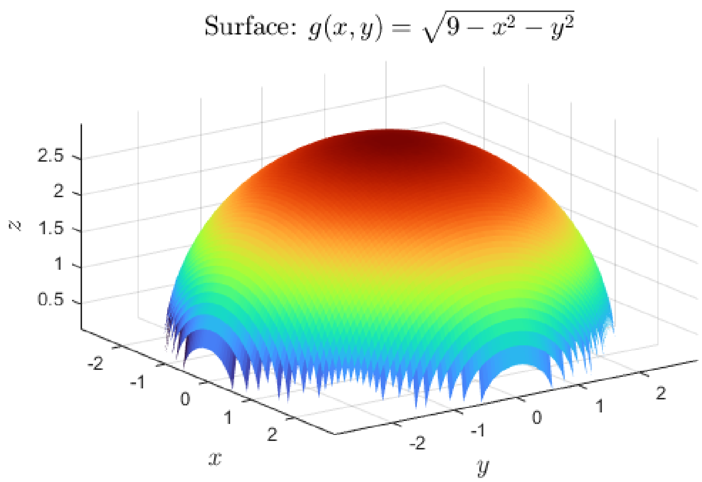
The minimum value occurs on the boundary, where $x^2 + y^2 = 9$, giving:

$$g(x, y) = \sqrt{0} = 0$$

Thus, given any value c between 0 and 3, we can find a set of points inside the domain of g) such that $g(x, y) = c$. That is, $\sqrt{9 - x^2 - y^2} = c$. Simplifying this yields $x^2 + y^2 = 9 - c^2$ which is greater than 0. This describes a circle where any point on it satisfies $g(x, y) = c$. The range is $[0, 3]$.

EXAMPLE 4.1 (CONTINUED)

Here is a mesh surface plot of the range:

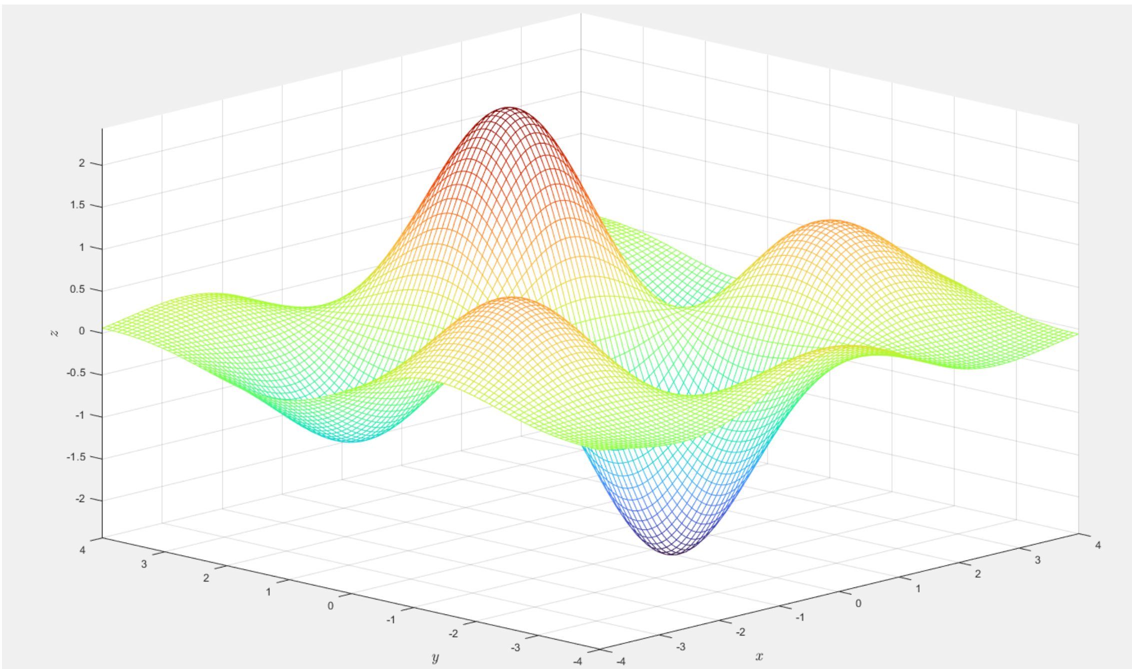


Please run the MATLAB code yourself and have a look!

```
[x, y] = meshgrid(linspace(-3, 3, 100));
z = sqrt(9 - x.^2 - y.^2);
z(imag(z) ~= 0) = NaN;  % Remove imaginary values

figure
surf(x, y, z, 'EdgeColor', 'none')
colormap turbo
axis equal
view(45, 30)

xlabel('$\it{x}$', 'Interpreter', 'latex', 'FontSize', 14)
ylabel('$\it{y}$', 'Interpreter', 'latex', 'FontSize', 14)
zlabel('$\it{z}$', 'Interpreter', 'latex', 'FontSize', 14)
title('Surface: $g(x, y) = \sqrt{9 - x^2 - y^2}$', ...
    'Interpreter', 'latex', 'FontSize', 14)
```

ex4point1.m

Graphing multivariable functions, even with a computer, can be difficult. For a function $z = f(x, y)$, every point in the plane has an ordered pair (x, y) associated with it. Formally, the graph of f with domain D is the set of all points x, y, z in $\mathbb{R}^3$ such that $z = f(x, y) \in D$ and $(x, y) \in D$. You can imagine the xy-plane like a map on a table. Then, every point z in the domain of the function tells you how far up ($z > 0$) or down (negative $z < 0$) you go from that point on the map. Over time, as you plot more points, you begin to trace out a two-dimensional *surface* that is the graph of f.



Mesh Surface Plot of $f(x, y) = e^{-0.1(x^2+y^2)} \cdot 3\cos(x)\sin(y)$

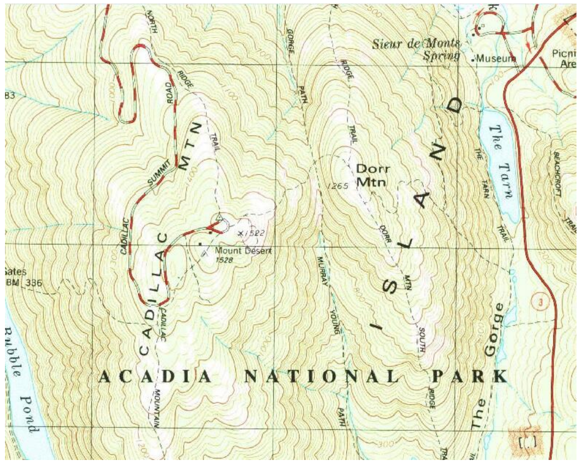
4.2 Level Curves

While mesh surface plots may look like mountains, the resemblance in **level curves** is even more striking.

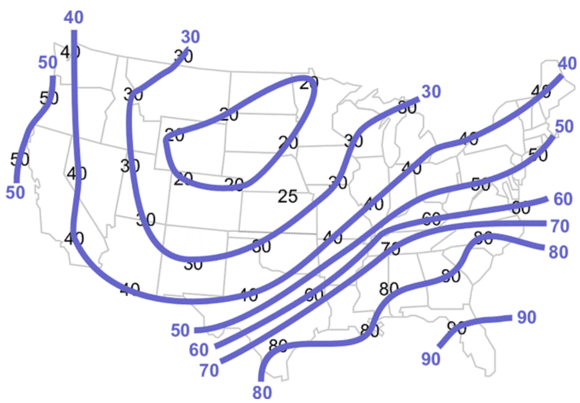
Let $f(x, y)$ with domain $D \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a real-valued function, and let z be be a constant number in the range of f.

The level curve L of f corresponding to the value z is the set of all points $(x, y) \in D$ such that

$$L_z(f) = \{(x, y) \in D \mid f(x, y) = z\}.$$



Topographic Map of Cadillac Mountain in Acadia National Park. Image credit: USGS

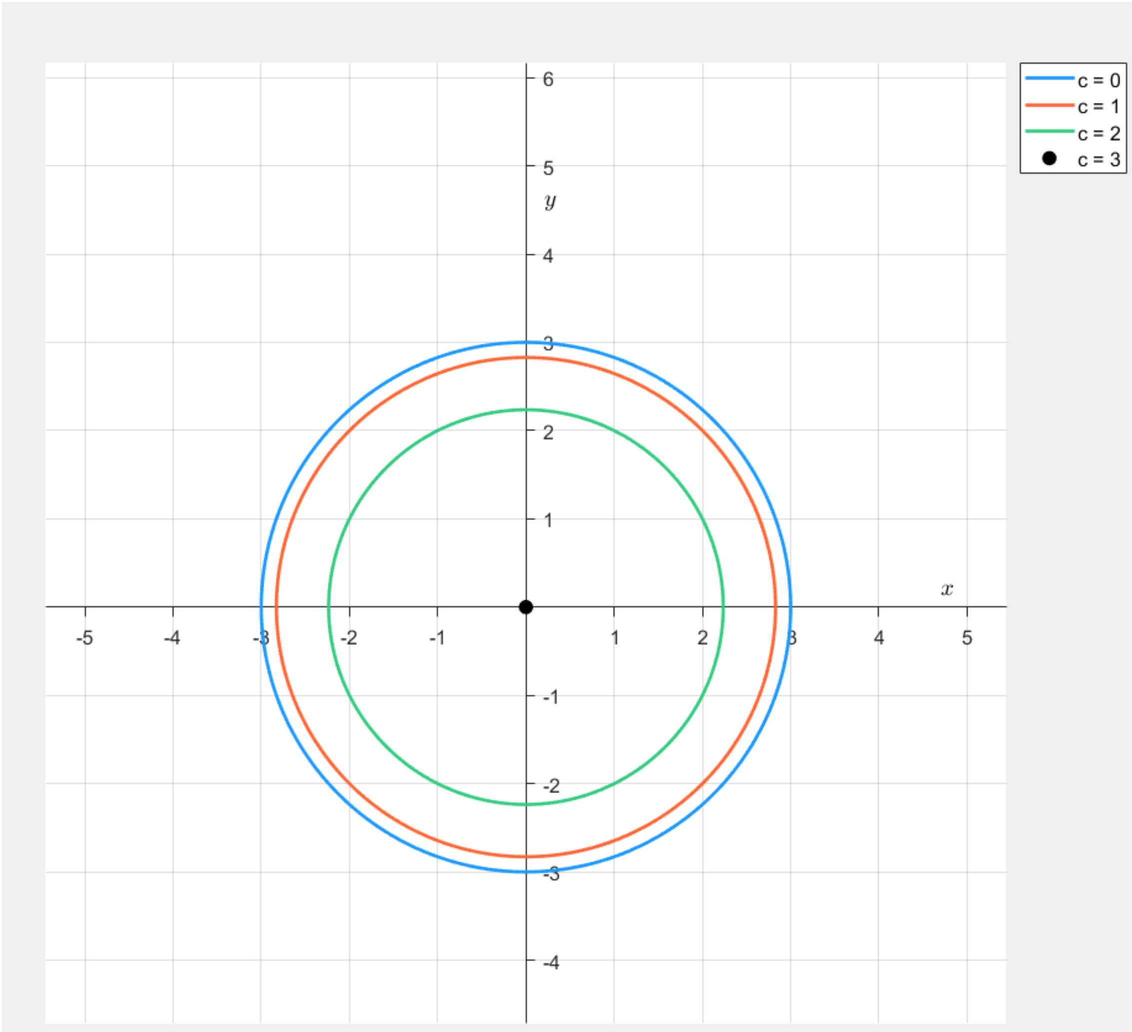


US Surface Temperature Map with Isotherms. Image Credit: NOAA

Recall the function $g(x, y) = \sqrt{9 - x^2 - y^2}$ with range $[0, 3]$. If we pick any number in this interval such as $c = 2$, the corresponding level curve is given by

$$\sqrt{9 - x^2 - y^2} = 2.$$

Solving this yields $x^2 + y^2 = 5$. This is the equation of a circle centered at the origin with radius $\sqrt{5}$. If we were to repeat the process we would have circle equations corresponding to $c = 0, 1, 2,$ and 3. Note that $c = 3$ yields $x^2 + y^2 = 0$ which is simply the origin. Graphing them would look like this:

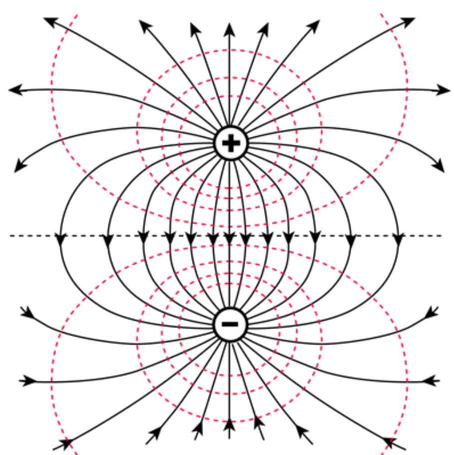


If you were to shade in the blue circle given by $c = 0$, that would give you a graph of the domain of $g(x, y)$. However, the graph here with various circles, or *level curves*, is called a *contour map*. You may recognize level curves in topographical maps, isotherms, isobars, or equipotential lines. The connection between all three of these is that each curve represents something of a constant value.

- In topographic maps, each contour represents constant elevation.
- For isotherms, each curve shows constant temperature.
- For isobars, the curves indicate constant pressure.
- For equipotential lines, they represent constant potential energy.

Contour maps are not a new idea either:

Isotherm Curves of the Northern Hemisphere, 1845. Image credit: Library of Congress



Equipotential lines showing the electric potential of a dipole. Image credit: HyperPhysics

EXAMPLE 4.2

Given the function $f(x, y) = \sqrt{8 + 8x - 4y - 4x^2 - y^2}$, create a contour map. Then, find the domain and range of f.

Solution:

To find the level curve for $c = 0$, we set $f(x, y)$ equal to 0 and solve. This gives $0 = \sqrt{8 + 8x - 4y - 4x^2 - y^2}$. Solving this yields

$$\frac{(x-1)^2}{4} + \frac{(y+2)^2}{16} = 1.$$

This describes an ellipse centered at $(1, -2)$.
Repeating the same process for an arbitrary c, we end up with

$$\frac{4(x-1)^2}{16-c^2} + \frac{(y+2)^2}{16-c^2} = 1.$$

Thus, the level curve for a fixed $c \in [0, 4)$ is an ellipse centered at $(1, -2)$. When $c = 4$, we can solve for the level curve as follows:

$$f(x, y) = \sqrt{8 + 8x - 4y - 4x^2 - y^2} \;\; = \sqrt{16 - 4(x-1)^2 - (y+2)^2} = 4$$

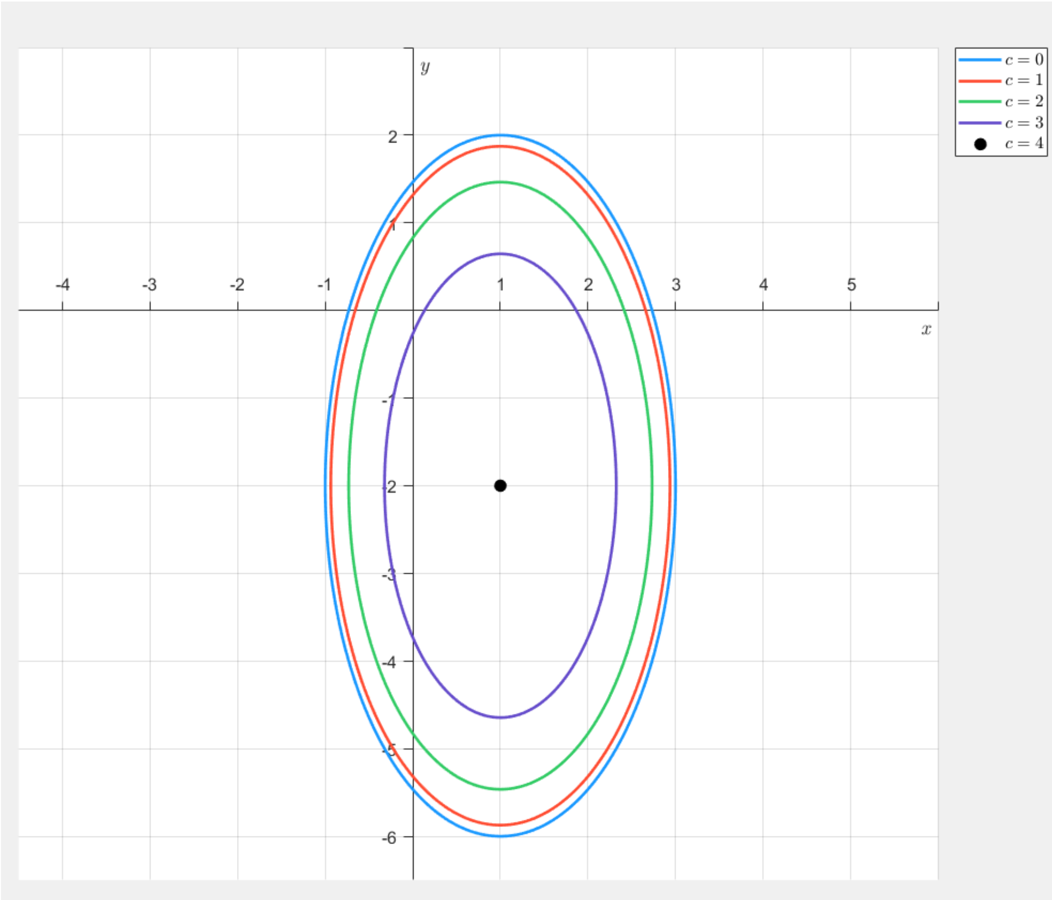
$$4(x-1)^2 + (y+2)^2 = 0$$

$$x = 1, \;\; y = -2$$

Thus, the level curve is the point $(-1, 2)$.

EXAMPLE 4.2 (CONTINUED)

Continue this process and then graph the level curves corresponding to $c = 0, 1, 2, 3$ and 4:



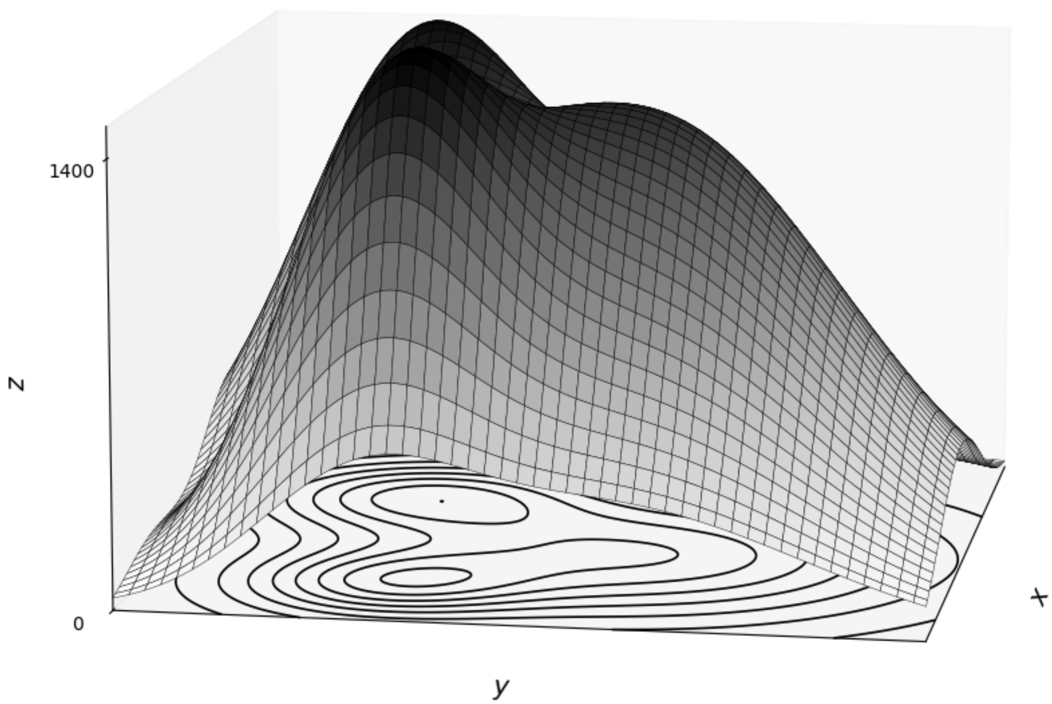
Let's find the domain. This is a square root function, so the radicand must be nonnegative. We have $8+8x-4y-4x^2-y^2 \geq 0$. Equivalently, we have $\frac{(x-1)^2}{4} + \frac{(y+2)^2}{16} \leq 1$. The domain is the ellipse given by $c = 0$ shaded in and centered at $(-1, 2)$. We can write this as

$$\left\{(x, y) \;\middle|\; \frac{(x-1)^2}{4} + \frac{(y+2)^2}{16} \leq 1\right\}.$$

As we can simplify $f(x, y)$ to $\sqrt{16 - 4(x-1)^2 - (y+2)^2}$ to make it easier to work with, the range is simple to find. At $(-1, 2)$, $(x-1)^2 = 0$ and $(y+2)^2 = 0$ which is where the maximum value of the radicand is. The value here is 16 therefore the maximum value is $\sqrt{16} = 4$. The minimum value is 0. Thus, the range of the function is $[0, 4]$.

For fun, here is the surface $f(x,y) = 1400\,e^{-0.02\left[(x-2)^2+(y-2)^2\right]} + 1300\,e^{-0.05\left[(x+4)^2+(y+5)^2\right]} + 1100\,e^{-0.08\left[(x-6)^2+(y+4)^2\right]}$ and its contour map. Feel free to explore using the Python code.

contourMapAndSurface.py



EXAMPLE 4.3

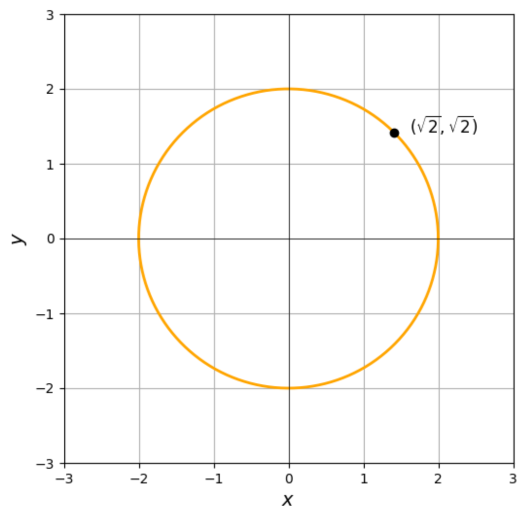
Find a tangent line to the level curve $z_0 = 4$ for the function $z = x^2 + y^2 = f(x, y)$.

Solution:

A level curve associated with $z_0 = 4$ is the subset of $\mathbb{R}^2$ given by

$$L_4(f) = \{(x, y) : f(x, y) = 4\} = \{(x, y) : x^2 + y^2 = 4\}.$$

The contour curve would be given by $C = \{(x, y, 4) : f(x, y) = 4\}$. We will choose the point $(\sqrt{2}, \sqrt{2})$ to find the tangent line which is given by $y - y_1 = m(x - x_1)$:



Since the slope is the derivative of the curve, we are then looking for $m = \frac{dy}{dx}$ evaluated at x_1, y_1. We have to then implicitly differentiate $x^2 + y^2 - 4 = 0$:

$$\frac{d}{dx}[x^2 + y^2 - 4] = \frac{d}{dx}[0]$$
$$\frac{d}{dx}[x^2] + \frac{d}{dx}[y^2] - \frac{d}{dx}[4] = 0$$
$$2x + \frac{d}{dx}[(y(x))^2] - 0 = 0$$

EXAMPLE 4.3 (CONTINUED)

Using the fact that $\frac{d}{dx}[f(g(x))] = f'(g(x))g'(x)$, we now have

$$2x + 2 \cdot y(x) \cdot y'(x) = 0 \Rightarrow y' = \frac{-2x}{2y} = \frac{-x}{y}.$$

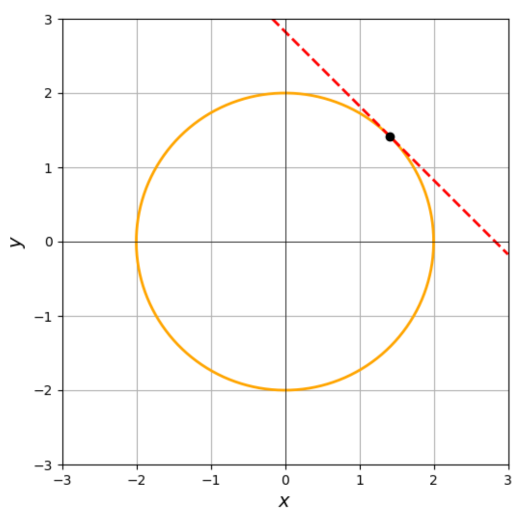
At $(\sqrt{2}, \sqrt{2})$, this gives

$$\frac{dy}{dx} = -\frac{\sqrt{2}}{\sqrt{2}} = -1.$$

So the equation of the tangent line at $(\sqrt{2}, \sqrt{2})$ is

$$y - \sqrt{2} = -1(x - \sqrt{2}) \Rightarrow y = -x + 2\sqrt{2}.$$

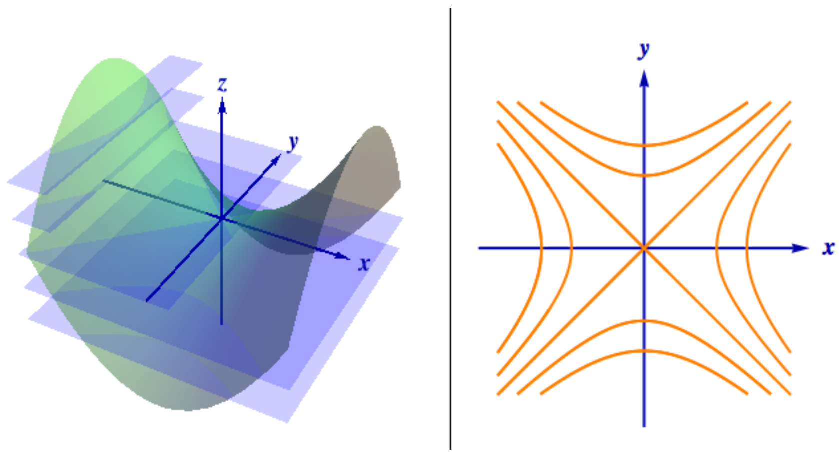
The direction vector of the tangent line is thus $\vec{v} = \langle 1, -1 \rangle$.



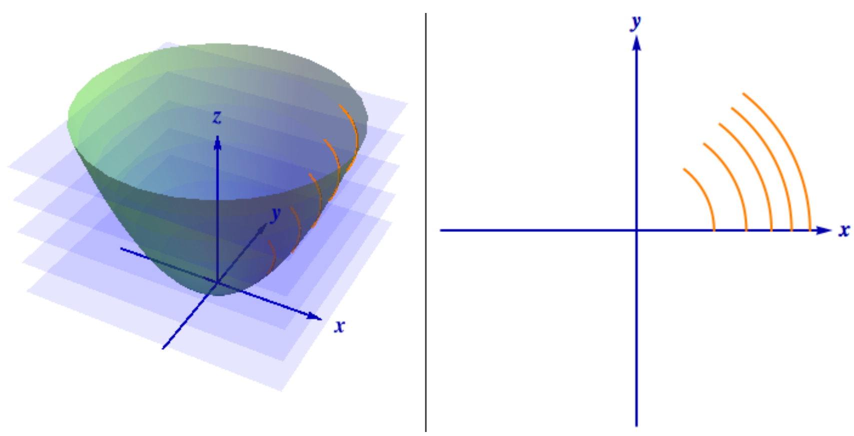
We can write the vector equation of the tangent line using vector form:

$$\vec{r}(t) = \vec{r}_0 + t\vec{v} = \langle \sqrt{2}, \sqrt{2} \rangle + t\langle 1, -1 \rangle = \langle \sqrt{2} + t, \sqrt{2} - t \rangle$$

Another method of visualizing multivariable functions is called a vertical trace. While level curves are graphed in the xy-plane, vertical traces are graphed in the xz- or yz-planes. For a function $z = f(x, y)$ with domain $D \subseteq \mathbb{R}^2$, a vertical trace of f is the set of points that satisfies the equation $f(a, y) = z$ for a given constant $x = a$ or a given constant $y = b$. If you fix $x = a$, you are cutting the surface with a vertical plane parallel to the yz-plane which shows how z changes as y changes at x. Likewise, if you fix $y = b$, you are slicing with a plane parallel to the xz-plane which shows how z changes as x changes at y.



A hyperbolic paraboloid as a surface and its contour map. Image credit: UT Austin



A paraboloid as a surface and its contour map. Image credit: UT Austin

EXAMPLE 4.4

Find vertical traces for the function $f(x, y) = \sin x \cos y$ corresponding to $x = -\frac{\pi}{4}, 0, \frac{\pi}{4}$ and $y = -\frac{\pi}{4}, 0, \frac{\pi}{4}$.

Solution:

We begin with traces parallel to the xz-plane. That is, fix $x = c$. First, we set $x = -\frac{\pi}{4}$:

$$z = \sin\left(-\frac{\pi}{4}\right)\cos y = -\frac{\sqrt{2}}{2}\cos y$$

This gives a cosine curve scaled by $-\frac{\sqrt{2}}{2}$ in the plane $x = \frac{\pi}{4}$. The results are summarized in the following table:

$x = c$	$z = \sin c \cos y$
$x = -\dfrac{\pi}{4}$	$z = -\dfrac{\sqrt{2}}{2}\cos y$
$x = 0$	$z = 0$
$x = \dfrac{\pi}{4}$	$z = \dfrac{\sqrt{2}}{2}\cos y$

Let's graph:

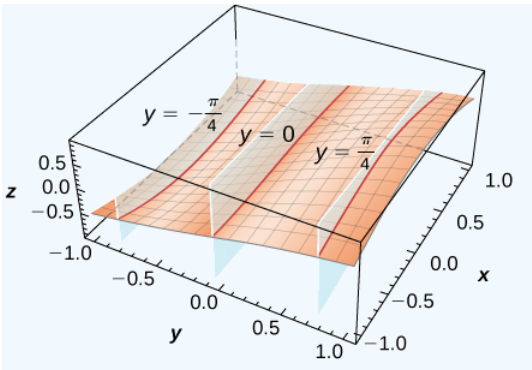


Image credit: Strang & Herman

EXAMPLE 4.4 (CONTINUED)

Now let's find traces parallel to the yz-plane. That is, fix $y = d$. Setting $y = -\frac{\pi}{4}$ yields

$$z = \sin x \cos\left(-\frac{\pi}{4}\right) = \sin x \cdot \frac{\sqrt{2}}{2}$$

This gives a sine curve scaled by $\frac{\sqrt{2}}{2}$ in the plane $y = \frac{\pi}{4}$. The results are summarized in the following table:

$y = d$	$z = \sin x \cos d$
$y = -\frac{\pi}{4}$	$z = \frac{\sqrt{2}}{2} \sin x$
$y = 0$	$z = \sin x$
$y = \frac{\pi}{4}$	$z = \frac{\sqrt{2}}{2} \sin x$

Let's graph:

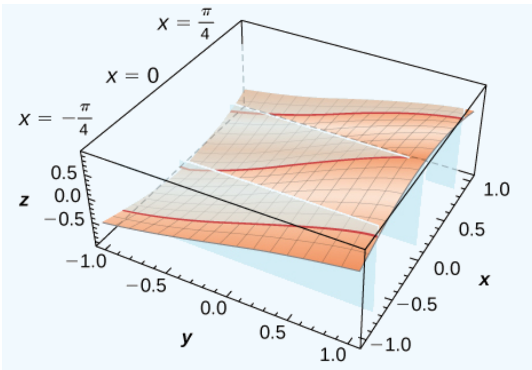


Image credit: Strang & Herman

A function of three variables $f(x, y, z)$ assigns to each ordered triple (x, y, z) in a domain $D \subseteq \mathbb{R}^3$ a unique real number. That is,

$$f : \mathbb{R}^3 \to \mathbb{R},\ f(x, y, z)$$

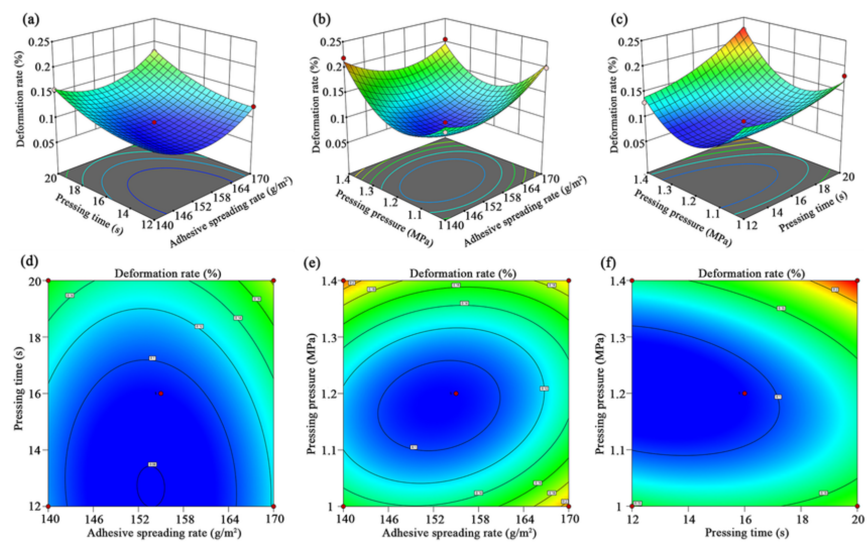
For instance, if we were trying to write a function calculate the reaction rate of an industrial chemical process, a function could take into account three variables: the concentration of the reactants, temperature, and the physical state of the reactants.

Let $f(x, y, z)$ with domain $D \subseteq \mathbb{R}^3 \to \mathbb{R}$ be a real-valued function, and let c be a constant number in the range of f.

The **level surface** S of f corresponding to the value c is the set of all points $(x, y, z) \in D$ such that

$$L_c(f) = \{(x, y, z) \in D \mid f(x, y, z) = c\}.$$

Level surfaces for three variable functions are formed in the same fashion as level curves for two variable functions. You simply set the function equal to a constant c and solve.



Deformation rate of wood flooring. Image credit: Huixiang Wang, Shaanxi Normal University

We can generalize functions so that any number of variables can be considered:

A function of n variables assigns a real number z to an n-tuple $x_1, x_2, ..., x_n)$ of real inputs:

$$z = f(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n,$$

where the domain $D \subseteq \mathbb{R}^n$ and the output lies in $\mathbb{R}$.

This is often written more compactly in vector form as

$$f(\vec{x}) = \vec{c} \cdot \vec{x},$$

where $\vec{c} = \langle c_1, c_2, \ldots, c_n \rangle$ and $\vec{x} = \langle x_1, x_2, \ldots, x_n \rangle$, and the dot product gives the output.

There are three common ways to view a function f defined on a subset of $\mathbb{R}^n$:

1. As a function of n real variables $x_1, x_2, \ldots, x_n$
2. As a function of a point in $\mathbb{R}^n$ $(x_1, x_2, \ldots, x_n)$
3. As a function of a vector $\vec{x} \in \mathbb{R}^n$ $\vec{x} = \langle x_1, x_2, ..., x_n \rangle$

4.3 Limits and Continuity

In single-variable calculus, recall that the limits from both the right-hand and left-hand limits must agree for the limit to exist. We could approach a from $a - \delta$ or from $a + \delta$. In the context of multivariable limits, it's much more complex than that. There are more than two possible directions to approach a from.

We have the *imprecise*, or *rough* definition* of a limit for a function in two variables:

A function $f(x, y)$ of two variables has a limit L as $P(x, y)$ approaches a fixed point $P_0(a, b)$ if

$$|f(x, y) - L| < \varepsilon$$

can be made arbitrary small by forcing the point $P(x, y)$ to be sufficiently close to the point $P_0(a, b) \in D$. If such a limit exists, we write

$$\lim_{(x,y)\to(a,b)} f(x, y) = \lim_{P(x,y)\to P_0(a,b)} f(x, y) = L.$$

*A more rigorous definition that holds up to formal proofs would require real analysis.

In other words, for the multivariable function $f(x, y)$, we say that

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

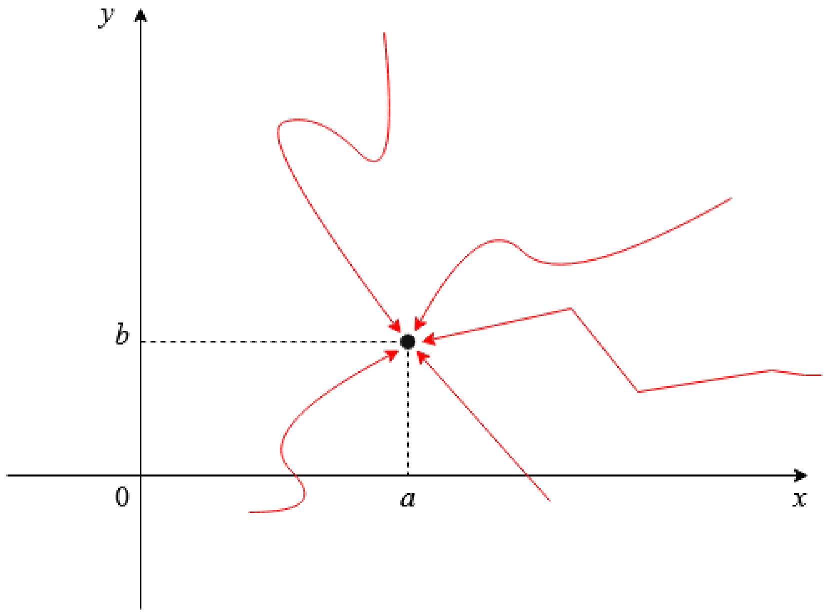
if and only if $f(x, y)$ gets sufficiently close to L as (x, y) gets sufficiently close to (a, b).

How do we define *sufficiently close* enough? We begin with the distance formula:

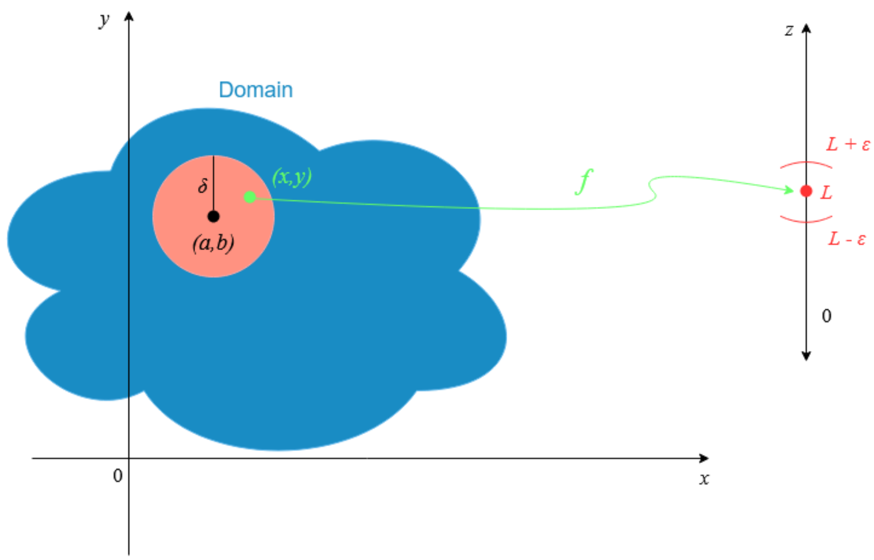
$$\sqrt{(x - a)^2 + (y - b)^2} < \delta$$

The point (x, y) lies inside a circle of radius δ centered at (a, b). In $\mathbb{R}^2$, we interpret "sufficiently close" as being within an open disk (also called a *neighborhood*) around the point (a, b). Geometrically, instead of approaching along a line from the left or right as in single-variable calculus, we allow (x, y) to approach (a, b) from *any direction* within this disk. Note that a disk with the neighborhood boundary included is called a closet set.

This behavior is known as **path independence**. Given that the requirements hold, the limit $\lim_{(x,y)\to(a,b)} f(x, y) = L$ exists regardless of the path we take towards (a, b).



There are an infinite number of directions from which we can approach (a, b).



f maps all the points in the neighborhood, denoted by an orange circle of radius δ, around $(a, b) \subset D$ except (a, b) into the interval $(L - \varepsilon, L + \varepsilon)$.

Limit Laws for Functions of Two Variables

Let $f(x,y)$ and $g(x,y)$ be defined for all $(x,y) \neq (a,b)$ on a neighborhood around (a,b) and suppose

$$\lim_{(x,y)\to(a,b)} f(x,y) = L \quad \text{and} \quad \lim_{(x,y)\to(a,b)} g(x,y) = M,$$

where $L, M \in \mathbb{R}$ and $c \in \mathbb{R}$ is a constant. Then the following rules apply:

Constant Law:

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

Identity Laws:

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

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

Sum Rule:

$$\lim_{(x,y)\to(a,b)} [f(x,y) + g(x,y)] = L + M$$

Limit Laws for Functions of Two Variables (CONTINUED)

Difference Rule:

$$\lim_{(x,y)\to(a,b)}[f(x,y) - g(x,y)] = L - M$$

Constant Multiple Rule:

$$\lim_{(x,y)\to(a,b)}[c \cdot f(x,y)] = cL$$

Product Rule:

$$\lim_{(x,y)\to(a,b)}[f(x,y) \cdot g(x,y)] = LM$$

Quotient Rule:

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y)}{g(x,y)} = \frac{L}{M} \quad \text{if } M \neq 0$$

Power Rule:

$$\lim_{(x,y)\to(a,b)}[f(x,y)]^n = L^n \quad \text{for any integer } n$$

Root Rule:

$$\lim_{(x,y)\to(a,b)} \sqrt[n]{f(x,y)} = \sqrt[n]{L}$$

for all L if n is odd and positive, and for all $L \geq 0$ if n is even and positive.

EXAMPLE 4.5

Evaluate the limit

$$\lim_{(x,y)\to(2,8)} \left(3x^2y + \sqrt{xy}\right).$$

Solution:

We apply the limit laws by splitting the expression into two separate limits:

$$\lim_{(x,y)\to(2,8)} \left(3x^2y + \sqrt{xy}\right) = \lim_{(x,y)\to(2,8)} \left(3x^2y\right) + \lim_{(x,y)\to(2,8)} \sqrt{xy}$$

Using the constant multiple law, we pull out the 3:

$$= 3 \cdot \lim_{(x,y)\to(2,8)} (x^2y) + \lim_{(x,y)\to(2,8)} \sqrt{xy}$$

Now plug in $x = 2$, $y = 8$:

$$= 3 \cdot (2^2 \cdot 8) + \sqrt{2 \cdot 8} = 3 \cdot (4 \cdot 8) + \sqrt{16}$$
$$= 3 \cdot 32 + 4 = 96 + 4 = 100$$

There are cases where limits cannot be computed by direct substitution or may not exist at all.

Two-Path Test for Nonexistence of a Limit

If the multivariable function $f(x, y)$ approaches two different values as the input point (x, y) approaches (a, b) along two different paths in the domain of f, then the limit

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

does not exist.

As soon as you find two paths that disagree, you can conclude that the limit does not exist. There are five easy paths you may want to look at first:

1. $(x, b) \to (a, b)$ which is along the $y = b$ line
2. $(a, y) \to (a, b)$ which is along the $x = a$ line
3. $(x, y) \to (a, b)$ which is along any line with slope m such that $y = m(x - a) + b$
4. $(x, y) \to (a, b)$ which is along a parabola in x through point (a, b) with $y = (x - a)^2 + b$
5. $(x, y) \to (a, b)$ which is along a parabola in y through point (a, b) with $x = (y - b)^2 + a$

EXAMPLE 4.6

Show that the limit

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

does not exist.

Solution:

Path 1: Let $y = b = 0$. Then

$$\lim_{(x,0)\to(0,0)} \frac{x^2 - 0}{x^2 + 0} = \lim_{x\to 0} \frac{x^2}{x^2} = 1$$

Path 2: Let $x = 0$. Then

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

Since $1 \neq -1$, the limit does not approach a single value. Thus, by the two-path test,

$$\lim_{(x,y)\to(0,0)} \frac{x^2 - y^2}{x^2 + y^2} \text{ does not exist.}$$

EXAMPLE 4.7

Let $f(x,y) = \frac{xy^2}{x^2+y^4}$. Does

$$\lim_{(x,y)\to(0,0)} f(x,y)$$

exist?

Solution:

Be careful because the numerator and denominator both go to zero at the origin.

Path 1: Let $y = mx$ arbitrarily. Then

$$\lim_{(x,mx)\to(0,0)} \frac{x\cdot(mx)^2}{x^2+(mx)^4} = \lim_{x\to 0} \frac{m^2\cdot x^3}{x^2+m^4\cdot x^4} = \lim_{x\to 0} \frac{x^2}{x^2}\cdot\frac{m^2\cdot x}{1+m^4x^2} = 0$$

Path 2: Let $x = 0$. Then

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4} = \lim_{y\to 0} \frac{0\cdot y^2}{0^2+y^4} = \lim_{y\to 0} \frac{0}{y^4} = 0$$

So far, both paths give 0. But that isn't sufficient to prove the limit exists. Let's try a more general curve.

Path 3: Let $x = y^2$, which curves into the origin:

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

Since $0 \neq \frac{1}{2}$, the limit does not approach a single value. Thus, by the two-path test,

$$\lim_{(x,y)\to(0,0)} \frac{xy^2}{x^2+y^4} \text{ does not exist.}$$

Rightfully, you may be wondering: what happens if you perform several paths for a variety of values and get the same value every time? Well, you might think that this means the limit exists. In reality, the best you can say is that it *likely* exists. There are infinitely many paths to approach a point and it's impossible to check every single one. You can try are rewriting the limit in polar coordinates and using the squeeze theorem. Ideally, you'll discover that the limit does not exist before having to resort to that which is generally much easier than proving a limit exists.

The same conditions for a single-variable function being continuous must be met for a function of two variables. Let's take a look at that before we move on to a formal definition:

Conditions for Continuity of a Function of Two Variables

A function $f(x, y)$ is continuous at a point (a, b) in its domain if all of the following are true:

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

EXAMPLE 4.8

Show that the function $f(x, y) = \frac{3x+2y}{x+y+1}$ is continuous at the point $(5, -3)$.

Solution:

1. Does $f(5, -3)$ exist?

$$f(5, -3) = \frac{3(5) + 2(-3)}{5 + (-3) + 1} = \frac{15 - 6}{3} = \frac{9}{3} = 3$$

2. Does the limit exist?

   Since $f(x, y)$ is a rational function and the denominator is nonzero at $(5, -3)$, the function is continuous wherever it is defined. Additionally,

$$\lim_{(x,y)\to(5,-3)} f(x, y) = 3 = f(5, -3).$$

3. Are the function value and the limit equal? As seen previously, they are equal.

All three conditions are satisfied therefore the function is continuous at $(5, -3)$.

Here is the formal definition:

Continuity of a Function of Two Variables

A function $f(x, y)$ is continuous at a point (x_0, y_0) in an open region $R \subseteq \mathbb{R}^2$ if the following is true:

$$\lim_{(x,y)\to(x_0,y_0)} f(x, y) = f(x_0, y_0)$$

We say f is continuous on R if it is continuous at every point $(x_0, y_0) \in R$.

When we say a function is continuous, what we mean is that small changes in (x, y) translate to small changes in $f(x, y)$. This means that things like jumps and holes are absent; the graph should be smooth. If you build one large function using continuous terms, it follows that the new function will also be continuous. This is why polynomials are always continuous. A polynomial is simply a sum of terms of the form $cx^m y^n$ where c is a constant, $m \leq 0$, and $n \leq 0$. It's simply a sum of continuous things.

The Sum of Continuous Functions is Continuous

If $f(x, y)$ is continuous at (x_0, y_0) and $g(x, y)$ is continuous at (x_0, y_0), then the sum $f(x, y) + g(x, y)$ is also continuous at (x_0, y_0).

The Product of Continuous Functions is Continuous

If $g(x)$ is continuous at x_0 and $h(y)$ is continuous at y_0, then the function $f(x, y) = g(x)h(y)$ is continuous at (x_0, y_0).

Continuity of a Composite Function

Let g be a function of two variables with domain $D \subseteq \mathbb{R}^2$ and range $R \subseteq \mathbb{R}$. Suppose g is continuous at some point $(x_0, y_0) \in D$ and let $z_0 = g(x_0, y_0)$.

Let f be a function that maps $R \to \mathbb{R}$ such that z_0 is in the domain of f, and suppose f is continuous at z_0. Assume f is continuous at z_0.

Then, the composition $f \circ g$ is continuous at (x_0, y_0).

EXAMPLE 4.9

Show that the functions $f(x,y) = 4x^3y^2$ and $g(x,y) = \cos(4x^3y^2)$ are continuous everywhere.

Solution:

The function $f(x,y) = 4x^3y^2$ is a polynomial with two polynomial terms, and polynomials are continuous at every point in $\mathbb{R}^2$. Therefore, as $f(x)$ represents a product of two continuous functions, it is continuous everywhere.

Notice that $g(x,y) = \cos(f(x,y))$, which means we are just applying the cosine function to the output of f. $\cos x$ is continuous at every real number and we have already established that $f(x,y)$ is continuous at every point (x,y) in the xy-plane. Thus, as we are composing two continuous functions, $g(x,y)$ is continuous at every point (x,y) in the xy-plane.

When it comes to taking the limit of functions of three or more variables, we have to extend our disk of radius δ into more than two dimensions.

Continuity of a Function of Three Variables

A function $f(x,y,z)$ is continuous at a point (x_0, y_0, z_0) in an open region $R \subseteq \mathbb{R}^3$ if the following is true:

$$\lim_{(x,y,z)\to(x_0,y_0,z_0)} f(x,y,z) = f(x_0,y_0,z_0)$$

The function f is continuous on R if it is continuous at every point $(x_0,y_0,z_0) \in R$.

We will now reestablish our "sufficiently close" criteria. Let $f(x,y,z)$ be a function defined on a domain D in $\mathbb{R}^3$, and suppose we are interested in

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

The distance between the point (x,y,z) and (a,b,c) is given by

$$\sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2}.$$

If for every number $\varepsilon > 0$, there exists a number $\delta > 0$ such that $(x,y,z) \in D$ and $0 < \sqrt{(x-a)^2 + (y-b)^2 + (z-c)^2} < \delta$, then

$$|f(x,y,z) - L| < \varepsilon.$$

We extend the idea of the disk to a ball.

Let point $(x_0, y_0, z_0) \in \mathbb{R}^3$. A ball centered at (x_0, y_0, z_0) with radius δ in three dimensions consists of all points in $\mathbb{R}^3$ that are less than distance δ away from (x_0, y_0, z_0). That is,

$$\left\{(x, y, z) \in \mathbb{R}^3 \,\middle|\, \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} < \delta\right\}.$$

To define a ball in higher dimensions, simply add terms under the radical corresponding to each additional coordinate. For example, given a point

$$P = (w_0, x_0, y_0, z_0) \in \mathbb{R}^4,$$

a ball centered at P is

$$\left\{(w, x, y, z) \in \mathbb{R}^4 \,\middle|\, \sqrt{(w - w_0)^2 + (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} < \delta\right\}.$$

All the limit rules for functions of two variables work for functions of three or more variables.

4.4 Partial Derivatives

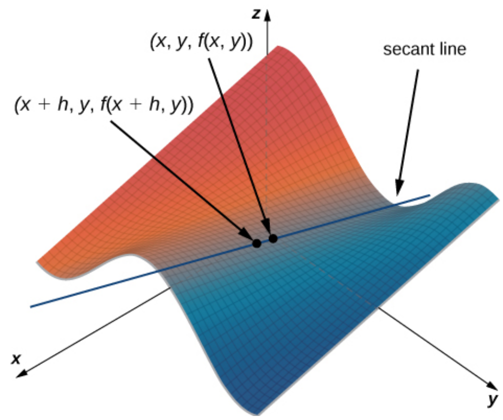
For single-variable functions, we write the derivative as y' which represents the instantaneous rate of change of y as a function of x. In Leibniz notation, we write that as $\frac{dy}{dx}$. For a function of two variables $z = f(x, y)$ which has two independent variables x and y and a dependent variable z, what does Leibniz notation look like? We use the symbol *partial* ∂:

Let $f(x, y)$ be a function of two variables. Then, the **partial derivative** of f with respect to x, written as $\frac{\partial f}{\partial x}$ or f_x, is defined as

$$\frac{\partial f}{\partial x} = f_x(x, y) = \lim_{h \to 0} \frac{f(x + h, y) - f(x, y)}{h}.$$

The partial derivative of f with respect to y, written as $\frac{\partial f}{\partial y}$ or f_y, is defined as

$$\frac{\partial f}{\partial y} = f_y(x, y) = \lim_{k \to 0} \frac{f(x, y + k) - f(x, y)}{k}.$$



Secant line passing through the points $(x, y, f(x, y))$ and $(x + h, y, f(x + h, y))$

There are many alternative notations for partial derivatives.

Just remember that to find f_x, you regard y as a constant and differentiate $f(x, y)$ with respect to x. And to find f_y, you regard x as a constant and differentiate $f(x, y)$ with respect to y.

EXAMPLE 4.10

Use the definition of the partial derivative to compute $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for the function $f(x, y) = x^2 - 3xy + 2y^2 - 4x + 5y - 12$.

Solution:

We are computing $\frac{\partial f}{\partial x}$ first. We begin by calculating $f(x + h, y) = x^2 + 2xh + h^2 - 3xy - 3hy + 2y^2 - 4x - 4h + 5y - 12$. Then we plug it in and simplify to get this:

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{h(2x + h - 3y - 4)}{h} = \lim_{h \to 0}(2x + h - 3y - 4) = 2x - 3y - 4$$

Now we compute $\frac{\partial f}{\partial y}$. We compute $f(x, y + h) = x^2 - 3xy - 3xh + 2y^2 + 4yh + 2h^2 - 4x + 5y + 5h - 12$. Then we plug it in and simplify to get this:

$$\frac{\partial f}{\partial y} = \lim_{h \to 0} \frac{h(-3x + 4y + 2h + 5)}{h} = \lim_{h \to 0}(-3x + 4y + 2h + 5) = -3x + 4y + 5$$

EXAMPLE 4.11

Let $f(x,y) = x^3 + x^2y^3 - 2y^2$. Find $f_x(2,1)$ and $f_y(2,1)$.

Solution:

We differentiate with respect to x and hold y constant:

$$f_x(x,y) = \frac{\partial}{\partial x}\left(x^3 + x^2y^3 - 2y^2\right) = 3x^2 + 2xy^3 \Rightarrow f_x(2,1) = 3 \cdot 2^2 + 2 \cdot 2 \cdot 1^3 = 12 + 4 = 16$$

We differentiate with respect to y and hold x constant:

$$f_y(x,y) = \frac{\partial}{\partial y}\left(x^3 + x^2y^3 - 2y^2\right) = 3x^2y^2 - 4y \Rightarrow f_y(2,1) = 3 \cdot 2^2 \cdot 1^2 - 4 \cdot 1 = 12 - 4 = 8$$

Now, we will think about how to geometrically interpret partial derivatives.

Let $z = f(x,y)$ represent a surface S in $\mathbb{R}^3$. The point $P = (a,b,c)$, given that $c = f(a,b)$, lies on this surface.

If we fix $y = b$, this slices the surface with a vertical plane parallel to the xz-plane. The intersection curve is C_1, and it traces the function $g(x) = f(x,b)$. The slope of the tangent line T_1 to this curve at point P is $f_x(a,b)$.

If we fix $x = a$, this slices the surface with a vertical plane parallel to the yz-plane. The intersection curve is C_2, and it traces the function $g(y) = f(a,y)$. The slope of the tangent line T_2 to this curve at point P is $f_y(a,b)$.

Thus, the partial derivatives $f_x(a,b)$ and $f_y(a,b)$ can be viewed as the slopes of tangent lines to these vertical traces of the surface in the planes $y = b$ and $x = a$, respectively.

We now move onward to a function of three variables such as $w = f(x,y,z)$:

Let $f(x,y,z)$ be a function of three variables. Then, the partial derivative of f with respect to x, written as $\frac{\partial f}{\partial x}$, or f_x, is defined to be

$$\frac{\partial f}{\partial x} = f_x(x,y,z) = \lim_{h\to 0} \frac{f(x+h,y,z)-f(x,y,z)}{h}.$$

The partial derivative of f with respect to y, written as $\frac{\partial f}{\partial y}$, or f_y, is defined to be

$$\frac{\partial f}{\partial y} = f_y(x,y,z) = \lim_{k\to 0} \frac{f(x,y+k,z)-f(x,y,z)}{k}.$$

The partial derivative of f with respect to z, written as $\frac{\partial f}{\partial z}$, or f_z, is defined to be

$$\frac{\partial f}{\partial z} = f_z(x,y,z) = \lim_{m\to 0} \frac{f(x,y,z+m)-f(x,y,z)}{m}.$$

When we want to calculate a partial derivative of a function of three variables, we use the same idea as we did for a function of two variables: we treat the other two independent variables as if they are constants and then differentiate with respect to whichever variable we are focusing on.

In general, if $w = f(x_1, x_2, \ldots, x_n)$, there are n partial derivatives denoted by

$$\frac{\partial w}{\partial x_k} = f_{x_k}(x_1, x_2, \ldots, x_n),$$

where $k = 1, 2, \ldots, n$.

EXAMPLE 4.12

Let $f(x, y, z) = e^{xy} \ln z$. Compute the partial derivatives f_x, f_y, and f_z.

Solution:

We first take the partial derivative with respect to x, holding y and z constant:

$$f_x = \frac{\partial}{\partial x}\left(e^{xy} \ln z\right) = \left(\frac{\partial}{\partial x} e^{xy}\right) \cdot \ln z = y e^{xy} \ln z$$

We then take the partial derivative with respect to y, holding x and z constant:

$$f_y = \frac{\partial}{\partial y}\left(e^{xy} \ln z\right) = x e^{xy} \ln z$$

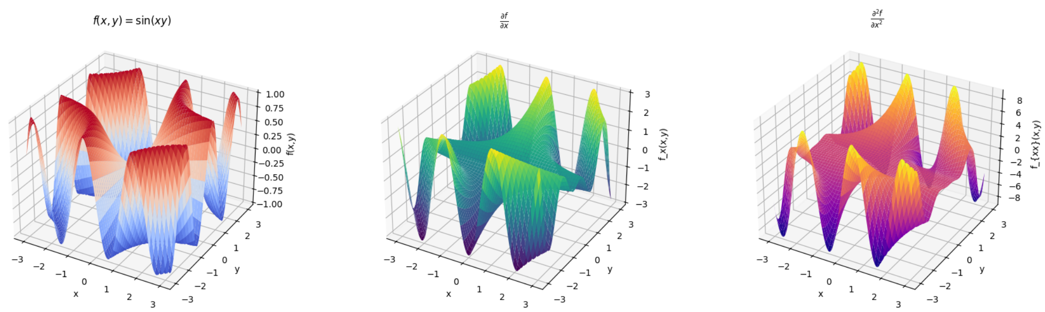
We then take the partial derivative with respect to z, holding x and y constant:

$$f_z = \frac{\partial}{\partial z}\left(e^{xy} \ln z\right) = e^{xy} \cdot \frac{1}{z} = \frac{e^{xy}}{z}$$

Let's now move on to higher derivatives. Take a look at the following table of second partial derivatives:

Leibniz Notation	**Subscript Notation**	**Pronunciation**
$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}$	$(f_x)_x = f_{xx}$	"d squared f over d x squared" or "f sub x x"
$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$	$(f_y)_y = f_{yy}$	"d squared f over d y squared" or "f sub y y"
$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$	$(f_y)_x = f_{yx}$	"d squared f over d x d y" or "f sub y x"
$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$	$(f_x)_y = f_{xy}$	"d squared f over d y d x" or "f sub x y"

Higher-Order Partial Derivative Notation and Pronunciation



Graphs of $f(x, y) = \sin(xy)$, $\frac{\partial f}{\partial x}$, and $\frac{\partial^2 f}{\partial x^2}$

EXAMPLE 4.12

Find the second partial derivatives of the function $f(x, y) = x^3 + x^2y^3 - 2y^2$ using the first partial derivatives we already found:

$$f_x(x, y) = 3x^2 + 2xy^3$$
$$f_y(x, y) = 3x^2y^2 - 4y$$

Solution:

We compute the second partials with respect to the same variable:

$$f_{xx} = \frac{\partial}{\partial x}(3x^2 + 2xy^3) = 6x + 2y^3$$

$$f_{yy} = \frac{\partial}{\partial y}(3x^2y^2 - 4y) = 6x^2y - 4$$

Now we compute the mixed partials. For f_{xy} we differentiate with respect to x first and then with respect to y and vice-versa for f_{yx}:

$$f_{xy} = \frac{\partial}{\partial y}(3x^2 + 2xy^3) = 6xy^2$$

$$f_{yx} = \frac{\partial}{\partial x}(3x^2y^2 - 4y) = 6xy^2$$

You may have noticed that $f_{xy} = f_{yx}$. Is this a coincidence? Let's test this using the general polynomial term $x^m y^n$:

$$\frac{\partial^2}{\partial x \partial y}(x^m y^n) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}(x^m y^n)\right) = \frac{\partial}{\partial x}(nx^m y^{n-1}) = mnx^{m-1}y^{n-1}$$

$$\frac{\partial^2}{\partial y \partial x}(x^m y^n) = \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}(x^m y^n)\right) = \frac{\partial}{\partial y}(mx^{m-1} y^n) = mnx^{m-1}y^{n-1}$$

It looks like it's not a coincidence after all.

Clairaut's Theorem: Assume that $f(x, y)$ is a multivariable function with domain $D \subseteq \mathbb{R}^2$. Suppose that f_{xy} and f_{yx} are continuous throughout D. Then, for all points in D, we have

$$f_{xy}(x, y) = \frac{\partial}{\partial y}\left[\frac{\partial}{\partial x}[f(x, y)]\right] = \frac{\partial}{\partial x}\left[\frac{\partial}{\partial y}[f(x, y)]\right] = f_{yx}(x, y).$$

In other words, if two functions are nice enough with continuous second mixed partial derivatives in some region around a point, the order in which you take partial derivatives doesn't matter. Now keep in mind that this is only true for *most* functions. Clairaut's theorem also holds true for any order of partial derivatives as long as they are continuous. For instance, assuming f has continuous third partial derivatives, then

$$f_{xyz} = f_{xzy} = f_{yzx} = f_{zxy} = f_{zyx}.$$

Let's talk about differentiability.

Suppose a multivariable function $f(x, y)$ has partial derivatives f_x and f_y defined on an open set containing the point (a, b), and both of these partial derivatives are continuous at that point. Then, f is differentiable at (a, b).

Moreover, if a function $f(x, y)$ is differentiable at a point (a, b), then it is also continuous at that point.

Continuity of the partial derivatives *implies* differentiability. And differentiability *implies* continuity. However, the reverse implications do not hold. A function can be continuous without being differentiable, and it can have partial derivatives without being continuous.

EXAMPLE 4.13

In a study of frost penetration, it was found that the temperature T at time t (in days) and depth x (in feet) is modeled by the function $T(x,t) = T_0 + T_1 e^{-\lambda x}\sin(\omega t - \lambda x)$ where $\omega = \frac{2\pi}{365}$ and λ is a positive constant.

(a) Find $\frac{\partial T}{\partial x}$. What is its physical significance?

(b) Find $\frac{\partial T}{\partial t}$. What is its physical significance?

Solution:

(a)

$$\frac{\partial T}{\partial x} = \frac{d}{dx}\left[T_0 + T_1 e^{-\lambda x}\sin(\omega t - \lambda x)\right]$$

$$= T_1 \cdot \frac{d}{dx}\left(e^{-\lambda x}\sin(\omega t - \lambda x)\right)$$

$$= T_1\left[\frac{d}{dx}(e^{-\lambda x}) \cdot \sin(\omega t - \lambda x) + e^{-\lambda x} \cdot \frac{d}{dx}(\sin(\omega t - \lambda x))\right]$$

$$= T_1\left[(-\lambda e^{-\lambda x}) \cdot \sin(\omega t - \lambda x) + e^{-\lambda x} \cdot (-\lambda\cos(\omega t - \lambda x))\right]$$

$$= -\lambda T_1 e^{-\lambda x}\left[\sin(\omega t - \lambda x) + \cos(\omega t - \lambda x)\right]$$

This measures the rate of change of temperature with respect to depth at time t. Because of the exponential decay, the oscillations in temperature get weaker as depth increases.

(b)

$$\frac{\partial T}{\partial t} = \frac{d}{dt}\left[T_0 + T_1 e^{-\lambda x}\sin(\omega t - \lambda x)\right]$$

$$= T_1 e^{-\lambda x} \cdot \frac{d}{dt}\left(\sin(\omega t - \lambda x)\right)$$

$$= T_1 e^{-\lambda x} \cdot \omega\cos(\omega t - \lambda x)$$

This measures the rate of change of temperature with respect to time at depth x. The cosine function reflects seasonal fluctuations in temperature and the amplitude again decreases with depth.

EXAMPLE 4.14

Express the volume of a right circular cylinder as a function V of two variables.

(a) Express it as a function of its radius r and height h.

(b) Show that the rate of change of the volume of the cylinder with respect to its radius is the product of its circumference multiplied by its height.

(c) Show that the rate of change of the volume of the cylinder with respect to its height is equal to the area of the circular base.

Solution:

(a) $V(r, h) = \pi r^2 h$

(b) Find the partial derivative of V with respect to r:

$$\frac{\partial V}{\partial r} = \frac{\partial}{\partial r}(\pi r^2 h) = 2\pi r h$$

This is the rate of change of volume with respect to radius. This shows that we are multiplying circumference $2\pi r$ by height h.

(c) Find the partial derivative of V with respect to h:

$$\frac{\partial V}{\partial h} = \frac{\partial}{\partial h}(\pi r^2 h) = \pi r^2$$

This is the rate of change of volume with respect to height and gives us area πr^2.

4.5 Chain Rule

The **chain rule** is likely one of the most powerful tools you used in single-variable calculus. Recall that it is written as

$$\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x).$$

We will now generalize the chain rule to multivariable functions. We begin where x and y are functions of one variable:

Let $z = f(x, y)$ be a differentiable function where $x = g(t)$ and $y = h(t)$ are both differentiable functions of t. Then z is a differentiable function of t, and we have

$$\frac{dz}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}$$

EXAMPLE 4.15

Let $z = x^2 y + 3xy^4$ where $x(t) = \sin(2t)$ and $y(t) = \cos(t)$. Find $\frac{dz}{dt}$ when $t = 0$.

Solution:

We will often write the chain rule like this:

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

Let's compute:

$$\frac{dz}{dt} = \left(2xy + 3y^4\right)(2\cos(2t)) + \left(x^2 + 12xy^3\right)(-\sin(t))$$

At $t = 0$, we have $x(0) = \sin(0) = 0$ and $y(0) = \cos(0) = 1$.
Evaluating yields

$$\left.\frac{dz}{dt}\right|_{t=0} = (2(0)(1) + 3(1)^{)}(2\cos(0)) + ((0)^2 + 12(0)(1)^3)(-\sin(0))$$

$$= 3 \cdot 2 + 0 = 6$$

We now move on to the chain rule where x and y are functions of two variables:

Let $z = f(x, y)$ be a differentiable function of x and y, where $x = g(s, t)$ and $y = h(s, t)$ are both differentiable functions of s and t. Then z is a differentiable function of s and t, and we have

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s}.$$

EXAMPLE 4.16

Let $z = e^x \sin y$ where $x = st^2$ and $y = s^2 t$. Find $\dfrac{\partial z}{\partial s}$ and $\dfrac{\partial z}{\partial t}$.

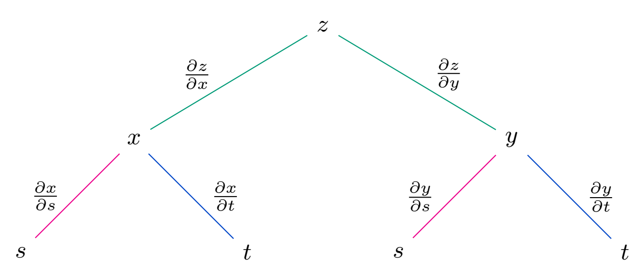
Solution:

We have $z = f(x, y)$ where x and y depend on s and t:

$$\frac{\partial z}{\partial s} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial s} = (e^x \sin y)\,(t^2) + (e^x \cos y)\,(2st)$$
$$= t^2 e^{st^2} \sin(s^2 t) + 2ste^{st^2} \cos(s^2 t)$$

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial z}{\partial y} \cdot \frac{\partial y}{\partial t} = (e^x \sin y)\,(2st) + (e^x \cos y)\,(s^2)$$
$$= 2ste^{st^2} \sin(s^2 t) + s^2 e^{st^2} \cos(s^2 t)$$

One method to remember and visualize the chain rule is through **tree diagrams**. We start by branching out from the z, the *dependent* variable. After that, we have x and y, the *intermediate* variables. This tells us that z is a function of x and y. Then, we branch out to s and t, the *independent* variables. The branches contain the partial derivative of the variable on the node it comes from with respect to the variable on the node it leads to.



If you wanted to find $\frac{\partial z}{\partial s}$, you have to follow two paths to s. For the first, you go from z to x to s and multiply together any partial derivatives you pass on the way. For the second, you go from z to y to s and multiply together any partial derivatives you pass on the way. Then, you add the two paths together.

Let's now look at the general version of the chain rule:

Suppose that u is a differentiable function of the n variables

$$x_1, x_2, \ldots, x_n,$$

where each x_j is a differentiable function of the m variables

$$t_1, t_2, \ldots, t_m.$$

That is, we have functions $u = f(x_1, x_2, ..., x_n)$ and $x_i = g_i(t_1, t_2, ..., t_m)$.

Then u is a function of $t_1, t_2, \ldots, t_m$, and for each $i = 1, 2, \ldots, m$, we have

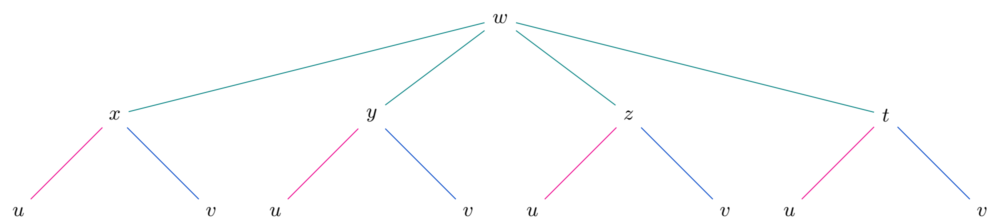
$$\frac{\partial u}{\partial t_i} = \frac{\partial u}{\partial x_1} \cdot \frac{\partial x_1}{\partial t_i} + \frac{\partial u}{\partial x_2} \cdot \frac{\partial x_2}{\partial t_i} + \cdots + \frac{\partial u}{\partial x_n} \cdot \frac{\partial x_n}{\partial t_i} = \sum_{j=1}^{n} \frac{\partial u}{\partial x_j} \cdot \frac{\partial x_j}{\partial t_i}.$$

EXAMPLE 4.17

Suppose $w = f(x, y, z, t)$ where $x = x(u, v)$, $y = y(u, v)$, $z = z(u, v)$, and $t = t(u, v)$. Draw the corresponding tree diagram and write out the expressions for the chain rule.

Solution:

Here is the corresponding tree diagram:



Using the tree diagram, we can easily acquire the chain rule expressions. For $n = 4$ and $m = 2$, we have the following:

$$\frac{\partial w}{\partial u} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial u} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial u} + \frac{\partial w}{\partial t} \cdot \frac{\partial t}{\partial u}$$

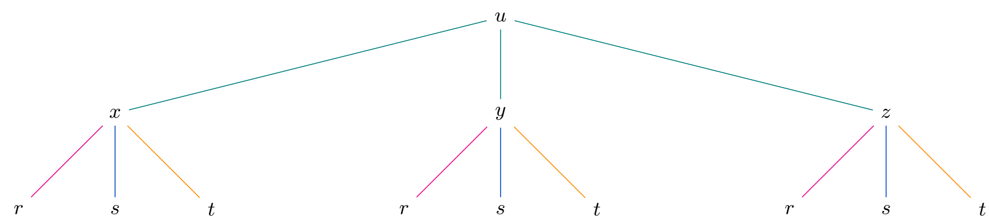
$$\frac{\partial w}{\partial v} = \frac{\partial w}{\partial x} \cdot \frac{\partial x}{\partial v} + \frac{\partial w}{\partial y} \cdot \frac{\partial y}{\partial v} + \frac{\partial w}{\partial z} \cdot \frac{\partial z}{\partial v} + \frac{\partial w}{\partial t} \cdot \frac{\partial t}{\partial v}$$

EXAMPLE 4.18

Suppose $u = x^4y + y^2z^3$ where $x = rse^t$, $y = rs^2e^{-t}$, and $z = r^2s\sin t$. Find the value of $\frac{\partial u}{\partial s}$ when $r = 2$, $s = 1$, and $t = 0$. Draw a tree diagram to help you.

Solution:

Here is the corresponding tree diagram:



With the help of the tree diagram, we apply the chain rule:

$$\frac{\partial u}{\partial s} = \frac{\partial u}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial u}{\partial y} \cdot \frac{\partial y}{\partial s} + \frac{\partial u}{\partial z} \cdot \frac{\partial z}{\partial s}$$

Now we compute:

$$\frac{\partial u}{\partial s} = (4x^3y)(re^t) + (x^4 + 2yz^3)(2rse^{-t}) + (3y^2z^2)(r^2\sin t)$$

At $r = 2, s = 1$, and $t = 0$, we find

$$\begin{aligned} x &= rse^t = 2 \cdot 1 \cdot 1 = 2 \\ y &= rs^2e^{-t} = 2 \cdot 1^2 \cdot 1 = 2 \\ z &= r^2s\sin t = 4 \cdot 1 \cdot 0 = 0 \end{aligned}$$

Now we plug them in:

$$\begin{aligned} \frac{\partial u}{\partial s} &= (4 \cdot 8 \cdot 2)(2 \cdot 1) + (16 + 2 \cdot 2 \cdot 0)(2 \cdot 2 \cdot 1) + (3 \cdot 4 \cdot 0)(4 \cdot 0) \\ &= (64)(2) + (16)(4) + (0)(0) = 192 \end{aligned}$$

In single-variable calculus, you used the chain rule to perform **implicit differentiation**, which is a method for finding $\frac{dy}{dx}$ when y is defined implicitly as a function of x. For instance, let's say we want to differentiate $x^2 + y^2 = 1$. As you can see, y is not isolated; it's not an explicit function of x. We would have to implicitly differentiate:

$$\frac{d}{dx}(x^2 + y^2) = \frac{d}{dx}(1)$$
$$2x + 2y \cdot \frac{dy}{dx} = 0$$
$$\frac{dy}{dx} = -\frac{x}{y}$$

You may think: why don't we just isolate the y? Well, we could have there. It just would have been a lot more difficult. We will now extend implicit differentiation to multivariable calculus:

Let the function F be differentiable on its domain and suppose $F(x, y) = 0$ defines y as a function of x. If $F_y \neq 0$, then

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{F_x}{F_y}.$$

Under the same conditions, for the function $F(x, y, z) = 0$, we have the following:

$$\frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z} = -\frac{F_x}{F_z}$$

$$\frac{\partial z}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z} = -\frac{F_y}{F_z}$$

Let's show how we used the chain rule to find $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$.

Suppose that z is given implicitly as a function $z = f(x, y)$ by an equation of the form $F(x, y, z) = 0$. This means that

$$F(x, y, f(x, y)) = 0$$

for all $(x, y) \in D$ where D is the domain of f. If F and f are differentiable, then we can use the chain rule to differentiate the equation $F(x, y, z) = 0$ as follows:

$$\frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial x} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial x} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = 0$$

Because $\frac{\partial}{\partial x}(x) = 1$ and $\frac{\partial}{\partial x}(y) = 0$, the equation becomes

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial x} = 0$$

If $\frac{\partial F}{\partial z} \neq 0$, we solve for $\frac{\partial z}{\partial x}$ and obtain $\frac{\partial z}{\partial x} = -\frac{\partial F/\partial x}{\partial F/\partial z}$. The formula for $\frac{\partial z}{\partial y}$ is obtained in a similar manner:

$$\frac{\partial F}{\partial x} \cdot \frac{\partial x}{\partial y} + \frac{\partial F}{\partial y} \cdot \frac{\partial y}{\partial y} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial y} = 0$$

$$0 + \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} \cdot \frac{\partial z}{\partial y} = 0$$

$$\frac{\partial z}{\partial y} = -\frac{\partial F/\partial y}{\partial F/\partial z}$$

EXAMPLE 4.19

Find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ if $F(x, y, z) = x^3 + y^3 + z^3 + 6xyz = 1$.

Solution:

We will z be defined implicitly as a function of x and y:

$$\frac{\partial z}{\partial x} = -\frac{F_x}{F_z} = -\frac{3x^2 + 6yz}{3z^2 + 6xy} = -\frac{x^2 + 2yz}{z^2 + 2xy}$$
$$\frac{\partial z}{\partial y} = -\frac{F_y}{F_z} = -\frac{3y^2 + 6xz}{3z^2 + 6xy} = -\frac{y^2 + 2xz}{z^2 + 2xy}$$

EXAMPLE 4.20

Wheat production W in a given year depends on the average temperature T and the annual rainfall R. Scientists estimate that the average temperature in the agricultural region of a country is rising at a rate of 0.15°C/year and rainfall is decreasing at a rate of 0.1 cm/year. They also estimate that, at current production levels, $\partial W/\partial T = -2$ and $\partial W/\partial T = 8$.

(a) What is the significance of the signs of these partial derivatives?
(b) Estimate the current rate of change of wheat production dW/dt.

Solution:

(a) The negative sign of $\partial W/\partial T$ means that as temperature increases, wheat production decreases assuming that annual rainfall remains constant. Conversely, the positive sign of $\partial W/\partial T$ indicates that more rainfall increases wheat production assuming that average temperature remains constant.

(b) We have temperature rate $dT/dt = 0.15$ and rainfall rate $dR/dt = -0.1$. Since $W = f(T, R)$ and both T and R depend on time t, we apply the chain rule:

$$\frac{dW}{dt} = \frac{\partial W}{\partial T} \cdot \frac{dT}{dt} + \frac{\partial W}{\partial R} \cdot \frac{dR}{dt} = (-2)(0.15) + (8)(-0.1) = -0.3 - 0.8 = -1.1.$$

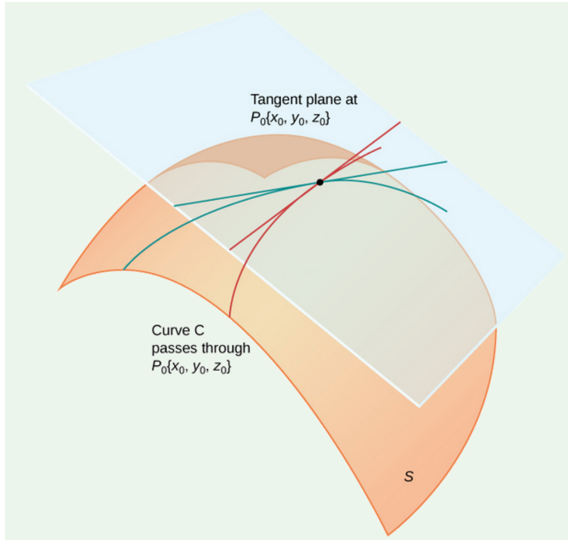
Wheat production is currently decreasing at a rate of 1.1 units per year.

5 Applications of Multivariable Differentiation

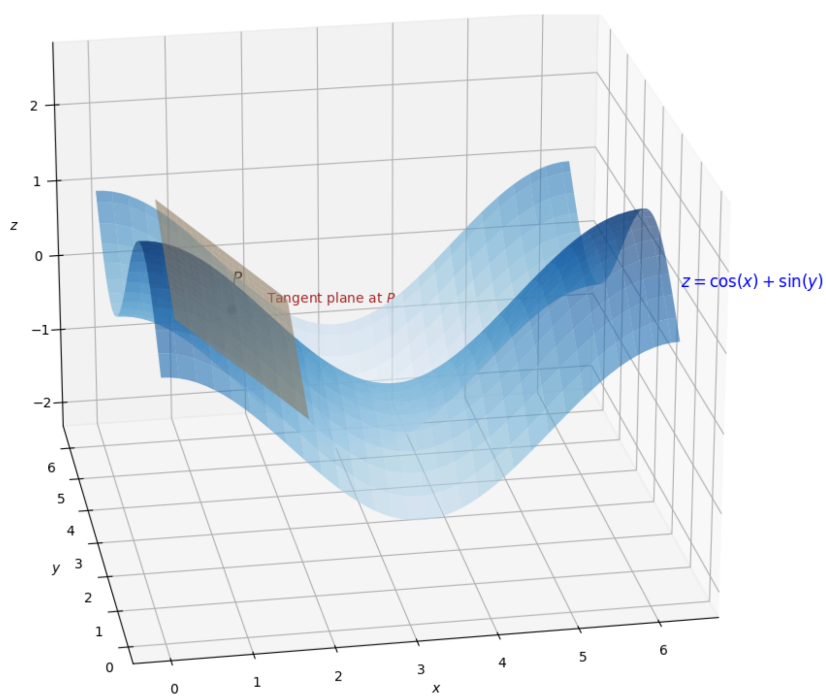
5.1 Tangent Planes and Linear Approximations

In single-variable calculus, you used the second derivative test to search for local minimums and maximums. You did this by finding where the slope of the tangent line to curves was equal to zero. You could also zoom in towards a point on a graph and approximate the function. Here, we will develop the same idea for multivariable calculus, where we can zoom in on a point on a surface and find a *tangent plane*. We begin with the geometric definition of a tangent plane:

Let $P_0 = (x_0, y_0, z_0)$ be a point on a surface S and let $C \subset S$ be a curve lying entirely that passes through P_0. If the tangent lines to any C at P_0 lie in the same plane, then we call this the tangent plane to the S at P_0. The tangent plane at P_0 is the plane that most closely approximates the surface near the P_0.



The tangent plane to S at P_0. Image credit: Strang & Herman



The tangent plane to $z = \cos x + \sin y$ at P

The equation of a plane passing through the point $P(x_0, y_0, z_0)$ can be written as

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

If we divide everything by C and let $a = -A/C$ and $b = -B/C$, we can rewrite it as

$$z - z_0 = a(x - x_a) + b(y - y_0),$$

which represents the tangent plane at P. Thus, its intersection with the plane $y = y_0$ must be the tangent line T_1. If we plug y_0 into the equation for the tangent plane, we get

$$z - z_0 = a(x - x_0),$$

where $y = y_0$. This is the point-slope form of a line with slope $a = f_x(x_0, y_0)$. For the tangent line T_2, we substitute $x = x_0$ into the equation of the tangent plane to get

$$z - z_0 = b(y - y_0)$$

where slope $b = f_y(x_0, y_0)$. Putting everything together, we get the equation of a tangent plane:

Suppose the function f has continuous partial derivatives. An equation of the tangent plane to the surface $z = f(x, y)$ at the point $P(x_0, y_0, z_0)$ is

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

EXAMPLE 5.1

Determine the tangent plane to the graph of $f(x,y) = x^3 + y^2 + 2x$ at the point $(-1, 2, f(-1,2))$.

Solution:

We first compute the partial derivatives:

$$\frac{\partial f}{\partial x} = f_x = 3x^2 + 2$$

$$\frac{\partial f}{\partial y} = f_y = 2y$$

We evaluate them at the point $(-1,2)$:

$$f_x\big|_{(-1,2)} = 3(-1)^2 + 2 = 5$$

$$f_y\big|_{(-1,2)} = 2(2) = 4.$$

The value of the function at this point is

$$f(-1,2) = (-1)^3 + 2^2 + 2(-1) = -1 + 4 - 2 = 1.$$

So the equation of the tangent plane is

$$z = f(-1,2) + f_x(-1,2)(x+1) + f_y(-1,2)(y-2) = 1 + 5(x+1) + 4(y-2)$$
$$z = 5x + 4y - 2.$$

We can also write out the equation of a tangent plane in vector form. Consider a surface S defined by the function $F(x,y,z) = 0$ and let point $P = (x_0, y_0, z_0) \in S$. Suppose $C \subset S$ is a curve that passes through P, defined by a vector-valued function

$$\vec{r}(t) = \langle x(t),\, y(t),\, z(t) \rangle$$

such that $F(x(t), y(t), z(t)) = 0$ for all t.

Differentiating both sides with respect to t using the chain rule, we obtain

$$\frac{d}{dt}F(x(t), y(t), z(t)) = F_x\, x'(t) + F_y\, y'(t) + F_z\, z'(t) = \nabla F \cdot \vec{r}\,'(t) = 0.$$

We evaluate at $P = (x_0, y_0, z_0)$ and this becomes

$$\nabla F(x_0, y_0, z_0) \cdot \vec{r}\,'(t_0) = 0,$$

where $\nabla F(x_0, y_0, z_0)$ is the *gradient* and $\vec{r}\,'(t)$ is the tangent vector. The gradient at P is orthogonal to the tangent vector of every curve on the surface through P. Therefore, the tangent plane to the surface at P is the plane through P that is perpendicular to the gradient. We have

$$\nabla F(x_0, y_0, z_0) \cdot \langle x - x_0,\, y - y_0,\, z - z_0\rangle = 0.$$

Writing this out in component form yields

$$\langle F_x(x_0, y_0, z_0),\, F_y(x_0, y_0, z_0),\, F_z(x_0, y_0, z_0)\rangle \cdot \langle x - x_0,\, y - y_0,\, z - z_0\rangle = 0,$$

and simplifying gives the scalar equation of the tangent plane:

$$F_x(x_0, y_0, z_0)(x - x_0) + F_y(x_0, y_0, z_0)(y - y_0) + F_z(x_0, y_0, z_0)(z - z_0) = 0.$$

This is another approach to thinking about tangent planes that ultimately leads to the same result.

EXAMPLE 5.2

We have the ellipsoid $F(x, y, z) = \frac{x^2}{9} + \frac{y^2}{25} + z^2 - 1 = 0$.

(a) Find the equation for the tangent plane at $P = (0, 4, \frac{3}{5})$.
(b) Find any points on the ellipsoid with a horizontal tangent plane.
(c) Graph parts (a) and (b).

Solution:

We verify that the point lies on the surface:

$$F(0, 4, \tfrac{3}{5}) = \frac{0^2}{9} + \frac{4^2}{25} + \left(\frac{3}{5}\right)^2 - 1 = 0.$$

We compute the gradient of F:

$$\nabla F(x, y, z) = \left\langle \frac{2x}{9}, \frac{2y}{25}, 2z \right\rangle$$

$$\nabla F(0, 4, \tfrac{3}{5}) = \left\langle 0, \frac{8}{25}, \frac{6}{5} \right\rangle$$

We compute the equation of the tangent plane:

$$0 = \nabla F(a, b, c) \cdot \langle x - 0,\ y - 4,\ z - \frac{3}{5} \rangle = 0 \cdot (x - 0) + \frac{8}{25}(y - 4) + \frac{6}{5}\left(z - \frac{3}{5}\right)$$

$$\frac{8}{25}y + \frac{6}{5}z - 2 = 0$$

EXAMPLE 5.2 (CONTINUED)

(b) A tangent plane is horizontal if the normal vector to the tangent plane points in the z-direction. That is,

$$\nabla F(x, y, z) = c \cdot \langle 0, 0, 1 \rangle$$

for some scalar multiple $c \in \mathbb{R}$. We then have

$$\left\langle \frac{2x}{9}, \frac{2y}{25}, 2z \right\rangle = \langle 0, 0, c \rangle$$

$$\frac{2x}{9} = 0 \Rightarrow x = 0$$

$$\frac{2y}{25} = 0 \Rightarrow y = 0$$

$$\frac{2x}{9} = 0 \Rightarrow z = \frac{c}{2}$$

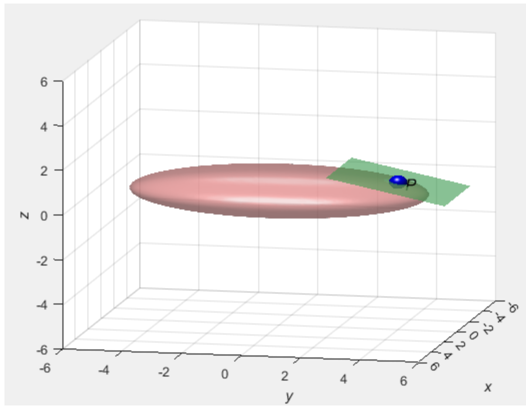
We want to find points where $(x, y, z) = (0, 0, z)$. Setting $F(0, 0, z) = 0$ and solving yields $z = \pm 1$. Thus, the two points on the surface with a horizontal tangent plane are $P_1 = (0, 0, 1)$ and $P_2 = (0, 0, -1)$. We will now find the tangent plane equation at $P_1 = (0, 0, 1)$:

$$\nabla F(0, 0, 1) \cdot \langle x - 0, y - 0, z - 1 \rangle = \langle 0, 0, 2 \rangle \cdot \langle x, y, z - 1 \rangle = 2z - 2 = 0$$

At $P_2 = (0, 0, -1)$, we have $2z + 2 = 0$.

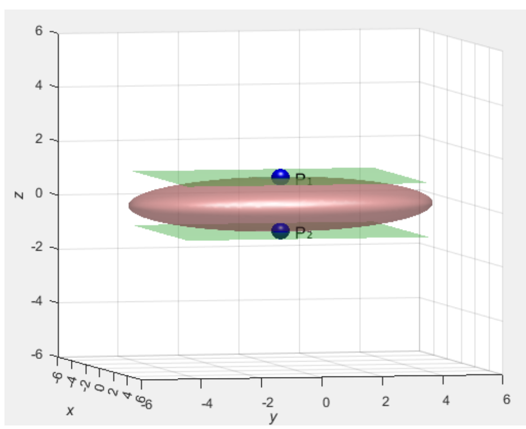
EXAMPLE 5.2 (CONTINUED)

(c) I have graphed both of the results. Please run the MATLAB code yourself and have a look!!



Part (a) graph

ex5point2a.m



Part (b) graph

ex5point2b.m

Let's formalize the fact that tangent planes can also be used to approximate multivariable functions:

In general, we know that an equation of the tangent plane to the graph of a function f of two variables 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 that this graph represents,

$$L(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b),$$

is called the **linearization** of f at (a, b). The approximation

$$f(x, y) \approx f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

is called the **linear approximation** of f at (a, b).

The idea is that if f is differentiable at (a, b), then for values of (x, y) close to (a, b), the tangent plane gives a good estimate of the actual value of $f(x, y)$. This is especially useful when evaluating the exact function is difficult and the partial derivatives are easier to compute.

Recall how differentiability works in single-variable calculus. For a function $y = f(x)$, the *increment* of y as x changes from a to $a + \Delta x$ is $\Delta y = f(a + \Delta x) - f(a)$. And if f is differentiable at a, then $\Delta y = f'(a)\Delta x + \varepsilon \Delta x$ where $\varepsilon \to 0$ as $\Delta x \to 0$.

Now consider a function of two variables, $z = f(x, y)$, and suppose x and y change from a, b to $(a + \Delta x, b + \Delta y)$. Then the increment of z is defined as

$$\Delta z = f(a + \Delta x,\, b + \Delta y) - f(a, b),$$

which represents the change in f as $(a, b) \to (a + \Delta x, b + \Delta y)$.

This brings us to the definition of differentiability for functions of two variables. The idea remains the same, which is that we are approximating the function by its linear portion and an error term:

If $z = f(x, y)$, then f is differentiable at (a, b) if Δz can be expressed in the form

$$\Delta z = f_x(a, b)\,\Delta x + f_y(a, b)\,\Delta y + \varepsilon_1 \Delta x + \varepsilon_2 \Delta y,$$

where $\varepsilon_1 \to 0$ and $\varepsilon_2 \to 0$ as $(\Delta x,\ \Delta y) \to (0, 0)$.

That is, error ε_1 and ε_2 can become sufficiently small when (x, y) is near (a, b). This means that the tangent plane should approximate the graph of f pretty accurately near the point of tangency. However, it can be hard to use this definition in practice. Thus, we have a more convenient one:

If the partial derivatives f_x and f_y exist near (a, b) and are continuous at (a, b), then f is differentiable at (a, b).

EXAMPLE 5.3

Find the linear approximation of the function $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ at the point $(3, 2, 6)$, and use it to approximate $\sqrt{(3.02)^2 + (1.97)^2 + (5.99)^2}$. Find the percent error.

Solution:

We want the linear approximation

$$f(x, y, z) \approx f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c),$$

where $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$ and $(a, b, c) = (3, 2, 6)$.
We first compute the partial derivatives:

$$f_x = \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \; f_y = \frac{y}{\sqrt{x^2 + y^2 + z^2}}, \; f_z = \frac{z}{\sqrt{x^2 + y^2 + z^2}}.$$

And then we evaluate at $3, 2, 6$:

$$f(3, 2, 6) = \sqrt{9 + 4 + 36} = \sqrt{49} = 7,$$

$$f_x(3, 2, 6) = \frac{3}{7}, \; f_y(3, 2, 6) = \frac{2}{7}, \; f_z(3, 2, 6) = \frac{6}{7}.$$

Thus, the linear approximation is

$$f(x, y, z) \approx 7 + \frac{3}{7}(x - 3) + \frac{2}{7}(y - 2) + \frac{6}{7}(z - 6).$$

We set $x = 3.02$, $y = 1.97$, and $z = 5.99$ and get

$$f(3.02, 1.97, 5.99) \approx 7 + \frac{3}{7}(0.02) + \frac{2}{7}(-0.03) + \frac{6}{7}(-0.01) = 6.9914.$$

Let's now find percent error:

$$\left| \frac{6.9914 - \sqrt{(3.02)^2 + (1.97)^2 + (5.99)^2}}{\sqrt{(3.02)^2 + (1.97)^2 + (5.99)^2}} \right| \times 100\% = 0.0018\%$$

EXAMPLE 5.4

Heat index (perceived temperature) I can be written as a function of actual temperature T and relative humidity H. Use the following table from the National Weather Service to find a linear approximation for the heat index $I = f(T, H)$ when T is near 96°F and H is near 70%. Then, use it to estimate the heat index when $T = 99$°F and the relative humidity is 67%.

Relative humidity (%)

Actual temperature (°F)

T \ H	50	55	60	65	70	75	80	85	90
90	96	98	100	103	106	109	112	115	119
92	100	103	105	108	112	115	119	123	128
94	104	107	111	114	118	122	127	132	137
96	109	113	116	121	125	130	135	141	146
98	114	118	123	127	133	138	144	150	157
100	119	124	129	135	141	147	154	161	168

Solution:
We will set $f(96, 70) = 125$ as our starting value. We can approximate $f_T(96, 70)$ by using values around the table. The average of $\frac{f(98,70)-f(96,70)}{2} = 4$ and $\frac{f(94,70)-f(96,70)}{-2} = 3.5$ is 3.75, so we have $f_T(96, 70) \approx 3.75$. We average out $\frac{f(96,75)-f(96,70)}{5} = 1$ and $\frac{f(96,95)-f(96,70)}{-5} = 0.8$ to get $f_H(96, 70) \approx 0.9$.

Thus, our linear approximation is

$$\begin{aligned} f(T, H) &\approx f(96, 70) + f_T(96, 70)(T - 96) + f_H(96, 70)(H - 70) \\ &\approx 125 + 3.75(T - 96) + 0.9(H - 70). \end{aligned}$$

We will use this to estimate the desired heat index:

$$f(99, 67) \approx 125 + 3.75(99 - 96) + 0.9(67 - 70) = 133.55°\text{F}$$

For a differentiable function of one variable, $y = f(x)$, we define the **differential** dx to be an independent variable; we can assign it any real number. The corresponding change in the function's output is approximated by the differential dy:

$$dy = f'(x)dx$$

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 an amount $dx = \Delta x$.

And now for the multivariable definition:

For a differentiable function of two variables, $y = f(x, y)$, we define both dx and dy as independent variables:

$$dz = f_x(x, y)\,dx + f_y(x, y)\,dy = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy$$

Suppose $dx = \Delta x = x - a$ and $dy = \Delta y = y - b$. Then, we can write the differential of z as

$$dz = f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

Thus, a linear approximation can be rewritten using a differential:

$$f(x, y) \approx f(a, b) + dz$$

Let's now rewrite this chapter's formulas for three or more variables. First off, we have linear approximation:

$$f(x, y, z) \approx f(a, b, c) + f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c)$$

and the linearization $L(x, y, z)$ is the right side of this expression.

If $w = f(x, y, z)$, then 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, dz of the independent variables by

$$dw = \frac{\partial w}{\partial x}\,dx + \frac{\partial w}{\partial y}\,dy + \frac{\partial w}{\partial z}\,dz$$

EXAMPLE 5.5

(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:

(a) We first compute the partial derivatives:

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

So the differential dz is

$$dz = \frac{\partial z}{\partial x}\,dx + \frac{\partial z}{\partial y}\,dy = (2x + 3y)\,dx + (3x - 2y)\,dy$$

(b) We plug our givens $x = 2,\ y = 3,\ dx = \Delta x = 0.05$, and $dy = \Delta y = -0.04$ in:

$$\begin{aligned} dz &= [2(2) + 3(3)](0.05) + [3(2) - 2(3)](-0.04) \\ &= [4 + 9](0.05) + [6 - 6](-0.04) \\ &= 13(0.05) + 0 = 0.65 \end{aligned}$$

Now we compute the increment of z which gives the actual change:

$$\Delta z = f(2.05, 2.96) - f(2, 3) = ((2.05)^2 + 3(2.05)(2.96) - (2.96)^2) - (4 + 18 - 9) = 0.6449$$

We have $dz \approx \Delta z$, meaning that the differential was a good approximation.

EXAMPLE 5.6

Use differentials to estimate the amount of tin in a closed tin can with a diameter of 8 cm and height of 12 cm if the tin is 0.04 cm thick.

Solution:

The volume V of the tin is $V = \pi r^2 h$. Thus the amount of tin can be approximated by the differential dV. Here, we have

$$dV = \frac{\partial V}{\partial r}\,dr + \frac{\partial V}{\partial h}\,dh = 2\pi r h\,dr + \pi r^2\,dh$$

We have $dr = 0.04$ cm is the thickness of the tin contributing to the side walls (increase in radius) and $dh = 0.08$ cm accounts for tin on both the top and bottom surfaces. We now plug everything in:

$$dV = 2\pi r h\,dr + \pi r^2\,dh = 2\pi(4\,\text{cm})(12\,\text{cm})(0.04\,\text{cm}) + \pi(4\,\text{cm})^2(0.08\,\text{cm}) = 16.08\,\text{cm}^3$$

5.2 Directional Derivatives and the Gradient Vector

To visualize derivatives, you begin by finding two points on the curve and drawing a secant line through them. After measuring the slope of this secant line, you transform the secant line into a tangent line using limits. Then, you set the derivative of the function at your desired point equal to the limit of the slope at the secant line. Here is an overview of this process for a single-variable and real-valued function $y = f(x)$:

That is, $F : D \subseteq \mathbb{R} \to \mathbb{R}$.

We define the derivative of $F(x)$ at a point $x = a$ as

$$F'(a) = \lim_{x \to a} \frac{F(x) - F(a)}{x - a}.$$

Let $h = x - a$, so that $x = a + h$. Then the limit becomes

$$= \lim_{h \to 0} \frac{F(a+h) - F(a)}{h}.$$

Then we evaluate at $x = a$:

$$= \frac{d}{dx}\left[F(x)\right]\bigg|_{x=a}$$

And now we will do it for partial derivatives. Let $z = f(x, y)$ be a real-valued function of two variables. For this, we are representing the curve from the $y = b$ trace.

That is, $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$.

$$\frac{\partial f}{\partial x}(a, b) = \lim_{x \to a} \frac{f(x, b) - f(a, b)}{x - a}$$

$$= \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

$$= \frac{\partial}{\partial x}\left[f(x, y)\right]\bigg|_{(x,y)=(a,b)}$$

For the $x = a$ trace, it would look as follows:

$$\frac{\partial f}{\partial y}(a, b) = \lim_{y \to b} \frac{f(a, y) - f(a, b)}{y - b}$$

$$= \lim_{h \to 0} \frac{f(a, b + h) - f(a, b)}{h}$$

$$= \frac{\partial}{\partial y}\left[f(x, y)\right]\bigg|_{(x,y)=(a,b)}$$

For $\frac{\partial}{\partial x}$, we traveled along the $y = b$ path. That is,

$$\vec{r}(t) = \vec{r}_0 + t \cdot \mathbf{u},$$

where $\mathbf{u}$ is the unit vector in the x-direction.

Continuing on,

$$\vec{r}(t) = \langle a, b\rangle + t \cdot \langle 1, 0\rangle = \langle a + t, b\rangle.$$

For $\frac{\partial}{\partial y}$, we traveled along the $x = a$ path. That is,

$$\vec{r}(t) = \vec{r}_0 + t \cdot \mathbf{u},$$

where $\mathbf{u}$ is the unit vector in the y-direction.

Continuing on,

$$\vec{r}(t) = \langle a, b\rangle + t \cdot \langle 0, 1\rangle = \langle a, b + t\rangle.$$

So we know how to use partial derivatives to measure how a function changes along the coordinate axes. But how can we generalize this? We would have find the slope of the tangent line to a surface $z = f(x, y)$ at input $P_0(a, b)$ in a general direction. This is known as the **directional derivative**. Let's try to outline this process:

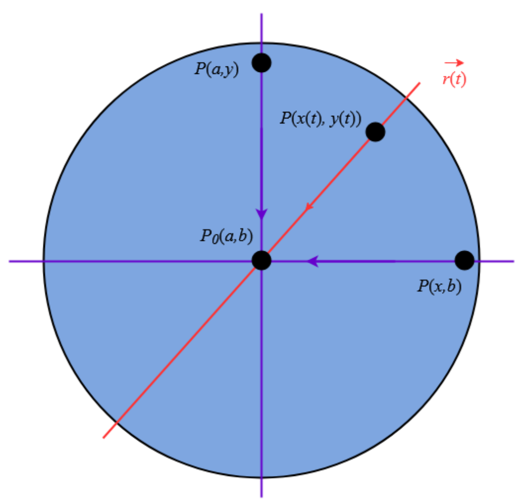
We have

$$\vec{r}(t) = \vec{r_0}(t) + t \cdot \mathbf{u}$$

where $\mathbf{u}$ is the unit vector in any direction.

Continuing on,

$$\vec{r}(t) = \langle a, b\rangle + t \cdot \langle u_1, u_2\rangle = \langle a + tu_1, b + tu_2\rangle.$$



With the partial derivatives (purple), we were limited to the direction of the coordinate axes. With directional derivatives (red), we can come from any direction.

To find the slope of a tangent line to a curve C at a given point, you can compute the rate of change of the function in the direction of a unit vector by taking its limit:

> The directional derivative of f at the point (a, b) in the direction of a unit vector $\mathbf{u} = \langle u_1, u_2 \rangle$ is defined by
>
> $$D_{\mathbf{u}} f(a, b) = \lim_{h \to 0} \frac{f(a + hu_1,\ b + hu_2) - f(a, b)}{h},$$
>
> given that the limit exists.

The partial derivatives expressed using limits from earlier are really just special cases of directional derivatives where

$$D_{\mathbf{i}} f(a, b) = \frac{\partial f}{\partial x}(a, b),$$

or

$$D_{\mathbf{j}} f(a, b) = \frac{\partial f}{\partial y}(a, b).$$

To avoid computing the limit definition of the directional derivative directly, we will define a new single-variable function $g(t)$ that captures how the multivariable function $f(x, y)$ behaves along a straight line in the direction of a given unit vector:

$$g(t) = f((\vec{l}(t))$$

where $\vec{l}(t) = \vec{P}_0 + t\mathbf{u} = \langle a + tu_1,\ b + tu_2 \rangle = \langle x(t),\ y(t) \rangle$.

Here, $\vec{P}_0 = \langle a, b \rangle$ is the base point, and $\vec{u} = \langle u_1, u_2 \rangle$ is a unit direction vector. So $\vec{l}(t)$ traces a straight line in the domain of $f(x, y)$, and $g(t)$ gives the corresponding z-value on the surface.

Then, the directional derivative of f at the point (a, b) in the direction of $\vec{u}$ is defined as the derivative of $g(t)$ at $t = 0$:

$$D_{\mathbf{u}} f(a, b) = g'(0) = \frac{d}{dt}\left[g(t)\right]\bigg|_{t=0} = \frac{d}{dt}\left[f\left(\vec{l}(t)\right)\right]\bigg|_{t=0} = \frac{d}{dt}\left[f\left(x(t),\ y(t)\right)\right]\bigg|_{t=0}$$

Now we apply the chain rule:

$$\frac{d}{dt}f(x(t),y(t)) = \frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt}$$

At $t = 0$, we have $x(0) = a$, $y(0) = b$, $\frac{dx}{dt} = u_1$, and $\frac{dy}{dt} = u_2$, so

$$D_{\mathbf{u}}f(a,b) = f_x(a,b)\cdot u_1 + f_y(a,b)\cdot u_2.$$

This is equivalent to the dot product of the gradient vector and the direction vector. Let's now write out the formal definition of the dot product formula for the directional derivative:

For function of two variables $f : \mathbb{R}^2 \to \mathbb{R}$ that is differentiable at point (a, b), the directional derivative of f that points in the direction of the unit $\mathbf{u} = \langle u_1, u_2\rangle$ in the xy-plane is given by

$$D_{\mathbf{u}}f(a,b) = \nabla f(a,b)\cdot\mathbf{u} = \langle f_x(a,b),\ f_y(a,b)\rangle\cdot\langle u_1,\ u_2\rangle.$$

Or more simply,

$$D_{\mathbf{u}}f(a,b) = f_x(a,b)\,u_1 + f_y(a,b)\,u_2.$$

EXAMPLE 5.7

Consider the paraboloid $z = f(x, y) = \frac{1}{4}(x^2 + 2y^2) + 2$. Let $P_0 = (3, 2)$ and let

$$\mathbf{u} = \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle \quad \text{and} \quad \mathbf{v} = \left\langle \frac{1}{2}, -\frac{\sqrt{3}}{2} \right\rangle.$$

Compute $D_{\mathbf{u}}f(3, 2)$ and $D_{\mathbf{v}}f(3, 2)$.

Solution:

We first compute the partial derivatives

$$f_x(x, y) = \frac{\partial}{\partial x}\left[\frac{1}{4}(x^2 + 2y^2) + 2\right] = \frac{x}{2}$$

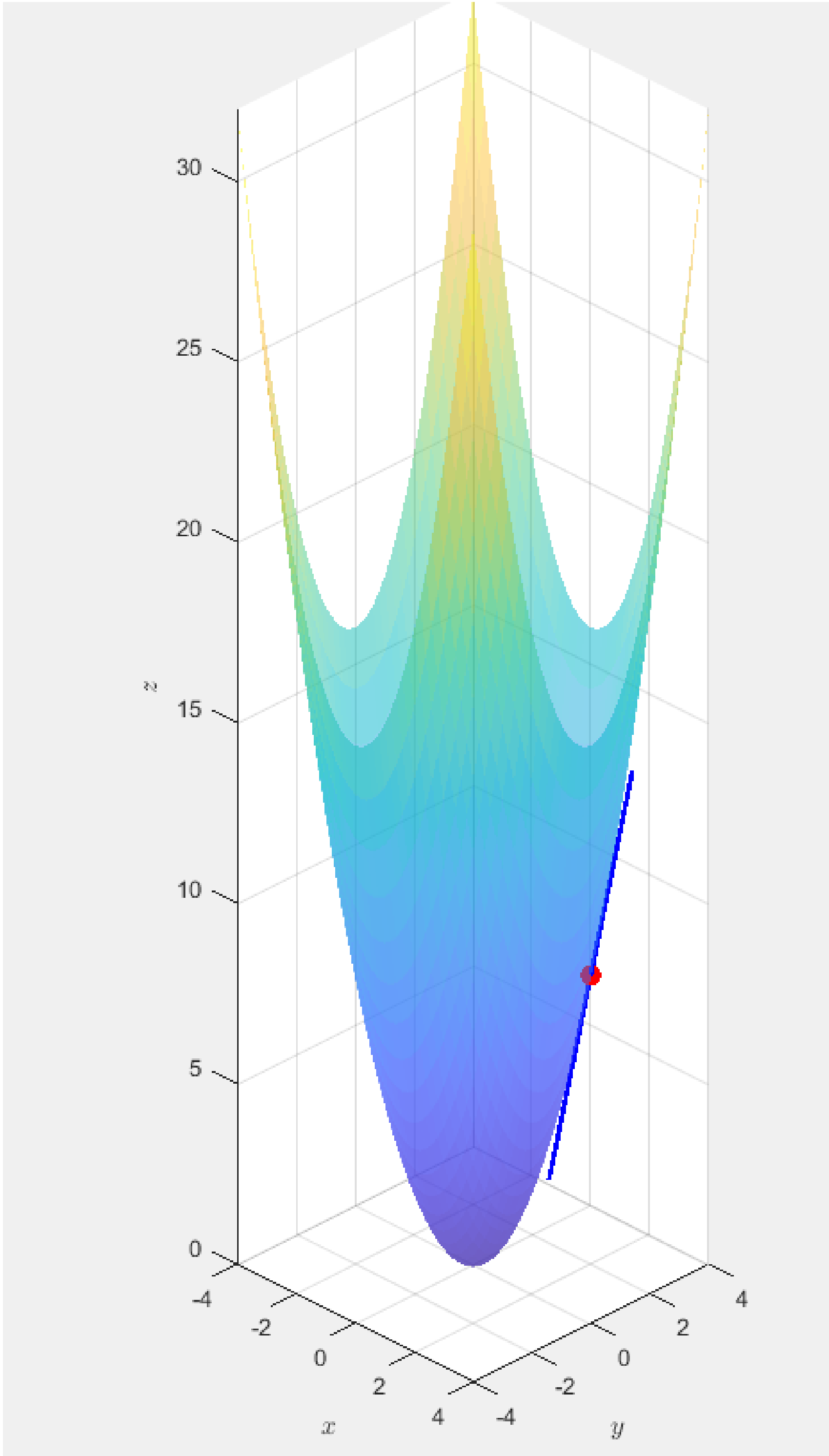
$$f_y(x, y) = \frac{\partial}{\partial y}\left[\frac{1}{4}(x^2 + 2y^2) + 2\right] = y$$

At the point $(3, 2)$, we have $f_x(3, 2) = \frac{3}{2}$ and $f_y(3, 2) = 2$. Now compute the directional derivative in the direction of $\mathbf{u}$:

$$\begin{aligned} D_{\mathbf{u}}f(3, 2) = \nabla f(3, 2) \cdot \mathbf{u} &= \left\langle \frac{3}{2}, 2 \right\rangle \cdot \left\langle \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\rangle \\ &= \frac{3}{2} \cdot \frac{1}{\sqrt{2}} + 2 \cdot \frac{1}{\sqrt{2}} = \frac{3 + 4}{2\sqrt{2}} = \frac{7}{2\sqrt{2}} \end{aligned}$$

Next, compute the directional derivative in the direction of $\mathbf{v}$:

$$D_{\mathbf{v}}f(3, 2) = \left\langle \frac{3}{2}, 2 \right\rangle \cdot \left\langle \frac{1}{2}, -\frac{\sqrt{3}}{2} \right\rangle = \frac{3}{2} \cdot \frac{1}{2} + 2 \cdot \left(-\frac{\sqrt{3}}{2}\right) = \frac{3}{4} - \sqrt{3}$$



Tangent line to $f(x, y) = x^2 + y^2$ at $(2, 2)$

EXAMPLE 5.8

Find the directional derivative of $f(x, y) = 3x^2 - 2y^2$ at the point $\left(-\frac{3}{4}, 0\right)$ in the direction from $P = \left(-\frac{3}{4}, 0\right)$ to $Q = (0, 1)$.

Solution:

First we find a vector in the specified direction and then turn it into a unit vector:

$$\overrightarrow{PQ} = \vec{v} = \langle 0 - (-\tfrac{3}{4}),\ 1 - 0\rangle = \left\langle \tfrac{3}{4},\ 1\right\rangle$$

$$\|\vec{v}\| = \sqrt{\left(\tfrac{3}{4}\right)^2 + 1^2} = \sqrt{\tfrac{25}{16}} = \tfrac{5}{4}$$

$$\mathbf{u} = \frac{\vec{v}}{\|\vec{v}\|} = \left\langle \tfrac{3}{4},\ 1\right\rangle \cdot \frac{4}{5} = \left\langle \tfrac{3}{5},\ \tfrac{4}{5}\right\rangle$$

We find the partial derivatives $f_x(x, y) = 6x$ and $f_y(x, y) = -4y$ and then use them to compute the gradient vector:

$$\nabla f(x, y) = \langle 6x,\ -4y\rangle$$

$$\nabla f\left(-\tfrac{3}{4}, 0\right) = \left\langle 6 \cdot \left(-\tfrac{3}{4}\right),\ -4 \cdot 0\right\rangle = \left\langle -\tfrac{9}{2}, 0\right\rangle$$

We now compute the directional derivative:

$$D_{\mathbf{u}} f\left(-\tfrac{3}{4}, 0\right) = \nabla f\left(-\tfrac{3}{4}, 0\right) \cdot \mathbf{u} = \left\langle -\tfrac{9}{2}, 0\right\rangle \cdot \left\langle \tfrac{3}{5}, \tfrac{4}{5}\right\rangle$$

$$= \left(-\tfrac{9}{2} \cdot \tfrac{3}{5}\right) + \left(0 \cdot \tfrac{4}{5}\right) = -\tfrac{27}{10}$$

Let's now generalize to functions in three dimensions:

The directional derivative of f at the point (a, b) in the direction of a unit vector $\mathbf{u} = \langle u_1, u_2 \rangle$ is defined by

$$D_{\mathbf{u}} f(a, b) = \lim_{h \to 0} \frac{f(a + hu_1,\ b + hu_2) - f(a, b)}{h}$$

given that the limit exists.

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be a differentiable function. The gradient of f at (x, y, z) is the vector:

$$\nabla f(x, y, z) = \langle f_x(x, y, z),\ f_y(x, y, z),\ f_z(x, y, z) \rangle$$

Now we will discuss the gradient vector ∇f geometrically. Suppose $f(x, y)$ is differentiable at a point (a, b), and let $\mathbf{u}$ be any unit vector.

Then the directional derivative is given by

$$D_{\mathbf{u}} f(a, b) = \nabla f(a, b) \cdot \mathbf{u}.$$

Using the dot product formula, we get

$$\nabla f(a, b) \cdot \mathbf{u} = \|\nabla f(a, b)\| \cdot \|\mathbf{u}\| \cdot \cos(\theta),$$

where θ is the angle between $\nabla f(a, b)$ and $\mathbf{u}$. Since $\|\mathbf{u}\| = 1$, we simplify:

$$D_{\mathbf{u}} f(a, b) = \|\nabla f(a, b)\| \cos(\theta)$$

The directional derivative is *maximized* with respect to θ when $\cos(\theta) = 1$, which occurs when $\theta = 0$. In this case, the direction vector $\mathbf{u}$ points in the same direction as $\nabla f(a, b)$. Thus,

$$D_{\mathbf{u}} f(a, b) = \|\nabla f(a, b)\|.$$

This means the function increases most rapidly when you move in the direction of the gradient. The gradient vector points in the direction of steepest ascent.

The directional derivative is *minimized* with respect to θ when $\cos(\theta) = -1$, which occurs when $\theta = \pi$. In this case, the direction vector $\mathbf{u}$ points in the opposite direction of the gradient. Thus,

$$D_{\mathbf{u}}f(a,b) = -\|\nabla f(a,b)\|.$$

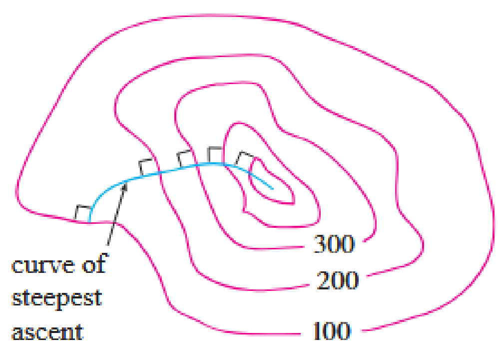
This represents the direction of steepest descent. The gradient vector points opposite the direction of the fastest decrease.

This leads to two theorems:

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable at a point (a,b), with $\nabla f(a,b) \neq \vec{0}$.

1. f is maximized when $\mathbf{u}$ points in the direction of $\nabla f(a,b)$. In this case,
$$D_{\mathbf{u}}f(a,b) = \|\nabla f(a,b)\|.$$
2. f is minimized when $\mathbf{u}$ points in the direction of $-\nabla f(a,b)$. In this case,
$$D_{\mathbf{u}}f(a,b) = -\|\nabla f(a,b)\|.$$
3. The directional derivative is zero in any direction orthogonal to $-\nabla f(a,b)$.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable at a point (a,b). Then the gradient vector $\nabla f(a,b)$ is perpendicular to the level curve of f passing through (a,b), provided $\nabla f(a,b) \neq \vec{0}$. Equivalently, the tangent line to the level curve of f at (a,b) is orthogonal to the gradient vector.



For a topographical map of a hill, the curve of steepest ascent is perpendicular to all of the contour lines. Image credit: Stewart

EXAMPLE 5.9

Suppose the temperature at a point (x, y, z) in space is given by

$$T(x, y, z) = \frac{80}{1 + x^2 + 2y^2 + 3z^2}$$

where T is in degrees Celsius and position is measured in meters. In which direction does the temperature increase most rapidly at the point $(1, 1, -2)$? What is the maximum rate of increase?

Solution:

Compute the gradient ∇T:

$$\nabla T(x, y, z) = \left\langle \frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z} \right\rangle$$

$$= \left\langle \frac{-160x}{(1 + x^2 + 2y^2 + 3z^2)^2}, \frac{-320y}{(1 + x^2 + 2y^2 + 3z^2)^2}, \frac{-480z}{(1 + x^2 + 2y^2 + 3z^2)^2} \right\rangle$$

Now, let's evaluate ∇T at $(1, 1, -2)$:

$$\nabla T(1, 1, -2) = \frac{160}{(1 + 1 + 2 + 12)^2} \cdot \langle -1,\ -2,\ 6 \rangle$$

$$= \frac{160}{256} \cdot \langle -1,\ -2,\ 6 \rangle = \frac{5}{8} \cdot \langle -1,\ -2,\ 6 \rangle$$

$$= \left\langle -\frac{5}{8}, -\frac{5}{4}, \frac{15}{4} \right\rangle$$

And thus the maximum rate of increase is the length of the gradient vector in the direction of the gradient vector:

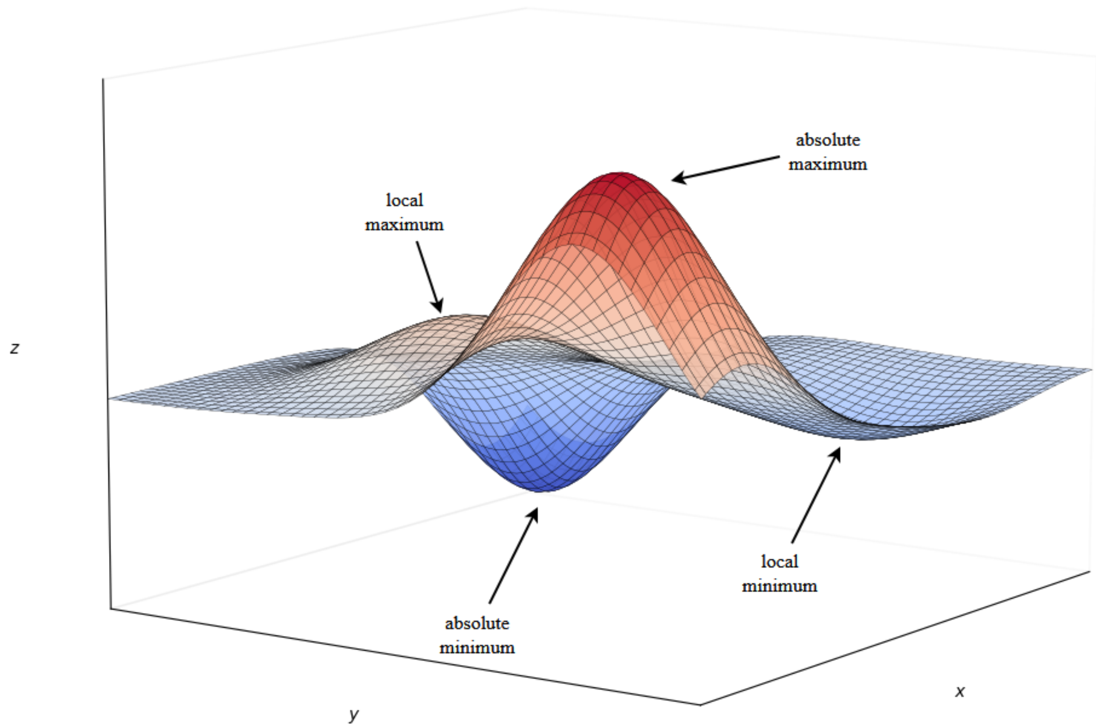
$$\|\nabla T(1, 1, -2)\| = \left\| \frac{5}{8} \cdot \langle -1,\ -2,\ 6 \rangle \right\|$$

$$= \frac{5}{8} \cdot \sqrt{1^2 + 2^2 + 6^2} = \frac{5}{8} \cdot \sqrt{41}$$

$$= \frac{5}{8} \cdot \sqrt{41}\ ^\circ\mathrm{C/m}$$

5.3 Maxima and Minima

In single-variable calculus, you found critical points using the first derivative and then the second derivative to classify them. We will now learn how to do this for multivariable functions.



A function of two variables has a **local maximum** at (a, b) if

$$f(x, y) \leq f(a, b)$$

for all points (x, y) near (a, b). In other words, $f(x, y) \leq f(a, b)$ for all points (x, y) in some disk centered at (a, b). In this case, the value $f(a, b)$ is called a local maximum value.

Similarly, f has a **local minimum** at (a, b) if

$$f(x, y) \geq f(a, b)$$

for all points (x, y) near (a, b). Then, $f(a, b)$ is a local minimum value.

If the inequality conditions above hold for all points in the domain of f, then we call $f(a,b)$ an **absolute maximum** or **absolute minimum**.

Now, we have our criteria for local extrema:

Critical points: If f has a local maximum or minimum at (a,b), and if the first-order partial derivatives of f exist at that point, then $f_x(a,b) = 0$ and $f_y(a,b) = 0$. Alternatively, this is where $\nabla f(a,b) = 0$.

EXAMPLE 5.10

Find the local and absolute minimum of the function $f(x,y) = x^2 + y^2 - 2x - 6y + 14$.

Solution:

We compute the partial derivatives:

$$f_x(x,y) = 2x - 2, \ f_y(x,y) = 2y - 6.$$

Setting the derivatives equal to zero, we solve

$$2x - 2 = 0 \Rightarrow x = 1, \ 2y - 6 = 0 \Rightarrow y = 3.$$

Thus, the only critical point is $(1,3)$. To classify the critical point, we complete the square:

$$f(x,y) = x^2 - 2x + y^2 - 6y + 14 = (x-1)^2 + (y-3)^2 + 4.$$

Since $(x-1)^2 \geq 0$ and $(y-3)^2 \geq 0$, we conclude that $f(x,y) \geq 4$ for all values of x and y. Therefore, $f(1,3) = 4$ is a local minimum, and also the absolute minimum of f.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables. Suppose that $f(x,y)$ is twice differentiable on an open disk centered at the point (a,b), where $\nabla f(a,b) = \vec{0}$. We will define the discriminant of f to be the function

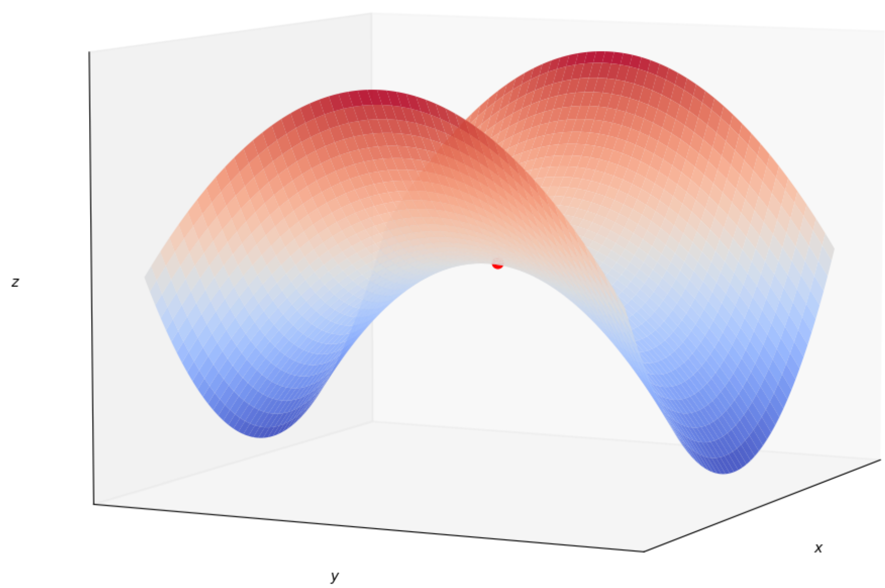
$$D(x,y) = f_{xx}(x,y)f_{yy}(x,y) - (f_{xy}(x,y))^2$$

Then we can use this to make the following conclusions:

- If $D(a,b) > 0$ and $f_{xx}(a,b) < 0$, then f has a local maximum at (a,b).
- If $D(a,b) > 0$ and $f_{xx}(a,b) > 0$, then f has a local minimum at (a,b).
- If $D(a,b) < 0$, then f has a saddle point at (a,b).
- If $D(a,b) = 0$, the test is inconclusive.

Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables. The function f is said to have a saddle point at the critical point (a,b) if and only if, in every disk centered at (a,b) the following holds:

There is at least one point (x,y) such that $f(x,y) > f(a,b)$ and at least one additional point (x,y) such that $f(x,y) < f(a,b)$.



A surface with a saddle point in red

EXAMPLE 5.11

Find and classify the critical points of the function $f(x, y) = 10x^2y - 5x^2 - 4y^2 - x^4 - 2y^4$.

Solution:

First, compute the partial derivatives:

$$f_x(x, y) = 20xy - 10x - 4x^3,\ \ f_y(x, y) = 10x^2 - 8y - 8y^3$$

To find the critical points, we solve the following:

$$f_x(x, y) = 2x(10y - 5 - 2x^2) = 0$$

$$f_y(x, y) = 10x^2 - 8y - 8y^3 = 0$$

From the first equation, we see that when $x = 0$, then $f_y = -8y(1 + y^2) = 0 \Rightarrow y = 0$, giving the critical point $(0, 0)$. If $x \neq 0$, we get $10y - 5 - 2x^2 = 0 \Rightarrow x^2 = 5y - 2.5$. We now substitute into f_y:

$$10(5y - 2.5) - 8y - 8y^3 = 0 \Rightarrow 50y - 25 - 8y - 8y^3 = 0 \Rightarrow 4y^3 - 42y + 25 = 0$$

Solving this yields $y = -2.5452$, $y = 0.6468$, $y = 1.8984$. Using $x^2 = 5y - 2.5$, we can find the corresponding x-values. For $y = -2.5452$, we get no real values. For $y = 0.6468$, we get $x = \pm 0.8567$. For $y = 1.8984$, we get $x = \pm 2.6442$. Finally, for these points, we will use the second derivative test:

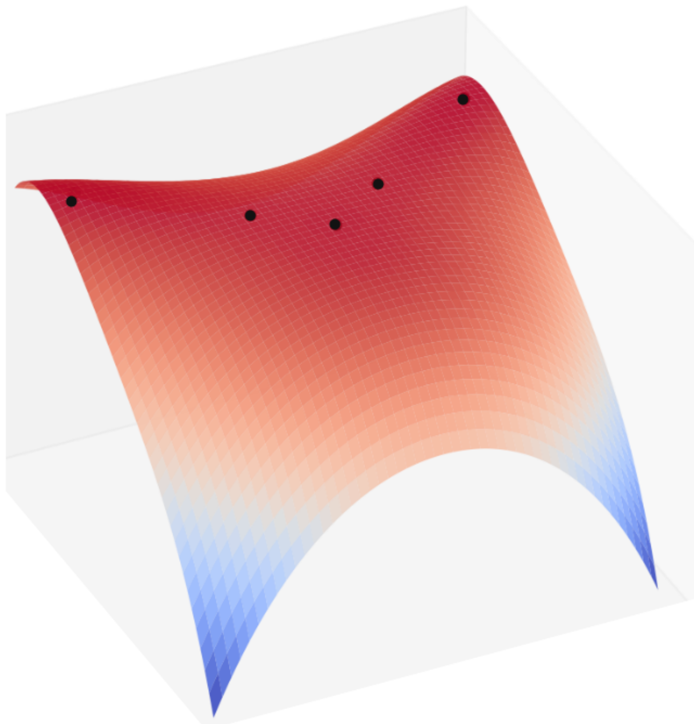
$$D = f_{xx}f_{yy} - (f_{xy})^2$$

We find the following:

Point	$f(x, y)$	f_{xx}	D	Conclusion
$(0, 0)$	0.00	-10.00	80.00	Local maximum
$(\pm 2.64, 1.90)$	8.50	-55.93	2488.72	Local maximum
$(\pm 0.86, 0.65)$	-1.48	-5.87	-187.64	Saddle point

EXAMPLE 5.11 (CONTINUED)

Here we have a graph of the surface with the critical points on it:



Here we have the critical points shown on a contour map:

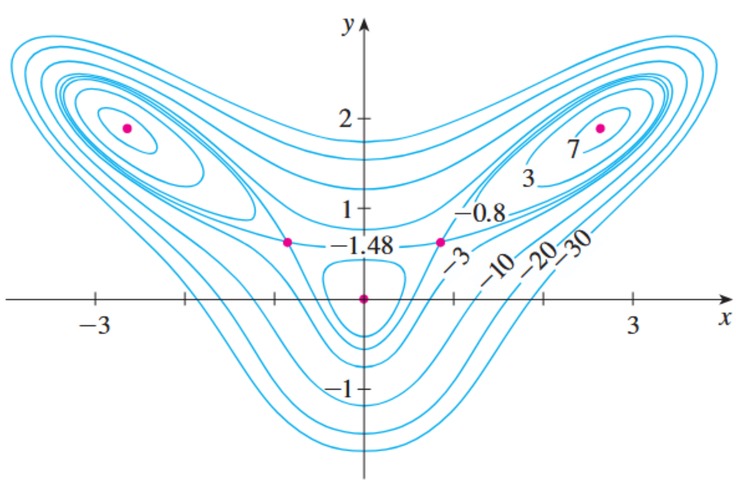


Image credit: Strang & Herman

EXAMPLE 5.12

Find the shortest distance from a general point $(1, 0, -2)$ to the plane $x + 2y + z = 4$.

Solution:

The distance d from any point (x, y, z) to the point $(1, 0, -2)$ is

$$d = \sqrt{(x-1)^2 + y^2 + (z+2)^2}.$$

If (x, y, z) lies on the plane $x + 2y + z = 4$, then $z = 4 - x - 2y$, and so

$$d = \sqrt{(x-1)^2 + y^2 + (6-x-2y)^2}.$$

To minimize d, we must minimize $f(x, y) = (x-1)^2 + y^2 + (6-x-2y)^2$. We first compute the partial derivatives:

$$f_x = 2(x-1) - 2(6-x-2y) = 4x + 4y - 14, \;\; f_y = 2y - 4(6-x-2y) = 4x + 10y - 24$$

We then set them equal to 0 and solve to yield the critical point $(\frac{11}{6}, \frac{5}{3})$.

Since $f_{xx} = 4$, $f_{xy} = 4$, and $f_{yy} = 10$, we have $D(x, y) = f_{xx}f_{yy} - (f_{xy})^2 = 24$. We thus have $D(x, y) > 0$ and $f_{xx} > 0$, meaning that f has a local minimum at $(\frac{11}{6}, \frac{5}{3})$. Given that this is the only critical point, it's also the absolute minimum.

Now we calculate the distance from $(1, 0, -2)$:

$$d = \sqrt{(x-1)^2 + y^2 + (6-x-2y)^2} = \sqrt{\left(\frac{5}{6}\right)^2 + \left(\frac{5}{3}\right)^2 + \left(\frac{5}{6}\right)^2} = \frac{5}{6}\sqrt{6}.$$

By the *extreme value theorem*, if f is continuous on a closed, bounded set $R \in \mathbb{R}^2$, then f has an absolute maximum value $f(x_1, y_1)$ and an absolute minimum value $f(x_2, y_2)$ at points $(x_1, y_1) \in D$ and $(x_2, y_2) \in D$. To find the values this theorem guarantees, we have the following:

To find the absolute maximum and minimum values of a continuous function $f : \mathbb{R}^2 \to \mathbb{R}$ on a closed, bounded set $R \subseteq \mathbb{R}^2$,

1. Find the output values of f at the critical points of f in R.
2. Find the maximum and minimum of f on the boundary of R.
3. The largest of the values from steps 1 and 2 is the absolute maximum value.
4. The smallest of these values from steps 1 and 2 is the absolute minimum value.

5.4 Lagrange Multipliers

Just in case you forgot, an *optimization* problem is where you minimize or maximize a function. We will sometimes be tasked to solve optimization problems with a constraint, which means there's a limit on how large or small a certain variable can get. To solve an optimization problem, the first thing you need to do is interpret the situation. This can include creating a visualization and modeling it through equations. Then, you want to differentiate your *objective function*, find critical points, and test. Optimization is one of the most important ideas in applied math. Every day, software engineers and mathematicians are likely working on further optimization of the very algorithms behind the search engine you are using to view this guide. Together, with economists, they also analyze different facets of the production of your computer, car, and phone with the goal of saving time, energy, and money through optimization.

EXAMPLE 5.13

A cardboard box without a lid is to have a volume of 32,000 cm^3. Find the dimensions that minimize the amount of cardboard used.

Solution:

Let the dimensions of the box be x, y, and z. The volume constraint is $xyz = 32{,}000$. The surface area to minimize is $f(x, y, z) = xy + 2xz + 2yz$. We solve the constraint for z and get $z = \frac{32{,}000}{xy}$.

Substituting this in, we get the equation

$$f(x, y) = xy + 2x\left(\frac{32{,}000}{xy}\right) + 2y\left(\frac{32{,}000}{xy}\right) = xy + \frac{64{,}000(x + y)}{xy}$$

Compute the partial derivatives:

$$f_x = y - \frac{64{,}000}{x^2}, \ f_y = x - \frac{64{,}000}{y^2}$$

Set these equal to zero to get

$$y = \frac{64{,}000}{x^2}, \ x = \frac{64{,}000}{y^2} = x = y$$

$$x^3 = 64{,}000 \Rightarrow x = y = 40$$

Then, we can solve to get $z = 20$. Lastly, the second derivative test confirms that this is is a local minimum. Thus, the dimensions of the box are $x = 40\,\text{cm}, \ y = 40\,\text{cm}, \ \text{and } z = 20\,\text{cm}$.

Suppose we want to find the extreme values of a differentiable function $f(x, y, z)$, subject to a constraint $g(x, y, z) = k$. This constraint forces us to remain on the surface S defined by the level set of g.

Let $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$ be a smooth curve that lies entirely on the surface $g(x, y, z) = k$ and passes through the point (x_0, y_0, z_0). We have

$$h(t) = f(x(t), y(t), z(t)),$$

which gives the values of f along the curve. If f has a maximum or minimum at the point (x_0, y_0, z_0), then $h(t)$ has a local extrema at $t = t_0$, and so $h'(t_0) = 0$.

Using the chain rule, we compute

$$h'(t_0) = f_x \frac{dx}{dt} + f_y \frac{dy}{dt} + f_z \frac{dz}{dt} = \nabla f(x_0, y_0, z_0) \cdot \vec{r}\,'(t_0).$$

This tells us that $\nabla f(x_0, y_0, z_0)$ is orthogonal to the tangent vector $\vec{r}\,'(t_0)$ of every curve on the surface. But since $\nabla g(x_0, y_0, z_0)$ is also orthogonal to such tangent vectors (because g is constant on the surface), the gradients must be parallel. Therefore, there exists a scalar λ such that

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0),$$

given that $\nabla g(x_0, y_0, z_0) > 0$. This is called a **Lagrange multiplier**.

Let the objective function $f : \mathbb{R}^2 \to \mathbb{R}$ and the constraint function $g : \mathbb{R}^2 \to \mathbb{R}$ be differentiable on a region in $\mathbb{R}^2$. Assume that $\nabla g(x, y) \neq \vec{0}$ on the curve C defined by the constraint $g(x, y) = 0$.

To find the maximum or minimum values of $f(x, y)$ subject to the constraint $g(x, y) = 0$, find all values of x, y, and λ that satisfy

$$\nabla f_x(x, y) = \nabla \lambda g_x(x, y)$$

and

$$g(x, y) = 0.$$

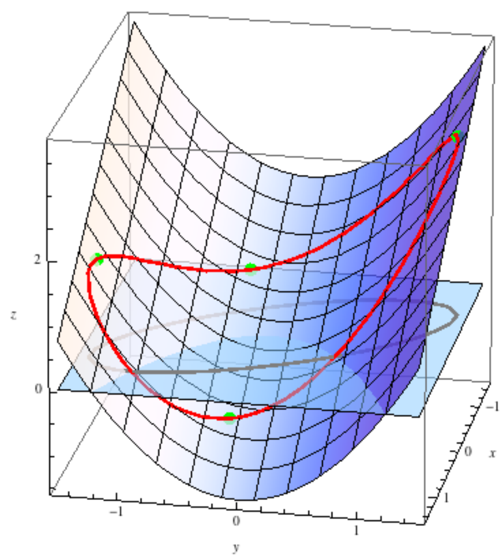
Among the points (x, y) found, evaluate $f(x, y)$ and identify the largest and smallest values. These are the maximum and minimum values of f subject to the given constraint.

The equation

$$\langle f_x(x, y), f_y(x, y) \rangle = \nabla f(x, y) = \lambda \nabla g(x, y) = \lambda \langle g_x(x, y), g_y(x, y) \rangle$$

is a vector equation. Therefore, the method of Lagrange multipliers in two variables involves solving these three equations:

1. $f_x(x, y) = \lambda g_x(x, y)$
2. $f_y(x, y) = \lambda g_y(x, y)$
3. $g(x, y) = 0$



The graph of $z = y^2 - x$ with the constraint $2x^2 + 2xy + y^2 = 1$ drawn in red and the extreme values in green. Image credit: UMich

EXAMPLE 5.14

Find the extreme values of the function $f(x, y) = x^2 + 2y^2$ on the circle $x^2 + y^2 = 1$.

Solution:

We are asked to find the extreme values of f subject to the constraint

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

Using Lagrange multipliers, we solve the equations $\nabla f = \lambda \nabla g$ and $g(x, y) = 1$.
which give:

1. $f_x = 2x = \lambda \cdot 2x$
2. $f_y = 4y = \lambda \cdot 2y$
3. $x^2 + y^2 = 1$

From the first equation, we have $x = 0$ or $\lambda = 1$.
If $x = 0$, then $x^2 + y^2 = 1 \Rightarrow y = \pm 1$.
If $\lambda = 1$, then from the second equation we get $y = 0 \Rightarrow x = \pm 1$.

Thus, the critical points are $(0, 1)$, $(0, -1)$, $(1, 0)$, and $(-1, 0)$. If we plug each of these into f, we get

$$f(0, 1) = 2,\ f(0, -1) = 2,\ f(1, 0) = 1,\ f(-1, 0) = 1.$$

Therefore, the maximum value of f on the circle 2. And the minimum value is 1.

We now expand the method of Lagrange multipliers to three variables:

Let the objective function $f : \mathbb{R}^3 \to \mathbb{R}$ and the constraint function $g : \mathbb{R}^3 \to \mathbb{R}$ be differentiable on a region in $\mathbb{R}^2$. Assume that $\nabla g(x, y, z) \neq \vec{0}$ on the curve C defined by the constraint $g(x, y, z) = 0$.

To find the maximum or minimum values of $f(x, y)$ subject to the constraint $g(x, y, z) = 0$, find all values of x, y, z, and λ that satisfy

$$\nabla f_x(x, y, z) = \nabla \lambda g_x(x, y, z)$$

and

$$g(x, y, z) = 0.$$

Among the points (x, y, z) found, evaluate $f(x, y, z)$ and identify the largest and smallest values. These are the maximum and minimum values of f subject to the given constraint.

To use this, we solve these four equations:

1. $f_x(x, y, z) = \lambda g_x(x, y, z)$
2. $f_y(x, y, z) = \lambda g_y(x, y, z)$
3. $f_z(x, y, z) = \lambda g_z(x, y, z)$
4. $g(x, y, z) = 0$

We now move on to two constraints:

Let $f(x, y, z)$ be the objective function, and suppose we have two constraints

$$g(x, y, z) = k \quad \text{and} \quad h(x, y, z) = c$$

To find the extreme values of f subject to both constraints, we solve the equation

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z)$$

Together with $g(x, y, z) = k$ and $h(x, y, z) = c$, this gives us a system of five equations:

$$\begin{gathered} f_x = \lambda g_x + \mu h_x \\ f_y = \lambda g_y + \mu h_y \\ f_z = \lambda g_z + \mu h_z \\ g(x, y, z) = k \\ h(x, y, z) = c \end{gathered}$$

Solving this system yields the candidate points for extrema of f on the intersection of the two constraint surfaces.

EXAMPLE 5.15

Find the maximum value of the function $f(x, y, z) = x + 2y + 3z$ on the curve of intersection of the plane $x - y + z = 1$ and the cylinder $x^2 + y^2 = 1$.

Solution:

We maximize $f(x, y, z)$ subject to the following constraints:

$$g(x, y, z) = x - y + z = 1, \; h(x, y, z) = x^2 + y^2 = 1$$

Using the method of Lagrange multipliers with two constraints, $\nabla f = \lambda \nabla g + \mu \nabla h$, we have to solve the following five equations:

$$1 = \lambda + 2x\mu$$
$$2 = -\lambda + 2y\mu$$
$$3 = \lambda$$
$$x - y + z = 1$$
$$x^2 + y^2 = 1$$

From the third equation, we have $\lambda = 3$. Substituting this gives

$$1 = 3 + 2x\mu \Rightarrow x = \frac{-1}{\mu}$$
$$2 = -3 + 2y\mu \Rightarrow y = \frac{5}{2\mu}$$

Substitute into the constraint:

$$\left(\frac{-1}{\mu}\right)^2 + \left(\frac{5}{2\mu}\right)^2 = 1 \Rightarrow \frac{1}{\mu^2} + \frac{25}{4\mu^2} = 1 \Rightarrow \frac{29}{4\mu^2} = 1 \Rightarrow \mu = \pm\frac{\sqrt{29}}{2}$$

EXAMPLE 5.15 (CONTINUED)

So we have

$$x = \pm\frac{2}{\sqrt{29}}, \ y = \pm\frac{5}{\sqrt{29}}, \text{ and } z = 1 - x + y = 1 \pm \frac{7}{\sqrt{29}}.$$

We then plug in to get $f(x, y, z) = x + 2y + 3z = 3 \pm \sqrt{29}$.

Here is the graph:

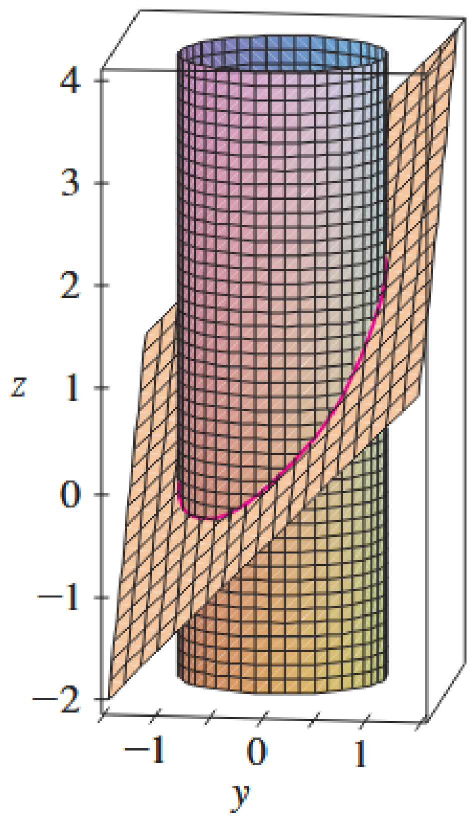


Image credit: Stewart

EXAMPLE 5.16

By investing x units of labor and y units of capital, a company produces shirts given by the function $f(x, y) = 40x^{3/5}y^{2/5}$. Determine the maximum number of shirts that can be produced on a budget of \$10,000 if labor costs \$100 per unit and capital costs \$200 per unit.

Solution:

We are to maximize $f(x, y) = 40x^{3/5}y^{2/5}$ with the constraint $g(x, y) = 100x+200y-10{,}000 = 0$. The maximum occurs at solutions of $\nabla f = \lambda\nabla g$ and $g(x, y) = 0$.
Compute the gradients:

$$f_x = \frac{\partial f}{\partial x} = 40 \cdot \frac{3}{5}x^{-2/5}y^{2/5}$$

$$f_y = \frac{\partial f}{\partial y} = 40 \cdot \frac{2}{5}x^{3/5}y^{-3/5}$$

Thus $\nabla f(x, y) = 8x^{-2/5}y^{-3/5} \cdot \langle 3y, 2x\rangle$ and $\nabla g(x, y) = \langle 100, 200\rangle$.
The critical points occur at solutions of

$$8x^{-2/5}y^{-3/5} \cdot 3y = 100\lambda$$

and

$$8x^{-2/5}y^{-3/5} \cdot 2x = 200\lambda.$$

This simplifies to

$$\frac{3y}{100} = \frac{2x}{200} \Rightarrow y = \frac{1}{3}x.$$

We substitute this into the constraint to get

$$100x + 200\left(\frac{1}{3}x\right) = 10{,}000 \Rightarrow x = 60 \Rightarrow y = 20.$$

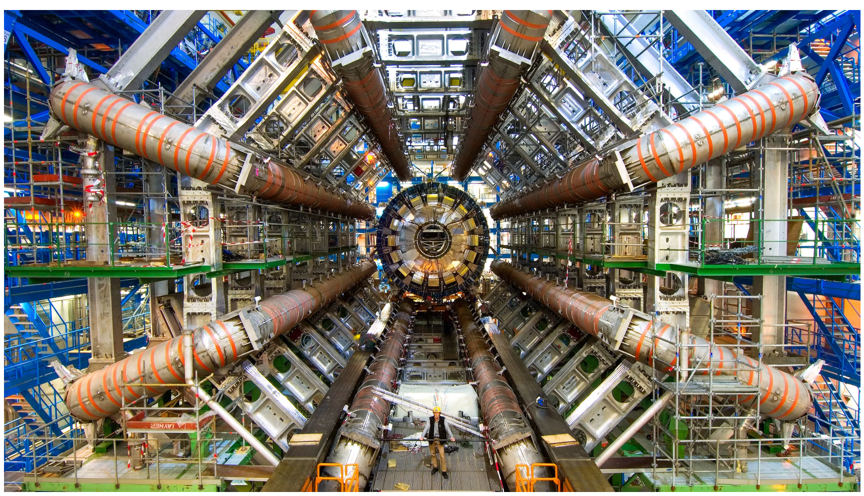
Lastly, $f(60, 20) = 40(60)^{3/5}(20)^{2/5} = 1546.55$. Thus, they can produce a maximum of 1546 shirts.

Part II

Multiple Integrals

In **Part II**, we will move beyond differentiation and focus on accumulation: computing areas, volumes, masses, and other quantities over regions in multiple dimensions. We will work with double and triple integrals over rectangular and general regions. We'll also work with them them over polar, cylindrical, and spherical coordinates which can simplify problems. Finally, we'll study the technique known as change of variables. By the end of this part, you will have learned about

- Setting up and evaluating double and triple integrals over different regions
- Using polar, cylindrical, and spherical coordinates to simplify integrals
- Applying multiple integrals
- Performing transformations and computing Jacobians
- Evaluating multiple integrals using a change of variables



The Large Hadron Collider (LHC) in Switzerland uses powerful magnetic fields to bend subatomic particles around a 27 km ring close to the speed of light. Maxwell's equations, which use surface and line integrals to govern electromagnetism, are responsible for our understanding of the science behind the LHC. Without multiple integrals, particle accelerators like the LHC and even much of modern technology would not exist. Image credit: CERN

6 Double Integrals

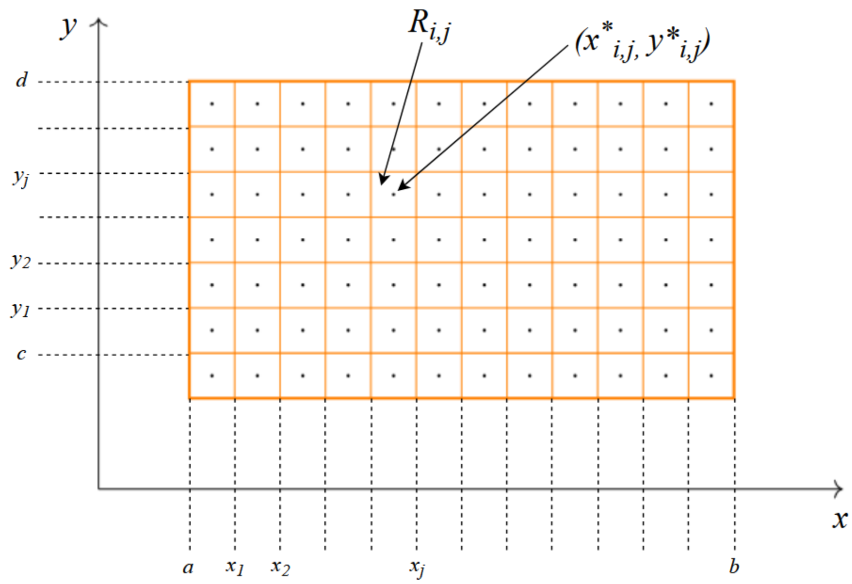
What happens if we have a function that accumulates over a region of a two dimensional plane? The truth is that when quantities accumulate in two dimensions, a single definite integral no longer is sufficient. Let's learn about what we can do instead.

6.1 Double Integration Over Rectangles

Let $z = f(x, y)$ be a nonnegative, explicit function representation of a surface in $\mathbb{R}^3$ with $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ where D is a rectangular region given by

$$D = \{(x, y) : a \leq x \leq b,\ c \leq y \leq d\} = \{(x, y) : x \in [a, b],\ y \in [c, d]\}.$$

We can divide D into subrectangles by dividing the interval into n subintervals using lines parallel to the x- and y-axes. Note that the lines and points in them do not need to be uniformly spaced:



If we choose any point $(x^*_{i,j}, y^*_{i.j})$ in each $R_{i,j}$, we can approximate the part of the surface that lies above each $R_{i,j}$ using a thin rectangular box. We will let Δx represent the width of each $R_{i,j}$ and Δy_k represent the height of each $R_{i,j}$. Thus, the area of the base is given by $\Delta A_k = \Delta x_k \Delta y_k$. Then we have $f(x^*_{i,j}, y^*_{i.j})$ representing the height of the kth box and ΔA_k representing the area of the base of the kth box for $1 \leq k \leq n$. Then, the volume of each box is given by

$$V_k = f(x^*_{i,j}, y^*_{i,j})\Delta A_k = f(x^*_{i,j}, y^*_{i,j})\Delta x_k \Delta y_k.$$

We can approximate the volume of the whole solid by adding up all of the boxes:

$$V_k \approx \sum_{i=1}^{m}\sum_{j=1}^{n} f(x^*_{i,j}, y^*_{i,j})\Delta x_k \Delta y_k,$$

where m denotes the number of subintervals $[x_{i-1}, x_i]$ and n denotes the number of subintervals $[y_{i-1}, y_i]$.

For a different division of D, the process might look like this:

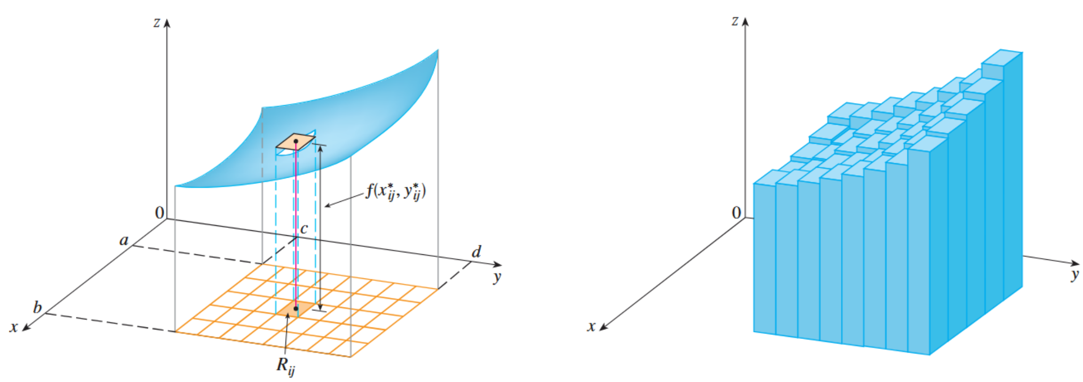


Image credit: Stewart

As the number of subintervals m and n increase, we get a better approximation. Thus, we can say that

$$V = \lim_{m,\,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} f(x^*_{i,j}, y^*_{i,j})\,\Delta A.$$

This is called a *double* Riemann sum.

Or more generally,

$$V = \iint_R f(x,y)\,dA.$$

The double Riemann sum is used to approximate double integrals.

EXAMPLE 6.1

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 $R_{i,j}$. Approximate the volume using a Riemann sum.

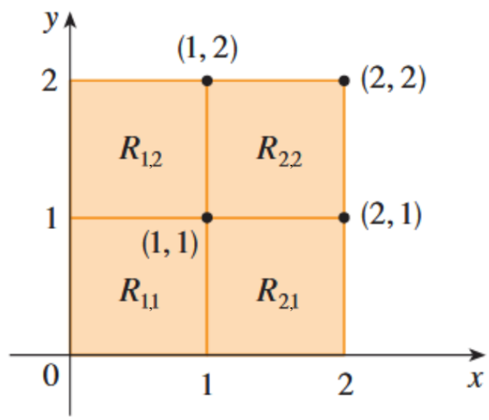
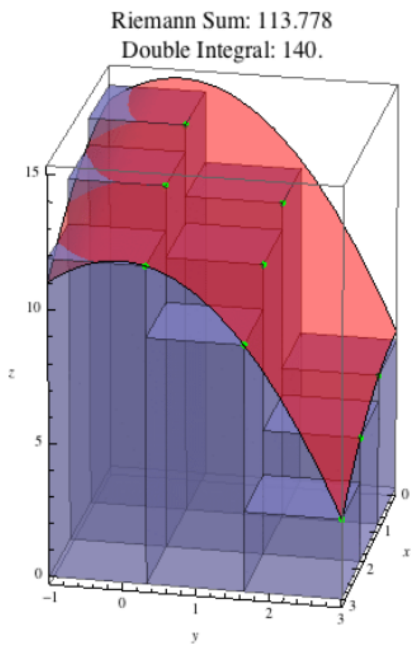


Image credit: Stewart

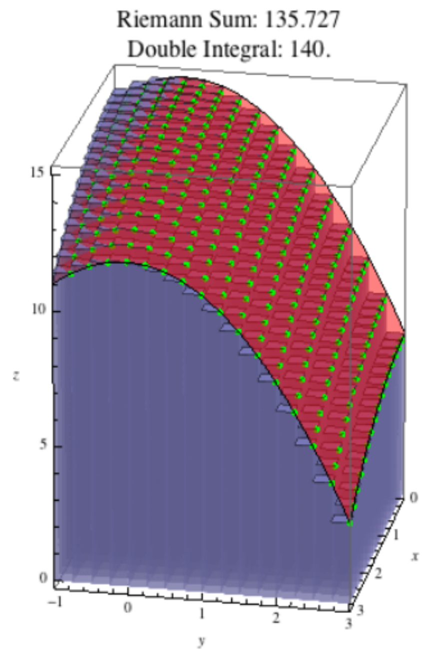
Solution:

The region is divided into four equal squares $m = n = 2$, so the area of each square is $\Delta A = 1$. We evaluate the Riemann sum using the sample points $(1,1),(1,2),(2,1),(2,2)$. The volume is approximately

$$\begin{aligned} V &\approx \sum_{i=1}^{2}\sum_{j=1}^{2} f(x_i, y_j)\,\Delta A \\ &= f(1,1)\Delta A + f(1,2)\Delta A + f(2,1)\Delta A + f(2,2)\Delta A \\ &= (13)(1) + (7)(1) + (10)(1) + (4)(1) \\ &= 34 \text{ units.} \end{aligned}$$



Less rectangles mean worse accuracy



More rectangles yield a more accurate sum

Image credit: UMich

The methods of approximating single integrals all have counterparts for double integrals. For instance, here is the midpoint rule for double integrals:

$$\iint_R f(x,y)\,dA \approx \sum_{i=1}^{m}\sum_{j=1}^{n} f(\bar{x}_i, \bar{y}_j)\,\Delta A$$

where $\bar{x}_i$ is the midpoint of $[x_{i-1}, x_i]$ and $\bar{y}_j$ is the midpoint of $[y_{j-1}, y_j]$.

EXAMPLE 6.2

Use the midpoint rule with $m = n = 2$ to estimate the value of the integral $\iint_R (x - 3y^2)\, dA$, where $R = \{(x, y) \mid 0 \le x \le 2,\ 1 \le y \le 2\}$.

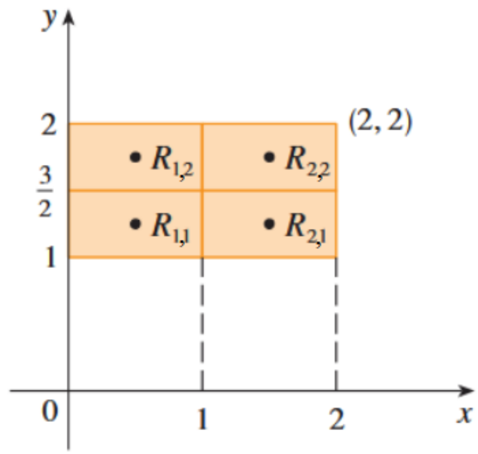


Image credit: Stewart

Solution:

Using the midpoint rule with $m = 2$, we evaluate $f(x, y) = x - 3y^2$ at the centers of the four subrectangles. The midpoints are

$$\bar{x}_1 = \frac{1}{2},\ \bar{x}_2 = \frac{3}{2},\ \bar{y}_1 = \frac{5}{4},\ \bar{y}_2 = \frac{7}{4}$$

and the area of each subrectangle is $\Delta A = \frac{1}{2}$. Thus we have

$$\iint_R (x - 3y^2)\, dA \approx \sum_{i=1}^{2} \sum_{j=1}^{2} f(\bar{x}_i, \bar{y}_j)\, \Delta A$$

$$= f(\bar{x}_1, \bar{y}_1)\, \Delta A + f(\bar{x}_1, \bar{y}_2)\, \Delta A + f(\bar{x}_2, \bar{y}_1)\, \Delta A + f(\bar{x}_2, \bar{y}_2)\, \Delta A$$

$$= f\left(\frac{1}{2}, \frac{5}{4}\right) \cdot \frac{1}{2} + f\left(\frac{1}{2}, \frac{7}{4}\right) \cdot \frac{1}{2} + f\left(\frac{3}{2}, \frac{5}{4}\right) \cdot \frac{1}{2} + f\left(\frac{3}{2}, \frac{7}{4}\right) \cdot \frac{1}{2}$$

$$= \left(-\frac{67}{16} \cdot \frac{1}{2}\right) + \left(-\frac{139}{16} \cdot \frac{1}{2}\right) + \left(-\frac{51}{16} \cdot \frac{1}{2}\right) + \left(-\frac{123}{16} \cdot \frac{1}{2}\right)$$

$$= -\frac{95}{8} = -11.875.$$

Suppose we are given a function $f(x, y) = 6 - 2x - y$ and a rectangular region $D = \{(x, y) \mid 0 \leq x \leq 1,\ 0 \leq y \leq 2\}$.

We want to evaluate the double integral

$$V = \iint_D f(x, y)\, dA. = \iint_D 6 - 2x - y\, dA.$$

Using the general slicing method, we can cut the solid using planes in the direction ($\langle 1, 0, 0\rangle$ (perpendicular to the x-axis). We will first use single-variable integration to find the area under each yz-trace as a function of x. Then, we will integrate over all values of $x \in D$ to find the total volume.

We begin by writing

$$V = \iint_D (6 - 2x - y)\, dA = \int_0^1 A(x)\, dx,$$

where

$$A(x) = \int_0^2 (6 - 2x - y)\, dy.$$

If we treat x as constant and find area under the yz-trace for $0 \leq x \leq 1$, then

$$A(x) = \int_0^2 (6 - 2x - y)\, dy = \left[6y - 2xy - \frac{y^2}{2}\right]_0^2 = 12 - 4x - 2 = 10 - 4x.$$

Now we integrate:

$$V = \int_0^1 (10 - 4x)\, dx = \left[10x - 2x^2\right]_0^1 = 10 - 2 = 8 \text{ units}^3.$$

We could have also sliced using y-direction cross sections instead. That is, using xz-traces for all y-values between $0 \leq y \leq 2$. We have

$$V = \iint_D f(x, y)\, dA = \int_0^2 A(y)\, dy,$$

where

$$A(y) = \int_0^1 (6 - 2x - y)\, dx.$$

If we treat y as constant, then

$$A(y) = \int_0^1 (6 - 2x - y)\, dx = \left[6x - x^2 - xy\right]_0^1 = 6 - 1 - y = 5 - y$$

Now integrate:

$$V = \int_0^2 (5 - y)\, dy = \left[5y - \frac{y^2}{2}\right]_0^2 = 10 - 2 = 8 \text{ units}^3.$$

We used two different approaches, but both yielded the same answer. Let's gather our thoughts by generalizing the process:

We were interested in computing the volume under a surface $z = f(x, y)$ over a rectangular region $D = \{(x, y) \mid 0 \leq x \leq 1,\ 0 \leq y \leq 2\}$. This volume is given by the double integral:

$$V = \iint_D f(x, y)\, dA$$

We can interpret this volume using the slicing method. First, fix a value of $x \in [a, b]$. For this fixed x, the vertical cross-section of the surface in the yz-plane gives a curve $f(x, y)$. The area under this curve for $y \in [c, d]$ is a function of x, which we define as

$$A(x) = \int_c^d f(x, y)\, dy.$$

This is the area under the trace of the surface for a fixed x. To recover the total volume, we integrate these area slices across $x \in [a, b]$:

$$V = \int_a^b A(x)\, dx = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx$$

Therefore, the volume under the surface can be written as an iterated integral:

$$V = \int_a^b \int_c^d f(x, y)\, dy\, dx$$

We evaluate the inner integral with respect to y first and hold x constant, which results in a function of x. Then we evaluate the outer integral.

We also sliced the solid in the other direction, fixing a value of $y \in [c, d]$ and considering the xz-trace of the surface. In this case, we define

$$A(y) = \int_a^b f(x, y)\,dx.$$

and integrate over $y \in [c, d]$:

$$V = \int_c^d A(y)\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx \right) dy$$

This gives the alternate iterated integral:

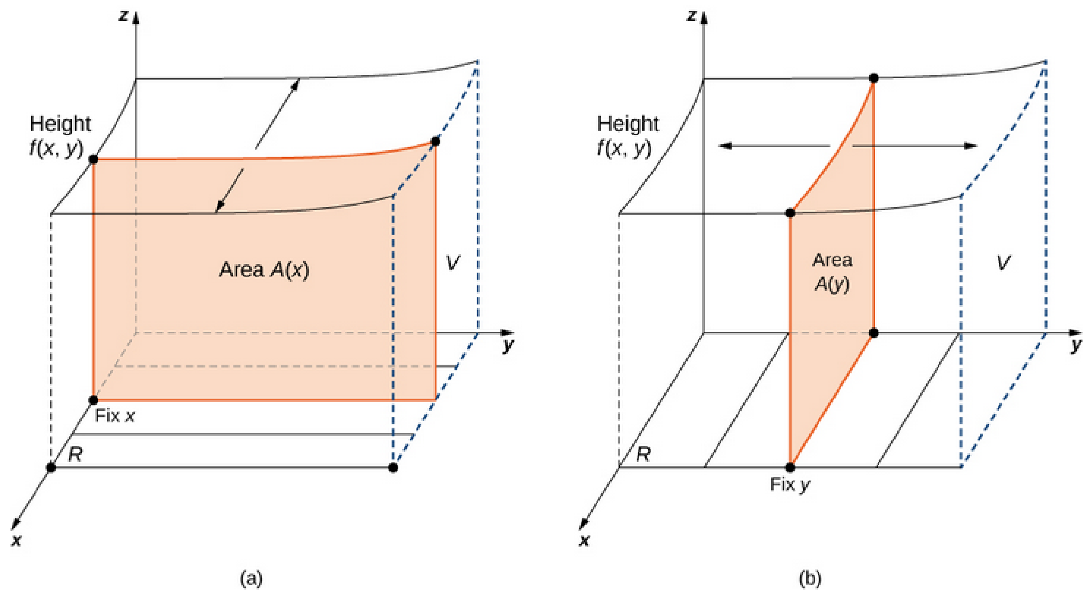
$$V = \int_c^d \int_a^b f(x, y)\,dx\,dy$$

We evaluate the inner integral with respect to x first and hold y constant, which results in a function of y. Then we evaluate the outer integral. Let's now write out our conclusion which is a property of double integrals known as Fubini's theorem.

The double integral over a rectangular region for a continuous function $f(x, y)$ over the rectangular regions $D = \{(x, y) \mid 0 \le x \le 1,\ 0 \le y \le 2\}$ can be computed by integrating in either order:

$$\iint_D f(x, y)\,dA = \int_a^b \int_c^d f(x, y)\,dy\,dx = \int_c^d \int_a^b f(x, y)\,dx\,dy$$

The differential area element dA is thus interpreted as either $dx\,dy$ or $dy\,dx$, depending on the method of slicing you pick.



This is a visualization of Fubini's theorem. In (a), we integrate first with respect to y and then with respect to x to find area $A(x)$. In (b), we integrate first with respect to x and then with respect to y to find area $A(y)$. Image credit: Strang & Herman

EXAMPLE 6.3

Evaluate the integral $\iint_D f(x,y)\,dA$, where $f(x,y) = ye^{xy}$ and $D = \{(x,y) \mid 0 \le x \le 1,\ 0 \le y \le \ln(2)\}$.

Solution:

By Fubini's theorem, we can evaluate this integral using either of the two iterated forms:

$$\iint_D ye^{xy}\,dA = \int_0^1 \left(\int_0^{\ln(2)} ye^{xy}\,dy \right) dx = \int_0^{\ln(2)} \left(\int_0^1 ye^{xy}\,dx \right) dy$$

We begin with the first form $\int_0^{\ln(2)} ye^{xy}\,dy$:
Use integration by parts: Let $u = y$, $dv = e^{xy}dy \Rightarrow v = \frac{1}{x}e^{xy}$, so

$$\int ye^{xy}\,dy = \frac{y}{x}e^{xy} - \frac{1}{x^2}e^{xy} = \frac{e^{xy}}{x^2}(xy-1)$$

Evaluate from $y = 0$ to $y = \ln(2)$:

$$\int_0^{\ln(2)} ye^{xy}\,dy = \frac{2^x}{x^2}(x\ln(2)-1) + \frac{1}{x^2}$$

So the outer integral becomes:

$$\int_0^1 \left[\frac{2^x}{x^2}(x\ln(2)-1) + \frac{1}{x^2} \right] dx$$

This is difficult to evaluate by hand.

```
format long
f = @(x) (2.^x ./ x.^2) .* (x * log(2) - 1) + 1 ./ x.^2;
I = integral(f, 0.00001, 1)  % avoid x = 0
```

```
I =
   0.306850417164929
```

Here is the result computed numerically in MATLAB. log is the natural logarithm.

```
In[1]:= Integrate[(2^x/x^2)*(x*Log[2] - 1) + 1/x^2, {x, 0, 1}]
Out[1]= 1 - Log[2]
```

Here is the result computed symbolically in Mathematica. log is the natural logarithm.

EXAMPLE 6.3 (CONTINUED)

And now we will evaluate the second form:

$$\int_0^1 ye^{xy}\,dx = y\int_0^1 e^{xy}\,dx$$

Letting $u = xy \Rightarrow du = y\,dx \Rightarrow dx = \frac{du}{y}$, we get

$$\int_0^1 e^{xy}\,dx = \frac{1}{y}\int_0^y e^u\,du = \frac{1}{y}(e^y - 1)$$

Thus,

$$y \cdot \left(\frac{1}{y}(e^y - 1)\right) = e^y - 1$$

Now we integrate:

$$\int_0^{\ln(2)} (e^y - 1)\,dy = [e^y - y]_0^{\ln(2)} = (2 - \ln(2)) - (1 - 0)$$

$$= 1 - \ln(2)$$

Recall from single-variable calculus that the average value of a function f on an interval $[a, b]$ is

$$f_{\text{avg}} = \frac{1}{b-a} \int_a^b f(x)\, dx.$$

Let's now define this for double integrals:

The average value of a function f of two variables over a rectangle R is defined to be

$$f_{\text{avg}} = \frac{1}{A(R)} \iint_R f(x, y)\, dA.$$

where $A(R)$ is the area of the region R. If $f(x, y) \geq 0$, this equation can be rearranged as:

$$A(R) \cdot f_{\text{avg}} = \iint_R f(x, y)\, dA.$$

EXAMPLE 6.4

A contour map is shown for a function f on the square region $R = [0, 4] \times [0, 4]$.

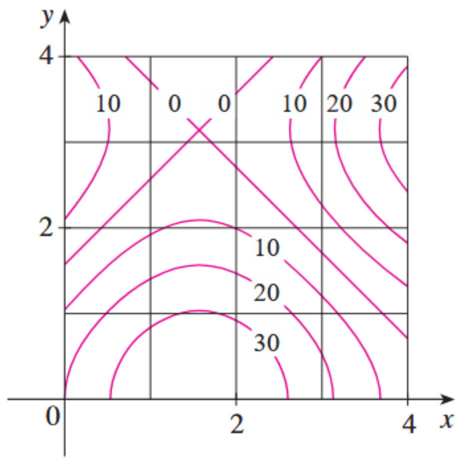


Image credit: Stewart

(a) Use the midpoint rule with $m = n = 2$ to estimate the value of $\iint_R f(x, y)\, dA$.
(b) Estimate the average value of f on R.

Solution:

We divide the region into four subrectangles, each with area $\Delta A = 4$, and estimate f at the midpoint of each:

$$\iint_R f(x, y)\, dA \approx \sum_{i=1}^{2}\sum_{j=1}^{2} f(\bar{x}_i, \bar{y}_j)\, \Delta A = \Delta A[f(1, 1) + f(1, 3) + f(3, 1) + f(3, 3)]$$

$$= \Delta A[27 + 4 + 14 + 17] = 4 \cdot 62 = 248$$

(b) The area of R is $A(R) = 16$, so the average value is

$$f_{\text{avg}} = \frac{1}{A(R)} \iint_R f(x, y)\, dA = \frac{1}{16}(248) = 15.5.$$

Lastly, we have some important properties of double integrals:

1. Linearity with respect to addition

$$\iint_R [f(x,y) + g(x,y)]\, dA = \iint_R f(x,y)\, dA + \iint_R g(x,y)\, dA$$

2. Linearity with respect to scalar multiplication

$$\iint_R cf(x,y)\, dA = c \iint_R f(x,y)\, dA \quad \text{where } c \text{ is a constant}$$

3. Monotonocity

If $f(x,y) \geq g(x,y)$ for all $(x,y) \in R$, then

$$\iint_R f(x,y)\, dA \geq \iint_R g(x,y)\, dA$$

6.2 Double Integration Over General Regions

Let $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ be a continuous function on a closed and bounded nonrectangular region D. We can partition D into rectangles just as we did with rectangular regions. The difference this time, however, is that we cannot cover a nonrectangular region perfectly with rectangles. We will only count the ones that lie completely within D. Also note that D can be enclosed by a rectangular region R.

For some surface $z = f(x, y)$ where $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, the net volume of the solid bounded by the z and D in the xy-plane can be approximated by the Riemann sum

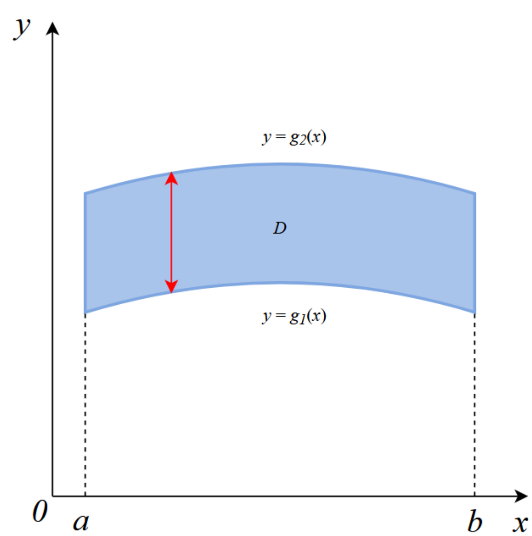
$$V \approx \sum_{k=1}^{n} f(x_k^*, y_k^*)\, \Delta A_k$$

where (x_k^*, y_k^*) is a sample point in the kth subrectangle of the partition inside D and ΔA_k is the area of the kth rectangle written as $\Delta x_k \Delta y_k$.

Type I nonrectangular regions are a region $D \subset \mathbb{R}^2$ that are known as y-simple:

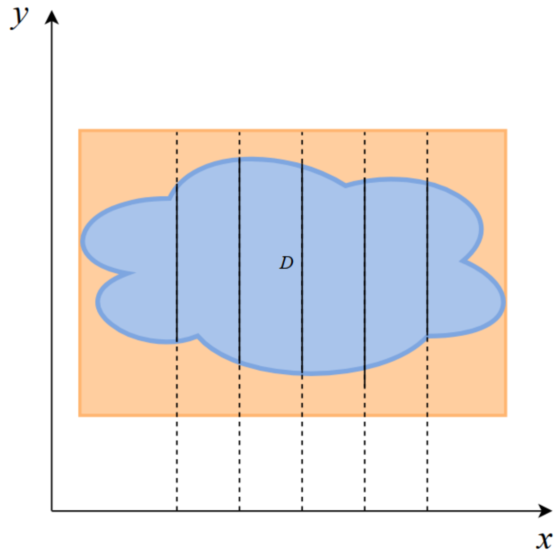
$$D = \left\{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ g_1(x) \leq y \leq g_2(x)\right\}$$

For each value of x, the vertical line through x intersects the region D in a segment between functions $y = g_1(x)$ and $y = g_2(x)$.This makes it possible to compute the area or volume using vertical slices.



A type I region lies between two vertical lines and two functions of x.

To evaluate a double integral over a type I region, we choose a rectangle $R = [a,b] \times [c,d]$ that contains D. It may look something like this:



We integrate as follows using Fubini's theorem:

$$\iint_D f(x,y)\,dA = \iint_R F(x,y)\,dA = \int_a^b \int_c^d F(x,y)\,dy\,dx$$

Notice that $F(x,y) = 0$ if $y < g_1(x)$ or $y > g_2(x)$ because (x,y) would then lie outside D. Therefore,

$$\int_c^d F(x,y)\,dy = \int_{g_1(x)}^{g_2(x)} F(x,y)\,dy = \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy,$$

because $F(x,y) = f(x,y)$ when $g_1(x) \le y \le g_2(x)$.

Thus we have the following formula that lets us evaluate the double integral as an iterated integral.

If f is continuous on a type I region D such that

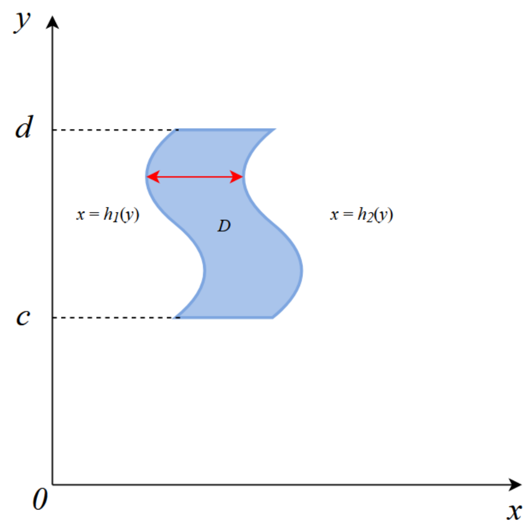
$$D = \{(x, y) \,|\, a \leq x \leq b,\ g_1(x) \leq y \leq g_2(x)\},$$

then

$$\iint_D f(x, y)\, dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x, y)\, dy\, dx.$$

Next, we have **type II** regions which are known as x-simple and satisfy

$$D = \{(x, y) \mid c \leq y \leq d,\ h_1(y) \leq x \leq h_2(y)\}.$$



A type II region lies between two horizontal lines and the graphs of two functions of y.

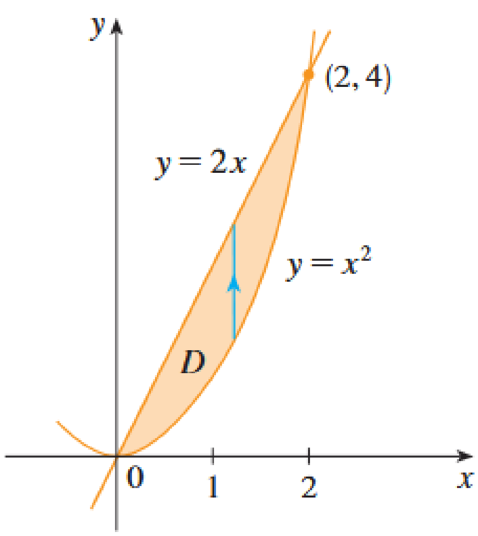
If f is continuous on a type II region D such that

$$D = \{(x, y) \mid c \le y \le d,\ h_1(y) \le x \le h_2(y)\},$$

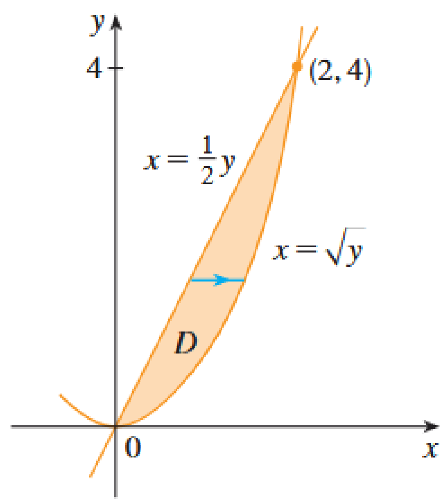
then

$$\iint_D f(x, y)\, dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x, y)\, dx\, dy,$$

where D is a type II region.



$D = \left\{(x, y) \mid 0 \le x \le 2,\ x^2 \le y \le 2x\right\}$ as a type I region



$D = \left\{(x, y) \mid 0 \le x \le 2,\ x^2 \le y \le 2x\right\}$ as a type II region

Image credit: Stewart

EXAMPLE 6.5

Evaluate the integral

$$\iint_D (x + 2y)\, dA,$$

where D is the region bounded by the parabolas $y = 2x^2$ and $y = 1 + x^2$.

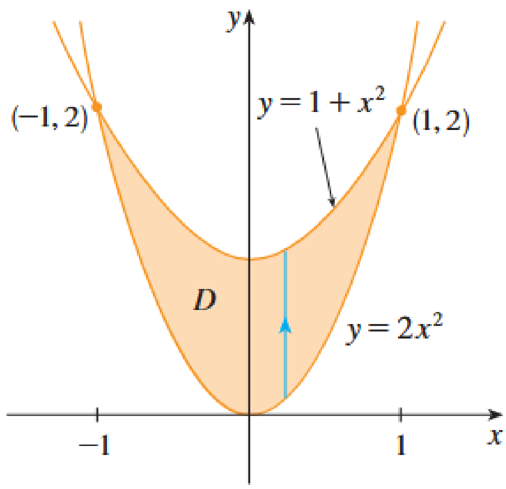


Image credit: Stewart

Solution:

The parabolas intersect when $2x^2 = 1 + x^2$, so $x^2 = 1 \Rightarrow x = \pm 1$. The region D is a type I region. We can write

$$D = \left\{(x, y) \;\middle|\; -1 \le x \le 1,\ 2x^2 \le y \le 1 + x^2\right\}.$$

EXAMPLE 6.5 (CONTINUED)

Since the lower boundary is $y = 2x^2$ and the upper boundary is $y = 1 + x^2$, we compute as follows:

$$\iint_D (x + 2y)\, dA = \int_{-1}^{1} \int_{2x^2}^{1+x^2} (x + 2y)\, dy\, dx$$

$$= \int_{-1}^{1} \left[xy + y^2\right]_{y=2x^2}^{y=1+x^2} \, dx$$

$$= \int_{-1}^{1} \left[x(1 + x^2) + (1 + x^2)^2 - x(2x^2) - (2x^2)^2\right] \, dx$$

$$= \int_{-1}^{1} \left(-3x^4 - x^3 + 2x^2 + x + 1\right) \, dx$$

$$= \left[-\frac{3x^5}{5} - \frac{x^4}{4} + \frac{2x^3}{3} + \frac{x^2}{2} + x\right]_{-1}^{1} = \frac{32}{15}$$

EXAMPLE 6.6

Let R be the region in the xy-plane bounded by the curves $y = x^3$ and $y = \sqrt{x}$:

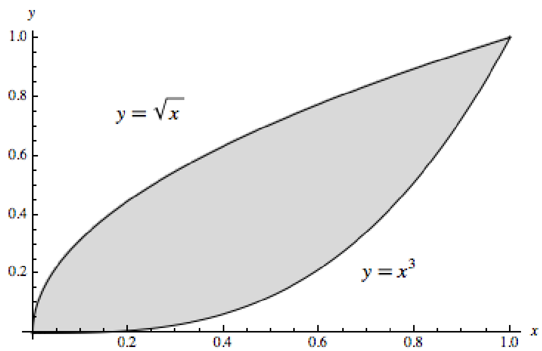


Image credit: UMich

Evaluate $\iint_R x^2 y \, dA$.

Solution:

We first identify the region R in the xy-plane. The curves $y = x^3$ and $y = \sqrt{x}$ intersect when $x^3 = \sqrt{x} \Rightarrow x(x^5 - 1) = 0$. Thus, the points of intersection occur at $x = 0$ and $x = 1$. Between these bounds, $y = x^3$ is the lower curve and $y = \sqrt{x}$ is the upper curve. Therefore, the region is of Type I:

$$R = \{(x, y) \mid 0 \leq x \leq 1,\ x^3 \leq y \leq \sqrt{x}\}.$$

We now write the double integral as an iterated integral:

$$\iint_R x^2 y \, dA = \int_0^1 \int_{x^3}^{\sqrt{x}} x^2 y \, dy \, dx$$

EXAMPLE 6.6 (CONTINUED)

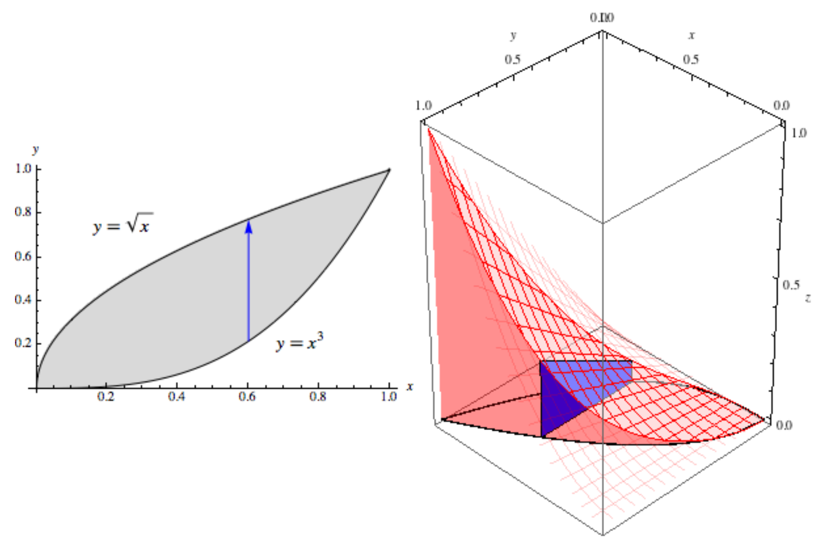
We evaluate the inner integral by holding x constant:

$$\int_{x^3}^{\sqrt{x}} x^2 y\,dy = x^2 \int_{x^3}^{\sqrt{x}} y\,dy = x^2 \left[\frac{1}{2}y^2\right]_{y=x^3}^{y=\sqrt{x}} = \frac{x^2}{2}\left(x - x^6\right) = \frac{1}{2}\left(x^3 - x^8\right)$$

Now integrate with respect to x:

$$\iint_R x^2 y\,dA = \int_0^1 \frac{1}{2}\left(x^3 - x^8\right)dx = \frac{1}{2}\left[\frac{x^4}{4} - \frac{x^9}{9}\right]_0^1 = \frac{1}{2}\left(\frac{1}{4} - \frac{1}{9}\right) = \frac{1}{2}\cdot\frac{5}{36} = \frac{5}{72}$$

The inner integral gives the area of the slice at any x-value. Let's visualize this:



The slice at $x = 0.6$. Image credit: UMich

As we let x go from 0 to 1, the slices will sweep out the entire volume of R

EXAMPLE 6.7

Find the area of the region $D \subseteq \mathbb{R}^2$, where the ordered pairs $(x, y) \in D$ satisfy the following inequalities:

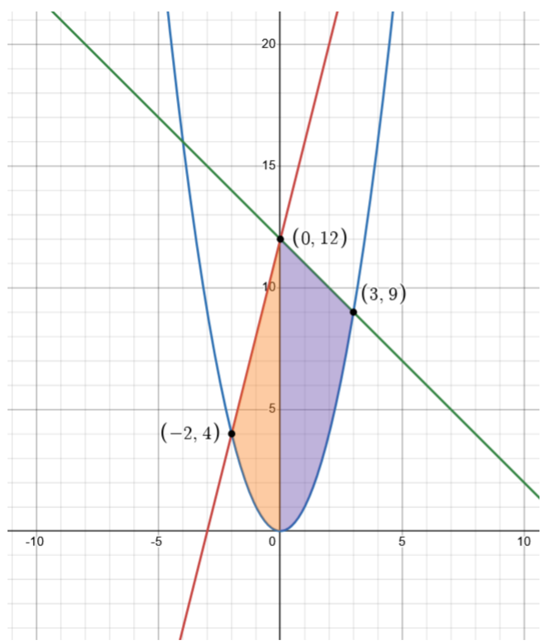
- $y \geq x^2$
- $y \leq 4x + 12$
- $y \leq 12 - x$

Solution:

We begin by analyzing the intersections of the curves:

- $x^2 = 4x + 12 \Rightarrow x^2 - 4x - 12 = 0 \Rightarrow x = -2$ or $x = 6$
- $4x + 12 = 12 - x \Rightarrow 5x = 0 \Rightarrow x = 0$
- $x^2 = 12 - x \Rightarrow x^2 + x - 12 = 0 \Rightarrow x = -4$ or $x = 3$

Let's visualize this through a plot featuring the two regions:



Here we have $y = 4x + 12$ in red, $y = x^2$ in blue, $y = 12 - x$ in green, region D_1 in orange, and region D_2 in purple.

We divide the region D into two simpler subregions D_1 and D_2, and integrate over each to get the expression for area A:

$$A(D) = A(D_1) + A(D_2) = \iint_D 1\, dA = \iint_{D_1} 1\, dA + \iint_{D_2} 1\, dA$$

EXAMPLE 6.7 (CONTINUED)

We begin with $D_1 = \{(x, y) \mid -2 \leq x \leq 0,\ x^2 \leq y \leq 4x + 12\}$:

$$A_1 = \int_{-2}^{0} \left[\int_{x^2}^{4x+12} 1\, dy \right] dx = \int_{-2}^{0} (4x + 12 - x^2)\, dx$$

$$= \int_{-2}^{0} (-x^2 + 4x + 12)\, dx = \left[-\frac{x^3}{3} + 2x^2 + 12x \right]_{-2}^{0} = (0) - \left(-\frac{-8}{3} + 8 - 24 \right) = \frac{40}{3}$$

And then $D_2 = \{(x, y) \mid 0 \leq x \leq 3,\ x^2 \leq y \leq 12 - x\}$:

$$A_2 = \int_{0}^{3} \left[\int_{x^2}^{12-x} 1\, dy \right] dx = \int_{0}^{3} (12 - x - x^2)\, dx$$

$$= \int_{0}^{3} (-x^2 - x + 12)\, dx = \left[-\frac{x^3}{3} - \frac{x^2}{2} + 12x \right]_{0}^{3} = \left(-9 - \frac{9}{2} + 36 \right) = \frac{45}{2}$$

Finally,

$$A(D) = A(D_1) + A(D_2) = \frac{40}{3} + \frac{45}{2} = \frac{215}{6} \text{ units}^2.$$

EXAMPLE 6.8

Evaluate the integral $I = \iint_D x\sqrt{1+y^3}\,dA$ where D is the triangular region bounded by the y-axis and the lines $y = \frac{1}{3}x$ and $y = 2$.

Solution:

We first sketch the region D. It lies between the line $y = \frac{1}{3}x$ and the horizontal line $y = 2$, bounded on the left by $x = 0$. Solving $y = \frac{1}{3}x \Rightarrow x = 3y$, so the right edge is $x = 6$ when $y = 2$.

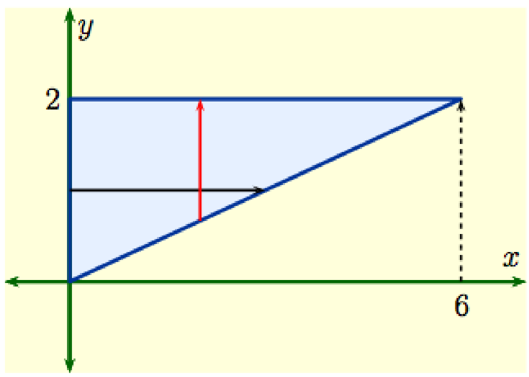


Image credit: UT Austin

If we integrate with respect to y first (vertical slices), we get

$$I = \int_0^6 \left(\int_{x/3}^2 x\sqrt{1+y^3}\,dy \right) dx$$

The inner integral $\int \sqrt{1+y^3}\,dy$ cannot be easily evaluated, so let's try the alternative. We *reverse* the order by integrating with respect to x first (horizontal slices):

$$I = \int_0^2 \left(\int_0^{3y} x\sqrt{1+y^3}\,dx \right) dy$$

Since y is constant in the inner integral,

$$= \int_0^2 \sqrt{1+y^3} \cdot \left[\tfrac{1}{2}x^2\right]_0^{3y} dy = \int_0^2 \tfrac{9}{2}y^2\sqrt{1+y^3}\,dy.$$

Let $u = 1 + y^3 \Rightarrow du = 3y^2\,dy \Rightarrow y^2\,dy = \frac{1}{3}du$.

$$I = \tfrac{9}{2} \cdot \tfrac{1}{3} \int_1^9 \sqrt{u}\,du = \tfrac{3}{2} \cdot \left[\tfrac{2}{3}u^{3/2}\right]_1^9 = \left[u^{3/2}\right]_1^9 = 9^{3/2} - 1 = 27 - 1 = 26$$

Before we move on, here is a summary of the properties of double integrals: Let $D \subset \mathbb{R}^2$ and assume that all of the following integrals exist.

1. **Linearity with respect to addition**

$$\iint_D [f(x,y) + g(x,y)]\, dA = \iint_D f(x,y)\, dA + \iint_D g(x,y)\, dA$$

2. **Linearity with respect to scalar multiplication**

$$\iint_D cf(x,y)\, dA = c \iint_D f(x,y)\, dA \quad \text{where } c \text{ is a constant}$$

3. **Monotonicity**
   If $f(x,y) \geq g(x,y)$ for all $(x,y) \in D$, then

$$\iint_D f(x,y)\, dA \geq \iint_D g(x,y)\, dA.$$

4. **Additivity over regions**
   If $D = D_1 \cup D_2$, where D_1 and D_2 do not overlap (except perhaps on their boundaries), then

$$\iint_D f(x,y)\, dA = \iint_{D_1} f(x,y)\, dA + \iint_{D_2} f(x,y)\, dA.$$

5. **Area via integration**
   The area of the region D for a constant function $f(x,y) = 1$ is given by

$$\iint_D 1\, dA = A(D).$$

6. **Bounds inequality**
   If $m \leq f(x,y) \leq M$ on D, then

$$mA(D) \leq \iint_D f(x,y)\, dA \leq MA(D).$$

EXAMPLE 6.9

Evaluate the integral $\iint_D x^2\,dA$ where D is the diamond-shaped region bounded by the lines $y = 1 - |x|$ and $y = |x| - 1$ within the square $[-1, 1] \times [-1, 1]$.

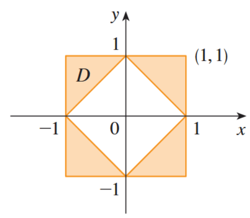


Image credit: Stewart

Solution:

We split D into two type I regions:

- Upper triangle: $D_1 = \{(x, y) \mid -1 \le x \le 1,\ -|x| + 1 \le y \le |x| - 1\}$
- Lower triangle: $D_2 = \{(x, y) \mid -1 \le x \le 1,\ |x| - 1 \le y \le -|x| + 1\}$

More precisely, we describe the regions as

$$D_1 = \{(x, y) \mid -1 \le x \le 0,\ x + 1 \le y \le -x + 1\} \cup \{(x, y) \mid 0 \le x \le 1,\ -x + 1 \le y \le x + 1\}$$
$$D_2 = \{(x, y) \mid -1 \le x \le 0,\ -x - 1 \le y \le x - 1\} \cup \{(x, y) \mid 0 \le x \le 1,\ x - 1 \le y \le -x + 1\}$$

Now compute the double integral:

$$\iint_D x^2\,dA = \int_{-1}^{0} \int_{x+1}^{-x+1} x^2\,dy\,dx + \int_{0}^{1} \int_{-x+1}^{x+1} x^2\,dy\,dx$$

Each inner integral evaluates to x^2 multipled by height. The height of the region in each case is $2(1 - |x|)$, so we have

$$\begin{aligned}
&= \int_{-1}^{0} x^2(-x + 1 - (x + 1))\,dx + \int_{0}^{1} x^2(x + 1 - (-x + 1))\,dx \\
&= \int_{-1}^{0} x^2(-2x)\,dx + \int_{0}^{1} x^2(2x)\,dx = -2\int_{-1}^{0} x^3\,dx + 2\int_{0}^{1} x^3\,dx \\
&= -2\left[\frac{x^4}{4}\right]_{-1}^{0} + 2\left[\frac{x^4}{4}\right]_{0}^{1} = -2\left(0 - \frac{1}{4}\right) + 2\left(\frac{1}{4} - 0\right) \\
&= 2 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = \frac{1}{2} + \frac{1}{2} = 1.
\end{aligned}$$

6.3 Double Integration in Polar Coordinates

There are going to be cases in which we try to integrate a function $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ where the domain $D \subseteq \mathbb{R}^2$ is expressed in polar coordinates. For instance, if the region is circular, it would be much easier to describe in polar coordinates. Recall that the polar coordinates (r, θ) of a point are related to the rectangular coordinates (x, y) by the following equations:

$$r^2 = x^2 + y^2$$
$$x = r\cos\theta$$
$$y = r\sin\theta$$

You may have used your intuition to conclude that working with polar coordinates won't be different from Cartesian coordinates because of your work in single-variable calculus. Indeed, the general methods do not change; instead, we simply encode information using a different coordinate system. As a matter of fact, we still use rectangles here. More specifically, we use **polar rectangles**. Polar rectangles are regions such that

$$R = \{(r, \theta) \mid a \le r \le b,\ \alpha \le \theta \le \beta\}.$$

In order to compute the double integral

$$\iint_R f(x, y)\, dA,$$

where R is a polar rectangle, we divide the interval $[a, b]$ into m subintervals $[r_{i-1}, r_i]$ of equal width $\Delta r = (b-a)/m$ and we divide the interval $[\alpha, \beta]$ into n subintervals $\theta_j - 1, \theta_j$ of equal width $\Delta\theta = (\beta - \alpha)/n$. Then the circles $r = r_i$ and the rays $\theta = \theta_j$ divide the polar rectangle R into the small polar rectangles $R_{i,j}$ as shown below:

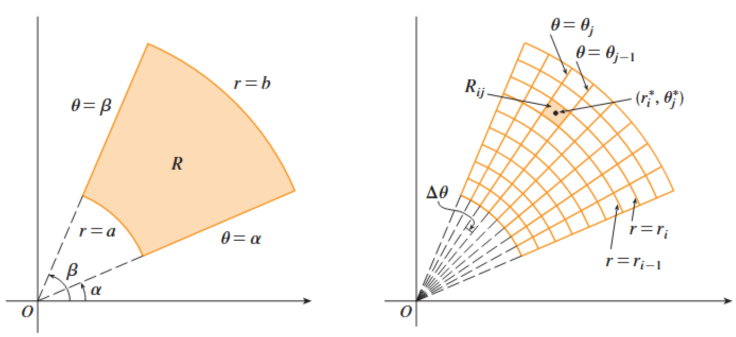


Image credit: Stewart

The center of the polar subrectangle,

$$R_{ij} = \{(r,\theta) \mid r_{i-1} \le r \le r_i,\ \theta_{j-1} \le \theta \le \theta_j\},$$

has polar coordinates $r_i^* = \frac{1}{2}(r_{i-1}+r_i)$ and $\theta_j^* = \frac{1}{2}(\theta_{j-1}+\theta_j)$. We can use the fact that the area of a sector of a circle with radius r and central angle θ is $\frac{1}{2}r^2\theta$. We subtract the areas of these two sectors with central angles $\Delta\theta = \theta_j - \theta_{j-1}$ to get the area of $R_{i,j}$:

$$\Delta A_i = \frac{1}{2}r_i^2\Delta\theta - \frac{1}{2}r_{i-1}^2\Delta\theta = \frac{1}{2}(r_i^2 - r_{i-1}^2)\Delta\theta = \frac{1}{2}(r_i + r_{i-1})(r_i - r_{i-1})\Delta\theta = r_i^*\Delta r\Delta\theta$$

We now start with a Riemann sum approximation over polar rectangles. Suppose R is a polar rectangle divided into subrectangles $R_{i,j}$, and the center of each subrectangle has polar coordinates (r_i^*, θ_j^*). Then, the rectangular coordinates are

$$x = r_i^* \cos\theta_j^*,\ y = r_i^* \sin\theta_j^*.$$

So a typical Riemann sum is

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(r_i^*\cos\theta_j^*, r_i^*\sin\theta_j^*) \cdot r_i^*\Delta r\Delta\theta.$$

If we define a new function $g(r,\theta) = f(r\cos\theta, r\sin\theta)\cdot r$, then the Riemann sum becomes

$$\sum_{i=1}^{m}\sum_{j=1}^{n} g(r_i^*, \theta_j^*)\,\Delta r\Delta\theta.$$

This is a Riemann sum for the double integral over the polar region R:

$$\int_a^b\int_\alpha^\beta g(r,\theta)\,d\theta\,dr = \int_a^b\int_\alpha^\beta f(r\cos\theta, r\sin\theta)\cdot r\,d\theta\,dr$$

Therefore, we conclude with

$$\iint_R f(x,y)\,dA = \int_\alpha^\beta\int_a^b f(r\cos\theta, r\sin\theta)\cdot r\,dr\,d\theta.$$

If f is continuous on a polar rectangle R given by

$$R = \{(r, \theta) \mid 0 \le a \le r \le b,\ \alpha \le \theta \le \beta\},$$

then the double integral becomes

$$\iint_R f(x, y)\, dA = \int_\alpha^\beta \int_a^b f(r\cos\theta, r\sin\theta) \cdot r\, dr\, d\theta.$$

Let's now discuss an important idea. The above formula says that we can convert from rectangular to polar coordinates in a double integral by writing $x = r\cos\theta$ and $y = r\sin\theta$ using the appropriate limits of integration for r and θ. When we do this, dA is not just $dr\, d\theta$. Notice how we actually have $rdr\, d\theta$. This extra r comes from what happens if we zoom in on a polar rectangle, the tiny wedge-shaped sector of a circle. In Cartesian coordinates, these rectangles have area $dx \cdot dy$. In polar coordinates, the height of the wedge is indeed dr, but the width is not solely $d\theta$. It is a *curved arc* given by $r\, d\theta$. Do not forget the r!

EXAMPLE 6.10

Evaluate the integral $I = \iint_D (x+y)\,dA$ where D is the region in the first quadrant bounded by the circle $x^2+y^2=9$, the x-axis, and the y-axis.

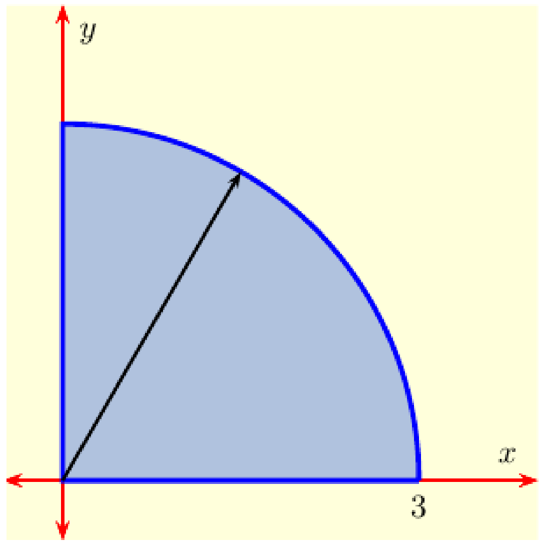


Image credit: UT Austin

Solution:

While we could work this using Cartesian coordinates, the region is much easier using polar coordinates. We simply have a circle of radius 3. Thus, our limits of integration are $0 \le r \le 3$ and $0 \le \theta \le \frac{\pi}{2}$. We will now convert the integrand using the substitutions $x = r\cos\theta$, $y = r\sin\theta$, and $dA = r\,dr\,d\theta$ to get

$$f(x,y) = x + y = r\cos\theta + r\sin\theta = r(\cos\theta + \sin\theta)$$

We now have the polar integral

$$\begin{aligned} I &= \int_0^{\pi/2} \int_0^3 r(\cos\theta + \sin\theta) \cdot r\,dr\,d\theta \\ &= \int_0^{\pi/2} \int_0^3 r^2(\cos\theta + \sin\theta)\,dr\,d\theta. \end{aligned}$$

EXAMPLE 6.10 (CONTINUED)

Let's evaluate the inner integral:

$$\begin{aligned}
&\int_0^3 r^2(\cos\theta + \sin\theta)\, dr \\
&= (\cos\theta + \sin\theta) \int_0^3 r^2\, dr \\
&= (\cos\theta + \sin\theta) \cdot \left[\frac{1}{3}r^3\right]_0^3 \\
&= 9(\cos\theta + \sin\theta)
\end{aligned}$$

And now we evaluate the outer integral:

$$\begin{aligned}
I &= \int_0^{\pi/2} 9(\cos\theta + \sin\theta)\, d\theta \\
&= 9\left(\int_0^{\pi/2} \cos\theta\, d\theta + \int_0^{\pi/2} \sin\theta\, d\theta\right) \\
&= 9(1+1) = 18
\end{aligned}$$

EXAMPLE 6.11

Find the volume beneath the surface $z = f(x, y) = 10 + xy$ and above the annular region

$$D = \{(r, \theta) \mid 2 \le r \le 4,\ 0 \le \theta < 2\pi\}.$$

Solution:

We are given a region described in polar coordinates, but the function is in rectangular form. So we use $x = r\cos\theta$ and $y = r\sin\theta$ to get

$$f(r, \theta) = 10 + r^2 \cos\theta \sin\theta.$$

Using the identity $\sin(2\theta) = 2\sin\theta\cos\theta$, we simplify:

$$f(r, \theta) = 10 + \frac{1}{2} r^2 \sin(2\theta)$$

To compute volume:

$$\begin{aligned}
V = \iint_D f(x, y)\, dA &= \iint_D \left(10 + \frac{1}{2} r^2 \sin(2\theta)\right) r\, dr\, d\theta \\
&= \int_0^{2\pi} \int_2^4 \left(10 + \frac{1}{2} r^2 \sin(2\theta)\right) r\, dr\, d\theta \\
&= \int_0^{2\pi} \left[\int_2^4 \left(10r + \frac{1}{2} r^3 \sin(2\theta)\right) dr\right] d\theta \\
&= \int_0^{2\pi} \left[5r^2 + \frac{1}{8} r^4 \sin(2\theta)\right]_2^4 d\theta \\
&= \int_0^{2\pi} \left(80 - 20 + \left[\frac{1}{8}(256 - 16)\right] \sin(2\theta)\right) d\theta \\
&= \int_0^{2\pi} (60 + 30\sin(2\theta))\, d\theta \\
&= 60 \cdot 2\pi + 30 \cdot \int_0^{2\pi} \sin(2\theta)\, d\theta \\
&= 120\pi.
\end{aligned}$$

6.4 Applications of Double Integrals

Density: Suppose a lamina (a thin plate) occupies a region $D \subset \mathbb{R}^2$, and has a density function $\rho(x, y)$ in units of mass per unit area.

Similarly to with volume, we divide a rectangle into subrectangles. We approximate the mass of each small rectangle as

$$\Delta m \approx \rho(x_i^*, y_j^*)\Delta A,$$

and then sum over all subrectangles:

$$m \approx \sum_{i=1}^{k} \sum_{j=1}^{l} \rho(x_i^*, y_j^*)\Delta A$$

Taking the limit as the partitions get finer yields

$$m = \lim_{k,l \to \infty} \sum_{i=1}^{k} \sum_{j=1}^{l} \rho(x_i^*, y_j^*)\Delta A = \iint_D \rho(x, y)\, dA.$$

This idea also works for other types of density. For example, if $\sigma(x, y)$ is the charge density, then the total charge is

$$Q = \iint_D \sigma(x, y)\, dA.$$

Center of mass: The moment of a lamina about the x-axis is

$$M_x = \iint_D y\, \rho(x, y)\, dA.$$

The moment about the y-axis is

$$M_y = \iint_D x\, \rho(x, y)\, dA.$$

The coordinates of the center of mass $(\bar{x}, \bar{y})$ are given by

$$\bar{x} = \frac{M_y}{m} = \frac{1}{m} \iint_D x\, \rho(x, y)\, dA$$

and

$$\bar{y} = \frac{M_x}{m} = \frac{1}{m} \iint_D y \, \rho(x, y) \, dA,$$

where the total mass m is

$$m = \iint_D \rho(x, y) \, dA.$$

The center of mass is an average of where particles lie in an object that you can safely use to approximate the position where most of the mass is concentrated. Thus this formula makes sense.

Here we try to balance a lamina on a thin wall given by the line $y = y_0$ where y_0 is constant. The lamina balances if and only if $y_0 = \bar{y}$:

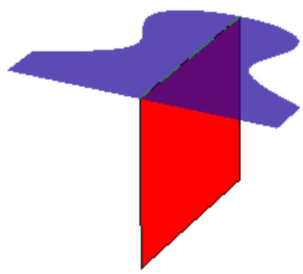


Image credit: UMich

If we want to balance the lamina on a specific point, it has to be on the center of mass $(\bar{x}, \bar{y})$:

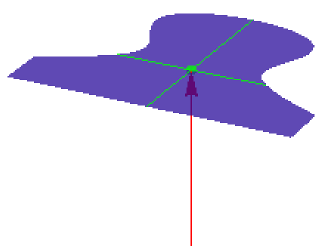


Image credit: UMich

EXAMPLE 6.12

The lamina L is the type I region bounded above by the semicircle $y = \sqrt{1-x^2}$ and below by the x-axis, over the interval $-1 \leq x \leq 1$. Assume constant density $\rho(x, y) = 1$. Compute the center of mass $(\bar{x}, \bar{y})$.

Solution:

First, the total mass of the lamina is:

$$m = \iint_L \rho(x, y)\, dA = \iint_L 1\, dA$$

Since the lamina is a semicircle of radius 1, its area is half the area of a full circle $\frac{\pi}{2}$.
Next, we compute $\bar{x}$. Because the lamina is symmetric about the y-axis, we expect $\bar{x} = 0$:

$$\begin{aligned} \bar{x} &= \frac{1}{m} \iint_L x\, dA = \frac{1}{m} \int_{-1}^{1} \int_0^{\sqrt{1-x^2}} x\, dy\, dx \\ &= \frac{1}{m} \int_{-1}^{1} x\, [y]_0^{\sqrt{1-x^2}}\, dx = \frac{1}{m} \int_{-1}^{1} x\sqrt{1-x^2}\, dx \\ &= -\frac{1}{3m} (1-x^2)^{3/2} \Big|_{-1}^{1} = 0 \end{aligned}$$

Now for $\bar{y}$:

$$\begin{aligned} \bar{y} &= \frac{1}{m} \iint_L y\, dA = \frac{1}{m} \int_{-1}^{1} \int_0^{\sqrt{1-x^2}} y\, dy\, dx \\ &= \frac{1}{m} \int_{-1}^{1} \left[\frac{1}{2} y^2\right]_0^{\sqrt{1-x^2}} dx = \frac{1}{m} \int_{-1}^{1} \frac{1}{2}(1-x^2)\, dx \\ &= \frac{1}{2m} \int_{-1}^{1} (1-x^2)\, dx = \frac{1}{2m} \left[x - \frac{x^3}{3}\right]_{-1}^{1} \\ &= \frac{1}{2m} \left(\left(1 - \frac{1}{3}\right) - \left(-1 + \frac{1}{3}\right) \right) = \frac{2}{3m} \end{aligned}$$

Substitute $m = \frac{\pi}{2}$ to get $\bar{y} = \frac{8}{3\pi}$. Therefore, the center of mass is at

$$(\bar{x}, \bar{y}) = \left(0, \frac{8}{3\pi}\right).$$

Moment of inertia: This is where a particle of mass m about an axis is defined as mr^2, where r is the distance from the particle to the axis. We have

$$I_x = \lim_{m,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} (y_{ij}^*)^2\, \rho(x_{i,j}^*, y_{i,j}^*)\, \Delta A = \iint_D y^2\, \rho(x,y)\, dA$$

$$I_y = \lim_{m,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} (x_{i,j}^*)^2\, \rho(x_{i,j}^*, y_{i,j}^*)\, \Delta A = \iint_D x^2\, \rho(x,y)\, dA$$

$$I_0 = \lim_{m,n\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n} \left[(x_{i,j}^*)^2 + (y_{i,j}^*)^2\right] \rho(x_{i,j}^*, y_{i,j}^*)\, \Delta A = \iint_D (x^2+y^2)\, \rho(x,y)\, dA$$

where x denotes the moment of inertia of the lamina about the x-axis, y denotes the moment of inertia of the lamina about the y-axis, and 0 denotes the moment of inertia about the origin. This is also called the polar moment of iniertia. Note that $I_0 = I_x + I_y$.

Surface area: Let $z = f(x,y)$ be a function defined over a region $D \subset \mathbb{R}^2$. We wish to compute the surface area of the graph of this function above the domain D.

To approximate the surface, we have a similar process of dividing D into small rectangles $R_{i,j}$ of area $\Delta A = \Delta x\, \Delta y$ with $P_{i,j} = (x_i, y_j, f(x_i, y_j))$ as a point on the surface above each subrectangle. We approximate the surface above each rectangle with a tangent plane at that point. Let's derive this:

Define two tangent vectors to the surface at $P_{i,j}$:

$$\begin{aligned}\vec{a} &= \Delta x\, \mathbf{i} + f_x(x_i, y_j)\Delta x\, \mathbf{k}\\ \vec{b} &= \Delta y\, \mathbf{j} + f_y(x_i, y_j)\Delta y\, \mathbf{k}\end{aligned}$$

The area of the parallelogram spanned by $\vec{a}$ and $\vec{b}$ is

$$\Delta T_{i,j} = \|\vec{a}\times\vec{b}\| = \sqrt{[f_x(x_i,y_j)]^2 + [f_y(x_i,y_j)]^2 + 1}\cdot \Delta A.$$

Summing over all rectangles and taking the limit gives the surface area:

$$A(S) = \lim_{\Delta x,\Delta y\to 0} \sum_{i=1}^{m}\sum_{j=1}^{n} \Delta T_{ij} = \iint_D \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dA$$

given that $f(x,y)$ has continuous partial derivatives on a region D.

This is a generalization of the arc length formula from single-variable calculus:

$$L = \int_a^b \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx$$

To better understand surface area, it's helpful to compare it to simpler geometric quantities:

- Length on the x-axis:

$$\int_a^b dx$$

- Arc length in the xy-plane for a curve $y = f(x)$:

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

- Area in the xy-plane:

$$\iint_R dA$$

- Surface area in space for a graph $z = f(x, y)$:

$$\iint_R dS = \iint_R \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dA$$

In other words, just as arc length adjusts horizontal length to account for slope, surface area adjusts flat area by accounting for slopes in both the x and y directions.

EXAMPLE 6.14

Compute the area of the region where $x \geq 0$ outside the circle $r = r_1(\theta) = \sqrt{2}$ and inside the lemniscate $r^2 = (r_2(\theta))^2 = 4\cos(2\theta)$.

Solution:

To compute the area of a region in $\mathbb{R}^2$, we set $f(r,\theta) = 1$. Now we must we determine the bounds for r and θ.

To find where the circle and lemniscate intersect, we solve:

$$2 = 4\cos(2\theta) \Rightarrow \cos(2\theta) = \frac{1}{2} \Rightarrow 2\theta = \pm\frac{\pi}{3} \Rightarrow \theta = \pm\frac{\pi}{6}.$$

Hence, the region lies between $r = \sqrt{2}$ and $r = 2\sqrt{\cos(2\theta)}$, and between $-\frac{\pi}{6} \leq \theta \leq \frac{\pi}{6}$:

$$D = \left\{(r,\theta) \mid \sqrt{2} \leq r \leq 2\sqrt{\cos(2\theta)},\ -\frac{\pi}{6} \leq \theta \leq \frac{\pi}{6}\right\}.$$

We now compute:

$$A = \iint_D f(r,\theta)\,dA = \iint_D 1 \cdot r\,dr\,d\theta.$$

Compute the inner integral:

$$\begin{aligned} A(\theta) = \int_{\sqrt{2}}^{2\sqrt{\cos(2\theta)}} r\,dr = \frac{1}{2}r^2\bigg|_{\sqrt{2}}^{2\sqrt{\cos(2\theta)}} &= \frac{1}{2}\left[(2\sqrt{\cos(2\theta)})^2 - (\sqrt{2})^2\right] \\ &= \frac{1}{2}(4\cos(2\theta) - 2) = 2\cos(2\theta) - 1. \end{aligned}$$

Integrate over θ:

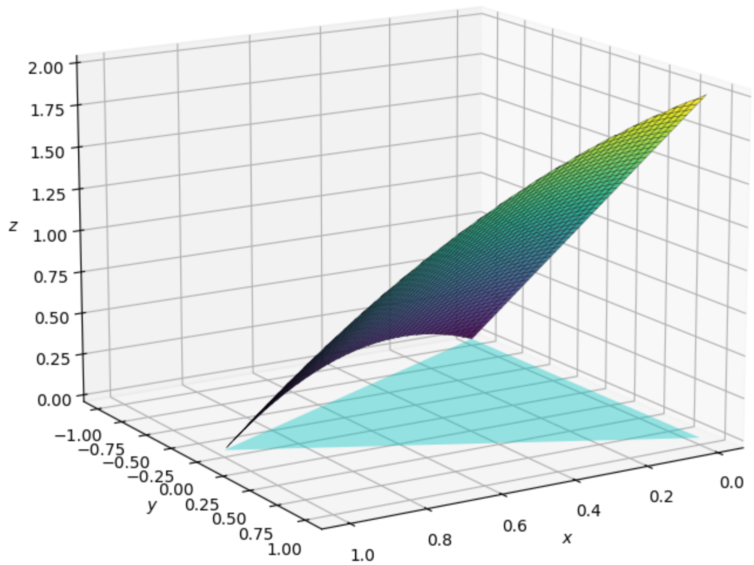
$$A = \int_{-\frac{\pi}{6}}^{\frac{\pi}{6}} (2\cos(2\theta) - 1)\,d\theta = \sin(2\theta) - \theta|_{-\frac{\pi}{6}}^{\frac{\pi}{6}}$$

$$= \left(\sin\left(\frac{\pi}{3}\right) - \frac{\pi}{6}\right) - \left(\sin\left(-\frac{\pi}{3}\right) + \frac{\pi}{6}\right) = 2\sin\left(\frac{\pi}{3}\right) - \frac{\pi}{3}.$$

$$= 2 \cdot \frac{\sqrt{3}}{2} - \frac{\pi}{3} = \sqrt{3} - \frac{\pi}{3}.$$

EXAMPLE 6.14

Find the area of the portion of the surface $f(x, y) = 1 - x^2 + y$ that lies above the triangular region with vertices $(1, 0), (0, -1), (0, 1)$. Please run the Python code for the following graph to visualize this problem for yourself.



ex6point14.py

Solution:

We will compute surface area using this formula:

$$A = \iint_R \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}\, dA$$

We compute the partial derivatives $\frac{\partial f}{\partial x} = -2x$ and $\frac{\partial f}{\partial y} = 1$ and plug them in:

$$A = \iint_R \sqrt{1 + 4x^2 + 1}\, dA = \iint_R \sqrt{2 + 4x^2}\, dA$$

From the region description, the bounds of integration are $0 \leq x \leq 1$ and $x - 1 \leq y \leq 1 - x$.

EXAMPLE 6.14 (CONTINUED)

Let's now integrate:

$$A = \int_0^1 \int_{x-1}^{1-x} \sqrt{2+4x^2}\, dy\, dx$$

Since $\sqrt{2+4x^2}$ is constant in y:

$$= \int_0^1 (2-2x)\sqrt{2+4x^2}\, dx = \int_0^1 \left(2\sqrt{2+4x^2} - 2x\sqrt{2+4x^2}\right) dx$$

We now integrate both terms separately. The antiderivative is

$$\int \left(2\sqrt{2+4x^2}\right) dx = x\sqrt{2+4x^2} + \ln\left(x+\sqrt{2+4x^2}\right) = \frac{(2+4x^2)^{3/2}}{6}.$$

Let's now put everything together:

$$A = \left[x\sqrt{2+4x^2} + \ln\left(x+\sqrt{2+4x^2}\right) - \frac{(2+4x^2)^{3/2}}{6}\right]_0^1$$

Evaluate at the bounds:

$$= \left(\sqrt{6} + \ln(2+\sqrt{6}) - \frac{6\sqrt{6}}{6}\right) - \left(0 + \ln\sqrt{2} - \frac{2\sqrt{2}}{6}\right)$$
$$= \ln(2+\sqrt{6}) - \ln\sqrt{2} + \frac{1}{3}\sqrt{2}$$

Therefore, the area is:

$$A = \ln\left(\frac{2+\sqrt{6}}{\sqrt{2}}\right) + \frac{1}{3}\sqrt{2} = 1.618 \text{ units}^2$$

Probability: In single-variable calculus, you may have worked with the *probability density function* $f(x)$ of a continuous random variable X which is a nonnegative function satisfying

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Then, for any interval $[a, b]$, the probability that X lies between a and b is given by

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx.$$

Now consider two continuous random variables X and Y, such as the height and weight of an individual, or the lifetimes of two machine parts. Their joint behavior is modeled by a *joint density function* $f(x, y)$ which satisfies:

$$f(x,y) \geq 0 \text{ for all } (x,y) \in \mathbb{R}^2 \quad \text{and} \quad \iint_{\mathbb{R}^2} f(x,y)\,dA = 1.$$

The first property of a joint density function makes sense because negative probability is not a thing. The second is because of the fact that if you add up every possibility, you would get 100% or 1.

Then, the probability that the pair (X, Y) lies within a region $D \subset \mathbb{R}^2$ is given by

$$P\left((X,Y) \in D\right) = \iint_D f(x,y)\,dA.$$

If the region D is a rectangular box defined by $a \leq x \leq b$ and $c \leq y \leq d$, then the probability that X lies between a and b becomes a double integral:

$$P(a \leq X \leq b,\ c \leq Y \leq d) = \int_a^b \int_c^d f(x,y)\,dy\,dx$$

We say that two probability distributions X with density function $f_1(x)$ and Y with density function $f_2(y)$ are independent of each other, or *independent random variables*, if their joint density function satisfies

$$f(x,y) = f_1(x) f_2(x).$$

A common real-world probability model is the exponential distribution. It is often used to model random waiting times, such as the time it takes for a radioactive isotope to decay or when a customer enters a store.

The *exponential density function* is given by

$$f(t) = \begin{cases} 0 & \text{if } t < 0 \\ \mu^{-1}e^{-t/\mu} & \text{if } t \geq 0 \end{cases}$$

where μ is the mean waiting time.

If X is a random variable with probability density function f, then its mean is

$$\mu = \int_{-\infty}^{\infty} xf(x)\,dx.$$

Now if X and Y are random variables with joint density function f, we define the X-mean and Y-mean, also called the *expected values* of X and Y, to be

$$\mu_1 = \iint_{\mathbb{R}^2} xf(x,y)\,dA$$

and

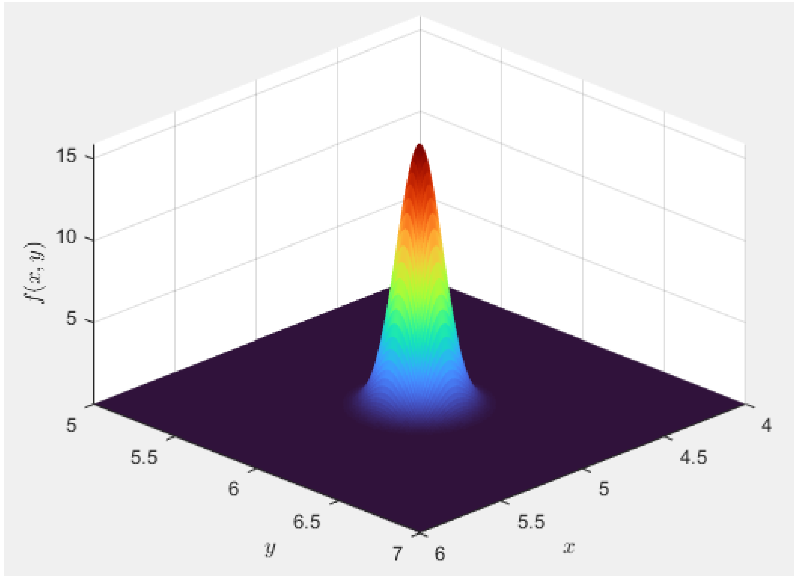
$$\mu_2 = \iint_{\mathbb{R}^2} yf(x,y)\,dA.$$

Notice how closely the expressions for μ_1 and μ_2 resemble the moments M_x and M_y of a lamina with density function ρ. In fact, we can think of probability as being like continuously distributed mass. This is because both rely on the idea of a density function. And because the total "probability mass" adds up to 1, the expressions for $\bar{x}$ and $\bar{y}$ show that we can think of the expected values of X and Y, μ_1 and μ_2, as the coordinates of the "center of mass" of the probability distribution.

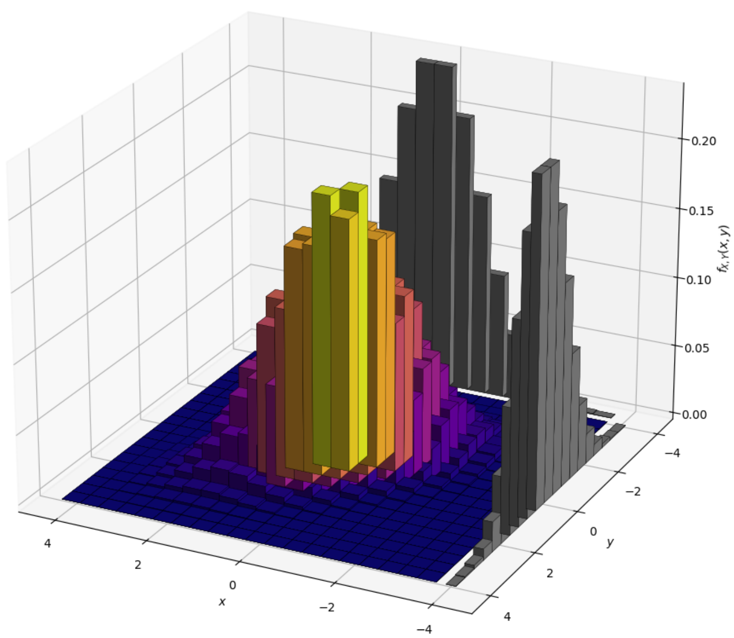
We say that a single random variable is *normally distributed* if its probability density function is of the form

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-\mu)^2/(2\sigma^2)}$$

where μ is the mean and σ is the standard deviation.



The graph of a bivariate normal joint density function



A visually interesting visualization of a bivariate normal joint density function with its marginal densities as histograms on the sides

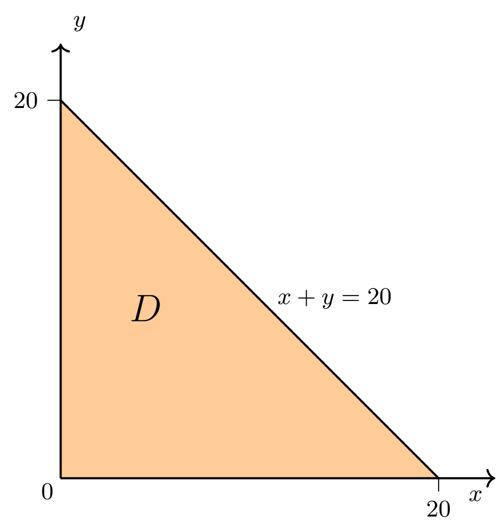
EXAMPLE 6.16

The manager of a movie theater determines that the average wait time for a moviegoer to buy a ticket is 10 minutes, and the average wait time to buy popcorn is 5 minutes. Assuming these are independent exponential wait times, what is the probability that the total wait time is less than 20 minutes?

Solution:

Let X be the time to buy a ticket and Y be the time to buy popcorn. These are modeled as such:

$$f_1(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{10}e^{-x/10} & \text{if } x \geq 0 \end{cases}, \quad f_2(y) = \begin{cases} 0 & \text{if } y < 0 \\ \frac{1}{5}e^{-y/5} & \text{if } y \geq 0 \end{cases}$$



Because X and Y are independent, their joint density is written as a product:

$$f(x,y) = f_1(x)f_2(y) = \begin{cases} \frac{1}{50}e^{-x/10}e^{-y/5} & \text{if } x \geq 0, y \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

We want to find the probability that $X + Y < 20$. This is the probability that the point (x, y) lies in the region

$$D = \left\{(x,y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0, x + y < 20\right\}.$$

EXAMPLE 6.16 (CONTINUED)

So we compute

$$P(X+Y<20) = \iint_D f(x,y)\,dA = \int_0^{20}\int_0^{20-x} \frac{1}{50} e^{-x/10} e^{-y/5}\,dy\,dx$$
$$= \frac{1}{50}\int_0^{20} e^{-x/10}\left[\int_0^{20-x} e^{-y/5}\,dy\right] dx.$$

Integrate the inner integral:

$$\int_0^{20-x} e^{-y/5}\,dy = -5e^{-y/5}\Big|_0^{20-x} = 5\left(1-e^{-(20-x)/5}\right)$$
$$= \frac{1}{50}\int_0^{20} e^{-x/10}\cdot 5\left(1-e^{-(20-x)/5}\right) dx = \frac{1}{10}\int_0^{20} e^{-x/10}\left(1-e^{-4+x/5}\right) dx$$
$$= \frac{1}{10}\int_0^{20}\left(e^{-x/10} - e^{-4}e^{-x/10}e^{x/5}\right) dx = \frac{1}{10}\int_0^{20}\left(e^{-x/10} - e^{-4}e^{x/10}\right) dx$$

Finally,

$$= \frac{1}{10}\left[-10e^{-x/10} - \frac{10}{e^4}e^{x/10}\right]_0^{20}$$
$$= \frac{1}{10}\left(-10e^{-2} - \frac{10}{e^4}e^2 + 10 + \frac{10}{e^4}\right) = 1 + e^{-4} - 2e^{-2} = 0.7476.$$

Therefore around 75% of the people at the theater will wait less than 20 minutes before getting to their seats.

7 Triple Integrals

We now extend to integration for functions of three variables that three-dimensional solids.

7.1 Triple Integration Over General Regions

The way we use area to interpret single and double integrals can actually get in the way of understanding triple integrals. Instead, try viewing integration as a weighted sum as we did in **Section 6.4**.

Let $w = f(x, y, z)$ be a real-valued function of three variables, defined over a region $D \subseteq \mathbb{R}^3$. That is,

$$f\colon D \subseteq \mathbb{R}^3 \to \mathbb{R}.$$

Assume that D is a closed and bounded region in $\mathbb{R}^3$. To construct the triple integral, we write

$$\int_D f\, dw = \iiint_D f(x, y, z)\, dV.$$

We move away from interpreting $f(x_k^*, y_k^*, z_k^*)$ as a "height" and instead treat it more generally as a weight assigned to a small volume.

We can partition the region $D \subseteq \mathbb{R}^3$ by slicing it with collections of planes:

- Planes parallel to the yz-plane in the direction $\vec{n}_1 = \alpha\langle 1, 0, 0\rangle$
- Planes parallel to the xz-plane in the direction $\vec{n}_2 = \alpha\langle 0, 1, 0\rangle$
- Planes parallel to the xy-plane in the direction $\vec{n}_3 = \alpha\langle 0, 0, 1\rangle$

This divides D into small rectangular boxes (subregions). We label each box with an index $k = 1, 2, \ldots, n$, where $n \in \mathbb{N}$, and enumerate only those boxes fully contained in D.

Let the kth box have side lengths Δx_k, Δy_k, and Δz_k. Then its volume is

$$\Delta V_k = \Delta x_k \cdot \Delta y_k \cdot \Delta z_k$$

We sample the function value within each box using a point (x_k^*, y_k^*, z_k^*) in the box and interpret $f(x_k^*, y_k^*, z_k^*)$ as the "weight" on that subregion.

We approximate the triple integral using the Riemann sum

$$\sum_{k=1}^{n} f(x_k^*, y_k^*, z_k^*) \cdot \Delta V_k.$$

To formalize convergence, define the diagonal length of the kth box as

$$d_k = \sqrt{(\Delta x_k)^2 + (\Delta y_k)^2 + (\Delta z_k)^2}.$$

Let $\Delta = \max\{d_1, d_2, \ldots, d_n\}$. Then the triple integral is defined as the limit

$$\iiint_D f(x, y, z)\, dV = \lim_{\Delta \to 0} \sum_{k=1}^{n} f(x_k^*, y_k^*, z_k^*) \cdot \Delta V_k.$$

As the maximum diagonal length $\Delta \to 0$, the number of boxes $n \to \infty$, and the approximation becomes more accurate.

The triple integral of f over the box B is

$$\iiint_B f(x, y, z)\, dV = \lim_{l,m,n \to \infty} \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} f(x_{ijk}^*, y_{ijk}^*, z_{ijk}^*)\, \Delta V$$

if this limit exists.

Just as with double integrals, Fubini's theorem also applies here:

If f is continuous on the rectangular box

$$B = [a, b] \times [c, d] \times [r, s],$$

then

$$\iiint_B f(x, y, z)\, dV = \int_r^s \int_c^d \int_a^b f(x, y, z)\, dx\, dy\, dz.$$

EXAMPLE 7.1

Evaluate the triple integral

$$\iiint_D (x + yz^2)\, dx\, dy\, dz \quad \text{where } D = [-1, 5] \times [2, 4] \times [0, 1]$$

Solution:

We integrate with respect to x, then y, then z:

$$\int_0^1 \int_2^4 \int_{-1}^5 (x + yz^2)\, dx\, dy\, dz$$

Integrate with respect to x:

$$= \int_0^1 \int_2^4 \left[\frac{x^2}{2} + xyz^2\right]_{x=-1}^{x=5} dy\, dz$$
$$= \int_0^1 \int_2^4 \left[12 + 6yz^2\right] dy\, dz$$

Integrate with respect to y:

$$= \int_0^1 \left[12y + 6 \cdot \frac{y^2}{2} z^2\right]_{y=2}^{y=4} dz$$
$$= \int_0^1 \left[24 + 36z^2\right] dz$$

Integrate with respect to z:

$$= \left[24z + 36 \cdot \frac{z^3}{3}\right]_{z=0}^{z=1}$$
$$= 24 + 12 = 36$$

EXAMPLE 7.2

Evaluate the triple integral

$$\iiint_B x^2yz\,dV \quad \text{where} \quad B = [-2,1]\times[0,3]\times[1,5].$$

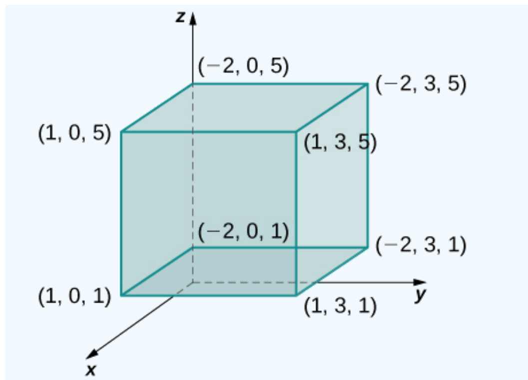


Image credit: Strang & Herman

Solution:

You can integrate in any order you want. I will pick to first integrate y, then x, and lastly z.

$$\begin{aligned}
\iiint_B x^2yz\,dV &= \int_1^5\int_{-2}^1\int_0^3 x^2yz\,dy\,dx\,dz \\
&= \int_1^5\int_{-2}^1 \left[x^2\cdot\frac{y^2}{2}\cdot z\right]_0^3 dx\,dz = \int_1^5\int_{-2}^1 \left(\frac{9}{2}x^2z\right)dx\,dz \\
&= \int_1^5 \frac{9z}{2}\left[\frac{x^3}{3}\right]_{-2}^1 dz = \int_1^5 \frac{9z}{2}\left(\frac{1^3-(-2)^3}{3}\right)dz = \int_1^5 \frac{9z}{2}\cdot\frac{9}{3}\,dz \\
&= \int_1^5 \frac{27z}{2}\,dz = \left[\frac{27z^2}{4}\right]_1^5 \\
&= \frac{27}{4}(25-1) = \frac{27\cdot 24}{4} = 162
\end{aligned}$$

So far, we've focused on triple integrals over rectangular boxes. But in practice, many solids do not have flat or rectangular boundaries. We will now generalize triple integrals to any three-dimensional solid.

We start by defining a new function F over a rectangular box B in a general region E such that

$$F(x,y,z) = \begin{cases} f(x,y,z) & \text{if } (x,y,z) \in E \\ 0 & \text{otherwise} \end{cases}.$$

Then we define the integral of f over E as

$$\iiint_E f(x,y,z)\,dV = \iiint_B F(x,y,z)\,dV.$$

This definition works as long as f is continuous and the boundary of E is reasonably smooth.

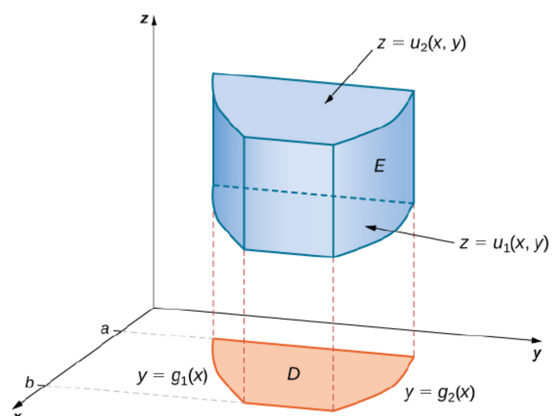
Type I Region: A solid region E is said to be type I if it lies between the graphs of two continuous functions of x and y:

$$E = \{(x,y,z) \mid (x,y) \in D,\ u_1(x,y) \le z \le u_2(x,y)\}$$

where D is the projection of E onto the xy-plane.

This is the general form of a triple integral over a type I region:

$$\iiint_E f(x,y,z)\,dV = \iint_D \left[\int_{u_1(x,y)}^{u_2(x,y)} f(x,y,z)\,dz\right] dA$$



A type I solid region. Image credit: Strang & Herman

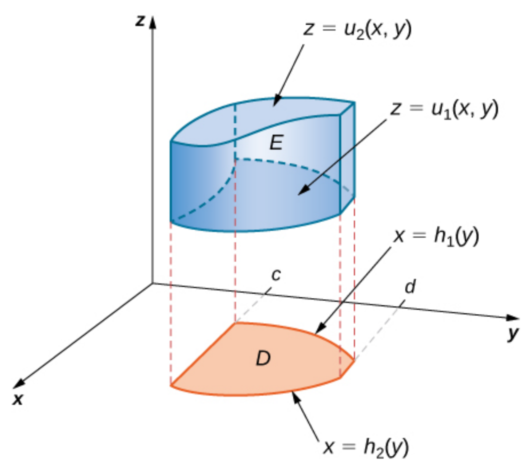
Type II Region: A solid region E is said to be type II if it lies between the graphs of two continuous functions of y and z:

$$E = \{(x, y, z) \mid (y, z) \in D,\ u_1(y, z) \le x \le u_2(y, z)\}$$

where D is the projection of E onto the yz-plane.

This is the general form of a triple integral over a type II region:

$$\iiint_E f(x, y, z)\, dV = \iint_D \left[\int_{u_1(y,z)}^{u_2(y,z)} f(x, y, z)\, dx \right] dA$$



A type II solid region. Image credit: Strang & Herman

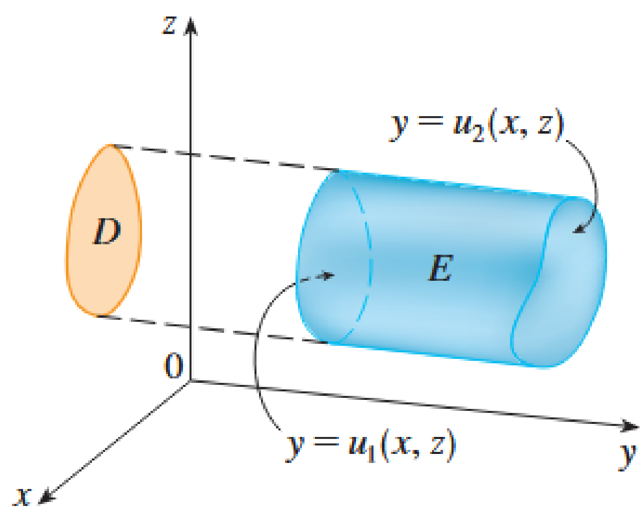
Type III Region: A solid region E is said to be type III if it lies between the graphs of two continuous functions of x and z:

$$E = \{(x, y, z) \mid (x, z) \in D,\ u_1(x, z) \leq y \leq u_2(x, z)\}$$

where D is the projection of E onto the xz-plane.

This is the general form of a triple integral over a type III region:

$$\iiint_E f(x, y, z)\, dV = \iint_D \left[\int_{u_1(x,z)}^{u_2(x,z)} f(x, y, z)\, dy \right] dA$$



A type III solid region. Image credit: Stewart

EXAMPLE 7.3

Evaluate the triple integral $\iiint_E (5x-3y)\,dV$, where E is the tetrahedron bounded by the coordinate planes and the plane $x+y+z=1$.

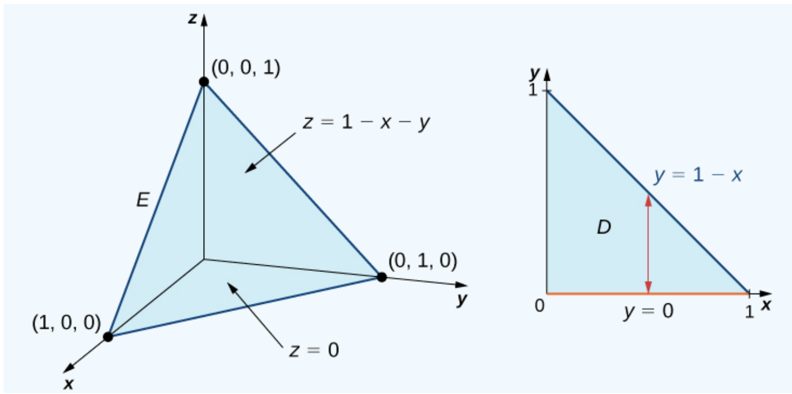


Image credit: Strang & Herman

Solution:

We first identify the projection of E onto the xy-plane. The region can be described as $E = \{(x,y,z) \mid 0 \le x \le 1,\ 0 \le y \le 1-x,\ 0 \le z \le 1-x-y\}$. Hence, the triple integral becomes

$$\iiint_E (5x-3y)\,dV = \int_0^1 \int_0^{1-x} \int_0^{1-x-y} (5x-3y)\,dz\,dy\,dx.$$

We begin by integrating with respect to z:

$$\int_0^{1-x-y} (5x-3y)\,dz = (5x-3y)(1-x-y)$$

Now integrate with respect to y:

$$\int_0^{1-x} (5x-3y)(1-x-y)\,dy = \frac{1}{2}(x-1)^2(6x-1)$$

We can integrate $\int_0^1 \frac{1}{2}(x-1)^2(6x-1)\,dx$ by expanding the integrand to $6x^3-13x^2+8x-1$:

$$\int_0^1 \frac{1}{2}(6x^3-13x^2+8x-1)\,dx = \frac{1}{2}\left[\frac{3x^4}{2} - \frac{13x^3}{3} + 4x^2 - x\right]_0^1 = \frac{1}{12}$$

Average Value of a Function of Three Variables

If $f(x, y, z)$ is integrable over a solid bounded region E with positive volume $V(E)$, then the average value of the function is

$$f_{\text{avg}} = \frac{1}{V(E)} \iiint_E f(x, y, z)\, dV.$$

Note that the volume is

$$V(E) = \iiint_E 1\, dV.$$

Many of the applications introduced in **Section 6.4**, particularly those related to physics and engineering, require triple integrals. If you have a physical quality that relies on volume, triple integrals are probably the way to go.

Let $D \subset \mathbb{R}^3$ be a solid region, and let the mass density at any point $(x, y, z) \in D$ be given by the function $\rho(x, y, z)$. Then the total mass of the solid is given by

$$m = \iiint_D \rho(x, y, z)\, dV.$$

Moments:

$$M_{yz} = \iiint_D x\, \rho(x, y, z)\, dV$$

$$M_{xz} = \iiint_D y\, \rho(x, y, z)\, dV$$

$$M_{xy} = \iiint_D z\, \rho(x, y, z)\, dV$$

Center of mass coordinates:

$$\bar{x} = \frac{M_{yz}}{m},\ \bar{y} = \frac{M_{xz}}{m},\ \bar{z} = \frac{M_{xy}}{m}$$

Moments of inertia:

$$I_x = \iiint_D (y^2 + z^2)\,\rho(x, y, z)\,dV$$

$$I_y = \iiint_D (x^2 + z^2)\,\rho(x, y, z)\,dV$$

$$I_z = \iiint_D (x^2 + y^2)\,\rho(x, y, z)\,dV$$

where $I_{xy} = I_x + I_y$, $I_{xz} = I_x + I_z$, and $I_{yz} = I_y + I_z$.

EXAMPLE 7.4

Find the moments of inertia about the x- and y-axes for the solid region lying between the hemisphere $z = \sqrt{4 - x^2 - y^2}$ and the xy-plane, given that the density at (x, y, z) is proportional to the distance between (x, y, z) and the xy-plane.

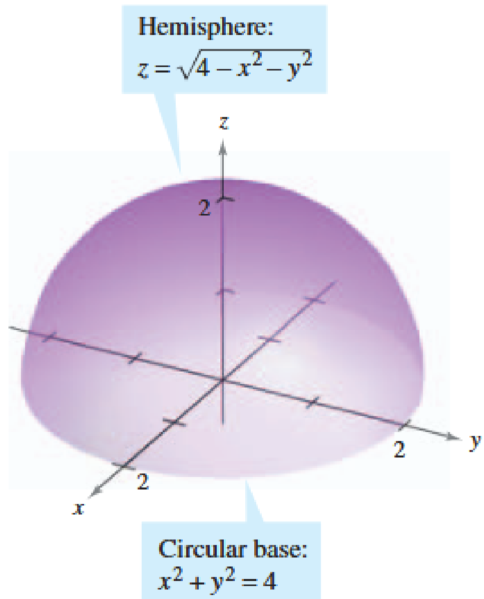


Image credit: Larson & Edwards

Solution.

The density of the region is given by $\rho(x, y, z) = kz$.
By symmetry, $I_x = I_y$, so we only need to compute one of them. We will go with the order dz, then dy, and lastly dx.

$$I_x = \iiint_Q (y^2 + z^2)\rho(x, y, z)\, dV = \iiint_Q (y^2 + z^2)(kz)\, dz\, dy\, dx$$

$$= k \int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \int_{0}^{\sqrt{4-x^2-y^2}} (y^2 + z^2)z\, dz\, dy\, dx$$

$$= k \int_{-2}^{2} \int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \left[\frac{y^2 z^2}{2} + \frac{z^4}{4} \right]_0^{\sqrt{4-x^2-y^2}} dy\, dx$$

EXAMPLE 7.4 (CONTINUED)

$$= k\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\left[\frac{y^2(4-x^2-y^2)}{2}+\frac{(4-x^2-y^2)^2}{4}\right]dy\,dx$$

$$= \frac{k}{4}\int_{-2}^{2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\left[(4-x^2)^2-y^4\right]dy\,dx$$

$$= \frac{k}{4}\int_{-2}^{2}\left[(4-x^2)^2y-\frac{y^5}{5}\right]_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}dx$$

$$= \frac{2k}{4}\int_{-2}^{2}\left[(4-x^2)^2\cdot\sqrt{4-x^2}-\frac{(4-x^2)^{5/2}}{5}\right]dx = \frac{4k}{5}\int_{-2}^{2}(4-x^2)^{5/2}dx$$

Using the substitution $x = 2\sin\theta$, we convert the integral

$$\int_{-2}^{2}(4-x^2)^{5/2}\,dx$$

into a cosine power integral:

$$= 4k\cdot\int_{0}^{2}(4-x^2)^{5/2}\,dx = 4k\cdot 64\int_{0}^{\pi/2}\cos^6\theta\,d\theta$$

We now apply the standard identity for even powers of cosine:

$$\int_{0}^{\pi/2}\cos^n\theta\,d\theta = \frac{(n-1)(n-3)\cdots 1}{n(n-2)\cdots 2}\cdot\frac{\pi}{2}$$

For $n = 6$, we get:

$$\int_{0}^{\pi/2}\cos^6\theta\,d\theta = \frac{5\cdot 3\cdot 1}{6\cdot 4\cdot 2}\cdot\frac{\pi}{2} = \frac{15}{48}\cdot\frac{\pi}{2} = \frac{5\pi}{32}$$

Finally,

$$I_x = I_y = 4k\cdot 64\cdot\frac{5\pi}{32} = \frac{256k}{5}\cdot\frac{5\pi}{32} = 8k\pi.$$

7.2 Triple Integration in Cylindrical and Spherical Coordinates

We can encode domain $D \subseteq \mathbb{R}^3$ with **cylindrical** coordinates. The reason we do this is because there are many solids that are far easier to represent with cylindrical coordinates as opposed to Cartesian or polar. Additionally, it also makes some triple integrals much easier to compute.

A point P in cylindrical coordinates is represented by the ordered triple:

$$P = (r, \theta, z)$$

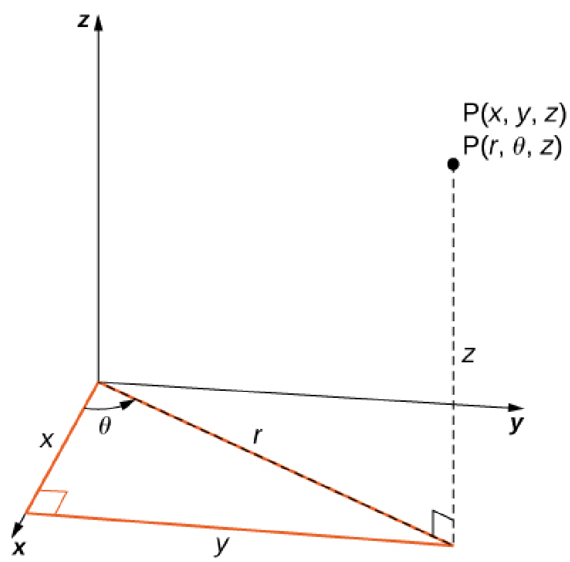
where $r \geq 0$ is the radial distance from the origin to the projection of P onto the xy-plane, $\theta \in [0, 2\pi)$ is the angle measured from the positive x-axis to the projection of P onto the xy-plane, and $z \in \mathbb{R}$ is the height of the point above the xy-plane.

We interpret this coordinate system as an extension of the polar coordinate system to three dimensions, where the first two coordinates (r, θ) describe the position of the projection P^* of the point onto the xy-plane, and the third coordinate z encodes the vertical height.

Thus, in cylindrical coordinates, a region $D \subset \mathbb{R}^3$ can be described as

$$D = \{(r, \theta, z) \mid r \in I_r,\ \theta \in I_\theta,\ z \in I_z\},$$

where I_r is the interval for radius r, I_θ is the interval for angle θ, and I_z is the interval for height z.



Cylindrical coordinates are essentially polar coordinates with an added z-component.

Let's use this system to create some regions in $\mathbb{R}^3$:

- All of space:

$$\mathbb{R}^3 = \{(r, \theta, z) \mid 0 \leq r < \infty,\ 0 \leq \theta < 2\pi,\ z \in \mathbb{R}\}$$

- A hollow cylinder centered at the z-axis with radius a and height $h_2 - h_1$:

$$C = \{(r, \theta, z) \mid r = a,\ h_1 \leq z \leq h_2,\ 0 \leq \theta \leq 2\pi)\}$$

- Cylindrical shell (annular region):

$$C_S = \{(r, \theta, z) \mid a < r \leq b,\ h_1 \leq z \leq h_2,\ 0 \leq \theta \leq 2\pi\}$$

- Positive yz-plane (i.e., a slice at $\theta = 0$ with $\vec{n} = \langle 1, 0, 0 \rangle$):

$$P_{yz} = \{(r, \theta, z) \mid \theta = 0,\ r \geq 0,\ z \in \mathbb{R}\}$$

- Positive xz-plane (i.e., a slice at $\theta = \frac{\pi}{2}$ with $\vec{n} = \langle 0, 1, 0 \rangle$):

$$P_{xz} = \left\{(r, \theta, z) \mid \theta = \frac{\pi}{2},\ r \geq 0,\ z \in \mathbb{R}\right\}$$

- Positive xy-plane (horizontal slice at $z = 0$ with $\vec{n} = \langle 0, 0, 1 \rangle$):

$$P_{xy} = \{(r, \theta, z) \mid z = 0\}$$

- Horizontal plane at height h with $\vec{n} = \langle 0, 0, 1 \rangle$ through point $(0, 0, h$:

$$P = \{(r, \theta, z) \mid z = h\}$$

- Vertical half-plane at fixed angle θ_0:

$$P = \{(r, \theta, z) \mid \theta = \theta_0\}$$

- Half-cone with vertex at the origin with linear height function $z = z(r) = ar$ where height is a function of radius:

$$C = \{(r, \theta, z) \mid z = ar,\ a \in \mathbb{R},\ r \geq 0,\ \ 0 \leq \theta \leq 2\pi\}$$

To convert between rectangular and cylindrical coordinates, we use the following relationships:

- From rectangular to cylindrical, use these:

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$z = z$$

- From cylindrical to rectangular, use these:

$$r^2 = x^2 + y^2$$
$$\tan\theta = \frac{y}{x}$$
$$z = z$$

This table summarizes common surfaces in Cartesian coordinates and their equivalents in cylindrical coordinates. These substitutions can help you evaluate triple integrals over regions bounded by these surfaces:

	Circular cylinder	**Circular cone**	**Sphere**	**Paraboloid**
Rectangular	$x^2 + y^2 = c^2$	$z^2 = c^2(x^2 + y^2)$	$x^2 + y^2 + z^2 = c^2$	$z = c(x^2 + y^2)$
Cylindrical	$r = c$	$z = cr$	$r^2 + z^2 = c^2$	$z = cr^2$

And now we introduce the formula:

Suppose that E is a type I region whose projection D onto the xy-plane is naturally described using polar coordinates. In particular, suppose that f is continuous and

$$E = \{(x, y, z) \mid (x, y) \in D,\ u_1(x, y) \le z \le u_2(x, y)\}$$

where D is given in polar coordinates as

$$D = \{(r, \theta) \mid \alpha \le \theta \le \beta,\ h_1(\theta) \le r \le h_2(\theta)\}.$$

Then the triple integral in cylindrical coordinates becomes

$$\iiint_E f(x, y, z)\, dV = \int_\alpha^\beta \int_{h_1(\theta)}^{h_2(\theta)} \int_{u_1(r\cos\theta, r\sin\theta)}^{u_2(r\cos\theta, r\sin\theta)} f(r\cos\theta, r\sin\theta, z)\, r\, dz\, dr\, d\theta.$$

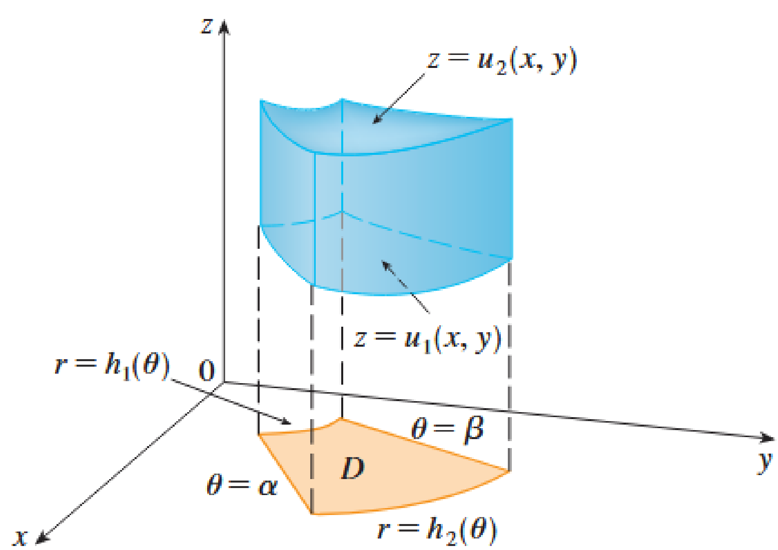


Image credit: Stewart

EXAMPLE 7.5

A solid region E lies within the cylinder $x^2 + y^2 = 1$, below the plane $z = 4$, and above the paraboloid $z = 1 - x^2 - y^2$. The density at any point is proportional to its distance from the z-axis. Find the mass of E.

Solution:

In cylindrical coordinates, the cylinder is described by $r = 1$, and the paraboloid becomes $z = 1 - r^2$. So we can express the region as

$$E = \left\{(r, \theta, z) \mid 0 \le \theta \le 2\pi,\ 0 \le r \le 1,\ 1 - r^2 \le z \le 4\right\}.$$

Since the density is proportional to the distance from the z-axis, we have

$$f(x, y, z) = K\sqrt{x^2 + y^2} = Kr.$$

where K is the proportionality constant.
Using the cylindrical form of the triple integral, the mass is

$$m = \iiint_E Kr\,dV = \int_0^{2\pi}\int_0^1\int_{1-r^2}^4 (Kr)\,r\,dz\,dr\,d\theta.$$

$$= \int_0^{2\pi}\int_0^1 Kr^2\left[4 - (1 - r^2)\right]dr\,d\theta = K\int_0^{2\pi} d\theta \int_0^1 (3r^2 + r^4)dr$$

$$= 2\pi K\left[\frac{r^3}{1} + \frac{r^5}{5}\right]_0^1 = \frac{12\pi K}{5}.$$

We now move on to another coordinate system in $\mathbb{R}^3$ that uses **spherical coordinates**. In spherical coordinates, every point is encoded as such:

$$P(\rho, \phi, \theta)$$

where ρ controls the distance from P to the origin (radius), ϕ represents the angle between the positive z-axis and the line segment connecting the origin to P, and θ measures rotation about the z-axis relative to the positive x-axis. While ρ and ϕ are new, θ is the same angle in cylindrical coordinates.

In spherical coordinates, all of space is defined as

$$\mathbb{R}^3 = \{(\rho, \phi, \theta) \mid 0 \leq \rho < \infty,\ 0 \leq \phi \leq \pi,\ 0 \leq \theta < 2\pi\}.$$

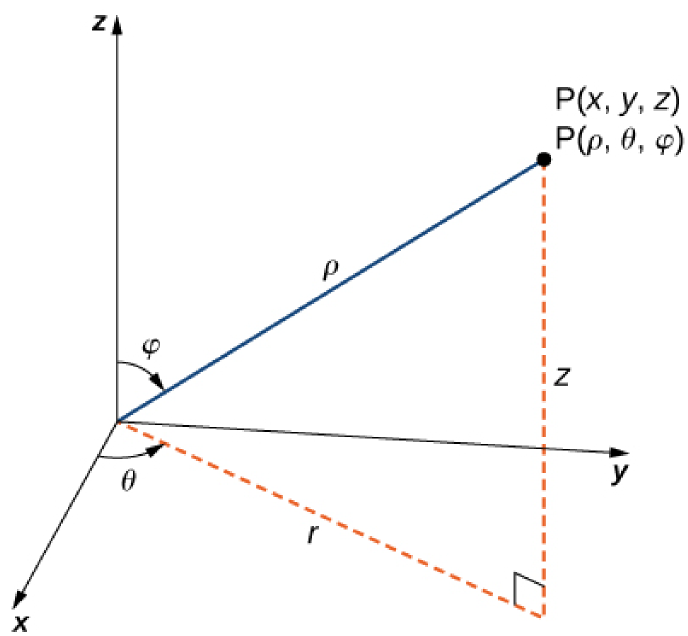


Image credit: Strang & Herman

For instance, a sphere centered at the origin with radius c has the equation $\rho = c$, $\theta = c$ represents a half-plane, and $\phi = c$ represents a half-cone.

- Use these for converting from spherical to rectangular:

$$x = \rho \sin\phi \cos\theta$$
$$y = \rho \sin\phi \sin\theta$$
$$z = \rho \cos\phi$$

- Use these for converting from rectangular to spherical:

$$\rho^2 = x^2 + y^2 + z^2$$
$$\theta = \tan^{-1}\left(\frac{y}{x}\right)$$
$$\phi = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right)$$

The distance formula gives $\rho^2 = x^2 + y^2 + z^2$, which is used for converting from rectangular to spherical coordinates.

Instead of a rectangular box, we use a spherical wedge:

$$E = \{(\rho, \theta, \phi) \mid a \leq \rho \leq b,\ \alpha \leq \theta \leq \beta,\ c \leq \phi \leq d\}$$

where $a \geq 0$ $\beta - \alpha \leq 2\pi$, and $d - c \leq \pi$. If we divide up E into small spherical wedges $E_{i,j,k}$ by means of equally spaced spheres $\rho = \rho$, half-planes $\theta = \theta_j$, and half-cones $\phi = \phi_k$. The spherical wedge can be approximated as a rectangular box with its dimensions being the arcs of circles.

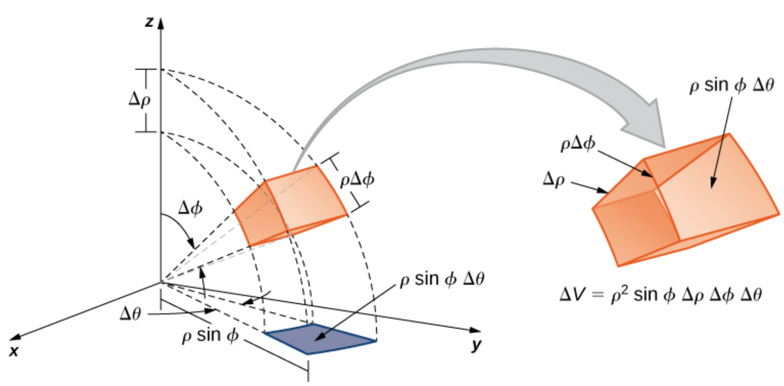


Image credit: Strang & Herman

The volume of each small wedge is:

$$\Delta V_{ijk} \approx (\Delta\rho)(\rho_i \Delta\phi)(\rho_i \sin\phi_k \Delta\theta) = \rho_i^2 \sin\phi_k \,\Delta\rho \,\Delta\theta \,\Delta\phi$$

or through the mean value theorem, the volume is given exactly by

$$\Delta V_{ijk} = \tilde{\rho}_i^2 \sin \tilde{\phi}_k^2 \, \Delta\rho \, \Delta\theta \, \Delta\phi$$

where $P(\tilde{\rho}i, \tilde{\phi}_i, \theta\phi_i)$ is some point in $E_{i,j,k}$. We will will denote the rectangular coordinate version of these as $P(x^*_{i,j,k}, y^*_{i,j,k}, z^*_{i,j,k})$.

So the triple integral becomes

$$\iiint_E f(x,y,z)\, dV = \lim_{m,n,p\to\infty} \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{p} f(x^*_{i,j,k}, y^*_{i,j,k}, z^*_{i,j,k})\Delta V_{i,j,k}$$

$$= \lim_{l,m,n\to\infty} \sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n} f\left(\tilde{\rho}_i \sin\tilde{\phi}_k \cos\tilde{\theta}_j, \tilde{\rho}_i \sin\tilde{\phi}_k \sin\tilde{\theta}_j, \tilde{\rho}_i \cos\tilde{\phi}_k\right) \tilde{\rho}_i^2 \sin\tilde{\phi}_k \, \Delta\rho \, \Delta\theta \, \Delta\phi.$$

which is a Riemann sum for

$$P(\rho,\theta,\phi) = f\left(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi\right)\rho^2 \sin\phi.$$

And now we will put this all together into the following formula:

Triple Integration in Spherical Coordinates

$$\iiint\limits_E f(x,y,z)\, dV = \int_c^d \int_\alpha^\beta \int_a^b f\left(\rho\sin\phi\cos\theta,\ \rho\sin\phi\sin\theta,\ \rho\cos\phi\right)\rho^2\sin\phi \, d\rho \, d\theta \, d\phi$$

where

$$E = \{(\rho,\theta,\phi) \mid a \le \rho \le b,\ \alpha \le \theta \le \beta,\ c \le \phi \le d\}.$$

Spheres are symmetric in all directions from the center which can make the spherical coordinate system convenient when the origin is the central point. On the other hand, cylindrical coordinates are best when symmetry is around the z-axis. In addition to the actual solid you are working with, keep this in mind when choosing to use cylindrical coordinates or spherical coordinates.

Fubini's theorem of course applies to integrals in spherical coordinates too.

EXAMPLE 7.6

Use (a) rectangular, (b) cylindrical, and (c) spherical coordinates to set up triple integrals for finding the volume of the region inside the sphere $x^2 + y^2 + z^2 = 4$ but outside the cylinder $x^2 + y^2 = 1$.

Solution:

(a)

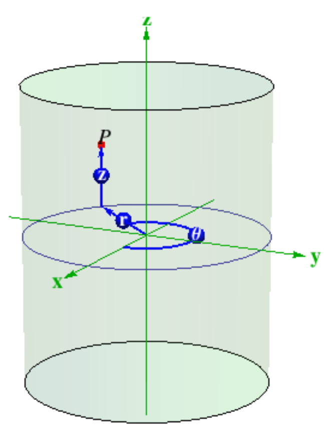
$$V = \int_{x=-2}^{2} \int_{y=-\sqrt{4-x^2}}^{\sqrt{4-x^2}} \int_{z=-\sqrt{4-x^2-y^2}}^{\sqrt{4-x^2-y^2}} dz\, dy\, dx - \int_{x=-1}^{1} \int_{y=-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{z=-\sqrt{4-x^2-y^2}}^{\sqrt{4-x^2-y^2}} dz\, dy\, dx$$

(b)

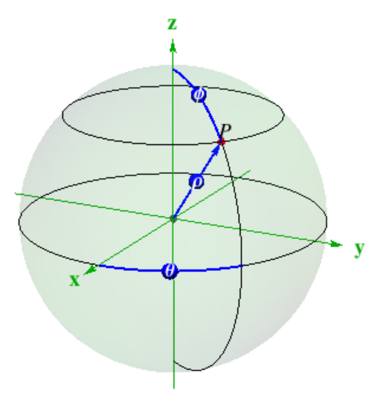
$$V = \int_{\theta=0}^{2\pi} \int_{r=1}^{2} \int_{z=-\sqrt{4-r^2}}^{\sqrt{4-r^2}} r\, dz\, dr\, d\theta$$

(c)

$$V = \int_{\theta=0}^{2\pi} \int_{\phi=\pi/6}^{5\pi/6} \int_{\rho=\csc\phi}^{2} \rho^2 \sin\phi\, d\rho\, d\phi\, d\theta$$



Cylindrical coordinates



Spherical coordinates

Image credit: UT Austin

EXAMPLE 7.7

Evaluate $\iiint_B e^{(x^2+y^2+z^2)^{1/2}}\,dV$, where B is the unit ball $B = \left\{(x,y,z)\,\middle|\,x^2+y^2+z^2 \le 1\right\}$.

Solution:

Since the boundary of B is a sphere, we use spherical coordinates:

$$B = \{(\rho,\theta,\phi)\,|\,0 \le \rho \le 1, \quad 0 \le \theta \le 2\pi, \quad 0 \le \phi \le \pi\}$$

In addition, spherical coordinates are appropriate because we have a solid of the form $x^2 + y^2 + z^2 = \rho^2$.
So the integral becomes

$$\iiint_B e^{(x^2+y^2+z^2)^{3/2}}\,dV = \int_0^{\pi}\int_0^{2\pi}\int_0^1 e^{(\rho^2)^{3/2}} \cdot \rho^2 \sin\phi\,d\rho\,d\theta\,d\phi.$$
$$= \int_0^{\pi} \sin\phi\,d\phi \cdot \int_0^{2\pi} d\theta \cdot \int_0^1 \rho^2 e^{\rho^3}\,d\rho.$$

Let $u = \rho^3$, so that $du = 3\rho^2 d\rho \Rightarrow d\rho = \frac{du}{3\rho^2}$. Then

$$\int_0^1 \rho^2 e^{\rho^3}\,d\rho = \frac{1}{3}\int_0^1 e^u\,du = \frac{1}{3}(e-1).$$
$$\int_0^{\pi} \sin\phi\,d\phi = [-\cos\phi]_0^{\pi} = -\cos\pi + \cos 0 = 2,$$
$$\int_0^{2\pi} d\theta = 2\pi.$$

Putting it all together, we get

$$\iiint_B e^{(x^2+y^2+z^2)^{3/2}}\,dV = 2 \cdot 2\pi \cdot \frac{1}{3}(e-1) = \frac{4\pi}{3}(e-1).$$

If we were to evaluate this using Cartesian coordinates, we would have to evaluate

$$\int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}\int_{-\sqrt{1-x^2-y^2}}^{\sqrt{1-x^2-y^2}} e^{(x^2+y^2+z^2)^{3/2}}\,dz\,dy\,dx.$$

EXAMPLE 7.8

Find the center of mass of the solid region Q of uniform density, bounded below by the upper nappe of the cone $z^2 = x^2 + y^2$ and above by the sphere $x^2 + y^2 + z^2 = 9$.

Solution:

Because the density is uniform, we can take $\rho(x, y, z) = k$. By symmetry, the center of mass lies on the z-axis, so it suffices to compute

$$\bar{z} = \frac{M_{xy}}{m},$$

where $m = kV$. We can use the equation of the sphere to find that $\rho = 3$. To find ϕ, we use the fact that the sphere and cone intersect when $(x^2 + y^2) + z^2 = (z^2) + z^2 = 9$. Solving this yields $z = \frac{3}{\sqrt{2}}$. Then, $z = \rho \cos\phi = \frac{3}{\sqrt{2}} \cdot \frac{1}{3} = \cos\phi \Rightarrow \phi = \frac{\pi}{4}$. Now we find V:

$$V = \iiint_Q dV = \int_0^{2\pi} \int_0^{\pi/4} \int_0^3 \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta$$

$$= \int_0^{2\pi} \int_0^{\pi/4} 9 \sin\phi \, d\phi \, d\theta = 9 \int_0^{2\pi} [-\cos\phi]_0^{\pi/4} \, d\theta = 9 \int_0^{2\pi} \left(1 - \frac{\sqrt{2}}{2}\right) d\theta$$

$$= 9 \cdot 2\pi \cdot \left(1 - \frac{\sqrt{2}}{2}\right) = 9\pi(2 - \sqrt{2}).$$

Thus, $m = kV = 9k\pi(2 - \sqrt{2})$. We now compute M_{xy}:

$$M_{xy} = \iiint_Q kz \, dV = k \iiint_Q \rho\cos\phi \cdot \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi$$

$$= k \int_0^{\pi/4} \int_0^{2\pi} \int_0^3 \rho^3 \cos\phi \sin\phi \, d\rho \, d\theta \, d\phi$$

$$= k \int_0^{\pi/4} \int_0^{2\pi} \left[\frac{\rho^4}{4}\right]_0^3 \cos\phi \sin\phi \, d\theta \, d\phi = \frac{81k}{4} \int_0^{\pi/4} \int_0^{2\pi} \cos\phi \sin\phi \, d\theta \, d\phi$$

$$= \frac{81k}{4} \int_0^{\pi/4} 2\pi \cos\phi \sin\phi \, d\phi = \frac{81k\pi}{2} \int_0^{\pi/4} \cos\phi \sin\phi \, d\phi$$

$$= \frac{81k\pi}{2} \cdot \left[\frac{1}{2}\sin^2\phi\right]_0^{\pi/4} = \frac{81k\pi}{2} \cdot \frac{1}{2} \cdot \sin^2\left(\frac{\pi}{4}\right) = \frac{81k\pi}{2} \cdot \frac{1}{2} \cdot \left(\frac{\sqrt{2}}{2}\right)^2$$

$$= \frac{81k\pi}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{81k\pi}{8}$$

Then, $\bar{z} = \frac{M_{xy}}{m} = \frac{81k\pi/8}{9k\pi(2-\sqrt{2})} = \frac{9(2+\sqrt{2})}{16} = 1.92$. The center of mass is $(0, 0, 1.92)$.

EXAMPLE 7.9

Surfaces of the form $\rho = 1 + \frac{1}{5}\sin(m\theta)\sin(n\phi)$ have been used as models for tumors. The "bumpy sphere" with $m = 6$ and $n = 5$ is shown. Set up the integral and then evaluate the integral numerically to find its volume.

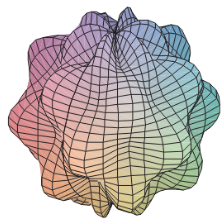


Image credit: Stewart

Solution:

The volume enclosed by a surface defined in spherical coordinates by $\rho = f(\theta, \phi)$ is given by the triple integral

$$V = \iiint_E \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta.$$

Since ρ is a function of θ and ϕ, we treat this as a variable upper bound. So the volume becomes

$$V = \int_0^{2\pi} \int_0^{\pi} \int_0^{1+\frac{1}{5}\sin(6\theta)\sin(5\phi)} \rho^2 \sin\phi \, d\rho \, d\phi \, d\theta.$$

We evaluate the innermost integral:

$$\int_0^{1+\frac{1}{5}\sin(6\theta)\sin(5\phi)} \rho^2 \, d\rho = \left[\frac{1}{3}\rho^3\right]_0^{1+\frac{1}{5}\sin(6\theta)\sin(5\phi)}.$$

So the final integral is

$$V = \int_0^{2\pi} \int_0^{\pi} \frac{1}{3}\left(1 + \frac{1}{5}\sin(6\theta)\sin(5\phi)\right)^3 \sin\phi \, d\phi \, d\theta$$

Here is the result in MATLAB:

```
f = @(phi, theta) (1/3) * (1 + (1/5)*sin(6*theta).*sin(5*phi)).^3 .* sin(phi);
V = integral2(f, 0, pi, 0, 2*pi);
disp(V)
```

```
4.3157
```

EXAMPLE 7.9

Find the volume of the spherical planetarium in l'Hemisfèric in Valencia, Spain, which has a radius of approximately 50 ft, using the equation $x^2 + y^2 + z^2 = r^2$.

Image credit: Visit Valencia

Solution:

We calculate the volume of the ball in the first octant, where $x \le 0$, $y \le 0$, and $z \le 0$, using spherical coordinates, and then multiply the result by 8 for symmetry. The range of the variables is:

$$E = \left\{ (\rho, \theta, \phi) \,\middle|\, 0 \le \rho \le r, \quad 0 \le \theta \le \frac{\pi}{2}, \quad 0 \le \phi \le \frac{\pi}{2} \right\}$$

Therefore,

$$\begin{aligned} V = \iiint_D dx\,dy\,dz &= 8 \int_0^{\pi/2} \int_0^{\pi/2} \int_0^r \rho^2 \sin\phi \, d\rho \, d\theta \, d\phi \\ &= 8 \left(\int_0^{\pi/2} \int_0^{\pi/2} d\theta \, \sin\phi \, d\phi \right) \left(\int_0^r \rho^2 \, d\rho \right) \\ &= 8 \left(\int_0^{\pi/2} \sin\phi \, d\phi \right) \left(\int_0^{\pi/2} d\theta \right) \left(\int_0^r \rho^2 \, d\rho \right) \\ &= 8 \cdot [-\cos\phi]_0^{\pi/2} \cdot [\theta]_0^{\pi/2} \cdot \left[\frac{\rho^3}{3} \right]_0^r = 8 \cdot (1) \cdot \left(\frac{\pi}{2} \right) \cdot \left(\frac{r^3}{3} \right) \\ &= \frac{4\pi r^3}{3} \end{aligned}$$

So for a sphere with a radius of approximately 50 ft, the volume is 523,600 ft^3.

8 Change of Variables in Multiple Integrals

8.1 The Jacobian

In single-variable calculus, one common method you used to evaluate integrals was u-substitution:

$$\int_a^b f(x)\,dx = \int_c^d f(g(u))\,g'(u)\,du$$

where $x = g(u)$, $a = g(c)$, $b = g(d)$.

With multiple integrals, you see change of variables when we convert from Cartesian to polar coordinates:

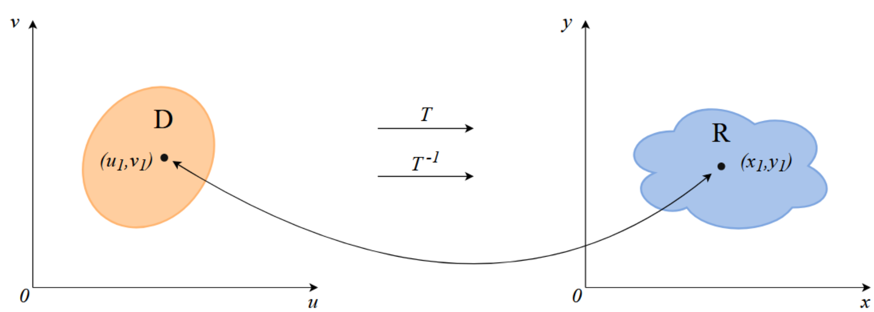
$$\iint_R f(x,y)\,dA = \iint_S f(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta$$

where $r = x\cos\theta$ and $y = r\sin\theta$. In this case, S is a region in the $r\theta$-plane that corresponds to the region R in the xy-plane.

We can generalize any of these processes as a *transformation* from the uv-plane to the xy-plane:

$$T(u,v) = (x,y)$$

where $x = g(u,v)$ and $y = h(u,v)$. This would be known as a C^{-1} transformation which means that g and h have continuous first-order partial derivatives. If $T(u_1,v_1) = (x_1,y_1)$, then we call point (x_1,y_1) the *image* of point (u_1,v_1). If there are no points on the domain that map to the same image, T is called *one-to-one*. Here, T transforms S into a region R which creates the image of S, consisting of the images of all points in S:



This would be written as $T : D \to R$.

If T is one-to-one, then it has an *inverse transformation* T^-1 from the xy-plane to the uv-plane.

EXAMPLE 8.1

A transformation is defined by the equations $x = u^2 - v^2$ and $y = 2uv$ Find the image of the square $S = \{(u, v) \mid 0 \leq u \leq 1,\ 0 \leq v \leq 1\}$.

Solution:

The transformation maps the boundary of S into the boundary of the image. We determine the image by examining each side of the square:

For the first side, S_1, we will let $v = 0$, with $0 \leq u \leq 1$. And because we have $x = u^2$ and $y = 0$, we have $0 \leq x \leq 1$. S_1 maps to the horizontal segment from $(0, 0)$ to $(1, 0)$ in the xy-plane.

For side S_2, we will Let $u = 1$, with $0 \leq v \leq 1$. Eliminating v from $x = 1 - v^2$ and $y = 2v$ yields $x = 1 - \frac{y^2}{4}$ with $0 \leq x \leq 1$ which is a parabolic arc that opens to the left.

For side S_3, we have $0 \leq u \leq 1$ from $v = 1$ and $x = \frac{y^2}{4} - 1$ with $-1 \leq x \leq 0$ which gives a parabolic arc that opens to the right.

For side S_4, we have $0 \leq v \leq 1$ from $u = 0$. Then we have $x = -v^2$ and $y = 0$ with $-1 \leq x \leq 0$. This maps to the horizontal segment from $(-1, 0)$ to $(0, 0)$.

Therefore, the image of the square S under the transformation is a region R in the xy-plane bounded by the x-axis and the two parabolic arcs:

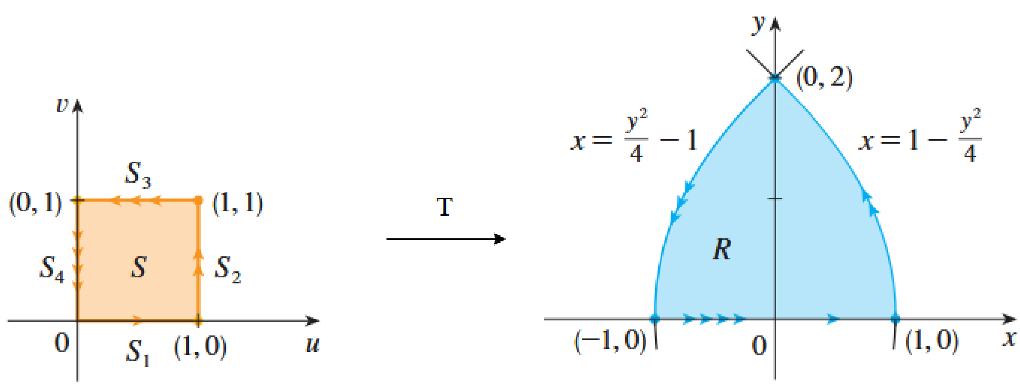


Image credit: Stewart

Disk of radius a	**Region**
Cartesian (disk)	$D = \left\{(x, y) : x^2 + y^2 \leq a^2\right\}$
Polar (rectangle)	$D^* = \{(r, \theta) : 0 \leq r \leq a,\ 0 \leq \theta \leq 2\pi\}$
Annulus between circles of radii a and b	**Region**
Cartesian (annulus)	$D = \left\{(x, y) : a^2 \leq x^2 + y^2 \leq b^2\right\}$
Polar (rectangle)	$D^* = \{(r, \theta) : a \leq r \leq b,\ 0 \leq \theta \leq 2\pi\}$

Examples of converting regions from rectangular to polar coordinates

Suppose a transformation T maps a region S in the uv-plane to a region R in the xy-plane:

$$T(u, v) = (x, y)$$

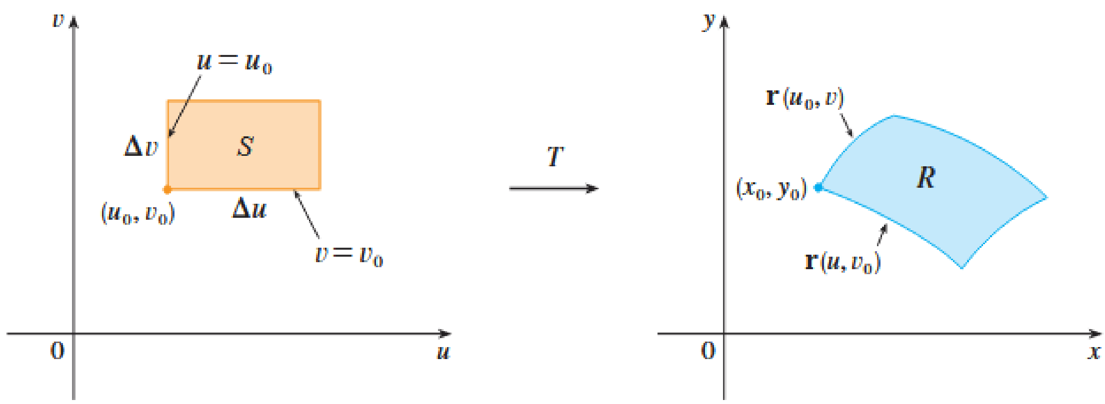


Image credit: Stewart

We can represent the position vector of the image of the point (u, v) as

$$\vec{r}(u, v) = x(u, v)\,\mathbf{i} + y(u, v)\,\mathbf{j}.$$

Now let's explain what happens in the image geometrically. The small rectangle in the uv-plane near a point (u_0, v_0), with width Δu and height Δv maps to a curved region in the xy-plane. For small Δu and Δv, we can approximate the image as a parallelogram.

The tangent vector at (u_0, v_0) is given by this partial derivative:

$$\vec{r}_u = g_u(u_0, v_0)\mathbf{i} + h_u(u_0, v_0)\mathbf{j} = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j}$$

The tangent vector at (x_0, y_0) is given by this partial derivative:

$$\vec{r}_v = g_v(u_0, v_0)\mathbf{i} + h_v(u_0, v_0)\mathbf{j} = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j}$$

Then the two adjacent sides of the image parallelogram are approximately:

$$\vec{a} = \vec{r}(u_0 + \Delta u, v_0) - \vec{r}(u_0, v_0)$$
$$\vec{b} = \vec{r}(u_0, v_0 + \Delta v) - \vec{r}(u_0, v_0)$$

Using the fact that

$$\vec{r}_u = \lim_{\Delta u \to 0} \frac{\vec{r}(u_0 + \Delta u, v_0) - \vec{r}(u_0, v_0)}{\Delta u},$$

we can say that

$$\vec{a} = \vec{r}(u_0 + \Delta u, v_0) - \vec{r}(u_0, v_0) \approx \Delta u\,\vec{r}_u$$

and

$$\vec{b} = \vec{r}(u_0, v_0 + \Delta v) - \vec{r}(u_0, v_0) \approx \Delta v\,\vec{r}_v.$$

The area of a parallelogram spanned by vectors $\vec{a}$ and $\vec{b}$ is given by the magnitude of their cross product. Thus, the area of the image region A is given by:

$$A \approx |\vec{r}_u \times \vec{r}_v| \cdot \Delta u \Delta v$$

In two dimensions, the cross product becomes the absolute value of the **Jacobian** determinant*:

$$\vec{r}_u \times \vec{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} & 0 \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} & 0 \end{vmatrix} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{vmatrix} \mathbf{k} = \left| \begin{array}{cc} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial v} \end{array} \right| \mathbf{k}$$

The **Jacobian** of the transformation T given by $x = g(u,v)$ and $y = h(u,v)$ is

$$\frac{\partial(x,y)}{\partial(u,v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \frac{\partial x}{\partial u} \cdot \frac{\partial y}{\partial v} - \frac{\partial x}{\partial v} \cdot \frac{\partial y}{\partial u}$$

* This is a function that you do not need to worry about for now! You will learn all about it in linear algebra.

This gives us the change of variables formula for area:

$$\Delta A \approx \left| \frac{\partial(x,y)}{\partial(u,v)} \right| \Delta u \, \Delta v$$

where the Jacobian would be evaluated at (u_0, v_0)

EXAMPLE 8.2

Find the Jacobian for the change of variables defined by

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta$$

Solution:

Using the definition of a Jacobian, we compute as follows:

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= r\cos^2\theta + r\sin^2\theta = r$$

$$= r$$

EXAMPLE 8.2

Let $\overline{D} \subseteq \mathbb{R}^2$ and define the transformation

$$\vec{T} : \overline{D} \subseteq \mathbb{R}^2 \longrightarrow \mathbb{R}^2.$$

Compute the Jacobian of the transformation

$$\vec{T}(r,\theta) = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ r\sin\theta \end{bmatrix}.$$

Solution:

The Jacobian determinant is given by:

$$J(r,\theta) = \det \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix}$$

We compute the needed partial derivatives:

$$x = r\cos\theta \Rightarrow \begin{cases} \frac{\partial x}{\partial r} = \cos\theta \\ \frac{\partial x}{\partial \theta} = -r\sin\theta \end{cases}$$

$$y = r\sin\theta \Rightarrow \begin{cases} \frac{\partial y}{\partial r} = \sin\theta \\ \frac{\partial y}{\partial \theta} = r\cos\theta \end{cases}$$

Thus the Jacobian of the polar transformation becomes

$$J(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= \cos\theta \cdot r\cos\theta - (-r\sin\theta)\cdot\sin\theta$$

$$= r\cos^2\theta + r\sin^2\theta$$

$$= r(\cos^2\theta + \sin^2\theta)$$

$$= r.$$

8.2 Change of Variables in Double and Triple Integration

Change of Variables in a Double Integral

Suppose T is a one-to-one C^1 transformation with nonzero Jacobian that maps a region S in the uv-plane to a region R in the xy-plane.

Let $T(u,v) = (x(u,v), y(u,v))$, and suppose $f(x,y)$ is continuous on R.

Then,

$$\iint_R f(x,y)\,dA = \iint_S f(x(u,v),y(u,v)) \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv.$$

The area element dA transforms as

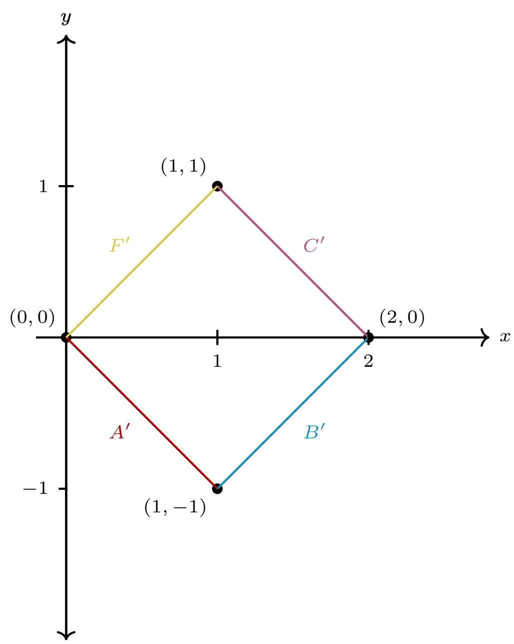
$$dA = \left|\frac{\partial(x,y)}{\partial(u,v)}\right| du\,dv.$$

Let's analyze a problem together. Suppose we are asked to evaluate the double integral

$$\iint_{\overline{D}} \sqrt{\frac{x-y}{x+y+1}}\,dA$$

where $\overline{D}$ is the square with vertices at $(0,0)$, $(1,-1)$, $(2,0)$, and $(1,1)$.

We begin by visualizing the region of integration. The region $\overline{D} \subseteq \mathbb{R}^2$ is a square oriented diagonally. The region $\overline{D}$ has sides of length $\sqrt{2}$ and a total area of 2 units^2:



This integral will be difficult to evaluate directly in x and y due to the complicated form of the integrand $f(x,y) = \sqrt{\frac{x-y}{x+y+1}}$. Moreover, the domain $\overline{D}$ is not aligned with the coordinate axes, which complicates the limits of integration.

To proceed, we partition the region $\overline{D}$ into two parts:

$$\overline{D} = \overline{D}_1 \cup \overline{D}_2$$

where

$$\overline{D}_1 = \{(x,y) : 0 \leq x \leq 1,\ -x \leq y \leq x\}$$

and

$$\overline{D}_2 = \{(x,y) : 1 \leq x \leq 2,\ -2+x \leq y \leq 2-x\}.$$

Given both the complexity of the integrand and the domain, we consider the following questions:

1. Can we map $\overline{D}$ onto a new region D that is easier to describe and integrate over?

2. Can we choose new variables to simplify the integrand?

Let's introduce the following change of variables:

$$u = u(x, y) = x - y, \ v = v(x, y) = x + y$$

This change of variables is a linear transformation that can also be written in matrix form:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

This transformation represents a rotation and dilation of the coordinate axes.

Let us now examine how the boundary of $\overline{D}$ transforms under this change of variables:

Boundary of $\overline{D}$ in the xy-plane	**Transformation**	**Boundary of $\overline{D}$ in the uv-plane**
A': $0 \leq x \leq 1, \ y = -x$	$A' \to A$	A: $0 \leq u \leq 2, \ v = 0$
B': $1 \leq x \leq 2, \ y = x - 2$	$B' \to B$	B: $u = 2, \ 0 \leq v \leq 2$
C': $1 \leq x \leq 2, \ y = 2 - x$	$C' \to C$	C: $0 \leq u \leq 2, \ v = 2$
F': $0 \leq x \leq 1, \ y = x$	$F' \to F$	F: $u = 0, \ 0 \leq v \leq 2$

The boundaries of $\overline{D}$ in the uv-plane are images. We will now compute each transformation.

We now compute how each side of the original region $\overline{D}$ transforms using the change of variable equations

From A' to A: We are given the edge A' defined by $y = -x$ and $0 \leq x \leq 1$. Substituting into the change of variables:

$$u = x - (-x) = 2x, \ v = x + (-x) = 0.$$

Solving for x in terms of u, we get

$$x = \frac{u}{2}.$$

Since $0 \leq x \leq 1$, we have $0 \leq u \leq 2$, and thus this edge maps to

$$0 \leq u \leq 2, \ v = 0.$$

From B' to B: The edge B' is given by $y = x - 2$ and $1 \leq x \leq 2$. Substituting yields $u = x - (x - 2) = 2$ so $u = 2$.

Also, $v = x + y = x + (x - 2) = 2x - 2$, so

$$x = \frac{v+2}{2}.$$

Since $1 \leq x \leq 2$, we find

$$1 \leq \frac{v+2}{2} \leq 2 \Rightarrow 0 \leq v \leq 2.$$

So this edge maps to

$$0 \leq v \leq 2, \ u = 2.$$

From C' to C: Here C' is the line segment $y = 2 - x$ with $1 \leq x \leq 2$. Then

$$v = x + (2 - x) = 2 \Rightarrow v = 2.$$

Also $u = x - y = x - (2 - x) = 2x - 2 \Rightarrow x = \frac{u+2}{2}$.

From $1 \leq x \leq 2$, we have

$$1 \leq \frac{u+2}{2} \leq 2 \quad \Rightarrow \quad 0 \leq u \leq 2.$$

So this maps to

$$0 \leq u \leq 2, \ v = 2.$$

From F' to F: The segment F' is the line $y = x$, $0 \leq x \leq 1$. Then

$$u = x - x = 0, \ v = x + x = 2x \Rightarrow x = \frac{v}{2}.$$

From $0 \leq x \leq 1$, we obtain

$$0 \leq \frac{v}{2} \leq 1 \Rightarrow 0 \leq v \leq 2.$$

Thus this side maps to

$$0 \leq v \leq 2, \ u = 0.$$

To verify that our transformation correctly maps the region $\overline{D} \subseteq \mathbb{R}^2$ to the new rectangular region $D \subseteq \mathbb{R}^2$, let's check our results by verifying the vertices match.

Vertex 1: $(x, y) = (0, 0)$

$$\begin{cases} u = 0 - 0 = 0 \\ v = 0 + 0 = 0 \end{cases} \Rightarrow (u, v) = (0, 0)$$

Vertex 2: $(x, y) = (1, -1)$

$$\begin{cases} u = 1 - (-1) = 2 \\ v = 1 + (-1) = 0 \end{cases} \Rightarrow (u, v) = (2, 0)$$

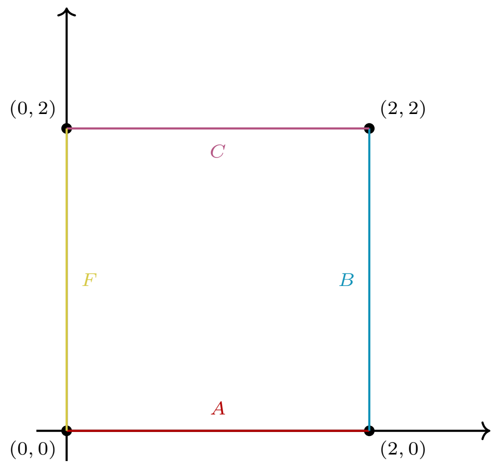
Vertex 3: $(x, y) = (2, 0)$

$$\begin{cases} u = 2 - 0 = 2 \\ v = 2 + 0 = 2 \end{cases} \Rightarrow (u, v) = (2, 2)$$

Vertex 4: $(x, y) = (1, 1)$

$$\begin{cases} u = 1 - 1 = 0 \\ v = 1 + 1 = 2 \end{cases} \Rightarrow (u, v) = (0, 2)$$

So the image of the region $\overline{D}$ under this transformation is a rectangle with corners at $(0, 0)$, $(2, 0)$, $(2, 2)$ and $(0, 2)$. Let's visualize our new transformed region:



Let's now compute. We were given the double integral

$$\iint_{\overline{D}} \sqrt{\frac{x-y}{x+y+1}}\, dw$$

where $\overline{D}$ is the square with vertices $(0,0),(1,-1),(2,0),(1,1)$. To evaluate this, we perform a change of variables to simplify the integrand.

We define the transformation:

$$x = x(u,v) \Rightarrow x = \frac{u+v}{2} \quad \text{and} \quad y = y(u,v) \Rightarrow y = \frac{v-u}{2}$$

Now compute the Jacobian:

$$J(u,v) = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{vmatrix} = \left(\frac{1}{2}\cdot\frac{1}{2}\right) - \left(-\frac{1}{2}\cdot\frac{1}{2}\right) = \frac{1}{4} - \left(-\frac{1}{4}\right) = \frac{1}{2}$$

Thus,

$$\iint_{\overline{D}} f(x,y)\, dw = \iint_D f(u,v)\cdot |J(u,v)|\; du\, dv.$$

Substitute the integrand $f(x,y) = \sqrt{\frac{x-y}{x+y+1}}$:

$$\frac{x-y}{x+y+1} = \frac{u}{v+1} \Rightarrow f(u,v) = \sqrt{\frac{u}{v+1}}$$

Therefore,

$$\iint_{\overline{D}} \sqrt{\frac{x-y}{x+y+1}}\, dw = \iint_D \sqrt{\frac{u}{v+1}} \cdot \left|\frac{1}{2}\right| du\, dv = \frac{1}{2} \iint_D \sqrt{\frac{u}{v+1}}\, du\, dv.$$

The transformed region D is the square with vertices $(0,0), (2,0), (2,2), (0,2)$, so we integrate

$$= \frac{1}{2} \int_0^2 \int_0^2 \sqrt{\frac{u}{v+1}}\, du\, dv.$$

We compute the inner integral first:

$$\int_0^2 \sqrt{\frac{u}{v+1}}\, du = \frac{1}{\sqrt{v+1}} \int_0^2 \sqrt{u}\, du = \frac{1}{\sqrt{v+1}} \cdot \left[\frac{2}{3} u^{3/2}\right]_0^2 = \frac{1}{\sqrt{v+1}} \cdot \frac{2}{3} \cdot (2)^{3/2} = \frac{1}{\sqrt{v+1}} \cdot \frac{2}{3} \cdot 2\sqrt{2} = \frac{4\sqrt{2}}{3\sqrt{v+1}}$$

Now the full integral becomes

$$\frac{1}{2} \int_0^2 \frac{4\sqrt{2}}{3\sqrt{v+1}}\, dv = \frac{2\sqrt{2}}{3} \int_0^2 (v+1)^{-1/2}\, dv = \frac{2\sqrt{2}}{3} \cdot \left[2(v+1)^{1/2}\right]_0^2 = \frac{2\sqrt{2}}{3} \cdot 2\left(\sqrt{3}-1\right) = \frac{4\sqrt{2}}{3}(\sqrt{3}-1)$$

$$\iint_{\overline{D}} \sqrt{\frac{x-y}{x+y+1}}\, dw = \frac{4\sqrt{2}}{3}(\sqrt{3}-1).$$

EXAMPLE 8.3

Find the area of a circle of radius R.

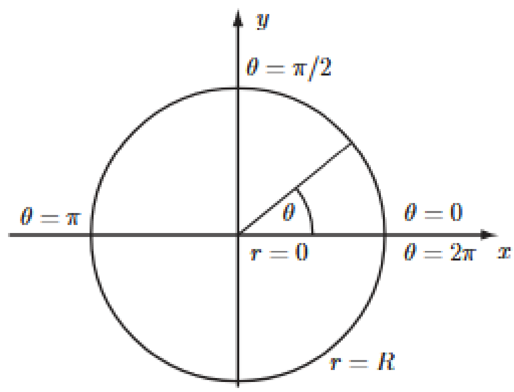


Image credit: Loughborough University

Solution:

Let C be the region bounded by a circle of radius R centered at the origin. Then the area A of this region is $A = \iint_C dC$. We change to polar coordinates using $x = r\cos\theta$ and $y = r\sin\theta$. We begin by computing the necessary partial derivatives:

$$\frac{\partial x}{\partial r} = \cos\theta, \quad \frac{\partial y}{\partial r} = \sin\theta, \quad \frac{\partial x}{\partial \theta} = -r\sin\theta, \quad \frac{\partial y}{\partial \theta} = r\cos\theta$$

Thus, the Jacobian determinant is

$$J(r,\theta) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix}$$

$$= \cos\theta \cdot r\cos\theta - (-r\sin\theta)\cdot\sin\theta$$

$$= r\cos^2\theta + r\sin^2\theta$$

$$= r.$$

Therefore, the area becomes

$$A = \iint_C dC = \int_0^{2\pi}\int_0^R r\,dr\,d\theta = \int_0^{2\pi}\left[\frac{1}{2}R^2\right]d\theta = \frac{1}{2}R^2\cdot 2\pi = \pi R^2.$$

By now, your intuition may prompt you to hypothesize that the change of variables for triple integrals is similar. And you would be correct. Let T be a transformation that maps a region S in uvw-space onto a region $R \subseteq \mathbb{R}^3$ use the equations

$$x = g(u, v, w),\ y = h(u, v, w),\ z = k(u, v, w).$$

Then the Jacobian of the transformation is

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} & \frac{\partial y}{\partial w} \\ \frac{\partial z}{\partial u} & \frac{\partial z}{\partial v} & \frac{\partial z}{\partial w} \end{vmatrix}.$$

Then the integral transforms as

$$\iiint_R f(x, y, z)\, dV = \iiint_S f\big(x(u, v, w),\ y(u, v, w),\ z(u, v, w)\big) \cdot \left| \frac{\partial(x, y, z)}{\partial(u, v, w)} \right| du\, dv\, dw$$

EXAMPLE 8.4

Use change of variables to derive the formula for volume in spherical coordinates.

Solution:

We will use these change of variable equations:

$$x = \rho \sin\phi \cos\theta,\ \ y = \rho \sin\phi \sin\theta,\ \ z = \rho \cos\phi$$

We first compute the Jacobian determinant:

$$\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)} = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial \rho} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \\ \cos\phi & 0 & -\rho\sin\phi \end{vmatrix}$$

We expand along the third row:

$$\det = \cos\phi \cdot \begin{vmatrix} -\rho\sin\phi\sin\theta & \rho\cos\phi\cos\theta \\ \rho\sin\phi\cos\theta & \rho\cos\phi\sin\theta \end{vmatrix} - 0 + (-\rho\sin\phi)\cdot \begin{vmatrix} \sin\phi\cos\theta & -\rho\sin\phi\sin\theta \\ \sin\phi\sin\theta & \rho\sin\phi\cos\theta \end{vmatrix}$$

Compute each 2×2 determinant:

$$\begin{aligned} &= \cos\phi\left(-\rho^2\sin\phi\cos\phi(\sin^2\theta + \cos^2\theta)\right) - \rho\sin\phi\left(\rho\sin^2\phi(\cos^2\theta + \sin^2\theta)\right) \\ &= -\rho^2\sin\phi\cos^2\phi - \rho^2\sin^3\phi = -\rho^2\sin\phi(\cos^2\phi + \sin^2\phi) = -\rho^2\sin\phi \end{aligned}$$

Since $0 \le \phi \le \pi$, we have $\sin\phi \ge 0$, so $\left|\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)}\right| = |-\rho^2\sin\phi| = \rho^2\sin\phi$. Finally, we have the formula that we used before:

$$\iiint_R f(x,y,z)\,dV = \iiint_S f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\cdot\rho^2\sin\phi\,d\rho\,d\theta\,d\phi$$

where $dV = \rho^2\sin\phi\,d\rho\,d\theta\,d\phi$.

EXAMPLE 8.5

Use spherical coordinates to find the volume of a sphere of radius R.

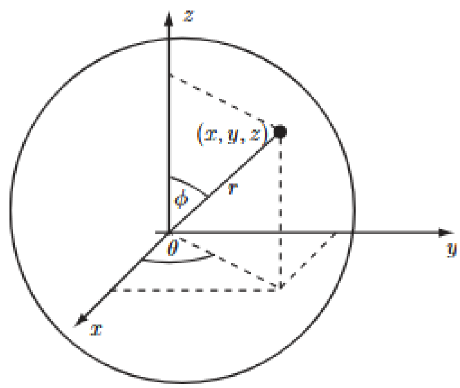


Image credit: Loughborough University

Solution:

Instead of using Cartesian coordinates, we switch to spherical coordinates, which are better suited to spheres. We have the change of variable equations $x = r\cos\theta\sin\phi$, $y = r\sin\theta\sin\phi$, and $z = r\cos\phi$.
We first compute the determinant of the Jacobian matrix formed by all partial derivatives:

$$J(r,\theta,\phi) = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \phi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \phi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \phi} \end{vmatrix} = \begin{vmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{vmatrix}$$

Because of the fact that $J(r,\theta,\phi)$ is negative for $0 \le \phi \le \pi$. Thus determinant evaluates to $J(r,\theta,\phi) = -r^2\sin\phi \Rightarrow |J(r,\theta,\phi)| = r^2\sin\phi$.

We now express the volume as a triple integral in spherical coordinates. The volume element becomes:

$$dV = |J|\ dr\, d\theta\, d\phi = r^2\sin\phi\, dr\, d\theta\, d\phi$$

So the volume of the sphere is

$$V = \iiint_E 1\, dV = \int_0^{\pi}\int_0^{2\pi}\int_0^{R} r^2\sin\phi\, dr\, d\theta\, d\phi.$$

EXAMPLE 8.5 (CONTINUED)

First, we evaluate the inner integral with respect to r:

$$\int_0^R r^2\,dr = \left[\frac{1}{3}r^3\right]_0^R = \frac{1}{3}R^3$$

And now we evaluate the middle integral with respect to θ:

$$\int_0^{2\pi} d\theta = 2\pi$$

Lastly, we evaluate the outer integral with respect to ϕ):

$$\int_0^{\pi} \sin\phi\,d\phi = [-\cos\phi]_0^{\pi} = -(-1) - (1) = 2$$

Finally, we get the volume:

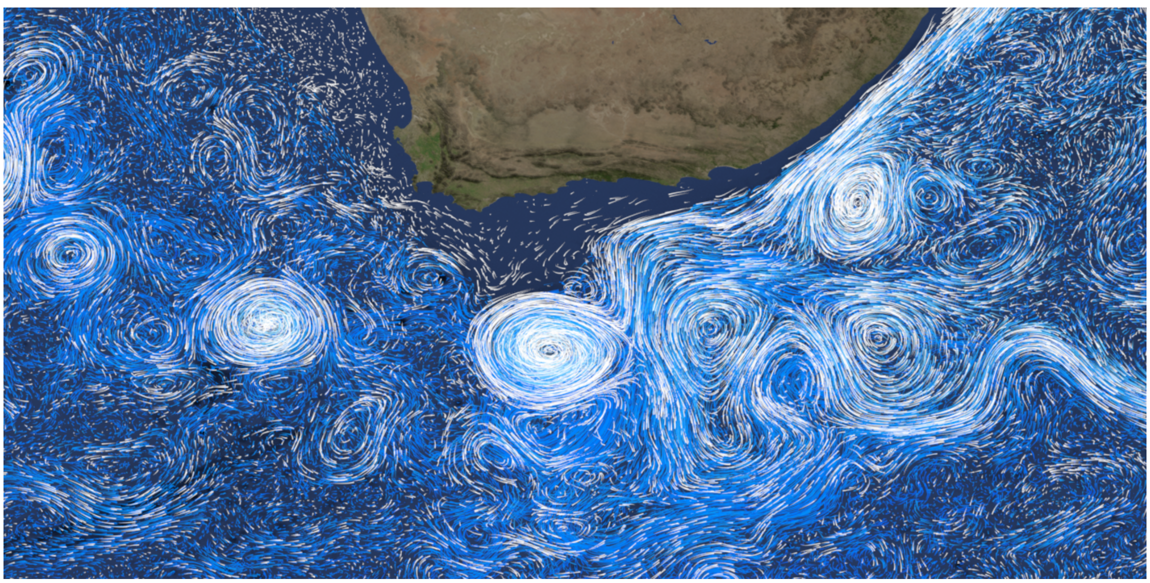
$$V = \left(\frac{1}{3}R^3\right)(2\pi)(2) = \frac{4}{3}\pi R^3$$

Part III

Vector Calculus

With knowledge of vectors, multivariable differentiation, and multiple integrals, we will now move on to vector calculus. In **Part III**, we will learn all about vector fields and the properties associated with them. We will learn about new, powerful integrals and work with them in higher dimensions. By the end of this part, you will have learned about

- The calculus of vector fields
- Line integrals
- Surface integrals
- Green's theorem, Stoke's theorem, and Divergence theorem



NASA and MIT have worked together on a project called Estimating the Circulation and Climate of the Ocean (ECCO). From millions of measurements of temperature, salinity, sea ice concentration, pressure, water height, and more, they have modeled the planet in gorgeous detail. And not only is their work stunning, but it has also enabled thousands of scientific discoveries. Data like these require both magnitude and direction, and vector fields let us represent them visually. Image credit: NASA

9 Foundations of Vector Analysis

At the heart of vector calculus lies vector fields. In this chapter, we will learn what they are and the different ways to analyze them.

9.1 Vector Fields

A vector field is a function that assigns a vector to every point in space on its domain. Here is the formal definition:

> Let $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ and $g : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ be multivariable, real-valued functions defined on a region $D \subseteq \mathbb{R}^2$. We define a vector field as a function
>
> $$\vec{F} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$$
>
> that assigns to each point $(x, y) \in D$ a two-dimensional vector $\vec{F}(x, y)$.

We can express a vector field $\vec{F}$ as

$$\vec{F}(x, y) = \langle f(x, y),\ g(x, y)\rangle = f(x, y)\,\mathbf{i} + g(x, y)\,\mathbf{j}.$$

A vector field $\vec{F} = \langle f, g\rangle$ is *continuous* on a region $D \subseteq \mathbb{R}^2$ if both component functions f and g are continuous on D.

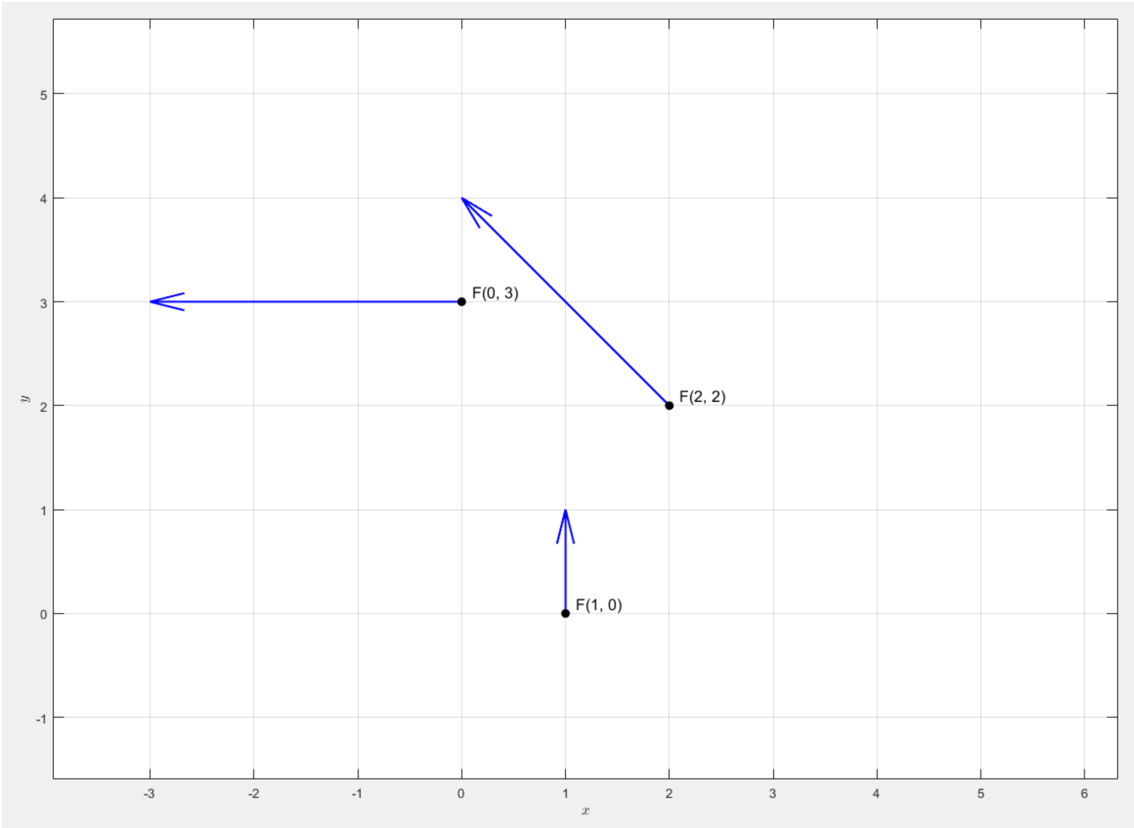
A vector field $\vec{F} = \langle f, g\rangle$ is *differentiable* on $D \subseteq \mathbb{R}^2$ if both f and g are differentiable on D.

Vector fields technically exist in four-dimensional spaces because there are two dimensions for the input and two dimensions for the output. We can't really draw in four dimensions by hand, so we draw them in two dimensions. For a selected input point $P(x, y)$, we plot the output vector $\vec{F}(x, y)$ with a tail at $P(x, y)$ and repeat for other points until the function is sufficiently represented.

Consider the vector field $\vec{F}(x, y) = -y\,\mathbf{i} + x\,\mathbf{j}$. This field assigns to each point a vector perpendicular to the position vector $\langle x, y\rangle$, causing the field to rotate counterclockwise around the origin.

To sketch the field, choose a grid of sample points and evaluate $\vec{F}(x, y)$ at each. For example:

(x, y)	$\vec{F}(x, y)$
$(0, 3)$	$-3\,\mathbf{i} = \langle -3,\ 0\rangle$
$(1, 0)$	$\mathbf{j} = \langle 0,\ 1\rangle$
$(2, 2)$	$-2\,\mathbf{i} + 2\,\mathbf{j} = \langle -2,\ 2\rangle$



These vectors all lie on the unit circle and have the same magnitude:

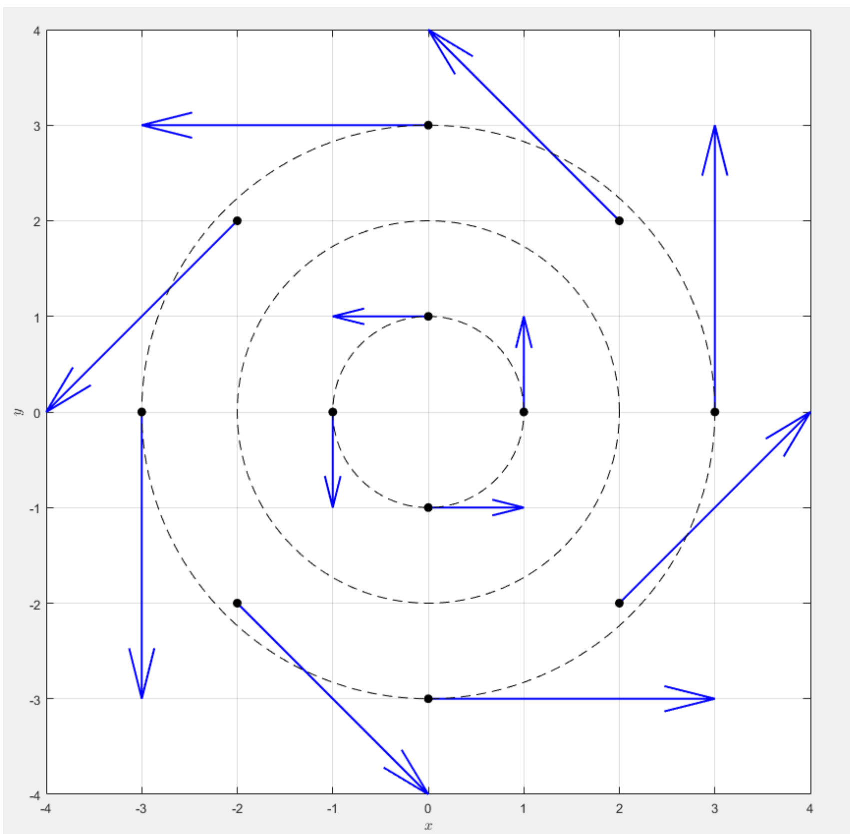
$$\|\vec{F}(x, y)\| = \sqrt{(-y)^2 + x^2} = \sqrt{x^2 + y^2}$$

Thus, vectors of constant magnitude trace out level curves in the plane.

Let's finish this plot by finding more points:

(x, y)	$\vec{F}(x, y)$	(x, y)	$\vec{F}(x, y)$
$(0, 3)$	$-3\mathbf{i} = \langle -3, 0 \rangle$	$(-3, 0)$	$-\mathbf{j} = \langle 0, -3 \rangle$
$(1, 0)$	$\mathbf{j} = \langle 0, 1 \rangle$	$(-2, -2)$	$2\mathbf{i} - 2\mathbf{j} = \langle 2, -2 \rangle$
$(2, 2)$	$-2\mathbf{i} + 2\mathbf{j} = \langle -2, 2 \rangle$	$(-1, 0)$	$-\mathbf{i} = \langle -1, 0 \rangle$
$(3, 0)$	$\mathbf{j} = \langle 0, 3 \rangle$	$(0, -3)$	$3\mathbf{i} = \langle 3, 0 \rangle$
$(0, -1)$	$\mathbf{i} = \langle 1, 0 \rangle$	$(2, -2)$	$2\mathbf{i} + 2\mathbf{j} = \langle 2, 2 \rangle$
$(-2, 2)$	$-2\mathbf{i} - 2\mathbf{j} = \langle -2, -2 \rangle$	$(2, 2)$	$-2\mathbf{i} + 2\mathbf{j} = \langle -2, 2 \rangle$

And then we finish the plot:

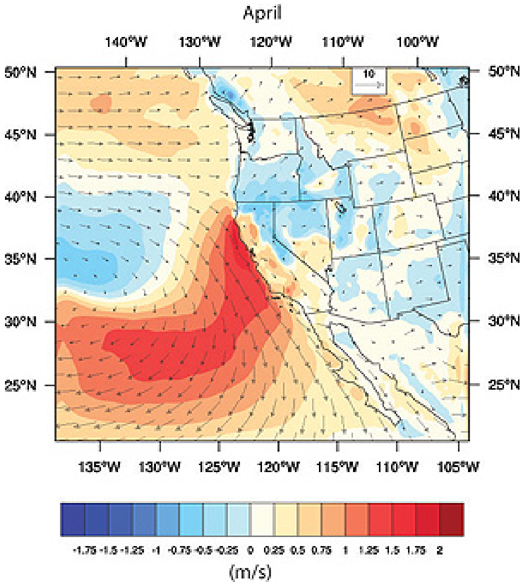


Please run the MATLAB code yourself and have a look!

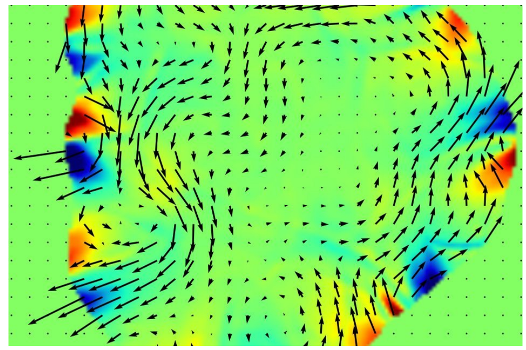
```
x = [1, 2, 3, 0, -2, 0, -1, -2, -3, 0, 2, 0];
y = [0, 2, 0, 1, 2, 3, 0, -2, 0, -1, -2, -3];

u = -y;
v = x;

quiver(x, y, u, v, 0, 'LineWidth', 1.5, 'Color', 'b')
hold on
plot(x, y, 'ko', 'MarkerFaceColor', 'k')

radii = [1, 2, 3];
theta = linspace(0, 2*pi, 200);
for r = radii
    xc = r * cos(theta);
    yc = r * sin(theta);
    plot(xc, yc, 'k--', 'LineWidth', 0.75)
end

axis equal
xlim([-4 4])
ylim([-4 4])
grid on
xlabel('$x$', 'Interpreter', 'latex')
ylabel('$y$', 'Interpreter', 'latex')
```

vectorFieldPlot_minusYplusX.m

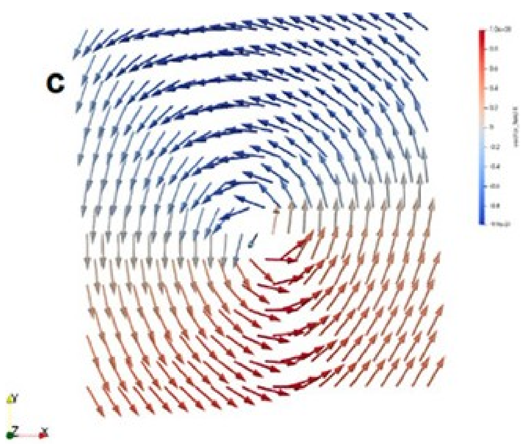
For your interest, here is a gallery of vector fields:



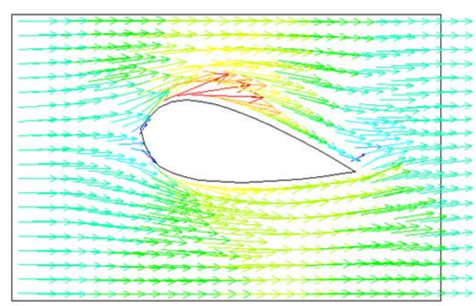
Coastal wind speeds off the coast of California from 2008. Image credit: UCSC



Hydrodynamic simulation of vorticity in quark-gluon plasma. The swirling arrows show rotational fluid motion arising from high-energy ion collisions. Image credit: Berkeley Lab



Magnetic spin textures in a synthetic material to visualize particle-like textures. The arrows show local magnetic spin orientations. Image credit: MagLab



Air flow around an airfoil in ANSYS. Image credit: CMU

Let $\vec{r}(x, y) = \langle x, y\rangle$ where $\vec{r} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$. Let $f(x, y)$ be a two-variable, real-valued function where $f : D \subseteq \mathbb{R}^2 \to \mathbb{R}$.

A vector field in the form

$$\vec{F}(x,y) = f(x,y) \cdot \vec{r} = f(x,y) \cdot \langle x, y \rangle$$

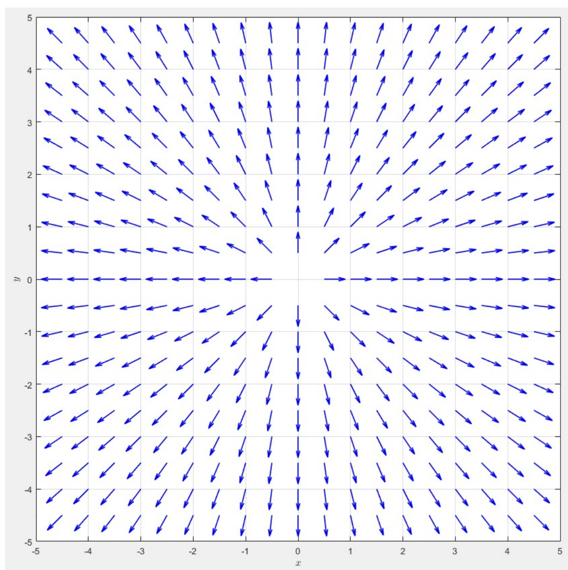
is called a **radial vector field**.

Of special interest are radial vector fields of the form

$$\vec{F}(x,y) = \frac{\vec{r}}{\|\vec{r}\|_2^p} = \frac{\langle x, y \rangle}{\|\vec{r}\|_2^p},$$

where $p \in \mathbb{R}$. At every point $(x,y) \in \mathbb{R}^2$, the vectors are pointed directly outward from the origin with

$$\|\vec{F}\| = \frac{1}{\|\vec{r}\|^{p-1}}.$$

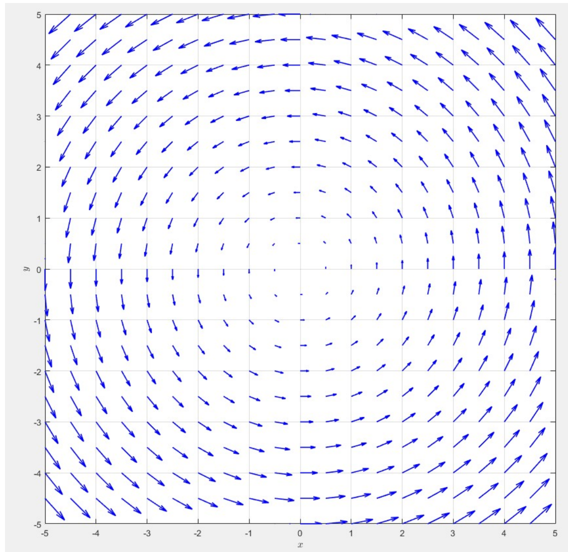


Radial vector field $\vec{F}(x,y) = \dfrac{x\,\mathbf{i} + y\,\mathbf{j}}{\sqrt{x^2+y^2}}$

On the other hand, A **rotational vector field** in $\mathbb{R}^2$ is a vector field of the form

$$\vec{F}(x,y) = \langle -y, x \rangle = -y\,\mathbf{i} + x\,\mathbf{j}.$$

This field assigns to each point a vector tangent to the circle centered at the origin passing through (x, y), inducing a counterclockwise rotation around the origin. Thus, the vector at a point (x, y) is tangent rather than perpendicular to a circle with radius $r = \sqrt{x^2 + y^2}$:



Rotational vector field

A **shear vector field** represents a linear distortion of space in one direction, but the amount of movement depends on "how far you are" in the perpendicular direction. For example, in a horizontal shear, the vectors point left and right, but their length changes depending on how far up or down you are.

Horizontal shear:

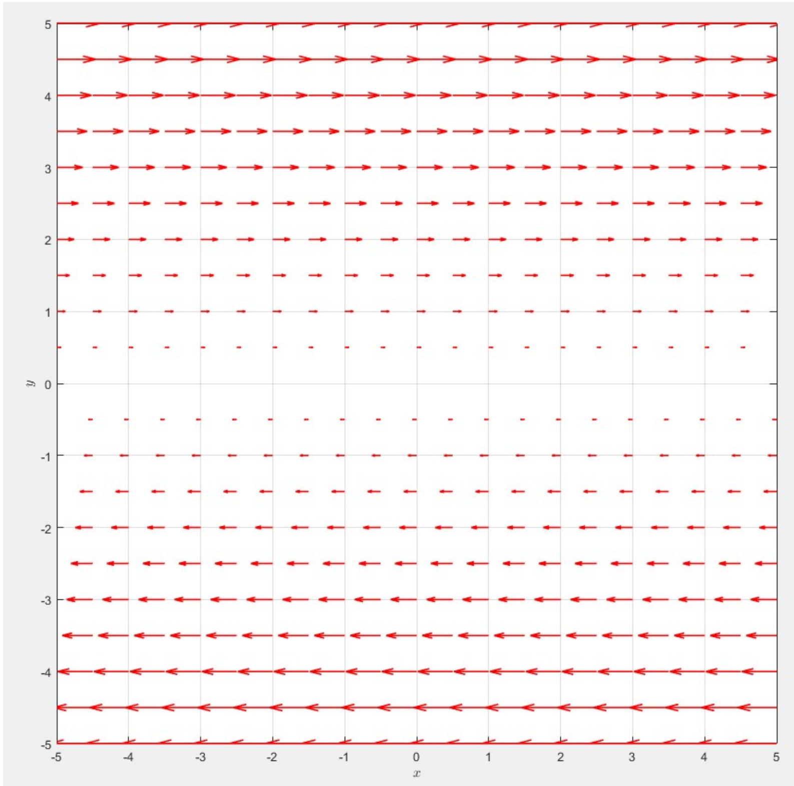
$$\vec{F}(x, y) = \langle y, 0 \rangle = y\,\mathbf{i}$$

Here, vectors point purely in the x-direction, with magnitude depending on the y-position. Horizontal layers of space are "slid" sideways.

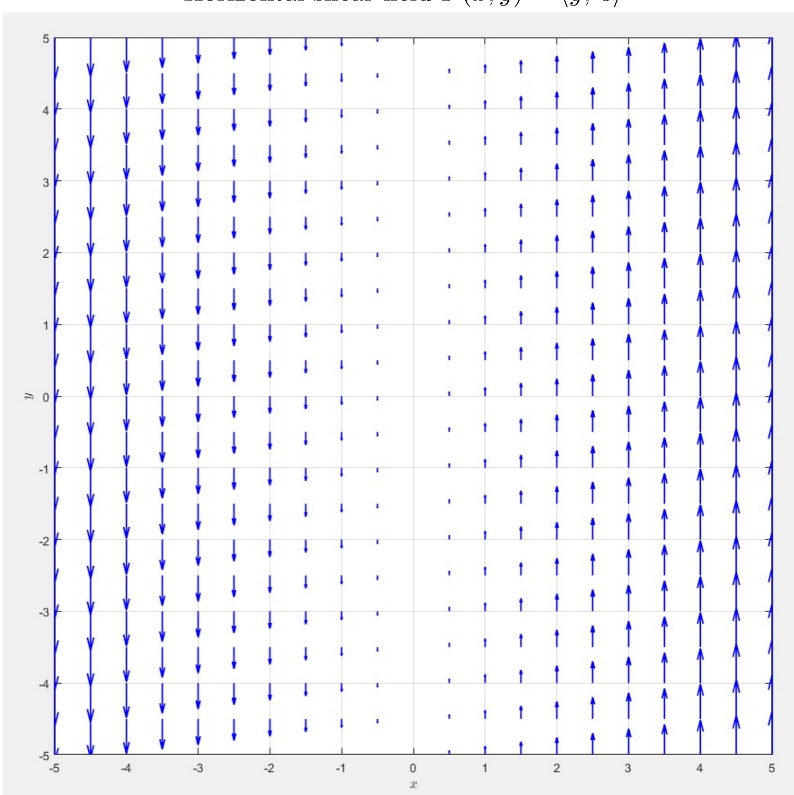
Vertical shear:

$$\vec{F}(x, y) = \langle 0, x \rangle = x\,\mathbf{j}$$

Vectors point purely in the y-direction, with magnitude depending on the x-position. Vertical layers are pushed upward or downward.



Horizontal shear field $\vec{F}(x, y) = \langle y,\, 0\rangle$



Vertical shear field $\vec{F}(x, y) = \langle 0,\, x\rangle$

Radial vector fields are great for modeling physical phenomena that radiate outward from a central point. For instance, you may use them to represent gravitational fields of planets. Rotational vector fields are great for physical phenomena that "swirl" such as fluids in a vortex or magnetic fields around circular wire loops. Shear vector fields are generally more niche, but can be used to capture a sort of "stretching." This includes solid mechanics and materials analysis.

A *velocity field* is a vector field that assigns a velocity vector to each point in space, typically used to describe the flow of a fluid. For example, imagine fluid moving steadily through a pipe: each point (x, y, z) inside the pipe has a corresponding velocity vector $\vec{V}(x, y, z)$ indicating the direction and speed of flow at that location. Velocity fields can also describe rotational motion, like the swirling of water around a drain or the counterclockwise rotation of a wheel.

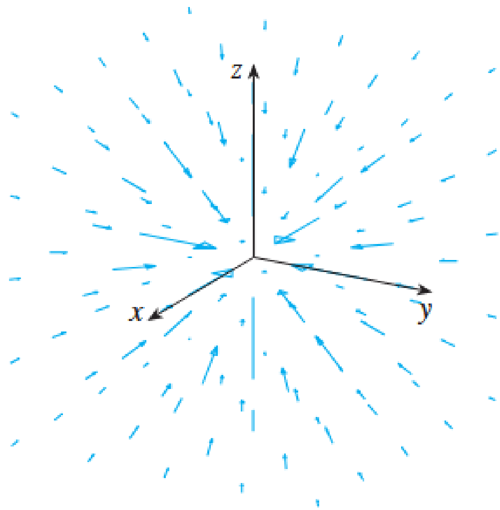
A *gravitational field* is another important example of a vector field. According to Newton's law of gravitation, an object of mass m located at position vector $\vec{x} = \langle x, y, z \rangle \in \mathbb{R}^3$ experiences a force due to a second object of mass M located at the origin. The force is given by the formula:

$$\vec{F}(\vec{x}) = -\frac{mMG}{|\vec{x}|^3}\vec{x} = -\frac{mMG}{r^3}\vec{r}$$

This field always points inward, toward the origin, because gravity acts as an attractive force. The farther away an object is, the weaker the force becomes. More specifically, the magnitude decays like $1/|\vec{x}|^2$. The vector field structure above is called an *inverse-square radial field*, since the force vector points along the radial direction and its strength decreases with the square of the distance.

We can write the gravitational field in terms of its component functions by using the fact that $\vec{x} = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}$ and $\|\vec{x}\| = \sqrt{x^2 + y^2 + z^2}$:

$$\vec{F}(x, y, z) = \frac{-mMGx}{(x^2 + y^2 + z^2)^{3/2}}\,\mathbf{i} + \frac{-mMGy}{(x^2 + y^2 + z^2)^{3/2}}\,\mathbf{j} + \frac{-mMGz}{(x^2 + y^2 + z^2)^{3/2}}\,\mathbf{k}.$$



Gravitational force field. Image credit: Stewart

Electric fields have a very similar structure. According to Coulomb's law, a point charge Q located at the origin exerts a force on another charge q at position $\vec{x}$ given by

$$\vec{F}(\vec{x}) = \frac{\varepsilon q Q}{|\vec{x}|^3}\vec{x}$$

where the constant, permittivity ε, depends on the medium. Like gravity, this force weakens with distance squared, but the direction can vary: if q and Q have the same sign, the force is repulsive (pointing away from the origin); if opposite, it's attractive. To simplify calculations, physicists often work with the electric field $\vec{E}$, which is the force per unit charge:

$$\vec{E}(\vec{x}) = \frac{\varepsilon Q}{|\vec{x}|^3}\vec{x}$$

This makes the electric field another example of a radial field, but one that can point outward or inward depending on the sign of Q.

Let $\phi : D \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a two-variable, real-valued function. Suppose we visualize the output of this function as a surface:

$$z = \phi(x, y)$$

We can visualize the behavior of this surface by graphing various level curves. Let a level curve be defined by

$$L_c(\phi) = \{(x, y) \in D : \phi(x, y) = c \in \mathbb{R}\}.$$

At the point (a, b) on a specific level curve, the gradient

$$\nabla\phi(a, b) = \langle \phi_x(a, b), \phi_y(a, b) \rangle$$

is orthogonal to the tangent line of the level curve at the point (a, b).

With this geometry in mind, one way to generate vector fields is to let

$$\vec{F}(x, y) = \nabla\phi(x, y) = \langle \phi_x(x, y), \phi_y(x, y) \rangle = \langle f(x, y), g(x, y) \rangle.$$

Such a vector field $\vec{F} = \nabla\phi$ is called a **gradient field**, since the field arises from taking the gradient of some scalar function.

The scalar function $\phi = \phi(x, y)$ is called a **potential function**.

Gradient fields are useful in many applications because many important physical quantities form gradients. One example is when molecules move from regions of high concentration to low concentration in cellular transport. Another example is in temperature. There is a law of physics which states that heat diffuses in the direction of the vector field

$$-\vec{F} = -\nabla\phi(x, y),$$

which points in the direction in which temperature decreases most rapidly. This idea governs how heat sinks work.

And now for the formal definition of the gradient:

If f is a scalar function of two variables, recall that its gradient is defined by

$$\nabla f(x, y) = f_x(x, y)\,\mathbf{i} + f_y(x, y)\,\mathbf{j}.$$

Therefore, ∇f is really a vector field on $\mathbb{R}^2$ and is called a **gradient vector field**. Likewise, if f is a scalar function of three variables, its gradient is a vector field on $\mathbb{R}^3$ given by

$$\nabla f(x, y, z) = f_x(x, y, z)\,\mathbf{i} + f_y(x, y, z)\,\mathbf{j} + f_z(x, y, z)\,\mathbf{k}.$$

9.2 Line Integrals

We begin by introducing line integrals of scalar-valued functions

$$f : D \subseteq \mathbb{R}^n \to \mathbb{R}$$

over a smooth, parameterized, oriented curve C.

Suppose $z = f(x, y)$ is a real-valued function of two variables, with

$$f : D \subseteq \mathbb{R}^2 \to \mathbb{R}.$$

Let $C \subseteq D$ be a parameterized curve contained in the domain D, where

$$C = \{\vec{r}(s) : a \le s \le b\} = \{\langle x(s), y(s)\rangle : a \le s \le b\},$$

and s represents the arc length parameter along the curve.

Consider the surface defined by the function values along the curve:

$$S_C = \{z = f(x, y) : (x, y) \in C\} = \{z = f(x(s), y(s)) : a \le s \le b\}.$$

Then, the area between the curve S_C on the surface and the embedding of C in the xy-plane is given symbolically by the line integral

$$\int_{S_C} f\,d\omega = \int_C f(x(s), y(s))\,ds = \lim_{\Delta\to 0} \sum_{k=1}^{n} f(x(s_k^*), y(s_k^*))\,\Delta s_k.$$

Line integrals generalize the concept of integration to curves in space. Instead of summing values over an interval, we sum values along a curve.

For a scalar field, we write

$$\int_C f(x, y)\, ds.$$

This represents the total weighted length along the curve, such as mass of a wire with density f.

For a vector field, we say

$$\int_C \vec{F} \cdot d\vec{r}.$$

This measures the total amount of the vector field that aligns with the direction of motion, like total work done by a force field along a path.

To define a line integral $\int_C \vec{F} \cdot d\vec{s}$ of a vector field $\vec{F} : D \subseteq \mathbb{R}^2 \to \mathbb{R}^2$, we start with a curve C.

Let $C = \{\vec{r}(t) \ : \ a \leq t \leq b\}$ be a smooth, oriented curve that lies entirely within the domain of the vector field. It is bounded by $\vec{r}(a) = \langle x(a), y(a)\rangle$ and $\vec{r}(b) = \langle x(b), y(b)\rangle$. Note that orientation determines the direction of the tangent line, meaning that increasing t gives the "positive" direction.

At any time $t_0 \in [a, b]$, we define the point on the curve as $\vec{r}_0 = \vec{r}(t_0) = \langle x(t_0), y(t_0)\rangle$ and the (not necessarily unit) tangent vector as $\vec{v} = \vec{r}\,'(t_0) = \langle x'(t_0), y'(t_0)\rangle$. The tangent line at $\vec{r}_0$ is then

$$\vec{r}(t) = \vec{r}_0 + t\vec{v} = \langle x_0, y_0\rangle + t\langle x'(t_0), y'(t_0)\rangle.$$

We now move on to a different idea. Let $C = \{\vec{r}(s) = \langle x(s), y(s)\rangle : a \leq s \leq b\}$ be a curve parameterized by arc length s, and let

$$\vec{F}(x, y) = \langle f(x, y), g(x, y)\rangle$$

be a vector field defined and continuous on a region D containing the curve C.

At a point $\vec{r}_0 = \vec{r}(s_0) = \langle x(s_0), y(s_0)\rangle$, we consider the tangent vector to the curve $\vec{v} = \vec{r}\,'(s_0) = \langle x'(s_0), y'(s_0)\rangle$, the unit tangent vector $\vec{T}(s_0) = \frac{\vec{v}}{\|\vec{v}\|} = \frac{\vec{r}\,'(s_0)}{\|\vec{r}\,'(s_0)\|}$, and the vector from the field $\vec{F}_0 = \vec{F}(\vec{r}_0) = \vec{F}(\vec{r}(s_0))$.

To understand the contribution of $\vec{F}$ to the line integral at $\vec{r}_0$, we consider the projection of $\vec{F}_0$ onto the unit tangent $\vec{T}(s_0)$. This projection gives the component of the field in the direction of motion:

$$\text{proj}_{\vec{T}}\vec{F} = (\vec{F} \cdot \vec{T})\vec{T}$$

Thus, the dot product $\vec{F} \cdot \vec{T}$ determines the amount of $\vec{F}$ that "pushes along" the path.

Now let θ be the angle between the field $\vec{F}$ and the unit tangent $\vec{T}$ at a point.

Then,

$$\vec{F} \cdot \vec{T} = \left\|\vec{F}\right\| \left\|\vec{T}\right\| \cos\theta = \left\|\vec{F}\right\| \cos\theta$$

since $\left\|\vec{T}\right\| = 1$.

The sign of $\vec{F} \cdot \vec{T}$ gives the nature of the contribution to the line integral:

Positive Contribution:

$$0 < \theta < 90^\circ \Rightarrow \vec{F} \cdot \vec{T} > 0$$

The vector field has a component pointing along the direction of the curve.

Zero Contribution:

$$\theta = 90^\circ \Rightarrow \vec{F} \cdot \vec{T} = 0$$

The field is perpendicular to the path and does no "work."

Negative Contribution:

$$90^\circ < \theta < 180^\circ \Rightarrow \vec{F} \cdot \vec{T} < 0$$

The field points against the direction of the curve and thus subtracts from the integral.

Let's now take a few steps back and go through the same approach we did previously for multiple integrals.

Line integrals let us integrate a function $f(x, y)$ along a curve C in the plane. You can think of it as summing up weighted contributions of f along small pieces of the curve, where each piece contributes based on its length and the value of the function at that location.

Suppose the curve C is given by a smooth parametrization:

$$x = x(t), \;\; y = y(t), \;\; \text{for } a \leq t \leq b,$$

or as a vector function

$$\vec{r}(t) = \langle x(t), y(t) \rangle.$$

We divide the interval $[a, b]$ into small subintervals, and approximate the curve by short segments. At each segment, we evaluate f and multiply by the segment's arc length Δs. This gives a Riemann-style sum:

$$\sum f(x_i^*, y_i^*)\, \Delta s_i.$$

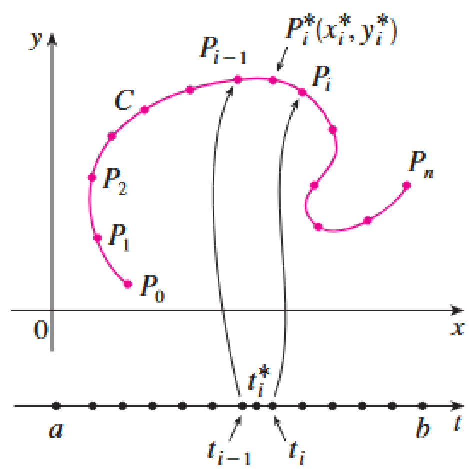


Image credit: Stewart

Taking the limit as the subintervals shrink, we define the scalar line integral:

If f is continuous on a smooth curve C, then the line integral of f along C is

$$\int_C f(x, y)\, ds = \lim_{n \to \infty} \sum_{i=1}^{n} f(x_i^*, y_i^*)\, \Delta s_i.$$

To evaluate this, we use the arc length formula from parametric curves:

$$ds = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.$$

So the line integral becomes

$$\boxed{\int_C f(x,y)\, ds = \int_a^b f(x(t), y(t)) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt.}$$

This is just a weighted integral along a curve!

If the curve C is a straight horizontal line from $(a, 0)$ to $(b, 0)$, then $x = x$ and $y = 0$. Thus, $ds = dx$. And the line integral becomes more familiar:

$$\int_C f(x,y)\, ds = \int_a^b f(x, 0)\, dx.$$

If $f(x, y) \geq 0$, then the line integral gives the area of a "fence" with base on the curve C, and height above each point equal to $f(x, y)$. You're summing the height of the fence along the curve, weighted by its arc length, similarly to what we did previous with multiple integrals.

The name "line integral" is a bit misleading. It actually would make more sense if they were called *curve integrals*.

EXAMPLE 9.1

Evaluate the line integral $\int_C 2x\,ds$, where C consists of two parts:

- C_1: the arc of the parabola $y = x^2$ from $(0,0)$ to $(1,1)$,
- C_2: the vertical line segment from $(1,1)$ to $(1,2)$.

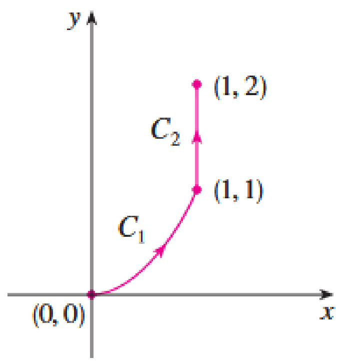


Image credit: Stewart

Solution:

We break the curve into two parts, $C = C_1 \cup C_2$, and compute each line integral separately. For C_1, we use x as the parameter since the curve is given by $y = x^2$. We also have $x = x$ and $0 \le x \le 1$.

The arc length element is:

$$ds = \sqrt{\left(\frac{dx}{dx}\right)^2 + \left(\frac{dy}{dx}\right)^2}\,dx = \sqrt{1 + (2x)^2}\,dx = \sqrt{1 + 4x^2}\,dx$$

So the line integral becomes

$$\int_{C_1} 2x\,ds = \int_0^1 2x\sqrt{1 + 4x^2}\,dx.$$

EXAMPLE 9.1 (CONTINUED)

We use substitution $u = 1 + 4x^2 \Rightarrow du = 8x\,dx$. Then, $x = 0 \Rightarrow u = 1$ and $x = 1 \Rightarrow u = 5$.

$$\int_0^1 2x\sqrt{1+4x^2}\,dx = \frac{1}{4}\int_1^5 u^{1/2}\,du = \frac{1}{4}\cdot\frac{2}{3}u^{3/2}\Big|_1^5$$
$$= \frac{1}{6}\left(5^{3/2} - 1\right) = \frac{5\sqrt{5}-1}{6}$$

We now move on to C_2. This is a vertical line, so we use y as the parameter:

$$x = 1,\ \ y = y,\ \ 1 \le y \le 2$$

Then,

$$ds = \sqrt{\left(\frac{dx}{dy}\right)^2 + \left(\frac{dy}{dy}\right)^2}\,dy = \sqrt{0+1}\,dy = dy.$$

Since $x = 1$, we have:

$$\int_{C_2} 2x\,ds = \int_1^2 2(1)\,dy = \int_1^2 2\,dy = 2$$

Finally,

$$\int_C 2x\,ds = \int_{C_1} 2x\,ds + \int_{C_2} 2x\,ds = \frac{5\sqrt{5}-1}{6} + 2$$

The value of a line integral $\int_C f(x,y)\,ds$ depends on the meaning of the physical function f. Suppose that $\rho(x,y)$ represents the linear density at a point (x,y) of a thin wire shaped like a curve C. Then the mass of the part of the wire from P_{i-1} to P_i is approximately $\rho(x_i^*, y_i^*)\,\Delta s_i$, and so the total mass of the wire is approximately $\sum \rho(x_i^*, y_i^*)\,\Delta s_i$. By taking more and more points on the curve, we obtain the mass m of the wire as the limiting value of these approximations:

$$m = \lim_{n\to\infty} \sum_{i=1}^{n} \rho(x_i^*, y_i^*)\, \Delta s_i = \int_C \rho(x, y)\, ds.$$

This comes from summing contributions of the form $\rho(x_i^*, y_i^*)\, \Delta s_i$ across the arc-length segments of the wire, then taking a limit as those segments shrink.

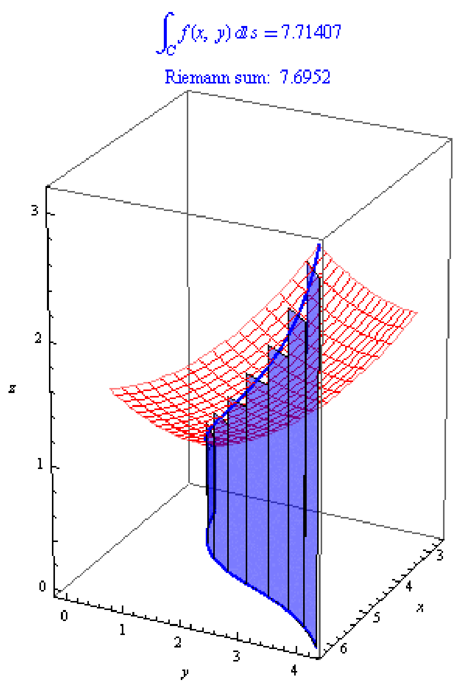
If $f(x, y) = 2 + x^2 y$ represents the density of a wire, then the line integral $\int_C f(x, y)\, ds$ gives the total mass of that wire.

The center of mass $(\bar{x}, \bar{y})$ of a wire with density function $\rho(x, y)$ is given by

$$\bar{x} = \frac{1}{m} \int_C x\rho(x, y)\, ds$$

and

$$\bar{y} = \frac{1}{m} \int_C y\rho(x, y)\, ds.$$



Visualization of Riemann sums for line integrals. Image credit: UMich

EXAMPLE 9.2

A wire takes the shape of the semicircle $x^2 + y^2 = 1$, $y \geq 0$, and is thicker near its base than near the top. Find the center of mass of the wire if the linear density at any point is proportional to its distance from the line $y = 1$.

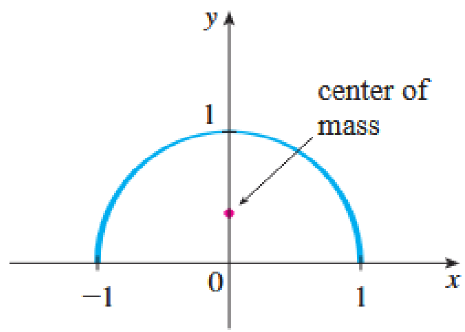


Image credit: Stewart

Solution:

We parametrize the semicircle as

$$x = \cos t,\ \ y = \sin t,\ \ 0 \leq t \leq \pi,$$

and note that $ds = dt$ since this is parameterized by arc length.

The linear density function is given by $\rho(x, y) = k(1 - y)$ where k is a constant.

We first find the mass of the wire:

$$m = \int_C k(1 - y)\, ds = \int_0^\pi k(1 - \sin t)\, dt = k\,[t + \cos t]_0^\pi = k(\pi - 2)$$

Then, center of mass in the y-direction is given by

$$\bar{y} = \frac{1}{m}\int_C y\rho(x, y)\, ds = \frac{1}{k(\pi - 2)}\int_0^\pi \sin t \cdot k(1 - \sin t)\, dt = \frac{1}{\pi - 2}\int_0^\pi (\sin t - \sin^2 t)\, dt.$$

EXAMPLE 9.2 (CONTINUED)

Use the identity $\sin^2 t = \frac{1}{2}(1 - \cos 2t)$:

$$\int_0^\pi (\sin t - \sin^2 t)\,dt = \int_0^\pi \left(\sin t - \frac{1}{2} + \frac{1}{2}\cos 2t\right) dt = \left[-\cos t - \frac{t}{2} + \frac{1}{4}\sin 2t\right]_0^\pi$$
$$= (-(-1) - \frac{\pi}{2} + 0) - (-1 - 0 + 0) = 1 - \frac{\pi}{2} + 1 = 2 - \frac{\pi}{2} = \frac{4-\pi}{2}$$

Hence,

$$\bar{y} = \frac{4-\pi}{2(\pi-2)}.$$

By symmetry, $\bar{x} = 0$, so the center of mass is

$$\left(0, \frac{4-\pi}{2(\pi-2)}\right) = (0, 0.38).$$

Sometimes, instead of integrating along a curve by arc length, it is more convenient to integrate with respect to one coordinate variable, x or y.

This means we approximate the line integral by summing values multiplied by small changes in x or y rather than by small arc length segments.

Formally, the line integrals with respect to x and y are defined by the limits of sums:

$$\int_C f(x,y)\,dx = \lim_{n\to\infty} \sum_{i=1}^n f(x_i^*, y_i^*)\Delta x_i$$

and

$$\int_C f(x,y)\,dy = \lim_{n\to\infty} \sum_{i=1}^n f(x_i^*, y_i^*)\Delta y_i,$$

where $\Delta x_i = x_i - x_{i-1}$ and $\Delta y_i = y_i - y_{i-1}$.

To evaluate these integrals, we use a parametrization of the curve C:

$$x = x(t),\ y = y(t),\ a \le t \le b.$$

Since $dx = x'(t)dt$ and $dy = y'(t)dt$, the integrals become

$$\int_C f(x,y)\,dx = \int_a^b f(x(t),y(t))x'(t)\,dt,$$

and

$$\int_C f(x,y)\,dy = \int_a^b f(x(t),y(t))y'(t)\,dt.$$

It is common for line integrals with respect to x and y to appear together. In such cases, the integral is abbreviated as

$$\int_C P(x,y)\,dx + \int_C Q(x,y)\,dy = \int_C P(x,y)\,dx + Q(x,y)\,dy.$$

To set up a line integral, we often need a parametric form of the curve. For a line segment starting at $\vec{r}_0$ and ending at $\vec{r}_1$, a natural parametrization is given by

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

which moves linearly from $\vec{r}_0$ to $\vec{r}_1$ as t goes from 0 to 1.

EXAMPLE 9.3

Let $f(x, y) = x^2 - y + 3$, and let C be the semicircle of radius 2 around the origin lying above the x-axis. Approximate $\int_C f(x, y)\,dx$ using a Riemann sum with 3 subdivisions and then evaluate the integral exactly.

Solution:

First, parametrize the curve C as

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

We divide the interval $[0, \pi]$ into 3 equal subdivisions $[0, \pi/3]$, $[\pi/3, 2\pi/3]$, and $[2\pi/3, \pi]$. The sample points are chosen as the midpoints:

$$t = \pi/6,\ t = \pi/2,\ t = 5\pi/6$$

For the Riemann sum approximation, the base length of the first rectangle is the distance between $(2\cos 0, 2\sin 0) = (2, 0)$ and $(2\cos\frac{\pi}{3}, 2\sin\frac{\pi}{3}) = (1, \sqrt{3})$ with length $\sqrt{(1-2)^2 + (\sqrt{3}-0)^2} = \sqrt{1+3} = 2$.

The height is the function value at the sample point $(2\cos\frac{\pi}{6}, 2\sin\frac{\pi}{6}) = (\sqrt{3}, 1)$:

$$f(\sqrt{3}, 1) = (\sqrt{3})^2 - 1 + 3 = 3 - 1 + 3 = 5.$$

Area of first rectangle is given by $2(5) = 10$. Then, the base length of the second rectangle is the distance between $(1, \sqrt{3})$ and $(-1, \sqrt{3})$ which is $\sqrt{(-1-1)^2 + (\sqrt{3}-\sqrt{3})^2} = \sqrt{4} = 2$. The height at the midpoint $(2\cos\frac{\pi}{2}, 2\sin\frac{\pi}{2}) = (0, 2)$ is $f(0, 2) = 0^2 - 2 + 3 = 1$.

EXAMPLE 9.3 (CONTINUED)

The area of the second rectangle is given by $2(1) = 2$. The base length of the third rectangle is the distance between $(-1, \sqrt{3})$ and $(-2, 0)$ which is $\sqrt{(-2+1)^2 + (0-\sqrt{3})^2} = \sqrt{1+3} = 2$. The height at $(2\cos\frac{5\pi}{6}, 2\sin\frac{5\pi}{6}) = (-\sqrt{3}, 1)$ is $f(-\sqrt{3}, 1) = (-\sqrt{3})^2 - 1 + 3 = 3 - 1 + 3 = 5$. The area of third rectangle is also given by $2(5) = 10$. Adding these areas gives the Riemann sum approximation:

$$10 + 2 + 10 = 22.$$

For the exact evaluation, first compute the derivatives $x'(t) = -2\sin t$ and $y'(t) = 2\cos t$. Thus,

$$\sqrt{x'(t)^2 + y'(t)^2} = \sqrt{4\sin^2 t + 4\cos^2 t} = \sqrt{4(\sin^2 t + \cos^2 t)} = 2.$$

The problem asks us to evaluate $\int_C f(x,y)\,dx$, so the line integral with respect to x is

$$\int_C f(x,y)\,dx = \int_0^\pi f(2\cos t, 2\sin t)x'(t)\,dt = \int_0^\pi (4\cos^2 t - 2\sin t + 3)(-2\sin t)\,dt.$$

EXAMPLE 9.3 (CONTINUED)

Expand the integrand:

$$= \int_0^\pi (-8\cos^2 t \sin t + 4\sin^2 t - 6\sin t)\, dt.$$

To simplify, apply the double-angle identity for cosine squared:

$$4\cos^2 t = 2\cos(2t) + 2,$$

So the integrand becomes

$$(2\cos(2t) - 2\sin t + 5) \cdot 2 = 4\cos(2t) - 4\sin t + 10.$$

This has antiderivative $2\sin(2t) + 4\cos t + 10t$. Using this, we finally get

$$\int_C f(x,y)\, dx = (2\sin(2t) + 4\cos t + 10t)\Big|_0^\pi = (0 - 4 + 10\pi) - (0 + 4 + 0) = 10\pi - 8 = 23.4.$$

Thus, our estimate of 22 was not horrible.

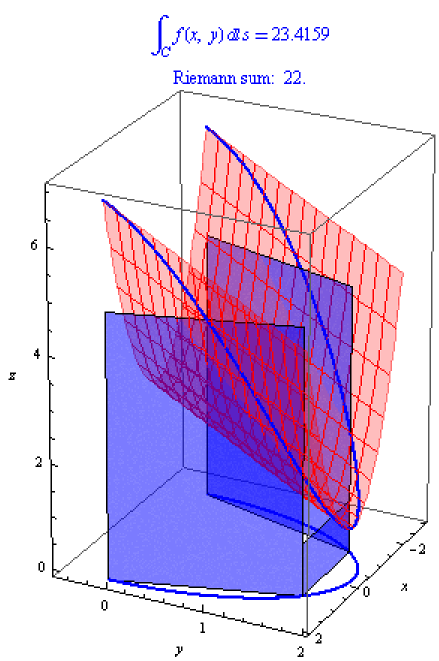


Image credit: UMich

Now that we've worked a little bit with line integrals in the plane, we will now move to line integrals in space. That is, line integrals in three-dimensional space.

Suppose C is a smooth space curve given by the parametric equations

$$x = x(t),\ y = y(t),\ z = z(t),\ a \le t \le b,$$

or equivalently by the vector equation

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}.$$

If f is a continuous function of three variables on a region containing C, then the line integral of f along C with respect to arc length is defined as

$$\int_C f(x, y, z)\, ds = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*, z_i^*)\, \Delta s_i.$$

This integral can be evaluated using the formula that generalizes the planar case by including the z-component:

$$\boxed{\int_C f(x, y, z)\, ds = \int_a^b f(x(t), y(t), z(t)) \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}\, dt.}$$

In vector notation, this can be expressed compactly as

$$\int_C f(\vec{r}(t)) \|\vec{r}\,'(t)\| dt.$$

For the special case $f(x, y, z) = 1$, the line integral measures the length L of the curve C:

$$\int_C ds = \int_a^b \|\vec{r}\,'(t)\| dt = L.$$

Line integrals along C with respect to x, y, and z can also be defined. For example,

$$\int_C f(x, y, z)\, dz = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i^*, y_i^*, z_i^*) \Delta z_i = \int_a^b f(x(t), y(t), z(t)) z'(t)\, dt$$

Therefore, as in the plane, we evaluate integrals of the form

$$\int_C P(x,y,z)\,dx + Q(x,y,z)\,dy + R(x,y,z)\,dz.$$

Previously, we went over line integrals with respect to x and y. We now add in the z-component:

$$\boxed{\begin{aligned}\int_C f(x,y,z)\,dx &= \int_a^b f(x(t),y(t),z(t))x'(t)\,dt,\\ \int_C f(x,y,z)\,dy &= \int_a^b f(x(t),y(t),z(t))y'(t)\,dt,\\ \int_C f(x,y,z)\,dz &= \int_a^b f(x(t),y(t),z(t))z'(t)\,dt,\end{aligned}}$$

where the curve C is parameterized by

$$x = x(t),\;\; y = y(t),\;\; z = z(t),\;\; a \le t \le b.$$

Combined, we have

$$\int_C P\,dx + Q\,dy + R\,dz = \int_C P(x,y,z)\,dx + \int_C Q(x,y,z)\,dy + \int_C R(x,y,z)\,dz.$$

EXAMPLE 9.4

Evaluate the line integral $\int_C y \sin z \, ds$, where C is the circular helix defined by

$$x = \cos t, \quad y = \sin t, \quad z = t, \quad 0 \leq t \leq 2\pi.$$

Solution:

Using the formula for scalar line integrals in space, we have:

$$\int_C y \sin z \, ds = \int_0^{2\pi} (\sin t) \sin t \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} \, dt.$$

$$\sqrt{(-\sin t)^2 + (\cos t)^2 + 1^2} = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{1+1} = \sqrt{2}.$$

Combining the integral together again yields

$$\int_0^{2\pi} \sin^2 t \cdot \sqrt{2} \, dt = \sqrt{2} \int_0^{2\pi} \sin^2 t \, dt.$$

Use the identity $\sin^2 t = \frac{1-\cos 2t}{2}$ to get

$$\sqrt{2} \int_0^{2\pi} \sin^2 t \, dt = \sqrt{2} \int_0^{2\pi} \frac{1 - \cos 2t}{2} \, dt = \frac{\sqrt{2}}{2} \int_0^{2\pi} (1 - \cos 2t) \, dt.$$

Evaluating the integral,

$$\int_0^{2\pi} 1 \, dt = 2\pi, \quad \int_0^{2\pi} \cos 2t \, dt = 0,$$

so

$$\frac{\sqrt{2}}{2}(2\pi - 0) = \sqrt{2}\pi.$$

We now come full circle and return to line integrals of vector fields.
Suppose a particle moves along a smooth curve C parameterized by

$$\vec{r}(t) = \langle x(t), y(t), z(t)\rangle,\ a \le t \le b.$$

The vector field $\vec{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ represents the force at each point in space.

We divide the curve C into small segments between points P_{i-1} and P_i, each with length Δs_i.

At each point $P_i^* = (x_i^*, y_i^*, z_i^*)$ in the ith segment, the particle approximately moves in the direction of the unit tangent vector

$$\mathbf{T}(t_i^*) = \frac{\vec{r}\,'(t_i^*)}{\|\vec{r}\,'(t_i^*)\|}.$$

The work done by the force $\vec{F}$ moving the particle through that segment is roughly

$$\vec{F}(x_i^*, y_i^*, z_i^*) \cdot \mathbf{T}(t_i^*)\, \Delta s_i,$$

which is the projection of the force onto the direction of motion times the length of the path segment.

Summing over all segments and taking the limit as the segment lengths approach zero gives the total work done by $\vec{F}$ along the curve:

$$W = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbf{F}(x_i^*, y_i^*, z_i^*) \cdot \mathbf{T}(t_i^*)\, \Delta s_i.$$

This limit defines the line integral of $\vec{F}$ along C:

$$\boxed{W = \int_C \vec{F} \cdot d\vec{r} = \int_C \vec{F} \cdot \mathbf{T}\, ds.}$$

Because the unit tangent vector is $\mathbf{T}(t) = \frac{\vec{r}\,'(t)}{\|\vec{r}\,'(t)|}\|$, we can rewrite the line integral using the parameter t as

$$\int_C \mathbf{F} \cdot d\vec{r} = \int_a^b \mathbf{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt.$$

Expanding in components:

$$\int_C \mathbf{F} \cdot d\vec{r} = \int_a^b [P(x(t), y(t), z(t))x'(t) + Q(x(t), y(t), z(t))y'(t) + R(x(t), y(t), z(t))z'(t)]\, dt.$$

This expression is often written as

$$\int_C P\, dx + Q\, dy + R\, dz,$$

where $\mathbf{F} = \langle P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}\rangle$.

Let $\mathbf{F}$ be a continuous vector field defined on a smooth curve C given by a vector function $\vec{r}(t)$ for $a \le t \le b$. Then the line integral of $\mathbf{F}$ along C is defined as

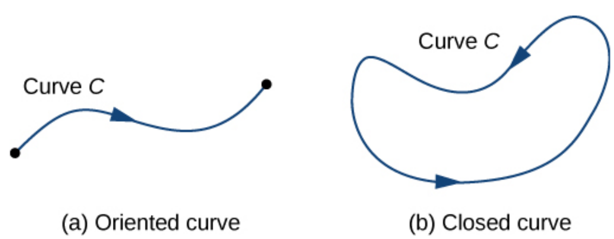
$$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt = \int_C \vec{F} \cdot \mathbf{T}\, ds,$$

where $\mathbf{T}$ is the unit tangent vector and ds is the differential arc length.

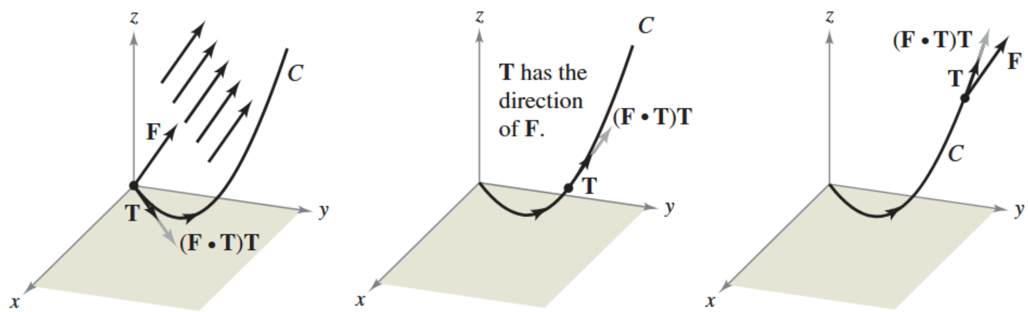
When computing the work done by $\mathbf{F}$ moving a particle along C, it is important to specify the direction of travel along the curve. A particle can move either forward or backward along C, and the work depends on this direction.

This specified direction along C is called the *orientation* of the curve. The positive direction along C is the specified orientation, while the opposite direction is negative. A curve with a chosen orientation is called an *oriented curve*.

A *closed curve* is one for which there exists a parameterization $\vec{r}(t)$ defined on $a \le t \le b$ and satisfying $\vec{r}(a) = \vec{r}(b)$ such that the curve is traversed exactly once. In other words, the parameterization is one-to-one on the open interval (a, b). We will discuss this more later.



(a) An oriented curve between two points. (b) A closed oriented curve. Image credit: Strang & Herman



At each point on C, the force in the direction of motion is $(\vec{F} \cdot \mathbf{T})\mathbf{T}$. Image credit: Larson & Edwards

Concept	**Formula**
Curve parameterization	$C : \vec{r}(t) = \langle x(t), y(t), z(t) \rangle, \quad a \le t \le b$
Length of curve	$L = \int_C ds = \int_a^b \lvert \vec{r}\,'(t) \rvert \, dt$
Arc length differential	$ds = \lVert \vec{r}\,'(t) \rVert dt = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2} dt$
Scalar line integral (general 3D)	$\int_C f(x, y, z) ds = \int_a^b f(\vec{r}(t)) \lVert \vec{r}\,'(t) \rVert dt$
Scalar line integral (2D special case)	$\int_C f(x, y) ds = \int_a^b f(x(t), y(t)) \sqrt{(x'(t))^2 + (y'(t))^2} dt$
Line integrals with respect to coordinates	$\int_C f(x,y,z)\,dx = \int_a^b f(\vec{r}(t))\,x'(t)\,dt$, $\int_C f(x,y,z)\,dy = \int_a^b f(\vec{r}(t))\,y'(t)\,dt$, $\int_C f(x,y,z)\,dz = \int_a^b f(\vec{r}(t))\,z'(t)\,dt.$
Vector line integral (general form)	$\int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t) dt$
Vector line integral (component form)	$\int_C \vec{F} \cdot d\vec{r} = \int_C P dx + Q dy + R dz, \quad \vec{F} = \langle P, Q, R \rangle$
Relation between scalar and vector line integrals	$\int_C \vec{F} \cdot d\vec{r} = \int_C \vec{F} \cdot \mathbf{T}\, ds, \quad \mathbf{T} = \frac{\vec{r}\,'(t)}{\lVert \vec{r}\,'(t) \rVert}$

Summary of Key Line Integral Formulas and Concepts

EXAMPLE 9.5

Evaluate the line integral $\int_C \vec{F} \cdot d\vec{r}$ where $\vec{F}(x, y, z) = xy\mathbf{i} + yz\mathbf{j} + zx\mathbf{k}$ and C is the twisted cubic given by

$$x = t,\ y = t^2,\ z = t^3,\ 0 \leq t \leq 1.$$

Solution:

$$\vec{r}(t) = t\mathbf{i} + t^2\mathbf{j} + t^3\mathbf{k}$$

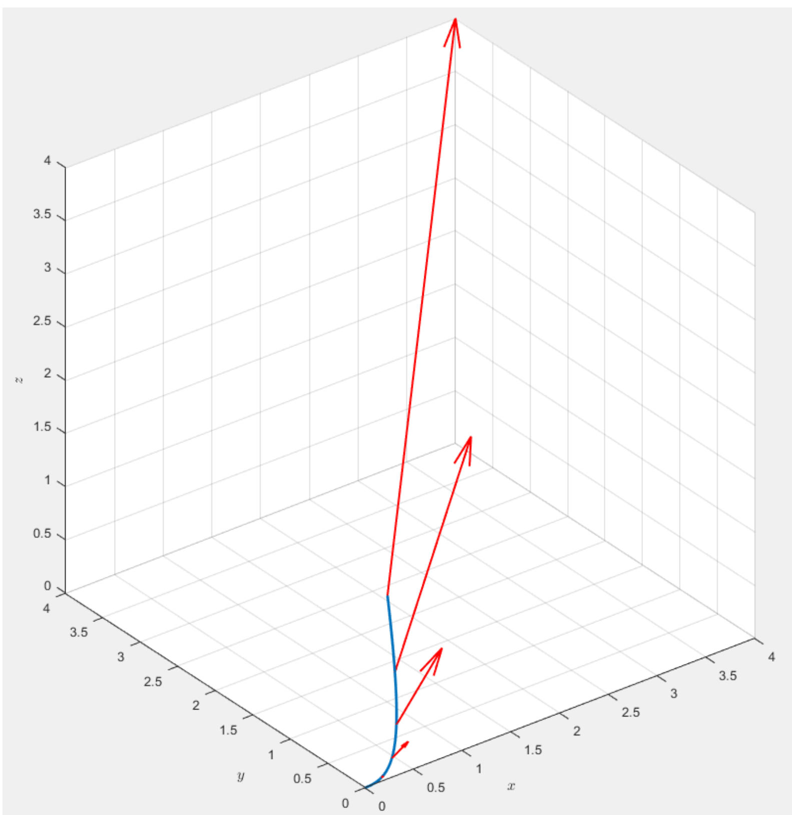
$$\vec{r}\,'(t) = \mathbf{i} + 2t\mathbf{j} + 3t^2\mathbf{k}$$

$$\vec{F}(\vec{r}(t)) = t^3\mathbf{i} + t^5\mathbf{j} + t^4\mathbf{k}$$

$$\vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t) = t^3 + 2t^6 + 3t^6 = t^3 + 5t^6$$

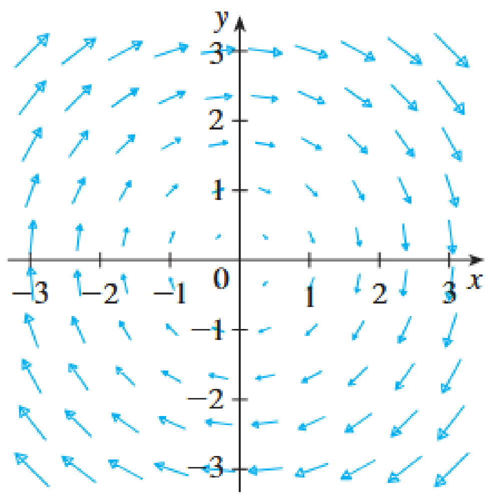
$$\int_C \vec{F} \cdot d\vec{r} = \int_0^1 (t^3 + 5t^6)dt = \left[\frac{t^4}{4} + \frac{5t^7}{7}\right]_0^1 = \frac{27}{28}$$

And here is a graph of the twisted cubic with some vectors acting along it:



EXAMPLE 9.6

Let $\vec{F}$ be the vector field shown in the figure.



If C_1 is the vertical line segment from $(-3,-3)$ to $(-3,3)$, determine whether

$$\int_{C_1} \vec{F} \cdot d\vec{r}$$

is positive, negative, or zero.

If C_2 is the counterclockwise-oriented circle with radius 3 centered at the origin, determine whether

$$\int_{C_2} \vec{F} \cdot d\vec{r}$$

is positive, negative, or zero.

Solution:

(a) Along the vertical line $x=-3$, the vectors of $\vec{F}$ have positive y-components. Since the path moves upward, the tangent vector $\vec{T}$ points up, making the dot product $\vec{F} \cdot \vec{T}$ always positive. Therefore,

$$\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_1} \vec{F} \cdot \vec{T}\, ds$$

is positive.

(b) Along the circle of radius 3, all the (nonzero) vectors of $\vec{F}$ point clockwise, which is opposite the counterclockwise orientation of the path. Thus, $\vec{F} \cdot \vec{T} < 0$ everywhere on C_2, and therefore,

$$\int_{C_2} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot \vec{T}\, ds$$

is negative.

EXAMPLE 9.7

A 160 lb man carries a 10 lb sack of grain up a helical staircase that wraps around a silo with radius 20 ft. The silo is 90 ft tall, and the man makes exactly three complete revolutions climbing to the top. How much work is done by the man against gravity?

Solution:

We have $\vec{F} = 160 + 10 = 170$ lb. We will parametrize the staircase with

$$x = 20\cos t,\ y = 20\sin t,\ z = \frac{90}{6\pi}t = \frac{15}{\pi}t,\ 0 \le t \le 6\pi$$

Thus,

$$\begin{aligned} W &= \int_C \mathbf{F} \cdot d\vec{r} \\ &= \int_0^{6\pi} \langle 0, 0, 170\rangle \cdot \langle -20\sin t, 20\cos t, \frac{15}{\pi}\rangle\, dt \\ &= \int_0^{6\pi} 170 \cdot \frac{15}{\pi} dt = 170 \cdot \frac{15}{\pi} \cdot 6\pi = 170 \times 15 \times 6 = 15300 \text{ ft} \cdot \text{lb}. \end{aligned}$$

9.3 The Fundamental Theorem for Line Integrals

Recall from single-variable calculus that the fundamental theorem of calculus states

$$\int_a^b F'(x)\, dx = F(b) - F(a),$$

where F' is continuous on $[a, b]$. This expresses that the integral of a rate of change equals the net change of the original function.

We can generalize this idea to functions of several variables and line integrals.

Let C be a smooth curve parameterized by a vector function $\vec{r}(t)$ that is defined on $a \le t \le b$. Let f be a differentiable scalar function of two or three variables whose gradient vector ∇f is continuous on C. Then,

$$\int_C \nabla f \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a)).$$

This theorem says that for a conservative vector field ∇f, the line integral depends only on the values of f at the endpoints of C, not on the path taken. This is known as *path independence*.

If f is a function of two variables and C is a plane curve with initial point $A(x_1, y_1)$ and terminal point $B(x_2, y_2)$, then

$$\int_C \nabla f \cdot d\vec{r} = f(x_2, y_2) - f(x_1, y_1).$$

If f is a function of three variables and C is a space curve from $A(x_1, y_1, z_1)$ to $B(x_2, y_2, z_2)$, then

$$\int_C \nabla f \cdot d\vec{r} = f(x_2, y_2, z_2) - f(x_1, y_1, z_1).$$

Using the definition of the line integral,

$$\int_C \nabla f \cdot d\vec{r} = \int_a^b \nabla f(\vec{r}(t)) \cdot \vec{r}\,'(t)\, dt,$$

we expand the dot product,

$$= \int_a^b \left(\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} \right) dt.$$

And then by the chain rule,

$$= \int_a^b \frac{d}{dt} f(\vec{r}(t))\, dt = f(\vec{r}(b)) - f(\vec{r}(a)).$$

This last step follows from the single-variable fundamental theorem of calculus.

Suppose C_1 and C_2 are two piecewise-smooth curves (paths) with the same initial point A and terminal point B. In general,

$$\int_{C_1} \vec{F} \cdot d\vec{r} \neq \int_{C_2} \vec{F} \cdot d\vec{r},$$

because the value of the line integral depends on how the field behaves along the path C.

However, one key implication of the fundamental theorem of line integrals is that for a conservative vector field $\vec{F} = \nabla f$, the line integral depends only on the endpoints. More precisely,

$$\int_{C_1} \nabla f \cdot d\vec{r} = \int_{C_2} \nabla f \cdot d\vec{r}$$

whenever ∇f is continuous.

In general, for any continuous vector field $\vec{F}$ defined on a domain D, the line integral

$$\int_C \vec{F} \cdot d\vec{r}$$

is said to be *independent of path* if

$$\int_{C_1} \mathbf{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$$

for any two paths C_1 and C_2 in D with the same initial and terminal points.

A curve C is called *closed* if its terminal point coincides with its initial point. That is, $\vec{r}(b) = \vec{r}(a)$.

If the line integral is independent of path in D, then for any closed curve $C \in D$, we have

$$\int_C \vec{F} \cdot d\vec{r} = 0.$$

Conversely, if and only if

$$\int_C \vec{F} \cdot d\vec{r} = 0,$$

for every closed curve $C \in D$, then the line integral is independent of path.

Finally, we have the fundamental characterization of conservative vector fields:

Suppose $\mathbf{F}$ is a continuous vector field defined on an open, connected region D. Then $\vec{F}$ is conservative if and only if the line integral $\int_C \vec{F} \cdot d\vec{r}$ is independent of path in D. Equivalently, there exists a scalar potential function f on D such that

$$\nabla f = \vec{F}.$$

Let's now learn when to classify vector fields as conservative.

Suppose $\vec{F} = P\mathbf{i} + Q\mathbf{j}$ is a conservative vector field, where P and Q have continuous first-order partial derivatives on a domain D. Then there exists a scalar function f such that $\vec{F} = \nabla f$, meaning

$$P = \frac{\partial f}{\partial x} \quad \text{and} \quad Q = \frac{\partial f}{\partial y}.$$

By Clairaut's theorem on the equality of mixed partial derivatives, we have

$$\frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial Q}{\partial x}.$$

This leads to the following important criterion:

If $\vec{F}(x, y) = P(x, y)\mathbf{i} + Q(x, y)\mathbf{j}$ is a conservative vector field, where P and Q have continuous first-order partial derivatives on a domain D, then throughout D we have

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}.$$

The converse of this theorem holds only for special types of domains. To understand this, we introduce the idea of the *simple curve*, which is a curve that does not intersect itself anywhere between its endpoints. For example, if $r(a) = r(b)$ but $r(t_1) \neq r(t_2)$ for $a < t_1 < t_2 < b$, the curve is simple and closed.

In the previous theorem, we required D to be an open connected region. For the converse, a stronger condition is necessary. A *simply-connected region* in the plane is a connected region D such that every simple closed curve in D encloses only points that are also in D. Intuitively, a simply-connected region contains no holes and cannot be split into two separate parts.

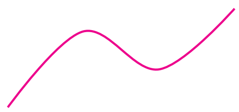
In terms of simply-connected regions, we can now state a version of the theorem with additional assumptions (also known as a partial converse) that provides a practical method for verifying whether a vector field on $\mathbb{R}^2$ is conservative:

Let $\vec{F} = P\mathbf{i} + Q\mathbf{j}$ be a vector field on an open, simply-connected region D. Suppose that P and Q have continuous first-order partial derivatives and satisfy

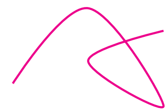
$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$$

throughout D. Then $\vec{F}$ is conservative.

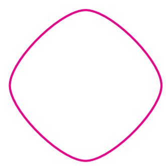
Here is a visualization of the types of curves:



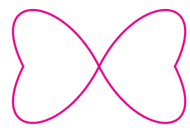
simple, not closed



not simple, not closed



simple, closed



not simple, closed

Here is a visualization of the types of regions:

simply-connected regions

connected regions that are not simply connected

regions that are not connected

EXAMPLE 9.8

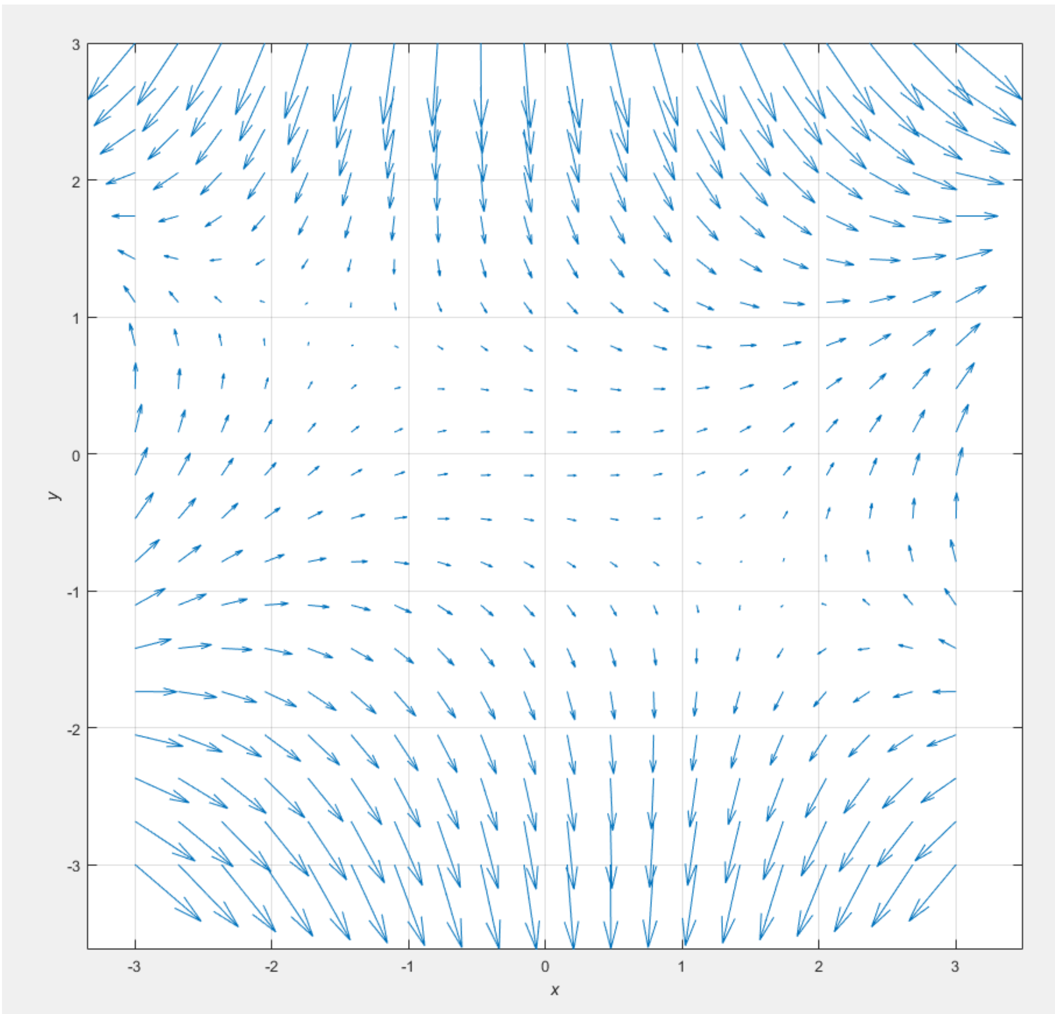
Determine whether or not the vector field $\vec{F}(x, y) = \big(3 + 2xy\big)\mathbf{i} + \big(x^2 - 3y^2\big)\mathbf{j}$ is conservative.

Solution:

Let $P(x, y) = 3 + 2xy$ and $Q(x, y) = x^2 - 3y^2$. Then compute the partial derivatives:

$$\frac{\partial P}{\partial y} = 2x, \ \frac{\partial Q}{\partial x} = 2x$$

Since $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$, and the domain of $\vec{F}$ is the entire plane $\mathbb{R}^2$, which is open and simply-connected, we see that $2x = 2x$ and thus $\vec{F}$ is conservative.



EXAMPLE 9.9

(a) Suppose $\vec{F}(x,y) = (3+2xy)\mathbf{i} + (x^2 - 3y^2)\mathbf{j}$ Find a function f such that $\vec{F} = \nabla f$.
(b) Evaluate the line integral $\int_C \vec{F} \cdot d\vec{r}$, where $\vec{r}(t) = e^t \sin t\mathbf{i} + e^t \cos t\mathbf{j}$ is defined on $0 \le t \le \pi$.

Solution:

From **Example 9.8**, we know $\vec{F}$ is conservative, so there exists a potential function f such that $\nabla f = \vec{F}$. That is,

$$f_x(x,y) = 3 + 2xy,$$
$$f_y(x,y) = x^2 - 3y^2.$$

Integrate f_x with respect to x:

$$f(x,y) = 3x + x^2 y + g(y),$$

where $g(y)$ is an unknown function of y. Differentiate this with respect to y:

$$f_y(x,y) = x^2 + g'(y).$$

Comparing with f_y above,

$$x^2 + g'(y) = x^2 - 3y^2 \Rightarrow g'(y) = -3y^2.$$

Integrate with respect to y:

$$g(y) = -y^3 + K,$$

where K is a constant.

Thus the potential function is

$$f(x,y) = 3x + x^2 y - y^3 + K.$$

EXAMPLE 9.9

(b) By the fundamental theorem for line integrals,

$$\int_C \vec{F} \cdot d\vec{r} = \int_C \nabla f \cdot d\vec{r} = f(\vec{r}(\pi)) - f(\vec{r}(0)).$$

Calculate the endpoints:

$$\vec{r}(0) = (0, 1)$$
$$\vec{r}(\pi) = (0, -e^{\pi})$$

Evaluate f at the endpoints, choosing $K = 0$:

$$f(0, -e^{\pi}) = 0 + 0 - (-e^{\pi})^3 = e^{3\pi}$$
$$f(0, 1) = 0 + 0 - 1 = -1.$$

Therefore,

$$\int_C \vec{F} \cdot d\vec{r} = e^{3\pi} - (-1) = e^{3\pi} + 1.$$

This method is much faster than directly evaluating the line integral.

Let $\vec{F}$ be a continuous force field moving an object along a path C parameterized by $\vec{r}(t)$, $a \leq t \leq b$, where $\vec{r}(a) = A$ and $\vec{r}(b) = B$.

According to Newton's second law, the force at a point on the curve relates to acceleration $\vec{a}(t) = \vec{r}''(t)$ by

$$\vec{F}(\vec{r}(t)) = m\vec{r}''(t),$$

where m is the mass.

The work done by the force is

$$W = \int_C \vec{F} \cdot d\vec{r} = \int_a^b \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\,dt = \int_a^b m\vec{r}\,''(t) \cdot \vec{r}\,'(t)\,dt.$$

Using the product rule, we can rewrite this as

$$W = m \int_a^b \frac{d}{dt}\left[\vec{r}\,'(t) \cdot \vec{r}\,'(t)\right] \frac{1}{2} dt = \frac{m}{2} \int_a^b \frac{d}{dt} \|\vec{r}\,'(t)\|^2 dt,$$

which evaluates to

$$W = \frac{m}{2}(\|\vec{r}\,'(b)\|^2 - \|\vec{r}\,'(a)\|^2).$$

Therefore,

$$W = \frac{1}{2}m\|\vec{v}(b)\|^2 - \frac{1}{2}m\|\vec{v}(a)\|^2,$$

where $\vec{v} = \vec{r}\,'$ is velocity.

The quantity $\frac{1}{2}m\|\vec{v}(t)\|^2$ is called the kinetic energy of the object.

We can rewrite the work as

$$W = P(A) + K(A) = K(B) - K(A),$$

which states that work done by the force along the path equals the change in kinetic energy.

Now assume $\vec{F}$ is a conservative force field. That is, $\vec{F} = \nabla f$.

In physics, the potential energy at a point (x, y, z) is

$$P(x, y, z) = -f(x, y, z),$$

so that

$$\vec{F} = -\nabla P.$$

Thus,

$$W = \int_C \vec{F} \cdot d\vec{r} = -\int_C \nabla P \cdot d\vec{r} = -[P(\vec{r}(b)) - P(\vec{r}(a))] = P(A) - P(B).$$

Comparing with the kinetic energy expression, we get the law of conservation of energy:

$$P(A) + K(A) = P(B) + K(B).$$

This means that the sum of potential and kinetic energy remains constant when moving from point A to point B under a conservative force field. This is why we call some vector fields *conservative*.

If a force is not conservative, then we cannot define a scalar potential function $f(x, y, z)$ such that $\vec{F} = \nabla f$. This is because work depends on the path taken, not just the endpoints:

$$\int_C \vec{F} \cdot d\vec{r} \neq f(B) - f(A)$$

In a non-conservative field, energy can be gained or lost as you move through the field.

And for a closed curve C, the line integral is generally nonzero:

$$\oint_C \vec{F} \cdot d\vec{r} \neq 0$$

EXAMPLE 9.10

We are given the vector field $\vec{F}(x,y) = \langle 2xy, x^2 \rangle$ and three curves that start at $(1,2)$ and end at $(3,2)$:

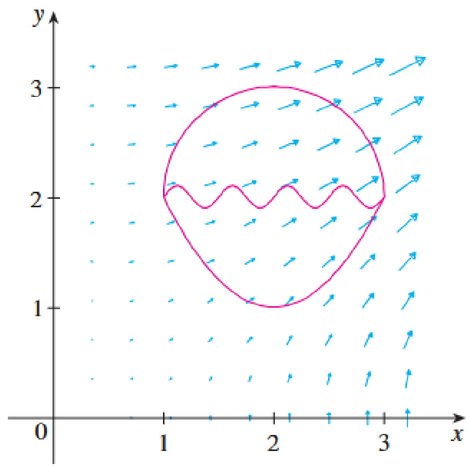


Image credit: Stewart

(a) Explain why $\int_C \vec{F} \cdot d\vec{r}$ has the same value for all three curves.
(b) What is this common value?

Solution:

(a) $\vec{F}$ is a conservative vector field because it has continuous first-order partial derivatives

$$\frac{\partial}{\partial y}(2xy) = 2x = \frac{\partial}{\partial x}(x^2)$$

on $\mathbb{R}^2$ which is open and simply-connected. Thus, $\vec{F}$ is independent of path. For conservative fields, the line integral depends only on the endpoints, not the path taken. This means that there exists a scalar potential function $f(x,y)$ such that $\vec{F} = \nabla f$.

EXAMPLE 9.10 (CONTINUED)

(b) To evaluate the line integral, we first find the potential function $f(x, y)$.

Integrate the first component of $\vec{F}$ with respect to x:

$$f(x, y) = \int 2xy\, dx = x^2 y + g(y),$$

where $g(y)$ is a function of y.

Differentiate with respect to y:

$$f_y = x^2 + g'(y)$$

We are told that $f_y = x^2$, so:

$$x^2 + g'(y) = x^2 \Rightarrow g'(y) = 0 \Rightarrow g(y) = C$$

where C is a constant.

So the potential function is $f(x, y) = x^2 y + C$. Since adding constants to potential functions does not affect the gradient, we can say that $f(x, y) = x^2 y + C \cong x^2 y$. We now apply the fundamental theorem for line integrals:

$$\int_C \vec{F} \cdot d\vec{r} = f(3, 2) - f(1, 2) = (9)(2) - (1)(2) = 18 - 2 = 16$$

Summary

Let $\vec{F}$ be a continuous vector field on a domain D.

1. $\vec{F}$ is called **conservative** if there exists a scalar potential function f such that $\vec{F} = \nabla f$.
2. The line integral $\int_C \vec{F} \cdot d\vec{r}$ is **independent of path** if for any two piecewise-smooth paths C_1 and C_2 in D with the same initial and terminal points. That is, $\int_{C_1} \vec{F} \cdot d\vec{r} = \int_{C_2} \vec{F} \cdot d\vec{r}$.
3. A path C is **closed** if its initial and terminal points coincide. For example, a circle is a closed path.
4. A path C is **simple** if it does not intersect itself. For example, a circle is simple; a figure-eight curve is not.
5. A region $R \subseteq D$ is **open** if it contains none of its boundary points.
6. A region $R \subseteq D$ is **connected** if any two points in R can be joined by a path lying entirely in R.
7. A region $R \subseteq D$ is **simply-connected** if it is connected and contains no holes.

Properties:

1. The line integral of a gradient field $\int_C \nabla f \cdot d\vec{r}$ is path independent.
2. If $\vec{F}$ is conservative, then $\int_C \vec{F} \cdot d\vec{r} = \int_C \nabla f \cdot d\vec{r}$, which is path independent.
3. If $\vec{F}$ is continuous on an open connected domain D and $\int_C \vec{F} \cdot d\vec{r}$ is independent of path for all paths in D, then $\vec{F}$ is conservative.
4. If $\int_C \vec{F} \cdot d\vec{r}$ is independent of path, then $\int_C \vec{F} \cdot d\vec{r} = 0$ for every closed path C.
5. Conversely, if $\int_C \vec{F} \cdot d\vec{r} = 0$ for every closed path C, then $\int_C \vec{F} \cdot d\vec{r}$ is independent of path.
6. $\oint_C \vec{F} \cdot d\vec{r}$ denotes a line integral around a closed curve. That is, the curve begins and ends at the same point. It measures the net circulation of a vector field around a loop. It is equal to 0 for a conservative vector field.

EXAMPLE 9.11

Evaluate the line integral $\oint_C \vec{F} \cdot d\vec{r}$ where $vecF(x, y) = \langle x, y\rangle$ along the curve $C = \vec{r}(t) = \langle 4\cos t, 4\sin t\rangle$ defined for $0 \leq t \leq 2\pi$.

Solution:

We use the parametrization and compute:

$$\vec{F}(x(t), y(t)) = \langle 4\cos t, 4\sin t\rangle, \quad \vec{r}\,'(t) = \langle -4\sin t, 4\cos t\rangle.$$

$$\vec{F} \cdot \vec{r}\,'(t) = (4\cos t)(-4\sin t) + (4\sin t)(4\cos t) = -16\sin t\cos t + 16\sin t\cos t = 0.$$

$$\oint_C \vec{F} \cdot d\vec{r} = \int_0^{2\pi} 0\,dt = 0.$$

Since the line integral around the closed curve is zero, we conclude that $\vec{F}$ is conservative.

To find the potential function $\phi(x, y)$, note that $\nabla\phi = \vec{F} = \langle x, y\rangle$.

Moving on,

$$\phi(x, y) = \int x\,dx = \frac{x^2}{2} + c(y).$$

$$\frac{\partial\phi}{\partial y} = c'(y) = y \Rightarrow c(y) = \frac{y^2}{2}.$$

So the potential function is

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

9.4 Green's Theorem

Green's theorem is a generalization of the fundamental theorem of calculus in two dimensions. It gives the relationship between a line integral around a closed curve and a double integral over the plane region bounded by the curve.

If $\vec{F}(x,y) = \langle P(x,y), Q(x,y)\rangle$ is a vector field defined on an open region that contains a positively oriented, piecewise-smooth, simple closed curve C, and if D is the region bounded by C, then

$$\oint_C \vec{F}\cdot d\vec{r} = \oint_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

We say that the curve C is positively oriented if it is traversed counterclockwise, so that the region D is always on the left as you move along C.

This theorem allows us to convert a difficult line integral into a potentially easier double integral over a region. The right-hand side measures the total "microscopic rotation" (curl) inside the region. Thus, the total circulation along the boundary, or the sum of curls inside is given by

$$\oint_C \vec{F}\cdot d\vec{r}.$$

As you can kind of see, there is a resemblance between the fundamental theorem of calculus,

$$\int_a^b f'(x)\,dx = f(b) - f(a),$$

and Green's theorem where we integrate derivatives across a two-dimensional region to recover a obtain a value over its boundary.

To prove Green's theorem for simple regions, we consider two key identities. Let D be a simple region:

$$D = \{(x,y) \mid a \le x \le b,\ g_1(x) \le y \le g_2(x)\}$$

Then

$$\iint_D \frac{\partial P}{\partial y}\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} \frac{\partial P}{\partial y}(x,y)\,dy\,dx = \int_a^b \left[P(x,g_2(x)) - P(x,g_1(x))\right]\,dx$$

which follows from the fundamental theorem of calculus.

To compute the left-hand side of

$$\oint_C P(x,y)\,dx,$$

we break the boundary C into four parts C_1, C_2, C_3 and C_4, which trace the edges of region D as shown:

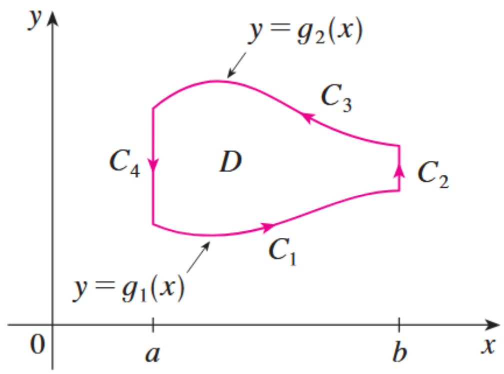


Image credit: Stewart

We now compute each segment. On C_1, where $y = g_1(x)$ and $x \in [a, b]$, we parametrize with x and get

$$\int_{C_1} P(x, y)\, dx = \int_a^b P(x, g_1(x))\, dx.$$

On C_3, the top boundary, the curve is traversed right to left, so its reverse $-C_3$ runs from left to right with $y = g_2(x)$. Therefore

$$\int_{C_3} P(x, y)\, dx = -\int_{-C_3} P(x, y)\, dx = -\int_a^b P(x, g_2(x))\, dx.$$

On C_2 and C_4, x is constant, so $dx = 0$, and thus

$$\int_{C_2} P(x, y)\, dx = \int_{C_4} P(x, y)\, dx = 0.$$

Adding all segments, we get

$$\begin{aligned}\oint_C P(x, y)\, dx &= \int_{C_1} P(x, y)\, dx + \int_{C_2} P(x, y)\, dx + \int_{C_3} P(x, y)\, dx + \int_{C_4} P(x, y)\, dx, \\ &= \int_a^b P(x, g_1(x))\, dx - \int_a^b P(x, g_2(x))\, dx = -\int_a^b [P(x, g_2(x)) - P(x, g_1(x))]\, dx.\end{aligned}$$

Comparing this with our earlier result,

$$\iint_D \frac{\partial P}{\partial y}\,dA = \int_a^b \left[P(x,g_2(x)) - P(x,g_1(x))\right] dx,$$

we conclude

$$\oint_C P(x,y)\,dx = -\iint_D \frac{\partial P}{\partial y}\,dA.$$

Similarly, expressing D as a type II region:

$$D = \{(x,y) \mid c \le y \le d,\ h_1(y) \le x \le h_2(y)\},$$

we could have proved

$$\iint_D \frac{\partial Q}{\partial x}\,dA = -\int_C Q\,dy.$$

Adding these two results yields Green's theorem:

$$\oint_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$$

EXAMPLE 9.12

Evaluate the line integral

$$\oint_C (3y - e^{\sin x})\, dx + \left(7x + \sqrt{y^4} + 1\right)\, dy,$$

where C is the circle $x^2 + y^2 = 9$, oriented counterclockwise.

Solution:

The region D bounded by C is the disk $x^2+y^2 \leq 9$. Since the boundary is closed and simple, we can apply Green's theorem:

$$\oint_C P\, dx + Q\, dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA,$$

where $P(x, y) = 3y - e^{\sin x}$ and $Q(x, y) = 7x + \sqrt{y^4} + 1 = 7x + y^2 + 1$.
Compute the partial derivatives:

$$\frac{\partial Q}{\partial x} = 7, \ \frac{\partial P}{\partial y} = 3.$$

Apply Green's theorem:

$$\oint_C (3y - e^{\sin x})\, dx + \left(7x + \sqrt{y^4} + 1\right)\, dy = \iint_D (7 - 3)\, dA = \iint_D 4\, dA.$$

Switch to polar coordinates:

$$\iint_D 4\, dA = \int_0^{2\pi} \int_0^3 4r\, dr\, d\theta = 4 \int_0^{2\pi} d\theta \int_0^3 r\, dr.$$
$$= 4(2\pi) \left[\frac{1}{2}r^2\right]_0^3 = 4(2\pi) \cdot \frac{9}{2} = 36\pi$$

Green's theorem can also be used in reverse to simplify calculations.

Suppose $P(x, y) = Q(x, y) = 0$ on the boundary curve C. Then by Green's theorem,

$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA = \oint_C P\,dx + Q\,dy = 0,$$

regardless of the values of P and Q in the interior of D.

Another important application of Green's Theorem is computing the area A of a region D. Since

$$A(D) = \iint_D 1\,dA,$$

we choose functions $P(x, y)$, $Q(x, y)$ such that

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 1.$$

There are many valid choices:

$$\begin{aligned} P(x,y) &= 0, & Q(x,y) &= x \\ P(x,y) &= -y, & Q(x,y) &= 0 \\ P(x,y) &= -\tfrac{1}{2}y, & Q(x,y) &= \tfrac{1}{2}x \end{aligned}$$

Applying Green's theorem with each of these pairs gives equivalent formulas for the area:

$$A = \oint_C x\,dy = -\oint_C y\,dx = \frac{1}{2}\oint_C x\,dy - y\,dx.$$

So far, we've worked with simply-connected regions. Well, what about connected regions that are not simply connected or regions that are not connected?

Finite Union of Simple Regions: Suppose a region D is the union of two simple regions D_1 and D_2 whose interiors do not overlap. Then Green's theorem holds over the whole region:

$$\oint_{C_1 \cup C_2} P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA,$$

where the boundary $C_1 \cup C_2$ is the positively oriented outer curve of $D_1 \cup D_2$. Internal boundaries cancel due to opposite orientation.

Region with a Hole: Let D be a region bounded by two closed curves: the outer curve C_1 and an inner hole C_2. If C_1 is oriented counterclockwise and C_2 is oriented clockwise, then

$$\oint_{C_1} P\,dx + Q\,dy + \oint_{C_2} P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

Or equivalently,

$$\oint_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA$$

where $C = C_1 \cup C_2$ and both are positively oriented with respect to the region. Thus the position direction is counterclockwise for the outer curve C_1 but clockwise for the inner curve C_2.

These extensions follow from applying Green's theorem to each simple piece and using the cancellation of integrals along shared internal boundaries.

Note that are two natural interpretations of Green's theorem, depending on the type of vector flow being measured: circulation and flux.

Circulation measures the tangential component of a vector field along a closed curve. It represents how much the field "swirls" or "spins" around the boundary. Mathematically, it is the line integral of the field projected onto the unit tangent vector $\mathbf{T}$. Circulation around C is given by

$$\oint_C \vec{F} \cdot \mathbf{T}\,ds = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

Flux measures the normal component of a vector field across a closed curve. It represents how much the field comes through the boundary. Mathematically, it is the line integral of the field projected onto the outward unit normal vector $\mathbf{N}$. Flux across C is given by

$$\oint_C \vec{F} \cdot \mathbf{N}\,ds = \iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dA.$$

We will discuss this further later on.

EXAMPLE 9.13

Let C be a circle of radius r centered at the origin, and let $\vec{F}(x,y) = \langle x, y \rangle$. Compute the total flux of $\vec{F}$ across C.

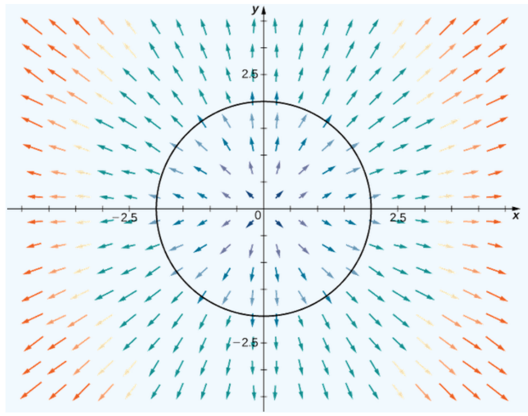


Image credit: Strang & Herman

Solution:

Let D be the disk enclosed by C. The flux across C is given by $\oint_C \vec{F} \cdot \mathbf{N}\, ds$ where $\mathbf{N}$ is the outward-pointing unit normal vector. Rather than computing this line integral directly, we apply Green's theorem in its flux form, which converts it into a double integral over the interior of the region:

$$\oint_C \vec{F} \cdot \mathbf{n}\, ds = \iint_D \nabla \cdot \vec{F}\, dA$$

Given $\vec{F}(x,y) = \langle x, y \rangle$, we compute the divergence:

$$\nabla \cdot \vec{F} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} = 1 + 1 = 2$$

This tells us that the field is uniformly expanding outward at every point. And by Green's theorem,

$$\oint_C \vec{F} \cdot \mathbf{N}\, ds = \iint_D 2\, dA = 2 \iint_D dA.$$

Since $\iint_D dA$ is the the area of the circle, it is equivalent to πr^2. Thus, we we conclude

$$\oint_C \vec{F} \cdot \mathbf{N}\, ds = 2\pi r^2.$$

EXAMPLE 9.14

Compute the work done by the force field $\vec{F}(x,y) = \langle x(x+y),\ xy^2\rangle$ that moves a particle from the origin along the x-axis to $(1,0)$, then to $(0,1)$, and returns to the origin along the y-axis.

Solution:

This represents a triangular path. Let C be the closed triangle traversed counterclockwise and D the region it encloses. The work is given by the line integral

$$W = \oint_C \vec{F} \cdot d\vec{r} = \oint_C x(x+y)\,dx + xy^2\,dy.$$

We will apply Green's theorem in circulation form:

$$W = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA.$$

Compute the partial derivatives for $P(x,y) = x(x+y) = x^2 + xy$ and $Q(x,y) = xy^2$:

$$\frac{\partial Q}{\partial x} = y^2, \quad \frac{\partial P}{\partial y} = x$$

We can now get the integrand, $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = y^2 - x$. Now integrate over the triangular region bounded by C, which lies below the line $y = 1 - x$:

$$W = \int_0^1 \int_0^{1-x} (y^2 - x)\,dy\,dx.$$

Evaluate the inner integral:

$$= \int_0^1 \left[\frac{1}{3}y^3 - xy\right]_{y=0}^{y=1-x} dx = \int_0^1 \left(\frac{1}{3}(1-x)^3 - x(1-x)\right) dx.$$

Simplify and integrate:

$$= \int_0^1 \left(\frac{1}{3}(1 - 3x + 3x^2 - x^3) - x + x^2\right) dx$$

$$= \int_0^1 \left(\frac{1}{3} - x + x^2 - \frac{1}{3}x^3\right) dx = \left[\frac{1}{3}x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{12}x^4\right]_0^1 = -\frac{1}{12}$$

And now for a more conceptual summary. Green's theorem tells us that the total circulation of a vector field around a closed curve C on the macroscopic level is the sum of all the tiny, local circulations inside the region D it encloses.

Each point inside the region contributes its own "microscopic circulation," which is measured by the scalar curl:

$$\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Green's theorem says

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA.$$

So instead of measuring rotation by walking around the entire boundary, we can simply sum up how much the field is spinning at each point inside. Mathematically, this lets us convert a line integral around a closed curve into a double integral over the region it encloses.

9.5 Curl and Divergence

Curl and **divergence** are two operators that can be performed on vector fields to reveal information about the structure of the field. They both are generally similar in that they represent differentiation, but there are differences both geometrically and mathematically. The easiest to remember is that curl produces a vector field, whereas divergence, produces a scalar field. We will begin with curl.

Let $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ be a vector field in $\mathbb{R}^3$, with continuous and differentiable components.

Then the **curl** of $\vec{F}$ is a vector field:

$$\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathbf{k} = \text{curl } \vec{F} = \nabla \times \vec{F}.$$

To remember this, you can use the symbolic determinant form:

$$\nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P & Q & R \end{vmatrix}$$

If f is a scalar function with continuous second-order partial derivatives, then

$$\operatorname{curl}(\nabla f) = \vec{0}.$$

Let's prove this. Using the determinant form,

$$\operatorname{curl}(\nabla f) = \nabla \times (\nabla f) = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{vmatrix}.$$

Each component becomes a difference of mixed partial derivatives:

$$= \left(\frac{\partial^2 f}{\partial y \partial z} - \frac{\partial^2 f}{\partial z \partial y} \right) \mathbf{i} + \left(\frac{\partial^2 f}{\partial z \partial x} - \frac{\partial^2 f}{\partial x \partial z} \right) \mathbf{j} + \left(\frac{\partial^2 f}{\partial x \partial y} - \frac{\partial^2 f}{\partial y \partial x} \right) \mathbf{k} = \vec{0}$$

by Clairaut's theorem. Thus $\operatorname{curl}(\nabla f) = \vec{0}$.

This tells us that if a vector field is conservative ($\vec{F} = \nabla f$), then

$$\operatorname{curl} \vec{F} = \vec{0}.$$

This provides a quick test. If

$$\operatorname{curl} \vec{F} \neq \vec{0},$$

then $\vec{F}$ is not conservative.

EXAMPLE 9.15

Let $\vec{F}(x,y,z) = xz\,\mathbf{i} + xyz\,\mathbf{j} - y^2\,\mathbf{k}$. Compute $\nabla \times \vec{F}$. Then determine whether or not it is conservative.

Solution:

We apply the determinant form of the curl:

$$\nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ xz & xyz & -y^2 \end{vmatrix}$$

We compute each component:

For $\mathbf{i}$, we have

$$\frac{\partial}{\partial y}(-y^2) - \frac{\partial}{\partial z}(xyz) = (-2y) - (xy) = -2y - xy.$$

For $\mathbf{j}$, we have

$$-\left(\frac{\partial}{\partial x}(-y^2) - \frac{\partial}{\partial z}(xz)\right) = -(0 - x) = x.$$

For $\mathbf{k}$, we have

$$\frac{\partial}{\partial x}(xyz) - \frac{\partial}{\partial y}(xz) = yz - 0 = yz.$$

And then,

$$\nabla \times \vec{F} = (-2y - xy)\,\mathbf{i} + x\,\mathbf{j} + yz\,\mathbf{k} = -y(2+x)\,\mathbf{i} + x\,\mathbf{j} + yz\,\mathbf{k}.$$

Since curl $\vec{F} \neq \vec{0}$, $\vec{F}$ is not conservative.

EXAMPLE 9.16

(a) Show that the vector field $\vec{F}(x,y,z) = y^2z^3\,\mathbf{i} + 2xyz^3\,\mathbf{j} + 3xy^2z^2\,\mathbf{k}$ is conservative.
(b) Find a scalar potential function $f(x,y,z)$ such that $\vec{F} = \nabla f$.

Solution:

We compute the curl of $\vec{F}$ using the determinant form:

$$\text{curl } \vec{F} = \nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ y^2z^3 & 2xyz^3 & 3xy^2z^2 \end{vmatrix}$$

We compute component **i**:

$$\frac{\partial}{\partial y}(3xy^2z^2) - \frac{\partial}{\partial z}(2xyz^3) = 6xyz^2 - 6xyz^2 = 0$$

We compute component **j**:

$$-\left(\frac{\partial}{\partial x}(3xy^2z^2) - \frac{\partial}{\partial z}(y^2z^3)\right) = -\left(3y^2z^2 - 3y^2z^2\right) = 0$$

We compute component **k**:

$$\frac{\partial}{\partial x}(2xyz^3) - \frac{\partial}{\partial y}(y^2z^3) = 2yz^3 - 2yz^3 = 0$$

Combining each components yields $\nabla \times \vec{F} = \vec{0}$. Therefore, $\vec{F}$ is a conservative vector field.

EXAMPLE 9.16 (CONTINUED)

(b)
Start with the x-component:

$$\frac{\partial f}{\partial x} = y^2 z^3 \Rightarrow f(x,y,z) = xy^2z^3 + g(y,z),$$

where $g(y,z)$ is an unknown function of y and z.
Differentiate with respect to y:

$$\frac{\partial f}{\partial y} = 2xyz^3 + \frac{\partial g}{\partial y}$$

This must match the given y-component of $\vec{F}$, which is $2xyz^3$. Therefore

$$\frac{\partial g}{\partial y} = 0 \Rightarrow g(y,z) = h(z).$$

Now differentiate with respect to z:

$$\frac{\partial f}{\partial z} = 3xy^2z^2 + \frac{dh}{dz}$$

This must match the given z-component of $\vec{F}$, which is $3xy^2z^2$. Thus

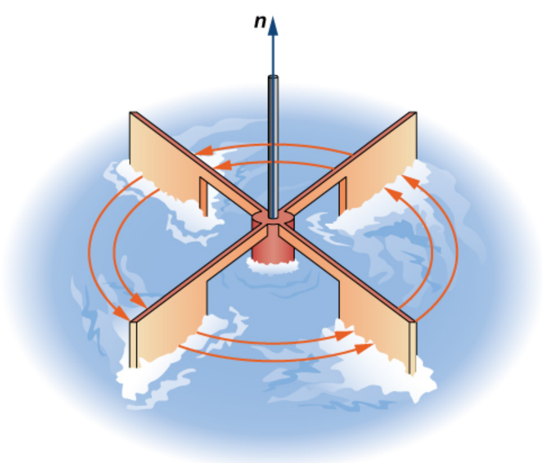
$$\frac{dh}{dz} = 0 \Rightarrow h(z) = C.$$

Therefore, the potential function is

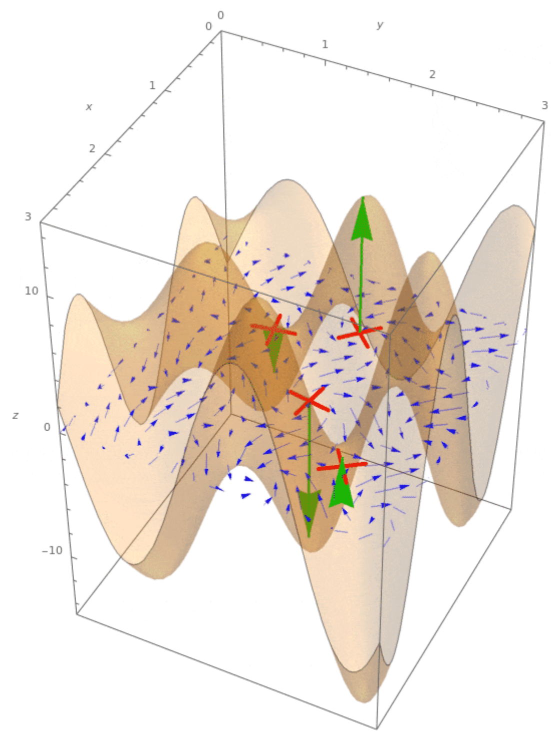
$$f(x,y,z) = xy^2z^3 + C.$$

The curl of a vector field measures its local rotational tendency.

Imagine placing a miniature paddlewheel in a fluid whose velocity field is described by $\vec{F}$. The wheel is free to spin but not to translate. If the fluid tends to swirl around the paddlewheel, it begins to rotate. The axis the wheel spins around points in the direction of curl $\vec{F}$, and the speed of rotation corresponds to the magnitude of the curl vector. If the paddle doesn't spin at all, then curl $\vec{F} = \vec{0}$, and the field is said to be irrotational at that point. This paddlewheel is just one of many in a field. Curl is a measure of how much a field causes each paddlewheel to spin at each point.

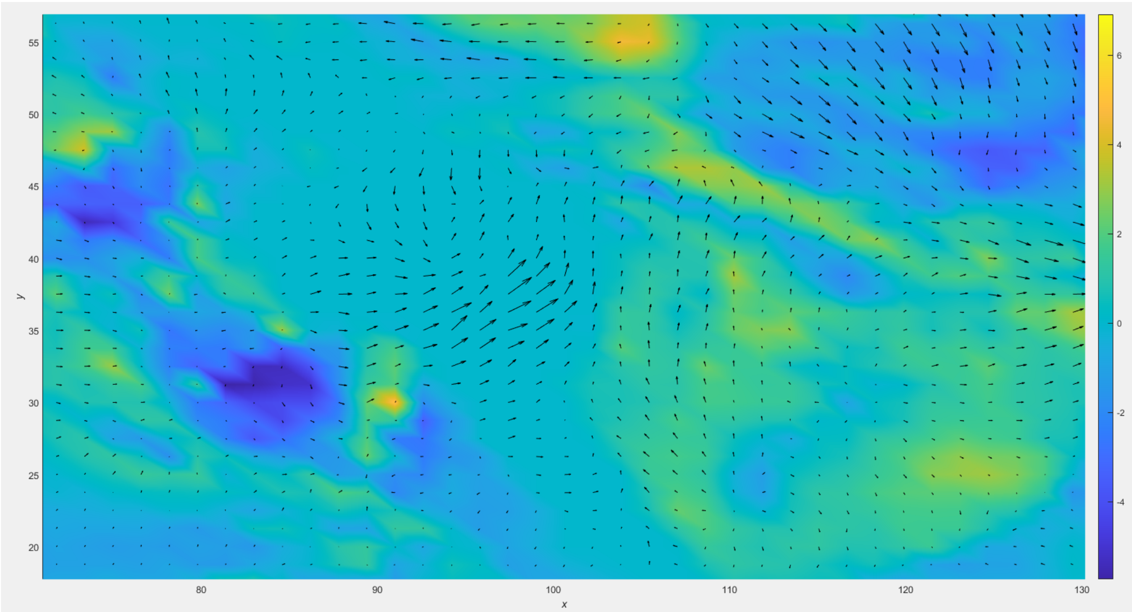


A single paddlewheel at one point.
Image credit: Strang & Herman

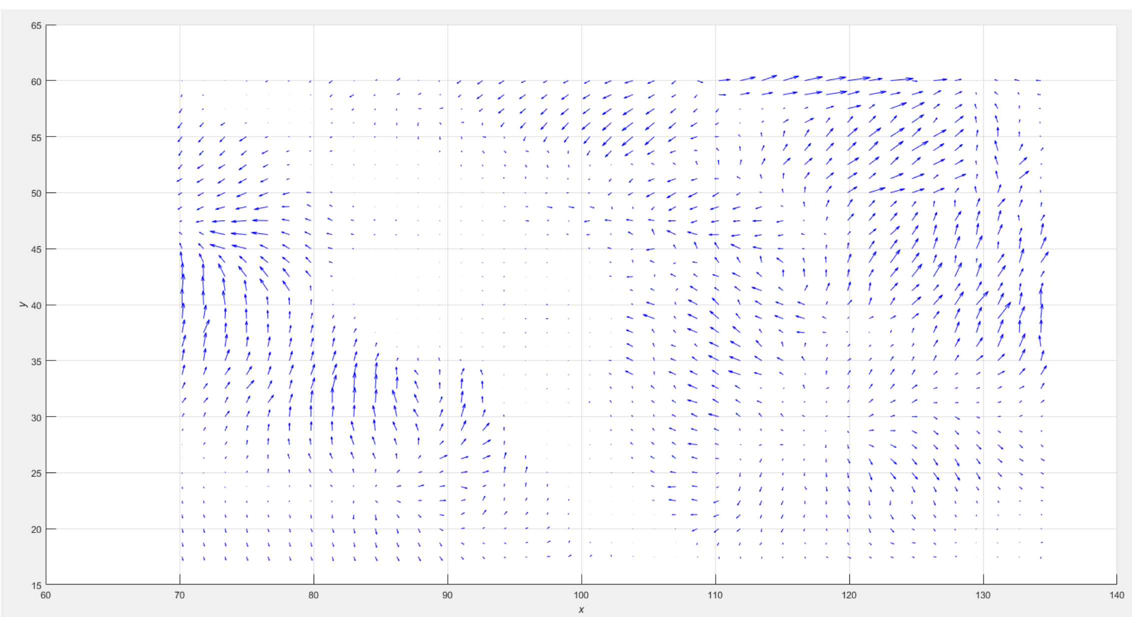


More of the paddlewheels, each of which can behave differently.
Image credit: UMich

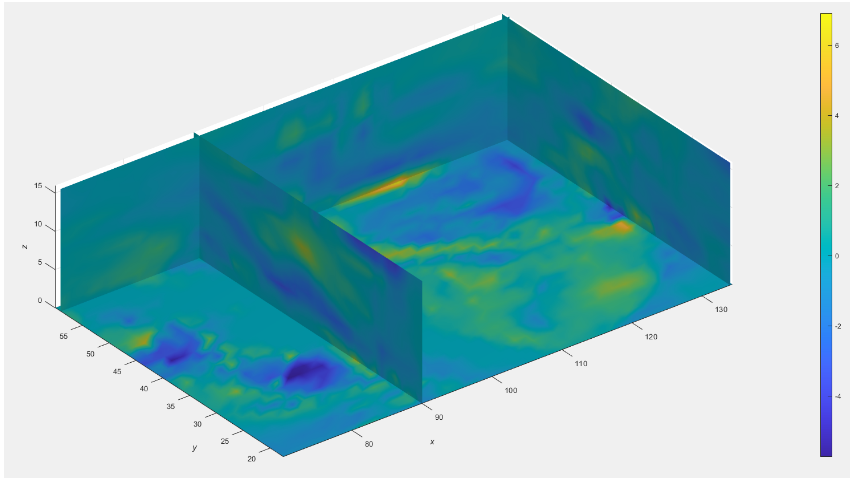
In two dimensions, the curl reduces to a scalar pointing in the **k**-direction, representing the net "spin" about a point on a surface. In three dimensions, the curl is a full vector field indicating the axis and strength of rotation.



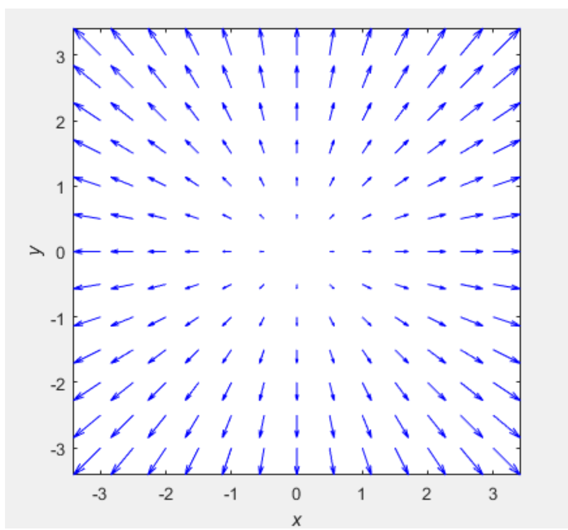
A 2D slice of angular velocity in a wind vector field. The color map represents the scalar curl. This plot tells you how much spin is happening at each point, but nothing about the direction.



A wind vector plot of the same curl field but in 3D and projected onto an xy-plane slice. This plot tells you the direction in which the fluid is twisting.



Visualization of curl in a 3D wind velocity field. Colors represent angular velocity, or spin, at each point in the fluid. More specifically, warmer colors (yellow) indicate counterclockwise rotation and cooler colors (blue) indicate clockwise rotation.



This vector field has 0 curl. Visually, you can see that the vectors just point outward. In MATLAB, this returns a matrix of 0s.

zeroCurlExample.m

EXAMPLE 9.17

Let $\vec{r}(x, y) = \langle \cos(x+y), \sin(x-y) \rangle$. Find the maximum magnitude of the curl in the region $0 \leq x \leq 2$, $0 \leq y \leq 2$.

Solution:

We compute the curl in two dimensions:

$$\nabla \times \vec{r} = \left\langle 0, 0, \frac{\partial}{\partial x}(\sin(x-y)) - \frac{\partial}{\partial y}(\cos(x+y)) \right\rangle = \langle 0, 0, \cos(x-y) + \sin(x+y) \rangle$$

Since the curl points entirely in the z-direction, its magnitude is the absolute value of $f(x, y) = \cos(x-y) + \sin(x+y)$.
To find local extrema of f, we compute the gradient:

$$\nabla f(x, y) = \langle -\sin(x-y) + \cos(x+y), \sin(x-y) + \cos(x+y) \rangle .$$

We want both components to be zero:

$$-\sin(x-y) + \cos(x+y) = 0 \sin(x-y) + \cos(x+y) = 0$$

Adding the equations gives

$$2\cos(x+y) = 0 \Rightarrow x + y = \frac{\pi}{2} + j\pi.$$

Subtracting them gives

$$2\sin(x-y) = 0 \Rightarrow x - y = k\pi.$$

So the only solution in our domain is $x = y = \frac{\pi}{4}$. At this point,

$$\text{curl} = \cos(0) + \sin\left(\frac{\pi}{2}\right) = 1 + 1 = 2.$$

Since $\cos(x-y) + \sin(x+y) \leq 2$, this is the maximum possible value. Therefore, the maximum magnitude of the curl is 2 at the point $\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$.

EXAMPLE 9.17 (CONTINUED)

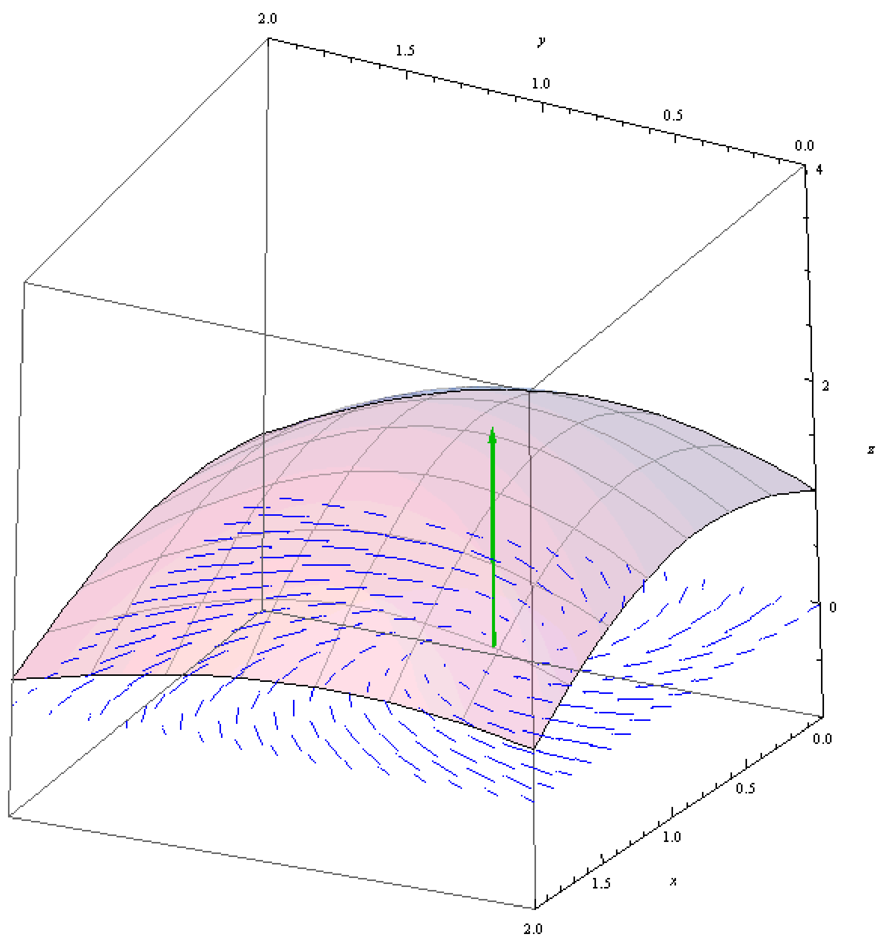


Image credit: UMich

The green arrow represents the curl vector at $(\frac{\pi}{4}, \frac{\pi}{4})$. The surrounding vector field visually resembles a whirlpool centered at this point. Intuitively, this suggests the rotation of the fluid is greatest at the center of the whirlpool—precisely matching the maximum curl we computed.

EXAMPLE 9.18

The gravitational field due to an object of mass m_1 at the origin, acting on a particle of mass m_2 at point (x, y, z), is

$$\vec{F}(x,y,z) = -Gm_1m_2 \left\langle \frac{x}{(x^2+y^2+z^2)^{3/2}}, \frac{y}{(x^2+y^2+z^2)^{3/2}}, \frac{z}{(x^2+y^2+z^2)^{3/2}} \right\rangle.$$

Show that the gravitational field has zero curl.

Solution:

Let

$$\begin{aligned} P(x,y,z) &= \frac{x}{(x^2+y^2+z^2)^{3/2}}, \\ Q(x,y,z) &= \frac{y}{(x^2+y^2+z^2)^{3/2}}, \text{ and} \\ R(x,y,z) &= \frac{z}{(x^2+y^2+z^2)^{3/2}}. \end{aligned}$$

Then, curl is given by

$$\nabla \times \vec{F} = -Gm_1m_2 \left[(R_y - Q_z)\,\mathbf{i} + (P_z - R_x)\,\mathbf{j} + (Q_x - P_y)\,\mathbf{k} \right].$$

Compute each partial derivative:

$$\begin{aligned} R_y &= \frac{-3yz}{(x^2+y^2+z^2)^{5/2}}, \; Q_z = \frac{-3yz}{(x^2+y^2+z^2)^{5/2}} \;\Rightarrow R_y - Q_z = 0 \\ P_z &= \frac{-3xz}{(x^2+y^2+z^2)^{5/2}}, \; R_x = \frac{-3xz}{(x^2+y^2+z^2)^{5/2}} \;\Rightarrow P_z - R_x = 0 \\ Q_x &= \frac{-3xy}{(x^2+y^2+z^2)^{5/2}}, \; P_y = \frac{-3xy}{(x^2+y^2+z^2)^{5/2}} \;\Rightarrow Q_x - P_y = 0 \end{aligned}$$

Thus, $\nabla \times \vec{F} = \vec{0}$. This makes sense because the force of gravity does not induce whirlpool-like behavior. It only pulls radially inward.

We now move on to **divergence**.

If $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ is a vector field on $\mathbb{R}^3$ and each component's partial derivative exists, then the divergence of $\vec{F}$ is the function of three variables defined by

$$\operatorname{div}\vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

where $\operatorname{div}\vec{F}$ is a scalar field.

In terms of the gradient operator

$$\nabla = \left(\frac{\partial}{\partial x}\right)\mathbf{i} + \left(\frac{\partial}{\partial y}\right)\mathbf{j} + \left(\frac{\partial}{\partial z}\right)\mathbf{k},$$

the divergence of $\vec{F}$ can be written symbolically as the dot product of ∇ and $\vec{F}$:

$$\operatorname{div}\vec{F} = \nabla \cdot \vec{F}$$

If $\vec{F}$ is a vector field on $\mathbb{R}^3$, then $\operatorname{curl}\vec{F}$ is also a vector field on $\mathbb{R}^3$. As such, we can compute its divergence. The next theorem shows that the result is 0:

Let $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ be a vector field on $\mathbb{R}^3$. If P, Q, and R have continuous second-order partial derivatives, then $\operatorname{curl}\vec{F} = 0$.

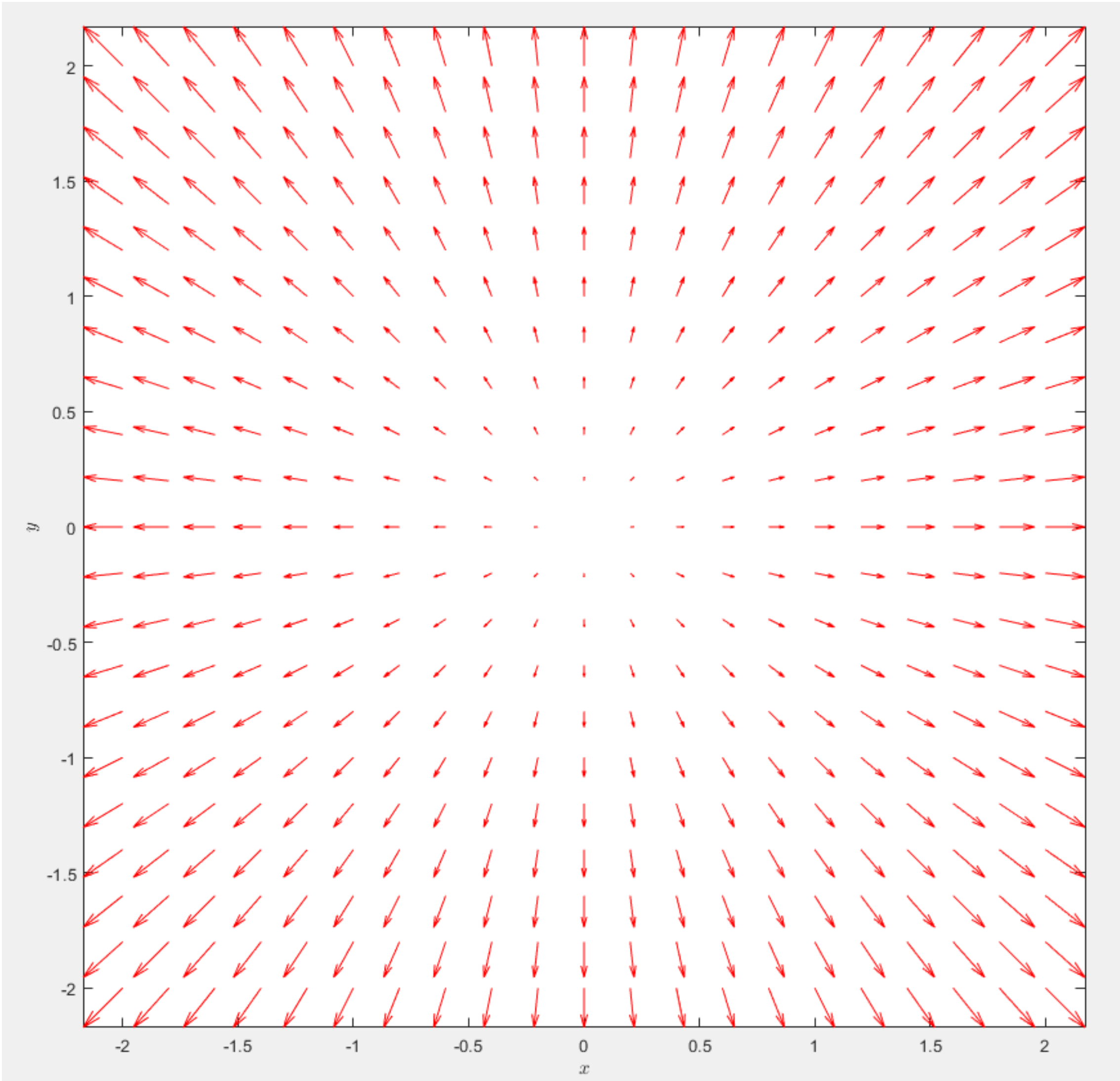
Using the definitions of divergence and curl, we have

$$\operatorname{div}(\operatorname{curl}\vec{F}) = \nabla \cdot (\nabla \times \vec{F}) = \frac{\partial}{\partial x}\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) + \frac{\partial}{\partial y}\left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) + \frac{\partial}{\partial z}\left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)$$

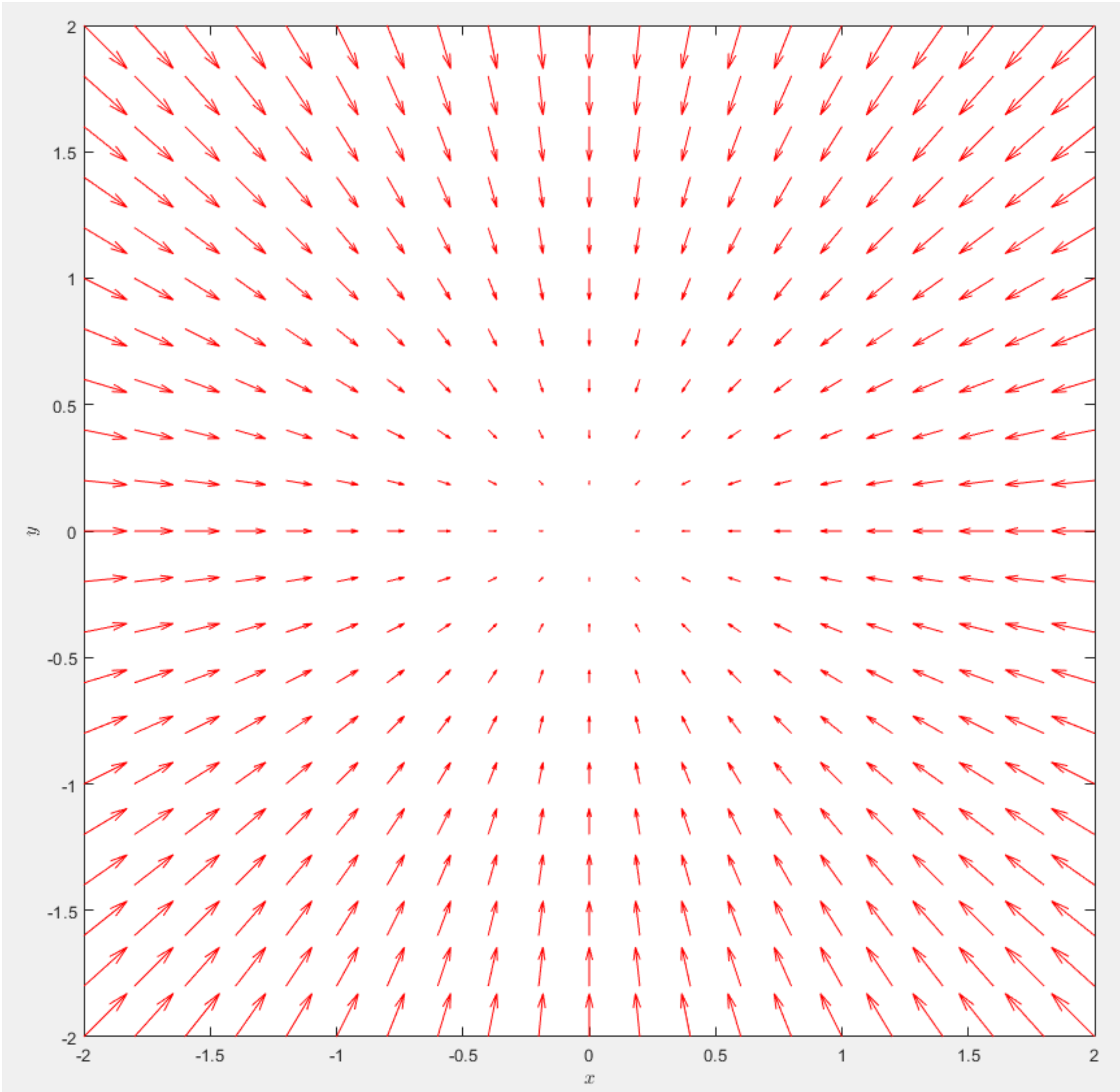
$$= \frac{\partial^2 R}{\partial x \partial y} - \frac{\partial^2 Q}{\partial x \partial z} + \frac{\partial^2 P}{\partial y \partial z} - \frac{\partial^2 R}{\partial y \partial x} + \frac{\partial^2 Q}{\partial z \partial x} - \frac{\partial^2 P}{\partial z \partial y} = 0$$

because the terms cancel in pairs by Clairaut's theorem. This really is a quite elegant identity.

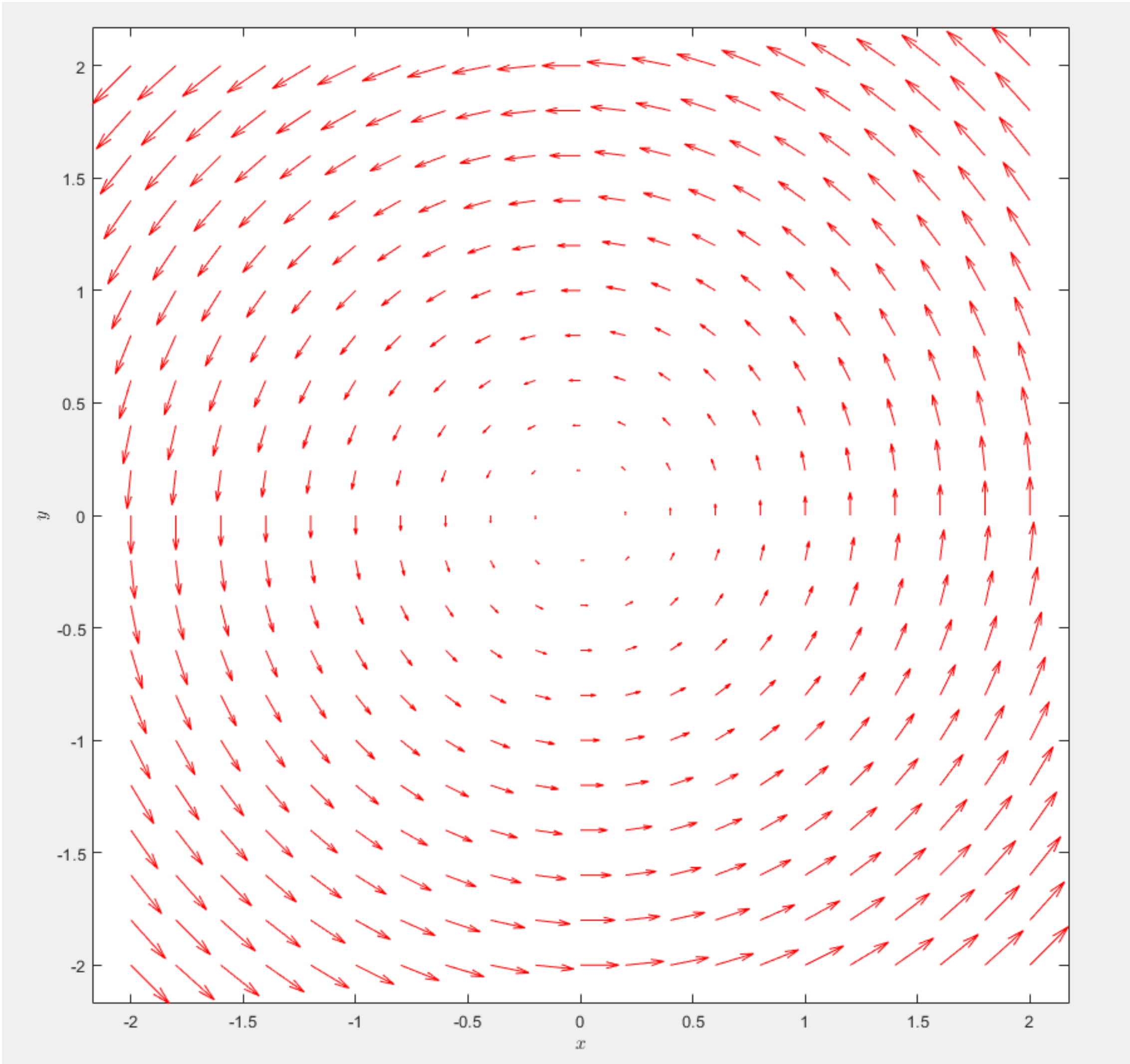
Let's visually interpret divergence through vector fields first.



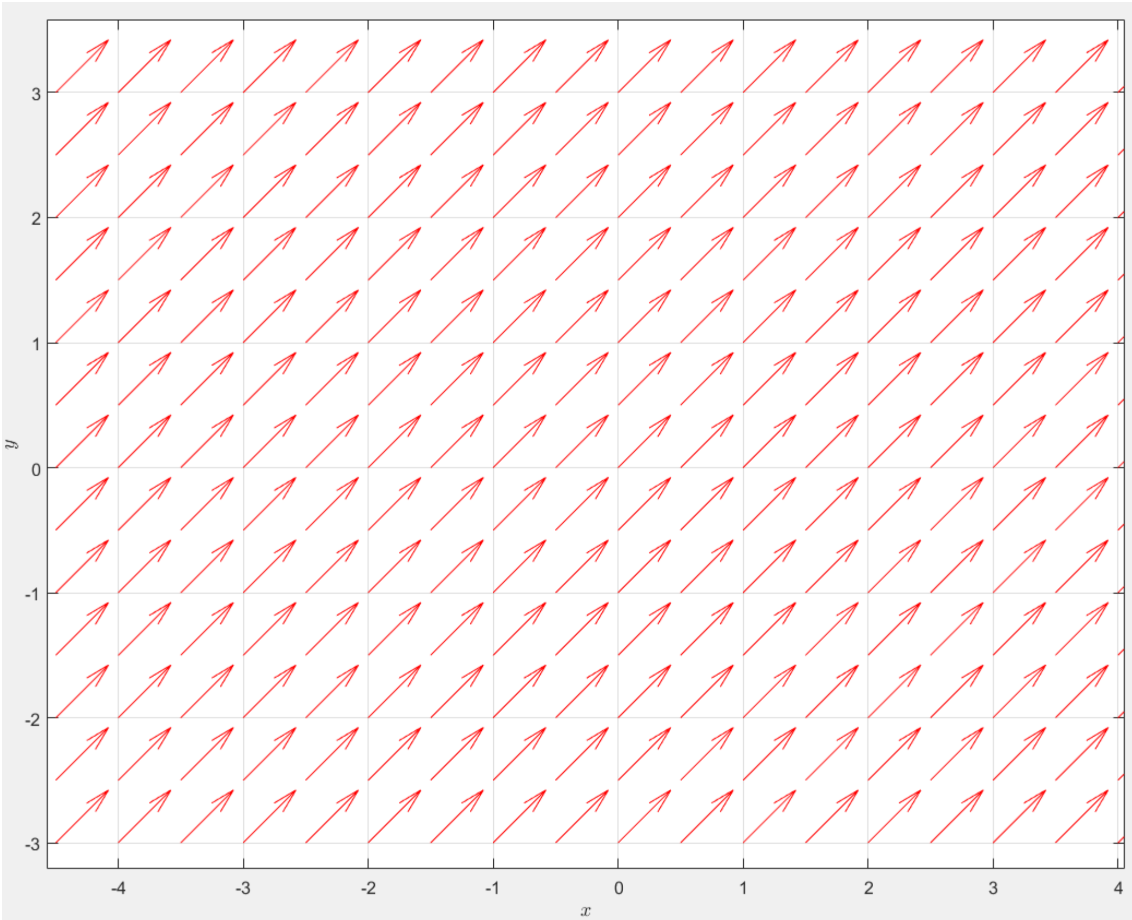
(a) $\nabla \cdot \vec{F} > 0$



(b) $\nabla \cdot \vec{F} < 0$



(c) $\nabla \cdot \vec{F} = 0$



(d) $\nabla \cdot \vec{F} = 0$

In vector field (a), the vectors all point outward. In vector field (b), the vectors all point inward. In vector field (c), the vectors are just moving around. In vector field (d), the vectors are all pointing in the same direction and that direction isn't inward or outward. There's no change in the net flow inward or outward.

You can think of each point on a vector field as a sink. Then, the vectors represent the speed and direction of the fluid. Divergence is a measure of the tendency of a fluid to flow in or out of the sink. Positive divergence means that more fluid is leaving the sink than not. Negative divergence means that more fluid is entering the sink than not. Zero divergence means that there's no net flow inward or outward. Zero divergence is known as *incompressible*

Suppose $f(x, y, z)$ is a scalar function. Then the gradient is

$$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right\rangle.$$

Taking the divergence of this gradient gives

$$\operatorname{div}(\nabla f) = \nabla \cdot (\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

This is called the Laplace operator, and we denote it by

$$\nabla^2 f = \nabla \cdot \nabla f.$$

It appears frequently in physics, particularly in Laplace's equation:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

We can also apply the Laplace operator component-wise to a vector field. For a vector field $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$, we have

$$\nabla^2 \vec{F} = \nabla^2 P\,\mathbf{i} + \nabla^2 Q\,\mathbf{j} + \nabla^2 R\,\mathbf{k}.$$

We now reinterpret Green's theorem using the operators of curl and divergence, allowing us to write it more compactly and geometrically. Suppose $\vec{F} = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j}$ is a vector field on a plane region $D \subseteq \mathbb{R}^2$, with boundary curve C oriented counterclockwise.

Green's theorem in its standard form is

$$\oint_C \vec{F} \cdot d\vec{r} = \oint_C P\,dx + Q\,dy = \iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA.$$

We now view $\vec{F}$ as a vector field in $\mathbb{R}^3$ with third component 0. Then the curl of $\vec{F}$ becomes

$$\operatorname{curl} \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P(x,y) & Q(x,y) & 0 \end{vmatrix} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \mathbf{k}.$$

Taking the dot product with $\mathbf{k}$, we have

$$(\operatorname{curl}\vec{F})\cdot\mathbf{k} = \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k}\cdot\mathbf{k} = \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}.$$

Thus, Green's theorem becomes the following elegant vector identity

$$\oint_C \vec{F}\cdot d\vec{r} = \iint_D (\operatorname{curl}\vec{F})\cdot\mathbf{k}\, dA.$$

This states that the line integral of the tangential component of $\vec{F}$ around C equals the double integral of the vertical component of the curl of $\vec{F}$ over the region D.

We can also write a second vector form by focusing on the flux of $\vec{F}$ through the boundary C.

Let the curve C be parametrized by

$$\vec{r}(t) = x(t)\,\mathbf{i} + y(t)\,\mathbf{j},\ \ a \le t \le b.$$

Then the unit tangent vector is

$$\mathbf{T}(t) = \frac{x'(t)}{\|\vec{r}\,'(t)\|}\,\mathbf{i} + \frac{y'(t)}{\|\vec{r}\,'(t)\|}\,\mathbf{j}.$$

Rotating this counterclockwise by 90°, we get the outward unit normal vector:

$$\mathbf{n}(t) = \frac{y'(t)}{\|\vec{r}\,'(t)\|}\,\mathbf{i} - \frac{x'(t)}{\|\vec{r}\,'(t)\|}\,\mathbf{j}$$

Using this, we compute the flux of $\vec{F}$ across C:

$$\oint_C \vec{F}\cdot\mathbf{n}\,ds = \int_a^b \vec{F}(\vec{r}(t))\cdot\mathbf{n}(t)\|\vec{r}\,'(t)\|\,dt$$

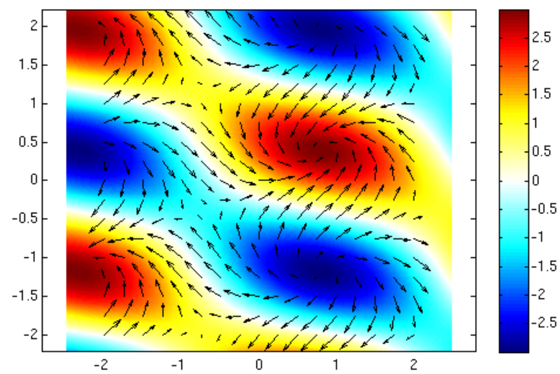
After simplifying, we find

$$\oint_C \vec{F}\cdot\mathbf{n}\,ds = \oint_C P\,dy - Q\,dx = \iint_D \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}\right) dA.$$

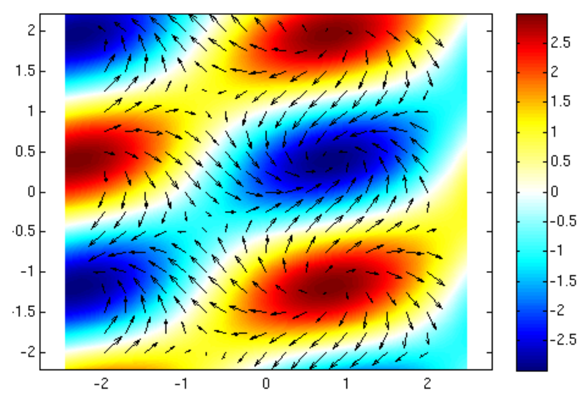
Since $\nabla\cdot\vec{F} = \operatorname{div}\vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y}$, this gives the second vector form of Green's theorem:

$$\oint_C \vec{F} \cdot \mathbf{n}\, ds = \iint_D \operatorname{div} \vec{F}(x, y)\ dA$$

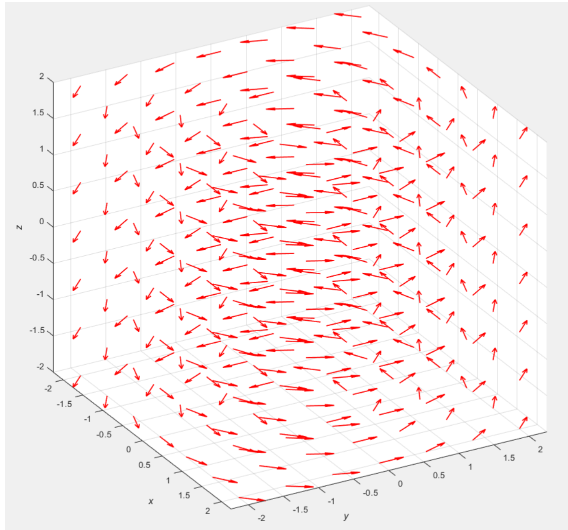
This version says that the total outward flux of $\vec{F}$ across the closed curve C, measured by the line integral of its normal component, is equal to the total divergence of $\vec{F}$ inside the region D enclosed by C which is given by the double integral.



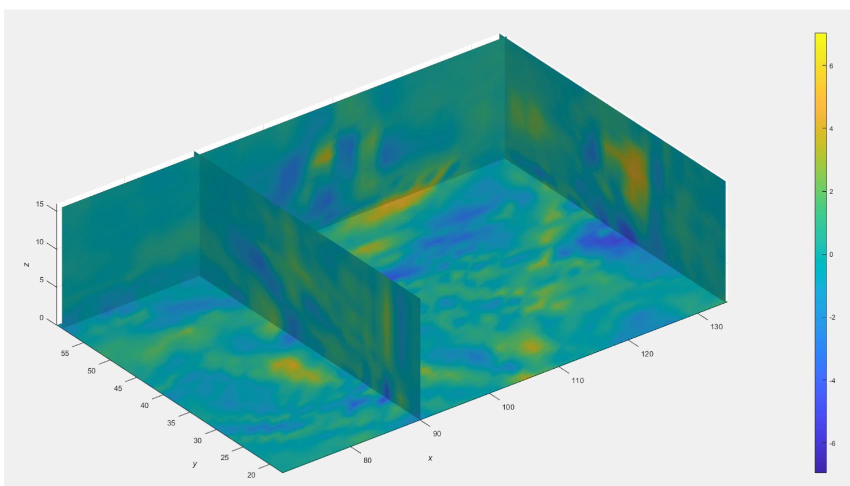
Visualization of curl. Arrows represent the vector field, while color represents the value of $\nabla \times \vec{F}$. Red indicates positive (counterclockwise) rotation; blue indicates negative (clockwise) rotation. Image credit: Von Petersdorff, UMD



Visualization of divergence. Arrows represent the vector field, and color shows $\nabla \cdot \vec{F}$. Red regions behave like sources (positive divergence), while blue regions behave like sinks (negative divergence). Image credit: Von Petersdorff, UMD



A vector field with 0 divergence in 3D



Visualization of divergence in a 3D wind velocity field. Colors represent net flow at each point in the fluid. More specifically, warmer colors (yellow) indicate positive divergence and cooler colors (blue) indicate negative divergence.

EXAMPLE 9.19

Consider the vector field $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{i} + R\,\mathbf{k}$ shown in the figure. It lies in the xy-plane and is identical in all horizontal planes (i.e., it is independent of z, and its z-component is 0).

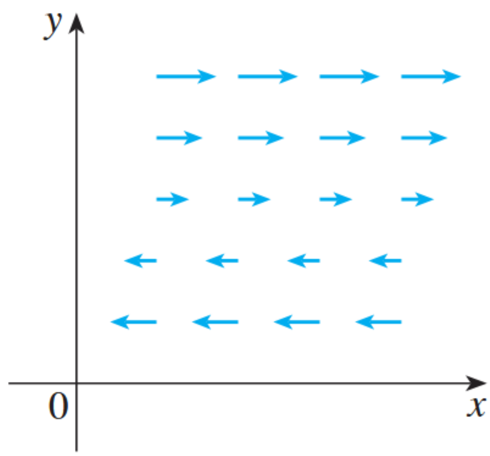


Image credit: Stewart

(a) Is div $\vec{F}$ positive, negative, or zero?
(b) Is $\nabla \times \vec{F} = \vec{0}$? If not, in what direction does curl $\vec{F}$ point?

Solution:

(a) The field vectors point strictly in the horizontal direction and get longer as we move up in the y-direction. This means the x-component of $\vec{F}$, which we call $P(x, y)$, depends on y. In particular, $\frac{\partial P}{\partial y} > 0$. However, $\vec{F}$ has no y-component, so $Q = 0$, and thus $\frac{\partial Q}{\partial x} = 0$. Since the field is constant in z, $\frac{\partial R}{\partial z} = 0$ as well. Thus we have

$$\operatorname{div} \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z} = 0 + 0 + 0 = 0.$$

(b) We compute the curl using the determinant form. Since $\vec{F} = P(x, y)\,\mathbf{i}$, with no j or k component, we have

$$\nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ P(x, y) & 0 & 0 \end{vmatrix} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathbf{i} + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathbf{j} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathbf{k} = -\frac{\partial P}{\partial y}\,\mathbf{k}$$

Since $\frac{\partial P}{\partial y} > 0$, $-\frac{\partial P}{\partial y}\,\mathbf{k}$ points in the negative z-direction.

EXAMPLE 9.20

Let a rigid body B rotate counterclockwise about the z-axis with angular velocity vector $\vec{w} = \omega\,\mathbf{k}$. The angular speed of ω is the tangential speed at any point $P \in B$ divided by the distance d from the axis of rotation. Let $\vec{r} = \langle x, y, z\rangle$ be the position vector of P.

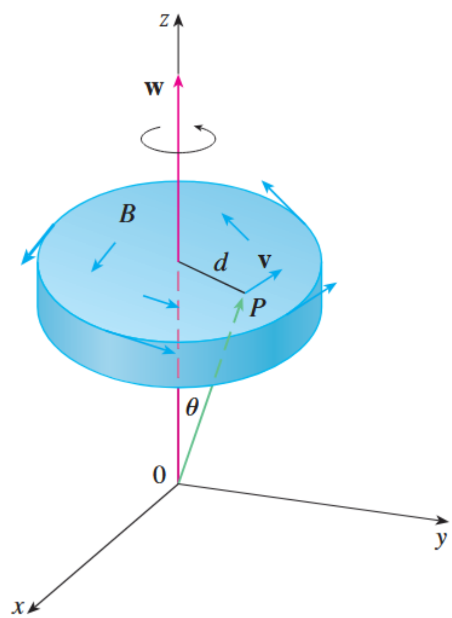


Image credit: Stewart

(a) Show that the velocity field is $\vec{v} = \vec{w} \times \vec{r}$.
(b) Show that $\vec{v} = -\omega y\,\mathbf{i} + \omega x\,\mathbf{j}$.
(c) Show that curl $\nabla \times \vec{v} = 2\vec{w}$.

Solution:

(a) The magnitude of the velocity is given by $v = \omega d = \omega r \sin\theta = \|\vec{w} \times \vec{r}\|$. The direction of $\vec{v}$ is perpendicular to both $\vec{w}$ and $\vec{r}$, which you can check with the right-hand rule. Therefore, the velocity vector at any point is $\vec{v} = \vec{w} \times \vec{r}$

(b) We have $\vec{w} = \langle 0, 0, \omega\rangle$. Then,

$$\vec{v} = \vec{w} \times \vec{r} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 0 & \omega \\ x & y & z \end{vmatrix} = (0 \cdot z - \omega y)\,\mathbf{i} + (\omega x - 0 \cdot z)\,\mathbf{j} + (0 \cdot y - x \cdot 0)\,\mathbf{k} = -\omega y\,\mathbf{i} + \omega x\,\mathbf{j}.$$

EXAMPLE 9.20 (CONTINUED)

(c) We compute the curl as follows:

$$\nabla \times \vec{v} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ -\omega y & \omega x & 0 \end{vmatrix}$$

$$= \left(\frac{\partial}{\partial y}(0) - \frac{\partial}{\partial z}(\omega x)\right)\mathbf{i} - \left(\frac{\partial}{\partial x}(0) - \frac{\partial}{\partial z}(-\omega y)\right)\mathbf{j} + \left(\frac{\partial}{\partial x}(\omega x) - \frac{\partial}{\partial y}(-\omega y)\right)\mathbf{k}$$

$$= (0-0)\,\mathbf{i} - (0-0)\,\mathbf{j} + (\omega+\omega)\,\mathbf{k} = 2\omega\,\mathbf{k} = 2\vec{\omega}$$

Thus, in rigid body dynamics, the curl of the velocity field encodes the local angular velocity.

To apply vector calculus meaningfully in the context of fluid flow, we have to adopt a mathematical model called a continuum model. We imagine a physical region of space that can be filled with a fluid like water, bounded by real surfaces, and modeled geometrically. Even if the fluid evaporated, the shape of the region remains, allowing us to assign coordinates to each point. We assume this fluid has mass, and that this mass is smoothly distributed. Not as isolated droplets (or, more formally, discrete particles), but as a dense collection of particles packed so closely that the fluid behaves like a continuous substance. Any small area in $\mathbb{R}^2$ or volume in $\mathbb{R}^3$ is assumed to contain a large number of molecules, so that quantities like density $\rho(x, y, z, t)$, velocity $\vec{v}(x, y, z, t)$, and pressure $p(x, y, z, t)$ can be defined at every point in space and time.

Under this view, a particle of fluid represents an infinitesimal piece of the continuous mass, which is the differential element used in our integrals. We can assign a vector field $\vec{v}(x, y, z, t) = \langle f, g, h\rangle$ to represent the velocity of the fluid, and track its behavior over time. A 2D vector field like $\vec{F}(x, y) = \langle f(x, y), g(x, y)\rangle$ can then be interpreted as a time-snapshot of motion across a fixed surface (like the surface of water). The divergence of this field measures how fluid accumulates or escapes from a region, a model of net outflow, while the curl represents local rotation. These operations only make sense under the assumption of continuity, which allows us to use calculus to describe and analyze the collective motion of infinitesimal fluid particles.

I know that calculus can be pretty dry at times. Depending on who you ask, it has arguably been "completed" for nearly 130 years. And even then, most of the work in the 20th century was simply the formalization of calculus by the likes of Weierstrass, Cauchy, Riemann, and Dedekind. Most of the content within the scale of multivariable calculus is even older. But I think this section is really quite beautiful. Up to this point in your mathematical journey, curl and divergence might be the most elegant material you've come across. I hope that you found it interesting.

10 Surface Integrals and Integral Theorems

Previously, we explored how vector fields behave over curves and regions in the plane using a variety of foundational tools. Then, we put it all together using curl and divergence. Now we will expand those ideas to surfaces in three dimensions by studying surface integrals and integral theorems.

10.1 Parametric Surfaces

Just as a space curve can be described by a vector-valued function $\vec{r}(t)$ of a single parameter t, a surface in $\mathbb{R}^3$ can be described by a vector-valued function of two parameters. A **parametric surface** is defined by a vector function

$$\vec{r}(u, v) = x(u, v)\,\mathbf{i} + y(u, v)\,\mathbf{j} + z(u, v)\,\mathbf{k}$$

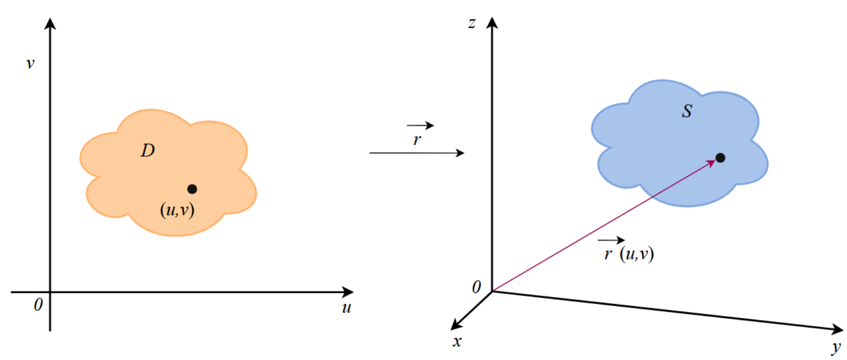
where $(u, v) \in D$ is a region in the uv-plane. The surface S is traced out by the tip of the position vector $\vec{r}(u, v)$ as (u, v) moves over D.

This is equivalent to defining three scalar functions:

$$x = x(u, v), \;\; y = y(u, v), \;\; z = z(u, v)$$

These are called the **parametric equations** of a surface.

Each choice of (u, v) gives a single point on the surface. Varying (u, v) over the entire domain D sweeps out all of S. That is, $S = \{\vec{r}(u, v) \mid (u, v) \in D\}$.



A parametric surface

EXAMPLE 10.1

Identify and sketch the surface defined by the vector function $\vec{r}(u, v) = 2\cos u\,\mathbf{i}+v\,\mathbf{j}+2\sin u\,\mathbf{k}$.

Solution:

The parametric equations are

$$x = 2\cos u,\;\; y = v,\;\; z = 2\sin u.$$

To identify the surface, eliminate the parameters. Since $x = 2\cos u$ and $z = 2\sin u$, we have the following for any point x, y, z on the surface:

$$x^2 + z^2 = 4\cos^2 u + 4\sin^2 u = 4$$

This equation describes a cylinder of radius 2 centered along the y-axis. Since $y = v$ varies freely, the surface is a circular cylinder of radius 2 extending along the y-direction.

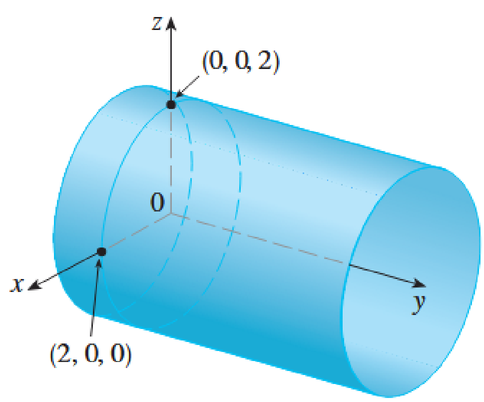


Image credit: Stewart

If a surface S is given by a vector function $\vec{r}(u, v)$, then there are two natural families of curves on the surface, called **grid curves**. One family is obtained by holding u constant and letting v vary, and the other by holding v constant and letting u vary. These correspond to vertical and horizontal lines in the uv-domain.

Fixing $u = u_0$, we obtain a curve C_1 on the surface given by $\vec{r}\,(u_0, v)$, which traces a curve in the v-direction. Fixing $v = v_0$, we get a curve C_2 on the surface given by $\vec{r}\,(u, v_0)$, which traces a curve in the u-direction. Together, these curves form a grid that helps visualize the geometry of the surface. This is also how computers graph surfaces.

For instance, if the parametric surface resembles a cylinder or cone, the grid curves may look like circles (when v is constant) and lines (when u is constant).

EXAMPLE 10.2

Find a parametric representation of the sphere $x^2 + y^2 + z^2 = a^2$.

Solution:

In spherical coordinates, the sphere is given by $\rho = a$. Let ϕ be the angle from the positive z-axis (colatitude) and θ the angle from the positive x-axis (longitude). Then the rectangular coordinate conversion gives:

$$x = a\sin\phi\cos\theta,\ \ y = a\sin\phi\sin\theta,\ \ z = a\cos\phi$$

The corresponding vector equation is:

$$\vec{r}(\phi,\theta) = a\sin\phi\cos\theta\,\mathbf{i} + a\sin\phi\sin\theta\,\mathbf{j} + a\cos\phi\,\mathbf{k}$$

The parameter domain is the rectangle $D = [0,\pi]\times[0,2\pi]$.
Grid curves with ϕ held constant are circles of constant latitude (including the equator).
Grid curves with θ held constant are meridians (vertical semicircles) connecting the poles.

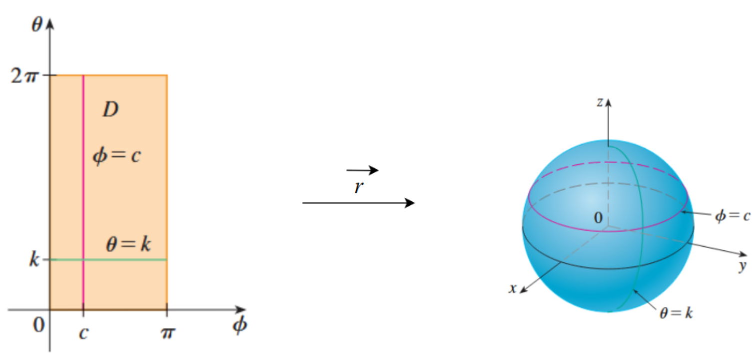


Image credit: Stewart

More generally, any surface that can be written as a function $z = f(x, y)$ can be treated as a parametric surface by taking:

$$x = x,\ y = y,\ z = f(x, y)$$

In this case, we use x and y as parameters and directly obtain the surface in vector form.

Note that a surface may have many possible parameterizations. The choice of parameter domain and coordinate expressions affects how the surface is traced out but not the surface itself.

A parameterization

$$\vec{r}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle,$$

is called a regular parameterization if

$$\vec{r}_u \times \vec{r}_v \neq \vec{0}$$

for all points (u, v) in the parameter domain.

If $\vec{r}(u, v)$ is regular, then its image is a two-dimensional object. Throughout this chapter, parameterizations are generally assumed to be regular.

Recall that a curve parameterization $\vec{r}(t)$, $a \leq t \leq b$, is smooth if $\vec{r}\,'(t)$ is continuous and $\vec{r}\,'(t) \neq \vec{0}$ for all $t \in [a, b]$. Visually, we see that a curve is smooth if it has no sharp corners. Similarly, a surface parameterization is smooth if the resulting surface has no sharp corners.

Formally, a surface parameterization

$$\vec{r}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle$$

is smooth if $\vec{r}_u \times \vec{r}_v \neq \vec{0}$ for any choice of u and v in the parameter domain.

A **surface of revolution** is formed by rotating a curve $y = f(x)$, where $f(x) \geq 0$, about the x-axis. Letting θ be the angle of rotation, a point on the surface has the following coordinates:

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

We use x and θ as parameters. These equations form the parametric representation of the surface. The domain is

$$a \leq x \leq b, \ 0 \leq \theta \leq 2\pi.$$

Given a parametric surface

$$\vec{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k}$$

the tangent vectors at a point $P_0 = \vec{r}(u_0, v_0)$ are given by partial derivatives. For C_1 we have

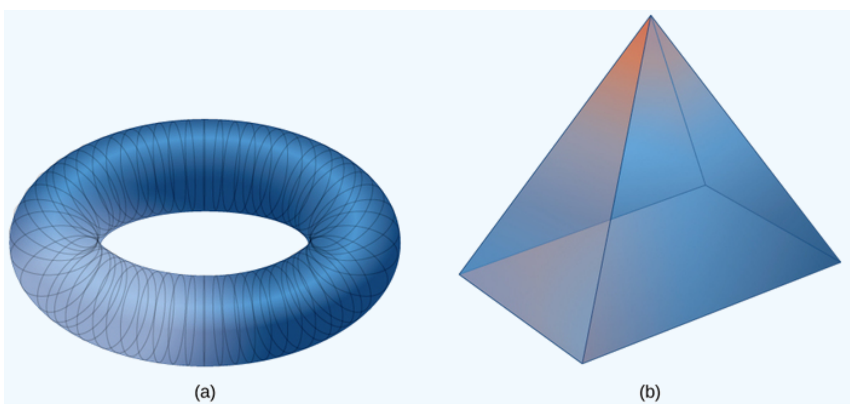
$$\vec{r}_v = \frac{\partial x}{\partial v}(u_0,v_0)\mathbf{i} + \frac{\partial y}{\partial v}(u_0,v_0)\mathbf{j} + \frac{\partial z}{\partial v}(u_0,v_0)\mathbf{k}.$$

And for C_2 we have

$$\vec{r}_u = \frac{\partial x}{\partial u}(u_0,v_0)\mathbf{i} + \frac{\partial y}{\partial u}(u_0,v_0)\mathbf{j} + \frac{\partial z}{\partial u}(u_0,v_0)\mathbf{k}.$$

If $\vec{r}_u \times \vec{r}_v \neq \vec{0}$, then the surface is smooth at P_0, and the tangent plane at that point contains both $\vec{r}_u$ and $\vec{r}_v$. A normal vector to the tangent plane is given by

$$\vec{n} = \vec{r}_u \times \vec{r}_v.$$



(a) is a smooth surface because it has no sharp corners. (b) has sharp corners, so directional derivatives do not exist at those locations. Thus, it has no smooth parameterization. That being said, it has four smooth faces, so it is piecewise smooth. Image credit: Strang & Herman

EXAMPLE 10.3

Find the tangent plane to the surface with parametric equations

$$x = u^2,\ y = v^2,\ z = u + 2v$$

at the point $(1, 1, 3)$.

Solution:

First compute the tangent vectors:

$$\vec{r}_u = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j} + \frac{\partial z}{\partial u}\,\mathbf{k} = 2u\,\mathbf{i} + 0\,\mathbf{j} + 1\,\mathbf{k} = 2u\,\mathbf{i} + \mathbf{k}$$

$$\vec{r}_v = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j} + \frac{\partial z}{\partial v}\,\mathbf{k} = 0\,\mathbf{i} + 2v\,\mathbf{j} + 2\,\mathbf{k} = 2v\,\mathbf{j} + 2\,\mathbf{k}$$

Now take the cross product to get the normal vector:

$$\vec{r}_u \times \vec{r}_v = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 2u & 0 & 1 \\ 0 & 2v & 2 \end{vmatrix} = (-2v)\,\mathbf{i} - (4u)\,\mathbf{j} + (4uv)\,\mathbf{k}$$

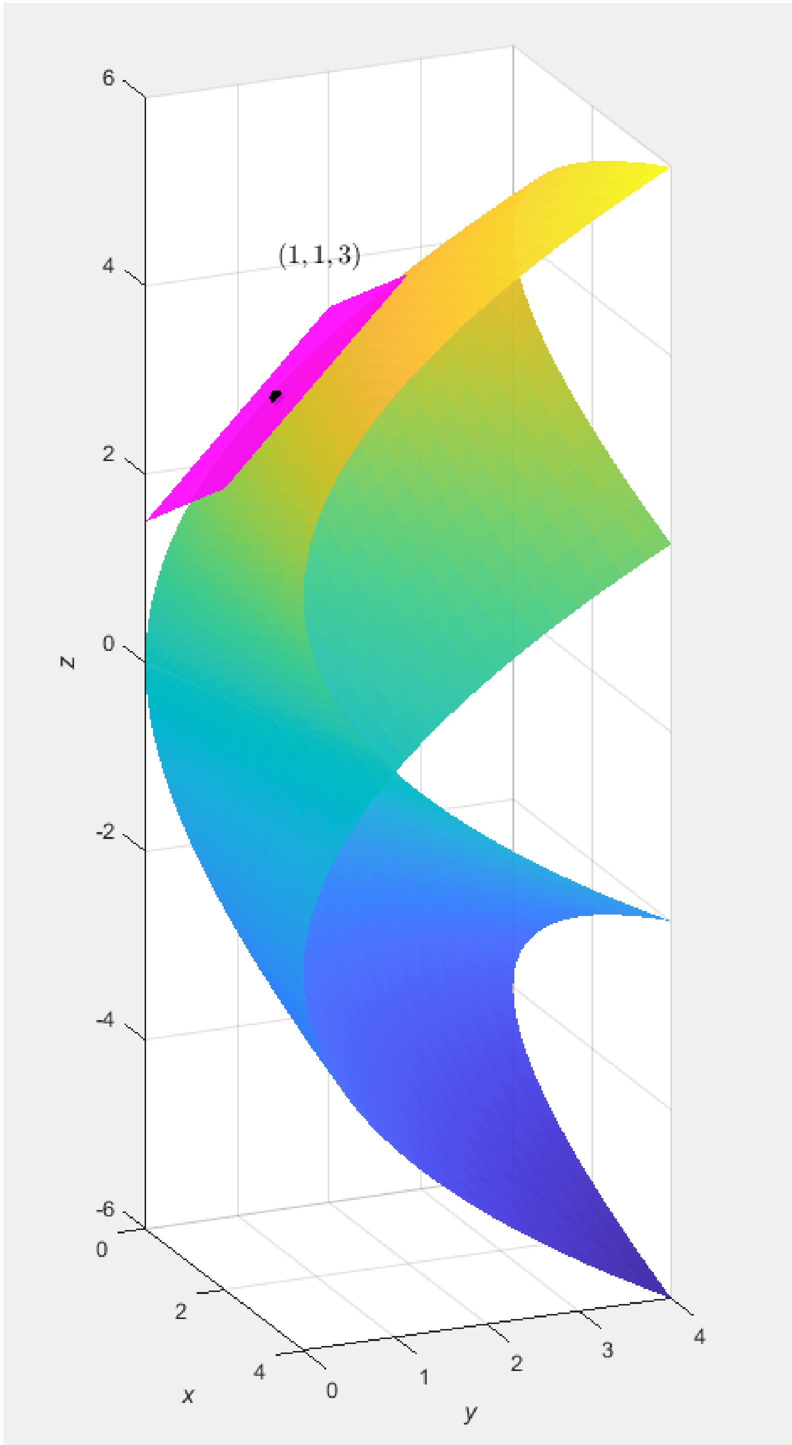
At the point $(1, 1, 3)$, we have $u = 1$ and $v = 1$, so the normal vector becomes

$$\vec{n} = -2\,\mathbf{i} - 4\,\mathbf{j} + 4\,\mathbf{k}.$$

The equation of the tangent plane at $(1, 1, 3)$ is given by:

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

EXAMPLE 10.3 (CONTINUED)



ex10point3.m

We now define the surface area of a general parametric surface given by

$$\vec{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k}$$

where $(u,v) \in D$ and D is a region in the uv-plane.

To approximate the area of this surface, we divide the domain D into a grid of small rectangles. Each rectangle $R_{i,j}$ is mapped by $\vec{r}(u,v)$ to a small curved patch $S_{i,j}$ on the surface. Each point $P_{i,j}$ corresponds to the lower left corner of a subrectangle. Then we say

$$\vec{r}^{\,*}_u = \vec{r}_u(u_i^*, v_j^*)$$

and

$$\vec{r}^{\,*}_v = \vec{r}_v(u^{\,*}_i, v^{\,*}_j)$$

are the tangent vectors at $P_{i,j}$.

These vectors span a parallelogram in space that approximates the patch $S_{i,j}$. The area of this parallelogram is given by the magnitude of their cross product:

$$\|(\Delta u_i\, \vec{r}^{\,*}_u) \times (\Delta v_i\, \vec{r}^{\,*}_v)\| = \|\vec{r}^{\,*}_u \times \vec{r}^{\,*}_v\|\, \Delta u_i\, \Delta v_i$$

Summing over all patches gives an approximation to the total surface area:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \|\vec{r}^{\,*}_u \times \vec{r}^{\,*}_v\|\, \Delta u\, \Delta v$$

Taking the limit as the grid becomes finer, this Riemann sum becomes a double integral:

$$A(S) = \iint_D \|\vec{r}_u \times \vec{r}_v\|\, dA$$

where $\vec{r}_u = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j} + \frac{\partial z}{\partial u}\,\mathbf{k}$ and $\vec{r}_v = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j} + \frac{\partial z}{\partial v}\,\mathbf{k}$.

EXAMPLE 10.4

Find the surface area of a sphere of radius a with $D = \{(\phi, \theta) \mid 0 \leq \phi \leq \pi,\ 0 \leq \theta \leq 2\pi\}$.

Solution:

We use the parametric representation from **EXAMPLE 10.2**:

$$x = a \sin\phi\cos\theta, \quad y = a\sin\phi\sin\theta, \quad z = a\cos\phi$$

We compute the cross product of the tangent vectors:

$$\begin{aligned}\vec{r}_\phi \times \vec{r}_\theta &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial x}{\partial \phi} & \frac{\partial y}{\partial \phi} & \frac{\partial z}{\partial \phi} \\ \frac{\partial x}{\partial \theta} & \frac{\partial y}{\partial \theta} & \frac{\partial z}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a\cos\phi\cos\theta & a\cos\phi\sin\theta & -a\sin\phi \\ -a\sin\phi\sin\theta & a\sin\phi\cos\theta & 0 \end{vmatrix} \\ &= a^2\sin^2\phi\cos\theta\,\mathbf{i} + a^2\sin^2\phi\sin\theta\,\mathbf{j} + a^2\sin\phi\cos\phi\,\mathbf{k}\end{aligned}$$

The magnitude of this vector is:

$$\begin{aligned}\|\vec{r}_\phi \times \vec{r}_\theta\| &= \sqrt{a^4\sin^4\phi\cos^2\theta + a^4\sin^4\phi\sin^2\theta + a^4\sin^2\phi\cos^2\phi} = a^2\sqrt{\sin^4\phi + \sin^2\phi\cos^2\phi} \\ &= a^2\sqrt{\sin^2\phi(\sin^2\phi + \cos^2\phi)} = a^2\sin\phi\end{aligned}$$

Our domain guarantees a nonnegative value of $\sin\phi$. Thus, we are safe to proceed. By the surface area formula,

$$\begin{aligned}A &= \iint_D \|\vec{r}_\phi \times \vec{r}_\theta\| dA = \int_0^{2\pi}\int_0^{\pi} a^2\sin\phi\, d\phi\, d\theta \\ &= a^2\left(\int_0^{2\pi} d\theta\right)\left(\int_0^{\pi}\sin\phi\, d\phi\right) = a^2(2\pi)(2) \\ &= 4\pi a^2.\end{aligned}$$

In the special case where a surface S is given as the graph of a function $z = f(x, y)$, we can interpret this surface by letting x and y serve as parameters. The corresponding parametric vector function is

$$\vec{r}(x, y) = x\,\mathbf{i} + y\,\mathbf{j} + f(x, y)\,\mathbf{k}.$$

The tangent vectors are obtained by taking partial derivatives with respect to x and y:

$$\vec{r}_x = \frac{\partial \vec{r}}{\partial x} = \mathbf{i} + \frac{\partial f}{\partial x}\,\mathbf{k},\ \ \vec{r}_y = \frac{\partial \vec{r}}{\partial y} = \mathbf{j} + \frac{\partial f}{\partial y}\,\mathbf{k}$$

To find the surface area, we compute the magnitude of the cross product $\vec{r}_x \times \vec{r}_y$. Using the determinant form:

$$\vec{r}_x \times \vec{r}_y = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{vmatrix} = -\frac{\partial f}{\partial x}\,\mathbf{i} - \frac{\partial f}{\partial y}\,\mathbf{j} + \mathbf{k}$$

Taking the magnitude of the cross product, we get

$$\|\vec{r}_x \times \vec{r}_y\| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1} = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}.$$

Hence, the surface area of the graph $z = f(x, y)$ over a region D is given by

$$A(S) = \iint_D \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\,dA.$$

To confirm that this surface area formula is consistent with the surface area formula you used in single-variable calculus, we consider the surface S formed by rotating the curve $y = f(x)$, where $a \le x \le b$, about the x-axis. Assume $f(x) \ge 0$ and $f'(x)$ is continuous. The parametric representation of the surface is

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

with domain

$$a \leq x \leq b, \quad 0 \leq \theta \leq 2\pi$$

We compute the tangent vectors:

$$\begin{aligned}
\vec{r}_x &= \frac{\partial \vec{r}}{\partial x} = \mathbf{i} + f'(x)\cos\theta\,\mathbf{j} + f'(x)\sin\theta\,\mathbf{k} \\
\vec{r}_\theta &= \frac{\partial \vec{r}}{\partial \theta} = -f(x)\sin\theta\,\mathbf{j} + f(x)\cos\theta\,\mathbf{k}
\end{aligned}$$

The cross product is:

$$\vec{r}_x \times \vec{r}_\theta = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & f'(x)\cos\theta & f'(x)\sin\theta \\ 0 & -f(x)\sin\theta & f(x)\cos\theta \end{vmatrix} = f(x)f'(x)\,\mathbf{i} - f(x)\cos\theta\,\mathbf{j} - f(x)\sin\theta\,\mathbf{k}.$$

Now compute the magnitude:

$$\|\vec{r}_x \times \vec{r}_\theta\| = \sqrt{[f(x)f'(x)]^2 + [f(x)\cos\theta]^2 + [f(x)\sin\theta]^2}$$

$$= \sqrt{f(x)^2[f'(x)^2 + \cos^2\theta + \sin^2\theta]} = \sqrt{f(x)^2[1 + f'(x)^2]} = f(x)\sqrt{1 + f'(x)^2}$$

Since $f(x) \geq 0$, the surface area is

$$A = \iint_D \|\vec{r}_x \times \vec{r}_\theta\|\,dA = \int_0^{2\pi} \int_a^b f(x)\sqrt{1 + f'(x)^2}\,dx\,d\theta = 2\pi \int_a^b f(x)\sqrt{1 + f'(x)^2}\,dx$$

This matches the formula for the surface area of a solid of revolution about the x-axis.

10.2 Surface Integrals

Surface integrals allow us to generalize the idea of integration to curved surfaces in space. Just as a line integral accumulates values along a curve, a surface integral accumulates values across a surface. Line integrals are to arc length as surface integrals are to surface area.

We now define the **surface integral** of a scalar function $f(x, y, z)$ over a parametric surface. Suppose a surface S is given by the vector equation:

$$\vec{r}(u,v) = x(u,v)\,\mathbf{i} + y(u,v)\,\mathbf{j} + z(u,v)\,\mathbf{k}$$

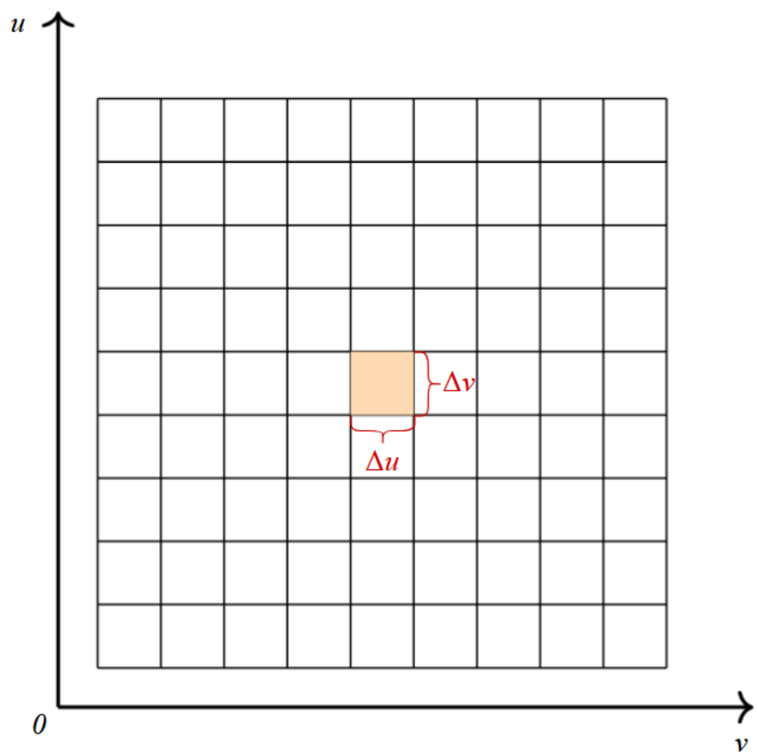
for $(u,v) \in D$.

We divide the domain D into rectangles $R_{i,j}$, and let each subrectangle map to a surface patch $S_{i,j}$ on S. The surface integral of f over S is approximated by the Riemann sum:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{i,j}^{*})\,\Delta S_{i,j}$$

As the number of subdivisions increases, this sum approaches the surface integral:

$$\iint_S f(x,y,z)\,dS = \lim_{\Delta u,\Delta v\to 0}\sum_{i=1}^{m}\sum_{j=1}^{n} f(P_{i,j}^{*})\,\Delta S_{i,j}$$



The orange square represents $R_{i,j} \subseteq D$.

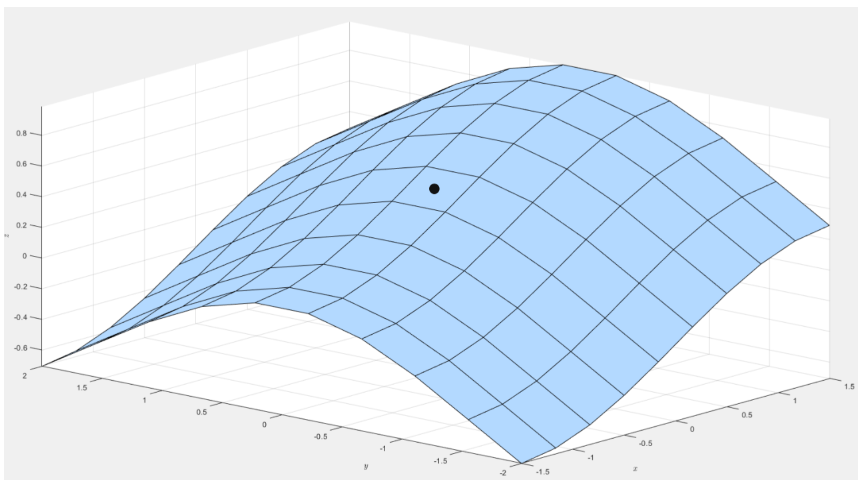
To approximate the area $\Delta S_{i,j}$ of each patch, we use the tangent vectors at $P_{i,j}$:

$$\vec{r}_u = \frac{\partial \vec{r}}{\partial u},\ \vec{r}_v = \frac{\partial \vec{r}}{\partial v}$$

Then the area of each patch is approximated as the area of the parallelogram spanned by $\vec{r}_u$ and $\vec{r}_v$:

$$\Delta S_{i,j} \approx \|\vec{r}_u \times \vec{r}_v\|\, \Delta u\, \Delta v$$

where $\vec{r}_u = \frac{\partial x}{\partial u}\,\mathbf{i} + \frac{\partial y}{\partial u}\,\mathbf{j} + \frac{\partial z}{\partial u}\,\mathbf{k}$ and $\vec{r}_v = \frac{\partial x}{\partial v}\,\mathbf{i} + \frac{\partial y}{\partial v}\,\mathbf{j} + \frac{\partial z}{\partial v}\,\mathbf{k}$ are the tangent vectors at a corner of $S_{i,j}$.



The point represents $P_{i,j} \in S_{i,j} \subseteq S$.

So, the surface integral becomes

$$\boxed{\iint_S f(x,y,z)\, dS = \iint_D f(x(u,v), y(u,v), z(u,v))\, \|\vec{r}_u \times \vec{r}_v\|\, dA.}$$

This formula lets us convert a surface integral into a double integral over the parameter domain D. For example, if $f(x,y,z) = 1$, the surface integral simply returns the surface area:

$$\iint_S 1\, dS = \iint_D \|\vec{r}_u \times \vec{r}_v\|\, dA = A(S)$$

If the surface S is a thin sheet with variable density $\rho(x,y,z)$, the total mass is

$$m = \iint_S \rho(x, y, z)\, dS.$$

The center of mass $(\bar{x}, \bar{y}, \bar{z})$ is given by the following:

$$\bar{x} = \frac{1}{m} \iint_S x\, \rho(x, y, z)\, dS$$
$$\bar{y} = \frac{1}{m} \iint_S y\, \rho(x, y, z)\, dS$$
$$\bar{z} = \frac{1}{m} \iint_S z\, \rho(x, y, z)\, dS$$

Any surface of the form $z = g(x, y)$ where z is a graph can be treated as a parametric surface by setting

$$x = x,\ \ y = y,\ \ z = g(x, y).$$

Then, the partial derivative vectors are $\vec{r}_x = \mathbf{i} + \left(\frac{\partial g}{\partial x}\right)\mathbf{k}$ and $\vec{r}_y = \mathbf{j} + \left(\frac{\partial g}{\partial y}\right)\mathbf{k}$.

The surface area element is determined from the cross product:

$$\vec{r}_x \times \vec{r}_y = -\frac{\partial g}{\partial x}\mathbf{i} - \frac{\partial g}{\partial y}\mathbf{j} + \mathbf{k}$$

and the magnitude of that vector is

$$\|\vec{r}_x \times \vec{r}_y\| = \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}.$$

So, the surface integral becomes

$$\boxed{\iint_S f(x, y, z)\, dS = \iint_D f(x, y, g(x, y)) \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 + 1}\, dA.}$$

A similar formula holds when projecting onto a different coordinate plane, such as if $y = h(x, z)$ and you're projecting onto the xz-plane:

$$\iint_S f(x,y,z)\,dS = \iint_D f(x,h(x,z),z)\sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial y}{\partial x}\right)^2 + 1}\,dA.$$

If S is not covered by a single parameterization and rather a finite union of smooth surfaces $S_1, S_2, ..., S_n$, the surface integral extends as a sum over the individual parts:

$$\iint_S f(x,y,z)\,dS = \iint_{S_1} f(x,y,z)\,dS + \cdots + \iint_{S_n} f(x,y,z)\,dS$$

We can thus break up a complex surface into multiple smooth regions that are easier to work with.

EXAMPLE 10.5

Evaluate the surface integral $\iint_S z\,dS$, where S is the closed surface composed of the cylindrical side S_1 of $x^2+y^2=1$, the disk S_2 in the plane $z=0$, and the top surface S_3 lying above the disk and defined by the plane $z=1+x$.

Solution:

We compute the surface integral $\iint_S z\,dS$ by breaking the surface into three parts:

$$\iint_S z\,dS = \iint_{S_1} z\,dS + \iint_{S_2} z\,dS + \iint_{S_3} z\,dS$$

We begin with side surface S_1 by parameterizing S_1 using cylindrical coordinates:

$$x = \cos\theta,\ \ y = \ \sin\theta,\ \ z = z$$

where $0 \le \theta \le 2\pi$ and $0 \le z \le 1+\cos\theta$.
We compute the normal vector via the cross product:

$$\vec{r}_\theta \times \vec{r}_z = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{vmatrix} = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}$$

$$\|\vec{r}_\theta \times \vec{r}_z\| = \sqrt{\cos^2\theta + \sin^2\theta} = 1$$

Now we evaluate the surface integral over S_1:

$$\iint_{S_1} z\,dS = \iint_D z\,\|\vec{r}_\theta \times \vec{r}_z\|\,dA$$

$$= \int_0^{2\pi}\int_0^{1+\cos\theta} z\,dz\,d\theta = \int_0^{2\pi} \frac{1}{2}(1+\cos\theta)^2\,d\theta$$

$$= \frac{1}{2}\int_0^{2\pi}\left(1+2\cos\theta+\cos^2\theta\right)d\theta = \frac{1}{2}\int_0^{2\pi}\left(1+2\cos\theta+\frac{1+\cos 2\theta}{2}\right)d\theta$$

$$= \frac{1}{2}\int_0^{2\pi}\left(\frac{3}{2}+2\cos\theta+\frac{1}{2}\cos 2\theta\right)d\theta = \frac{1}{2}\left[\frac{3}{2}\theta + 2\sin\theta + \frac{1}{4}\sin 2\theta\right]_0^{2\pi} = \frac{3\pi}{2}$$

EXAMPLE 10.5 (CONTINUED)

And now for the bottom surface S_2, which lies in the plane $z = 0$:

$$\iint_{S_2} z\, dS = \iint_{S_2} 0 \cdot dS = 0$$

The top surface S_3 is described by $z = 1 + x$, over the unit disk D in the xy-plane. We use the formula for graphs:

$$\iint_{S_3} z\, dS = \iint_D (1 + x)\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dA$$

Since $\frac{\partial z}{\partial x} = 1$, $\frac{\partial z}{\partial y} = 0$, we get

$$\iint_{S_3} (1 + x)\sqrt{2}\, dA = \sqrt{2} \iint_D (1 + x)\, dA$$

Switch to polar coordinates:

$$x = r\cos\theta,\ dA = r\, dr\, d\theta$$

Then,

$$\begin{aligned}
\iint_{S_3} z\, dS = \sqrt{2}\int_0^{2\pi}\int_0^1 (1 + r\cos\theta) r\, dr\, d\theta &= \sqrt{2}\int_0^{2\pi}\int_0^1 \left(r + r^2\cos\theta\right) dr\, d\theta \\
&= \sqrt{2}\int_0^{2\pi}\left(\frac{1}{2} + \frac{1}{3}\cos\theta\right) d\theta \\
&= \sqrt{2}\left[\frac{\theta}{2} + \frac{\sin\theta}{3}\right]_0^{2\pi} = \sqrt{2}\pi \\
&= \sqrt{2}\int_0^{2\pi}\left[\frac{1}{2} + \frac{1}{3}\cos\theta\right] d\theta = \sqrt{2}\,(\pi + 0) = \sqrt{2}\pi
\end{aligned}$$

Finally, we add each surface together:

$$\iint_S z\, dS = \iint_{S_1} z\, dS + \iint_{S_2} z\, dS + \iint_{S_3} z\, dS = \frac{3\pi}{2} + 0 + \sqrt{2}\pi = \left(\frac{3}{2} + \sqrt{2}\right)\pi.$$

Before we define the flux of a vector field across a surface, we must ensure that the surface is **orientable**. That is, we must be able to assign a continuous unit normal vector $\mathbf{n}$ at every point. Every orientable surface has two possible orientations: one given by $\mathbf{n}$, and the other by $-\mathbf{n}$.

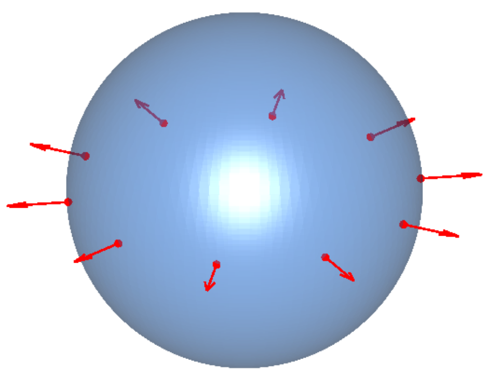
For a surface given as the graph $z = g(x, y)$, a natural upward-pointing unit normal vector (positive $\mathbf{k}$-component is

$$\mathbf{n} = \frac{-\frac{\partial g}{\partial x}\,\mathbf{i} - \frac{\partial g}{\partial y}\,\mathbf{j} + \mathbf{k}}{\sqrt{1 + \left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2}}.$$

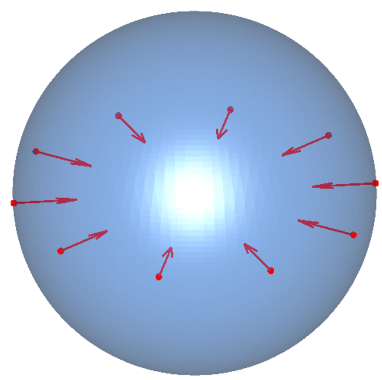
For a general parametric surface $\vec{r}(u, v)$, the orientation is given by the unit vector

$$\mathbf{n} = \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}$$

The opposite orientation is given by $-\mathbf{n}$. For closed surfaces, the standard is for it to be assigned the outward orientation.



(a) Positive orientation: outward-pointing normal vectors.



(b) Negative orientation: inward-pointing normal vectors.

Now, consider a vector field $\vec{F}$, such as a velocity field. The **flux** of $\vec{F}$ across a surface S measures how much of $\vec{F}$ passes through S. This is defined by the surface integral

$$\iint_S \vec{F} \cdot \vec{n}\, dS,$$

and interpreted as the rate of flow across the surface.

The flux of a continuous vector field $\vec{F}$ defined on an oriented surface S with unit normal vector $\mathbf{n}$ is given by the following surface integral:

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_S \vec{F} \cdot \vec{n}\, dS$$

This is called the flux of $\vec{F}$ across or over S.

For a surface S given in parametric form $\vec{r}(u, v)$, the unit normal vector is given by

$$\vec{n} = \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}.$$

So the flux of a vector field $\vec{F}$ across the surface becomes

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_S \vec{F} \cdot \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}\, dS.$$

We convert the surface integral into a double integral over the parameter domain D using $dS = \|\vec{r}_u \times \vec{r}_v\|\, dA$. This gives

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \left[\vec{F}(\vec{r}(u, v)) \cdot \frac{\vec{r}_u \times \vec{r}_v}{\|\vec{r}_u \times \vec{r}_v\|}\right] \|\vec{r}_u \times \vec{r}_v\|\, dA.$$

The magnitude cancels, so the formula simplifies to

$$\boxed{\iint_S \vec{F} \cdot d\vec{S} = \iint_D \vec{F}(\vec{r}(u, v)) \cdot (\vec{r}_u \times \vec{r}_v)\, dA.}$$

This equation expresses the flux through a surface S in terms of its parameterization. At every point, we compute the component of $\vec{F}$ in the direction normal to the surface (given by the cross product of tangent vectors $\vec{r}_u \times \vec{r}_v$), and then multiply by the area element of the surface. This captures how much of the vector field flows through each infinitesimal patch of the surface.

If the surface is given as a graph $z = g(x, y)$, we can also express the flux by using partial derivatives. Let

$$\vec{F}(x, y, z) = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$$

and recall that

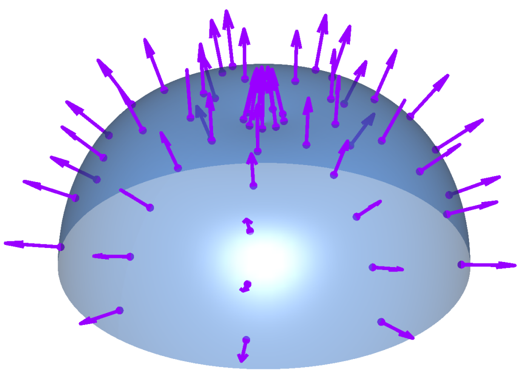
$$\vec{r}_x \times \vec{r}_y = -\frac{\partial g}{\partial x}\,\mathbf{i} - \frac{\partial g}{\partial y}\,\mathbf{j} + \mathbf{k}.$$

Then

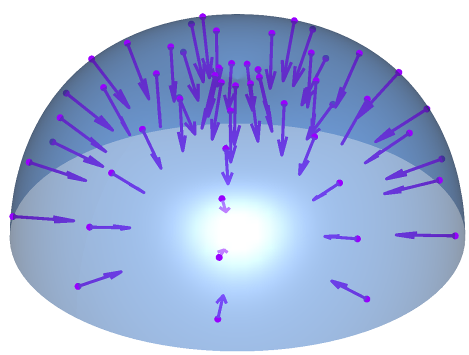
$$\vec{F} \cdot (\vec{r}_x \times \vec{r}_y) = -P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R$$

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \left(-P\frac{\partial g}{\partial x} - Q\frac{\partial g}{\partial y} + R\right) dA$$

This assumes an upward-pointing orientation; for downward-pointing orientation, multiply the entire expression by -1.



(a) Positive orientation: normal vectors point outward from the surface.



(b) Negative orientation: normal vectors point downward from the surface.

EXAMPLE 10.6

Compute the flux of the vector field $\vec{F}(x,y,z) = z\,\mathbf{i} + y\,\mathbf{j} + x\,\mathbf{k}$ across the unit sphere $x^2+y^2+z^2=1$.

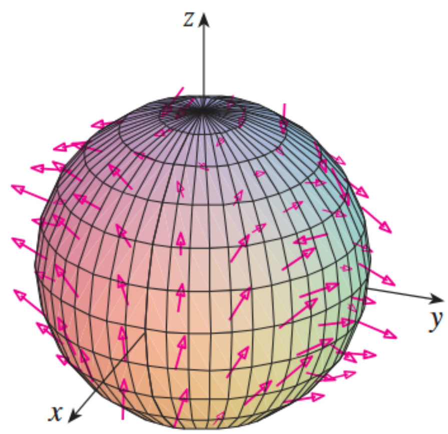


Image credit: Stewart

Solution:

We use the spherical parametrization:

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

where $0 \le \phi \le \pi$ and $0 \le \theta \le 2\pi$.

The vector field evaluated on the surface becomes $\vec{F}(\vec{r}(\phi,\theta)) = \cos\phi\,\mathbf{i} + \sin\phi\,\mathbf{j} + \sin\phi\cos\theta\,\mathbf{k}$. From **EXAMPLE 10.4**, we know that

$$\vec{r}_\phi \times \vec{r}_\theta = \sin^2\phi\cos\theta\,\mathbf{i} + \sin^2\phi\sin\theta\,\mathbf{j} + \sin\phi\cos\phi\,\mathbf{k}.$$

Now compute the dot product:

$$\vec{F}(\vec{r}(\phi,\theta)) \cdot (\vec{r}_\phi \times \vec{r}_\theta) = \cos\phi\,\sin^2\phi\,\cos\theta + \sin^3\phi\,\sin^2\theta + \sin^2\phi\,\cos\phi\,\cos\theta$$

EXAMPLE 10.6 (CONTINUED)

And now we integrate:

$$\begin{aligned}\iint_S \vec{F} \cdot d\vec{S} &= \iint_D \vec{F} \cdot (\vec{r}_\phi \times \vec{r}_\theta)\, dA \\ &= \int_0^{2\pi} \int_0^{\pi} \left(2\sin^2\phi\cos\phi\cos\theta + \sin^3\phi\sin^2\theta\right) d\phi\, d\theta \\ &= 2\int_0^{\pi} \sin^2\phi\cos\phi\, d\phi \int_0^{2\pi} \cos\theta\, d\theta + \int_0^{\pi} \sin^3\phi\, d\phi \int_0^{2\pi} \sin^2\theta\, d\theta \\ &= 0 + \int_0^{\pi} \sin^3\phi\, d\phi \int_0^{2\pi} \sin^2\theta\, d\theta \\ &= \frac{4\pi}{3}\end{aligned}$$

The surface integral of a vector field $\vec{F}$ over a surface S measures how much of the field "flows through" a surface. Let's go over some applications.

If $\vec{E}$ is an electric field, then the total electric field passing through a surface S is called the *electric flux*, defined as

$$\iint_S \vec{E} \cdot d\vec{S}.$$

Gauss' law tells us that for a closed surface, this flux equals the total enclosed charge divided by the vacuum permittivity:

$$Q = \varepsilon_0 \iint_S \vec{E} \cdot d\vec{S}$$

where $\varepsilon_0 = 8.8542 \times 10^{-12}\ \mathrm{C}^2/\mathrm{N} \cdot \mathrm{m}^2$ is a constant called the permittivity of free space.

If $u(x, y, z)$ is the temperature at a point in a solid, then the heat flux vector is:

$$\vec{F} = -K\nabla u$$

where K is the thermal conductivity. The rate of heat flow across a surface S is

$$\iint_S \vec{F} \cdot d\vec{S} = -K \iint_S \nabla u \cdot d\vec{S}.$$

The negative sign reverses the direction of ∇u, making $\vec{F}$ point from hot regions to cold regions. In other words, it's a gradient where heat flows from the hot region to the cold region. This is known as a heat sink.

Concept	**Formula**
Surface parameterization	$S : \vec{r}(u, v) = \langle x(u, v), y(u, v), z(u, v)\rangle, \quad (u, v) \in D$
Surface area	$A = \iint_S dS = \iint_D \|\vec{r}_u \times \vec{r}_v\| \, du \, dv$
Surface area differential	$dS = \|\vec{r}_u \times \vec{r}_v\| \, du \, dv$
Scalar surface integral (general 3D)	$\iint_S f(x, y, z) \, dS = \iint_D f(\vec{r}(u, v))\|\vec{r}_u \times \vec{r}_v\| \, du \, dv$
Scalar surface integral (for $z = g(x, y)$)	$\iint_S f(x, y, z) \, dx \, dy = \iint_D f(x, y, g(x, y))\sqrt{\left(\frac{\partial g}{\partial x}\right)^2 + \left(\frac{\partial g}{\partial y}\right)^2 + 1} \, dx \, dy$
Oriented surface element vector	$d\vec{S} = (\vec{r}_u \times \vec{r}_v) \, du \, dv$
Vector surface integral (general form)	$\iint_S \vec{F} \cdot d\vec{S} = \iint_D \vec{F}(\vec{r}(u, v)) \cdot (\vec{r}_u \times \vec{r}_v) \, du \, dv$
Vector surface integral (component form)	$\iint_S \vec{F} \cdot d\vec{S} = \iint_S P \, dy \, dz + Q \, dz \, dx + R \, dx \, dy \quad \text{for } \vec{F} = \langle P, Q, R\rangle$
Relation to flux and orientation	$\text{Flux} = \iint_S \vec{F} \cdot \vec{n} \, dS = \iint_S \vec{F} \cdot d\vec{S} \quad \text{(depends on orientation of } S\text{)}$

Summary of Key Surface Integral Formulas and Concepts

EXAMPLE 10.7

The temperature u in a metal ball is proportional to the square of the distance from the center. Find the rate of heat flow across a spherical surface S of radius a.

Solution:

We will assume the ball is centered at the origin. Let the temperature function be $u(x, y, z) = C(x^2 + y^2 + z^2)$. where C is a constant. Then the heat flux vector is

$$\vec{F}(x, y, z) = -K\nabla u = -KC(2x\,\mathbf{i} + 2y\,\mathbf{j} + 2z\,\mathbf{k}).$$

We factor to simplify:

$$\vec{F}(x, y, z) = -2KC(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k})$$

On the surface of a sphere of radius a, the unit normal vector is

$$\mathbf{n} = \frac{1}{a}(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}).$$

So the dot product becomes

$$\vec{F} \cdot \mathbf{n} = -2KC(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}) \cdot \frac{1}{a}(x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k}) = -\frac{2KC}{a}(x^2 + y^2 + z^2).$$

On S, we have $x^2 + y^2 + z^2 = a^2$, so

$$\vec{F} \cdot \mathbf{n} = -\frac{2KC}{a} \cdot a^2 = -2aKC.$$

Then the total heat flow is

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_S \vec{F} \cdot \mathbf{n}\,dS = -2aKC \iint_S dS = -2aKC(4\pi a^2)$$
$$= -8KC\pi a^3.$$

EXAMPLE 10.8

An infinitely long vertical wire along the z-axis carries a current I, generating a magnetic field

$$\vec{B} = \frac{\mu_0 I}{2\pi} \left(\frac{-y\,\mathbf{i} + x\,\mathbf{j}}{x^2 + y^2} \right)$$

Find the magnetic flux through a rectangle in the yz-plane where $y = 0$, bounded by $x_1 \leq x \leq x_2$ and $z_1 \leq z \leq z_2$.

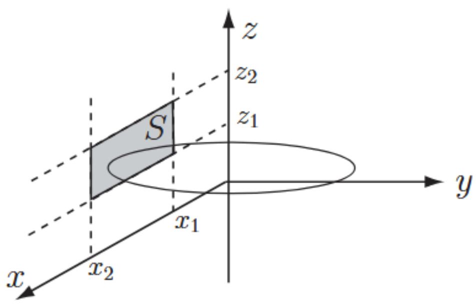


Image credit: Loughborough University

Solution.

On the plane $y = 0$, the magnetic field simplifies to $\vec{B} = \frac{\mu_0 I}{2\pi x}\,\mathbf{j}$. The surface lies in the xz-plane and is oriented with an outward unit normal vector $\mathbf{j}$. So $d\vec{S} = dx\,dz\,\mathbf{j}$.
Then the flux is given by:

$$\Phi = \iint_S \vec{B} \cdot d\vec{S} = \iint_{x_1}^{x_2} \int_{z_1}^{z_2} \frac{\mu_0 I}{2\pi x}\,dx\,dz$$

Since the integrand is independent of z, we can factor. We will integrate z first and then x:

$$\begin{aligned}
&= \int_{x=x_1}^{x_2} \left(\int_{z=z_1}^{z_2} \frac{\mu_0 I}{2\pi x}\,dz \right) dx \\
&= \int_{x=x_1}^{x_2} \left(\frac{\mu_0 I}{2\pi x}(z_2 - z_1) \right) dx \\
&= \frac{\mu_0 I (z_2 - z_1)}{2\pi} \int_{x=x_1}^{x_2} \frac{1}{x}\,dx \\
&= \frac{\mu_0 I (z_2 - z_1)}{2\pi} \left[\ln x\right]_{x_1}^{x_2} \\
&= \frac{\mu_0 I}{2\pi}(z_2 - z_1) \ln\left(\frac{x_2}{x_1} \right)
\end{aligned}$$

10.3 Stokes' Theorem

Stokes' theorem is a generalization of Green's theorem. It relates the circulation of a vector field $\vec{F}$ around a closed space curve C to the total curl of $\vec{F}$ passing through a surface S bounded by C:

$$\boxed{\oint_C \vec{F} \cdot d\vec{r} = \iint_S (\nabla \times \vec{F}) \cdot d\vec{S} = \iint_S \text{curl } \vec{F} \cdot d\vec{S}}$$

where C is positively oriented (counterclockwise with respect to the surface normal), S is an oriented surface with unit normal $\vec{n}$, $\vec{F}$ has continuous partial derivatives on an open region in $\mathbb{R}^3$, and $d\vec{S} = \vec{n}\,dS$.

This says that the line integral of $\vec{F}$ along the boundary equals the total normal component of curl across the surface.

If the surface lies in the xy-plane with upward normal $\vec{k}$, then Stokes' theorem becomes:

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S (\nabla \times \vec{F}) \cdot \vec{k}\,dA$$

This is exactly the vector form of Green's theorem.

Let S be the surface given by $z = g(x, y)$, where g has continuous second-order partial derivatives, $(x, y) \in D$, and let $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$ with continuously differentiable components.

Then

$$\iint_S (\nabla \times \vec{F}) \cdot d\vec{S} = \iint_D \left[-\left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z} \right) \frac{\partial z}{\partial x} - \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x} \right) \frac{\partial z}{\partial y} + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) \right] dA.$$

Let the boundary C be parameterized by:

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

Then

$$\begin{aligned}
\oint_C \vec{F} \cdot d\vec{r} &= \int_a^b \left(P\frac{dx}{dt} + Q\frac{dy}{dt} + R\frac{dz}{dt} \right) dt \\
&= \int_a^b \left[P\frac{dx}{dt} + Q\frac{dy}{dt} + R\left(\frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \right) \right] dt \\
&= \int_a^b \left[\left(P + R\frac{\partial z}{\partial x} \right)\frac{dx}{dt} + \left(Q + R\frac{\partial z}{\partial y} \right)\frac{dy}{dt} \right] dt \\
&= \int_C \left(P + R\frac{\partial z}{\partial x} \right) dx + \left(Q + R\frac{\partial z}{\partial y} \right) dy.
\end{aligned}$$

Now apply Green's theorem:

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_D \left[\frac{\partial}{\partial x}\left(Q + R\frac{\partial z}{\partial y} \right) - \frac{\partial}{\partial y}\left(P + R\frac{\partial z}{\partial x} \right) \right] dA$$

Use the product and chain rules:

$$\begin{aligned}
\oint_C \vec{F} \cdot d\vec{r} = \iint_D \Bigg[&\left(\frac{\partial Q}{\partial x} + \frac{\partial Q}{\partial z}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial x}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y} + R\frac{\partial^2 z}{\partial x \partial y} \right) \\
&- \left(\frac{\partial P}{\partial y} + \frac{\partial P}{\partial z}\frac{\partial z}{\partial y} + \frac{\partial R}{\partial y}\frac{\partial z}{\partial x} + \frac{\partial R}{\partial z}\frac{\partial z}{\partial y}\frac{\partial z}{\partial x} + R\frac{\partial^2 z}{\partial y \partial x} \right) \Bigg] dA
\end{aligned}$$

After simplifying, we have the surface integral formula:

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S (\nabla \times \vec{F}) \cdot d\vec{S}$$

EXAMPLE 10.9

Use Stokes' theorem to compute the integral

$$\iint_S (\nabla \times \vec{F}) \cdot d\vec{S}$$

where $\vec{F}(x, y, z) = xz\,\mathbf{i} + yz\,\mathbf{j} + xy\,\mathbf{k}$, and S is the part 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:

To find the boundary curve C, we solve the system

$$\begin{cases} x^2 + y^2 + z^2 = 4 \\ x^2 + y^2 = 1 \end{cases} \quad \Rightarrow z^2 = 3 \Rightarrow z = \sqrt{3}$$

So C is the circle $x^2 + y^2 = 1$ at height $z = \sqrt{3}$. A vector equation for the curve is:

$$\vec{r}(t) = \cos t\,\mathbf{i} + \sin t\,\mathbf{j} + \sqrt{3}\,\mathbf{k}, \quad 0 \le t \le 2\pi$$

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

We evaluate the vector field along $\vec{r}(t)$:

$$\vec{F}(\vec{r}(t)) = xz\,\mathbf{i} + yz\,\mathbf{j} + xy\,\mathbf{k} = \sqrt{3}\cos t\,\mathbf{i} + \sqrt{3}\sin t\,\mathbf{j} + \cos t \sin t\,\mathbf{k}$$

Now compute the circulation using Stokes' theorem:

$$\begin{aligned}
\iint_S (\nabla \times \vec{F}) \cdot d\vec{S} = \oint_C \vec{F} \cdot d\vec{r} = \int_0^{2\pi} \vec{F}(\vec{r}(t)) \cdot \vec{r}\,'(t)\,dt \\
&= \int_0^{2\pi} \left(\sqrt{3}\cos t\,\mathbf{i} + \sqrt{3}\sin t\,\mathbf{j} + \cos t \sin t\,\mathbf{k}\right) \cdot (-\sin t\,\mathbf{i} + \cos t\,\mathbf{j})\,dt \\
&= \int_0^{2\pi} \left(-\sqrt{3}\cos t \sin t + \sqrt{3}\sin t \cos t\right) dt \\
&= \sqrt{3}\int_0^{2\pi} (-\cos t \sin t + \sin t \cos t) dt \\
&= \sqrt{3}\int_0^{2\pi} 0\,dt = 0
\end{aligned}$$

If two surfaces S_1 and S_2 share the same positively oriented boundary curve C, and both satisfy the conditions of Stokes' theorem, then

$$\iint_{S_1} \nabla \times \vec{F} \cdot d\vec{S} = \oint_C \vec{F} \cdot d\vec{r} = \iint_{S_2} \nabla \times \vec{F} \cdot d\vec{S}$$

This identity shows that the value of the surface integral of $\nabla \times \vec{F}$ is completely determined by the circulation around the boundary C and not by the specific surface spanning it. This is useful when one surface is easier to integrate over than another.

For instance, let $\vec{v}$ be a vector field representing fluid velocity, and suppose C is an oriented closed curve in space. The circulation of $\vec{v}$ around C is defined by:

$$\oint_C \vec{v} \cdot d\vec{r} = \oint_C \vec{v} \cdot \vec{T}\, ds$$

where $\mathbf{T}$ is the unit tangent vector to the curve and ds is the arc length element.

The dot product $\vec{v} \cdot \mathbf{T}$ measures the component of $\vec{v}$ in the direction of the curve at each point. If this value is large and positive, the fluid moves along the curve's orientation. If negative, it flows opposite to the curve's direction.

Thus, $\oint_C \vec{v} \cdot d\vec{r}$ measures the net tendency of the fluid to circulate around C, and is called the **circulation** of $\vec{v}$ around C.

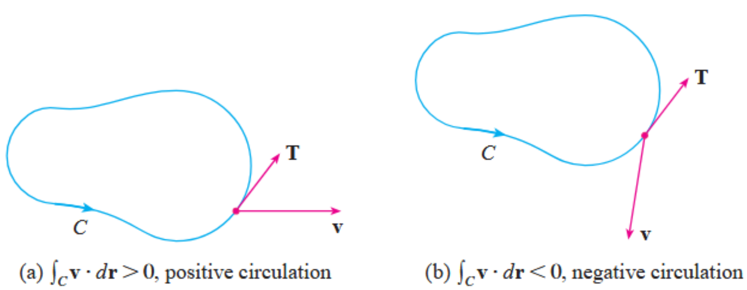
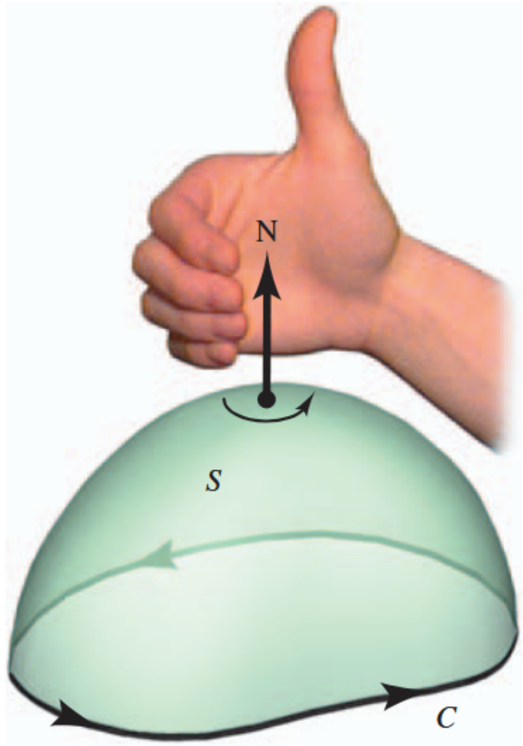


Image credit: Stewart



Direction along C is counterclockwise relative to the surface normal vector $\mathbf{N}$. Point your right thumb in the direction of $\mathbf{N}$ and then curl your fingers in the positive direction along C. Image credit: Larson & Edwards

Let $P_0 = (x_0, y_0, z_0)$ be a point in a fluid, and let S_a be a small disk of radius a centered at P_0. If $\vec{v}$ is a velocity field and $\mathbf{n}$ is the unit normal to S_a, then by Stokes' theorem and the continuity of $\nabla \times \vec{v}$, we approximate:

$$\oint_{C_a} \vec{v} \cdot d\vec{r} = \iint_{S_a} \nabla \times \vec{v} \cdot d\vec{S} = \iint_{S_a} \nabla \times \vec{v} \cdot \mathbf{n}\, dS \approx (\nabla \times \vec{v})(P_0) \cdot \mathbf{n}(P_0)\, \pi a^2$$

Taking the limit as $a \to 0$, we define the curl as circulation density:

$$(\nabla \times \vec{v})(P_0) \cdot \vec{n}(P_0) = \lim_{a \to 0} \frac{1}{\pi a^2} \oint_{C_a} \vec{v} \cdot d\vec{r}$$

This equation tells us that $\nabla \times \vec{v} \cdot \vec{n}$ measures the local rotational tendency of the fluid about the axis defined by $\vec{n}$. The more the field "swirls" around that point, the greater the value.

Stokes' theorem can also be used to prove that if a vector field has zero curl everywhere on a simply-connected domain, then it is conservative.

If $\nabla \times \vec{F} = \vec{0}$ throughout a region $\mathbb{R}^3$, then

$$\oint_C \vec{F} \cdot d\vec{r} = \iint_S \nabla \times \vec{F} \cdot d\vec{S} = \iint_S \vec{0} \cdot d\vec{S} = 0$$

for every closed curve C. Therefore, $\vec{F}$ is conservative on the domain. Thus, we can break any non-simple curve into a number of more simple curves where the integrals around each simple curve are 0.

EXAMPLE 10.10

A liquid is swirling around in a cylindrical container of radius 2, so that its motion is described by the velocity field

$$\vec{F}(x, y, z) = -y\sqrt{x^2+y^2}\,\mathbf{i} + x\sqrt{x^2+y^2}\,\mathbf{j}$$

as shown in the figure. Find $\iint_S (\nabla \times \vec{F}) \cdot \mathbf{N}\, dS$ where S is the upper surface of the cylindrical container.

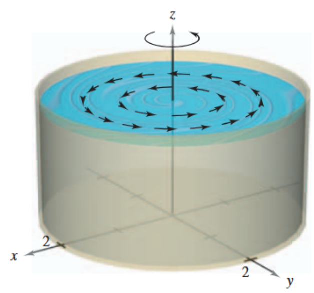


Image credit: Larson & Edwards

Solution:
The curl of $\vec{F}$ is given by

$$\nabla \times \vec{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ -y\sqrt{x^2+y^2} & x\sqrt{x^2+y^2} & 0 \end{vmatrix} = 3\sqrt{x^2+y^2}\,\mathbf{k}$$

Letting $\mathbf{N} = \mathbf{k}$, we compute the surface integral:

$$\begin{aligned} \iint_S (\nabla \times \vec{F}) \cdot \mathbf{N}\, dS &= \iint_R 3\sqrt{x^2+y^2}\, dA \\ &= \int_0^{2\pi} \int_0^2 3r \cdot r\, dr\, d\theta \\ &= \int_0^{2\pi} \int_0^2 3r^2\, dr\, d\theta \\ &= \int_0^{2\pi} \left[r^3\right]_0^2 d\theta \\ &= \int_0^{2\pi} 8\, d\theta \\ &= 16\pi \end{aligned}$$

10.4 The Divergence Theorem

Previously, we rewrote Green's theorem in a vector form as

$$\oint_C \vec{F} \cdot \mathbf{n}\, ds = \iint_D \operatorname{div} \vec{F}(x, y)\, dA,$$

where C is the positively oriented boundary of a planar region D. This relates the flux across a curve to the divergence inside the region.

Extending this idea to vector fields in $\mathbb{R}^3$, we are led to the divergence theorem:

$$\iint_S \vec{F} \cdot \mathbf{n}\, dS = \iiint_E \operatorname{div} \vec{F}(x, y, z)\, dV,$$

where S is the closed boundary surface of the solid region E, and the orientation of S is outward.

Let E be a solid region with a closed, orientable boundary surface S, and let $\vec{F}$ be a vector field whose components have continuous partial derivatives. If $\mathbf{n}$ is the outward unit normal vector on S, then

$$\iint_S \vec{F} \cdot \mathbf{n}\, dS = \iiint_E \nabla \cdot \vec{F}\, dV = \iiint_E \operatorname{div}\, \vec{F}\, dV.$$

The divergence theorem states that, under the appropriate conditions, the total flux of $\vec{F}$ across the closed surface S is equal to the triple integral of $\nabla \cdot \vec{F}$ over the solid region E enclosed by S. Let's prove this.

Let $\vec{F} = P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}$, so that

$$\nabla \cdot \vec{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

Then

$$\iiint_E \nabla \cdot \vec{F}\, dV = \iiint_E \frac{\partial P}{\partial x}\, dV + \iiint_E \frac{\partial Q}{\partial y}\, dV + \iiint_E \frac{\partial R}{\partial z}\, dV.$$

On the other hand,

$$\iint_S \vec{F} \cdot \mathbf{n}\, dS = \iint_S (P\,\mathbf{i} + Q\,\mathbf{j} + R\,\mathbf{k}) \cdot \mathbf{n}\, dS = \iint_S P\,(\mathbf{i} \cdot \mathbf{n})\, dS + \iint_S Q\,(\mathbf{j} \cdot \mathbf{n})\, dS + \iint_S R\,(\mathbf{k} \cdot \mathbf{n})\, dS.$$

To prove the theorem, we must verify that the following hold true:

$$\begin{aligned}
\iint_S P\,(\mathbf{i} \cdot \mathbf{n})\, dS &= \iiint_E \frac{\partial P}{\partial x}\, dV, \\
\iint_S Q\,(\mathbf{j} \cdot \mathbf{n})\, dS &= \iiint_E \frac{\partial Q}{\partial y}\, dV, \\
\iint_S R\,(\mathbf{k} \cdot \mathbf{n})\, dS &= \iiint_E \frac{\partial R}{\partial z}\, dV.
\end{aligned}$$

We will test the equation for the z-component. Suppose E is a type I solid bounded below by $z = u_1(x, y)$ and above by $z = u_2(x, y)$, with projection D onto the xy-plane. Then, by the fundamental theorem of calculus,

$$\iiint_E \frac{\partial R}{\partial z}\, dV = \iint_D \int_{u_1(x,y)}^{u_2(x,y)} \frac{\partial R}{\partial z}\, dz\, dA = \iint_D \left[R(x, y, u_2(x, y)) - R(x, y, u_1(x, y))\right] dA.$$

Now consider the surface integral

$$\iint_S R\,(\mathbf{k} \cdot \mathbf{n})\, dS = \iint_{S_1} R\,(\mathbf{k} \cdot \mathbf{n})\, dS + \iint_{S_2} R\,(\mathbf{k} \cdot \mathbf{n})\, dS + \iint_{S_3} R\,(\mathbf{k} \cdot \mathbf{n})\, dS.$$

On the vertical side surface S_3, $\mathbf{n}$ is horizontal, so $\mathbf{k} \cdot \mathbf{n} = 0 \Rightarrow \iint_{S_3} R\,(\mathbf{k} \cdot \mathbf{n})\, dS = 0$.

On the top surface S_2, where $z = u_2(x, y)$ and $\mathbf{n}$ points upward to align with $\mathbf{k}$, we have

$$\iint_{S_2} R\,(\mathbf{k} \cdot \mathbf{n})\, dS = \iint_D R(x, y, u_2(x, y))\, dA.$$

On the bottom surface S_1, where $z = u_1(x, y)$ and $\mathbf{n}$ points downward, we apply a negative sign:

$$\iint_{S_1} R\,(\mathbf{k} \cdot \mathbf{n})\, dS = -\iint_D R(x, y, u_1(x, y))\, dA.$$

Thus, combining all:

$$\iint_S R\,(\mathbf{k}\cdot\mathbf{n})\,dS = \iint_D \left[R(x,y,u_2(x,y)) - R(x,y,u_1(x,y))\right] dA,$$

which matches the triple integral expression, proving that

$$\iint_S R\,(\mathbf{k}\cdot\mathbf{n})\,dS = \iiint_E \frac{\partial R}{\partial z}\,dV.$$

EXAMPLE 10.11

Find the flux of the vector field $\vec{F}(x,y,z) = z\,\mathbf{i}+y\,\mathbf{j}+x\,\mathbf{k}$ over the unit sphere $x^2+y^2+z^2=1$.

Solution: First, compute the divergence of $\vec{F}$:

$$\operatorname{div}\vec{F} = \frac{\partial}{\partial x}(z) + \frac{\partial}{\partial y}(y) + \frac{\partial}{\partial z}(x) = 0+1+0 = 1.$$

The unit sphere S is the boundary of the unit ball B, given by

$$B = \left\{(x,y,z)\in\mathbb{R}^3 \mid x^2+y^2+z^2 \le 1\right\}.$$

By the divergence theorem,

$$\iint_S \vec{F}\cdot d\vec{S} = \iiint_B \operatorname{div}\vec{F}\,dV = \iiint_B 1\,dV = V(B) = \frac{4}{3}\pi(1)^3 = \frac{4\pi}{3}.$$

EXAMPLE 10.12

Evaluate $\iint_S \vec{F}\cdot d\vec{S}$, where

$$\vec{F}(x,y,z) = xy\,\mathbf{i} + (y^2 + e^{xz^2})\,\mathbf{j} + \sin(xy)\,\mathbf{k},$$

and S is the closed surface bounding the solid region E, enclosed by the parabolic cylinder $z = 1 - x^2$ and the planes $z = 0$, $y = 0$, and $y + z = 2$.

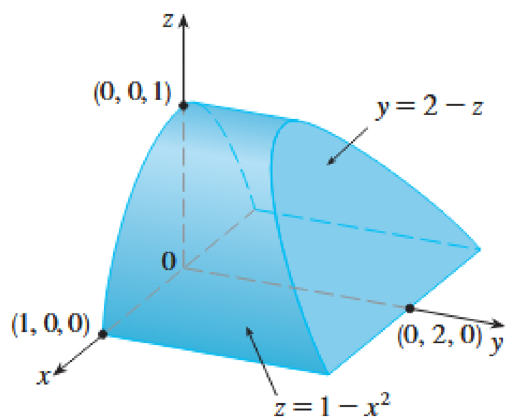


Image credit: Stewart

Solution:

The divergence of $\vec{F}$ is

$$\operatorname{div}\vec{F} = \frac{\partial}{\partial x}(xy) + \frac{\partial}{\partial y}(y^2 + e^{xz^2}) + \frac{\partial}{\partial z}(\sin(xy)) = y + 2y + 0 = 3y.$$

By the divergence theorem:

$$\iint_S \vec{F}\cdot d\vec{S} = \iiint_E 3y\,dV.$$

Express E as the region $E = \left\{(x,y,z)\;\middle|\; -1 \le x \le 1,\ 0 \le z \le 1 - x^2,\ 0 \le y \le 2 - z\right\}$:

$$\iiint_E 3y\,dV = 3\int_{-1}^{1}\int_0^{1-x^2}\int_0^{2-z} y\,dy\,dz\,dx = 3\int_{-1}^{1}\int_0^{1-x^2}\frac{(2-z)^2}{2}\,dz\,dx$$

$$= \frac{3}{2}\int_{-1}^{1}\int_0^{1-x^2}(4 - 4z + z^2)\,dz\,dx = \int_0^1 (x^6 + 3x^4 + 3x^2 - 7)\,dx = \frac{184}{35}$$

Although we originally proved the divergence theorem for simple solid regions, it can be extended to finite unions of such regions. Let E be the region bounded between closed surfaces S_1 and S_2, where S_1 lies inside S_2. Let $\mathbf{n}_1$ and $\mathbf{n}_2$ be outward-pointing normals to S_1 and S_2, respectively. Then the boundary of E is $S = S_1 \cup S_2$ with $\mathbf{n} = -\mathbf{n}_1$ on S_1, and $\mathbf{n} = \mathbf{n}_2$ on S_2. This ensures that $\mathbf{n}$ always points outward from E.

Applying the divergence theorem to S, we obtain:

$$\iiint_E \nabla \cdot \vec{F}\, dV = \iint_S \vec{F} \cdot \mathbf{n}\, dS = \iint_{S_1} \vec{F} \cdot (-\mathbf{n}_1)\, dS + \iint_{S_2} \vec{F} \cdot \mathbf{n}_2\, dS = -\iint_{S_1} \vec{F} \cdot dS + \iint_{S_2} \vec{F} \cdot dS.$$

This is useful when the interior surface S_1 is simpler to compute than the total surface S_2.

Say we are given the electric field due to a point charge at the origin:

$$\vec{E}(\vec{x}) = \frac{\varepsilon Q}{|\vec{x}|^3}\, \vec{x}, \text{ where } \vec{x} = \langle x, y, z \rangle.$$

Let S_2 be any closed surface enclosing the origin. We want to compute the flux of $\vec{E}$ across S_2. Since $\vec{E}$ is undefined only at the origin, we consider a small sphere S_1 of radius a centered at the origin, entirely inside S_2. Let D be the region bounded between S_1 and S_2.

By the divergence theorem,

$$\iint_{S_2} \vec{E} \cdot \mathbf{n}\, dS = \iint_{S_1} \vec{E} \cdot \mathbf{n}\, dS + \iiint_D \nabla \cdot \vec{E}\, dV.$$

Since $\nabla \cdot \vec{E} = 0$ in D everywhere except the origin, the triple integral vanishes:

$$\iint_{S_2} \vec{E} \cdot d\vec{S} = \iint_{S_1} \vec{E} \cdot \mathbf{n}\, dS.$$

Now compute the flux through the sphere S_1. At any point $\vec{x}$ on the sphere, the outward unit normal is

$$\mathbf{n} = \frac{\vec{x}}{\|\vec{x}\|}.$$

Then,

$$\vec{E} \cdot \mathbf{n} = \frac{\varepsilon Q}{\|\vec{x}\|^3} \vec{x} \cdot \frac{\|\vec{x}\|}{\|\vec{x}\|} = \frac{\varepsilon Q}{\|\vec{x}\|^2}.$$

Since the sphere has radius a, this simplifies to:

$$\vec{E} \cdot \mathbf{n} = \frac{\varepsilon Q}{a^2}.$$

The flux across S_1 becomes:

$$\iint_{S_1} \vec{E} \cdot d\vec{S} = \frac{\varepsilon Q}{a^2} \cdot A(S_1) = \frac{\varepsilon Q}{a^2} \cdot 4\pi a^2 = 4\pi\varepsilon Q.$$

This result shows that the total flux through any closed surface enclosing the origin depends only on the total charge Q, not on the shape of the surface. This is a direct consequence of the inverse-square nature of the electric field and symmetry. The contributions from different parts of the surface "balance out" to give a consistent total.

This is a special case of Gauss' law. If $\varepsilon = \frac{1}{4\pi\varepsilon_0}$, then this becomes the familiar

$$\iint_{S_2} \vec{E} \cdot d\vec{S} = \frac{Q}{\varepsilon_0}.$$

We will now revisit a previous idea now that we have more mathematical intuition.

Let $\vec{v}(x, y, z)$ be a velocity field and ρ a constant fluid density. Then $\vec{F} = \rho\vec{v}$ represents the rate of flow per unit area. Consider a small ball B_a of radius a centered at point $P_0 = (x_0, y_0, z_0)$. We can assume that div$\vec{F} \approx$ div $\vec{F}(P_0)$ for all $P \in B_a$ since div $\vec{F}$ is continuous. Then we approximate flux over the boundary sphere S_a as

$$\iint_{S_a} \vec{F} \cdot d\vec{S} = \iiint_{B_a} \nabla \cdot \vec{F}\, dV \approx \iiint_{B_a} \nabla \cdot \vec{F}(P_0)\, dV = \nabla \cdot \vec{F}(P_0)\, V(B_a)$$

As $a \to 0$, the approximation becomes better and suggests that

$$\nabla \cdot \vec{F}(P_0) = \lim_{a \to 0} \frac{1}{V(B_a)} \iint_{\partial B_a} \vec{F} \cdot \mathbf{n}\, dS.$$

This means that divergence measures the net outward flow per unit volume at a point. If $\nabla \cdot \vec{F}(P) > 0$, more flow is exiting than entering. Thus, P is a source. If $\nabla \cdot \vec{F}(P) < 0$m more flow is entering than exiting. Thus, P is a sink. If $\nabla \cdot \vec{F}(\vec{P}) = 0$, the flow is locally incompressible.

For instance, let $\vec{F}(x, y) = x^2\,\mathbf{i} + y^2\,\mathbf{j}$. Then

$$\nabla \cdot \vec{F} = \frac{\partial}{\partial x}(x^2) + \frac{\partial}{\partial y}(y^2) = 2x + 2y.$$

At $x + y > 0$, we have $\nabla \cdot \vec{F} > 0$. This is a source.

At$x + y < 0$, we have $\nabla \cdot \vec{F} < 0$. This is a sink.

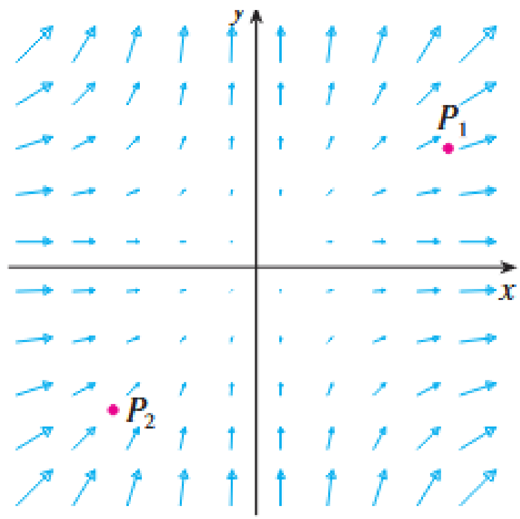


Image credit: Stewart

As you can see in the vector field, the vectors that end near P_1 are shorter than the vectors that start near P_1. Thus, the net flow is outward around P_1 and thus a source. Near P_2, the incoming vectors are longer and coming in, therefore this represents net inward flow.

Divergence measures a field's tendency to diverge from or converge towards a point in space. And the divergence theorem expands out and tells us that the total field's total expansion is equivalent to the net flux outward. It is a unification of the microscopic and the macroscopic perspectives that helps us represent the laws of our universe.

Fundamental Theorem	Statement
Fundamental Theorem of Calculus	$$\int_a^b F'(x)\,dx = F(b) - F(a)$$ Relates the rate of change of a scalar function to its net change over an interval
Fundamental Theorem for Line Integrals	$$\int_C \nabla f \cdot d\vec{r} = f(\vec{r}(b)) - f(\vec{r}(a))$$ Evaluates a line integral using only the values of the scalar field at the endpoints
Green's Theorem	$$\iint_D \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dA = \oint_C P\,dx + Q\,dy$$ Relates a double integral over a region to a line integral around its boundary
Stokes' Theorem	$$\iint_S (\nabla \times \vec{F}) \cdot d\vec{S} = \oint_C \vec{F} \cdot d\vec{r}$$ Relates the surface integral of curl to the circulation around the boundary curve
Divergence Theorem	$$\iiint_E (\nabla \cdot \vec{F})\,dV = \iint_S \vec{F} \cdot d\vec{S}$$ Relates the total divergence inside a region to the net outward flux across its boundary

Bibliography

- Briggs, Williams et al. *Calculus: Early Transcendentals*.
- Applied Linear Algebra, Jeff Anderson.
  https://www.appliedlinearalgebra.com
- HELM Workbooks, Loughborough University.
  https://www.lboro.ac.uk/departments/mlsc/student-resources/helm-workbooks/ *
- Larson, Ron and Edwards, Bruce. *Calculus: Early Transcendentals*.
- *M408M Multivariable Calculus Page*. UT Austin.
  https://web.ma.utexas.edu/users/m408m/Display.html
- *Math 215 Demonstrations*. UMich.
  https://instruct.math.lsa.umich.edu/lecturedemos/ma215/docs/
- Spivak, Michael. *Calculus*.
- Stewart, James. *Calculus*.
- Strang, Gilbert and Herman, Edward. *Calculus*. https://math.libretexts.org/Bookshelves/Calculus/Calculus_(OpenStax)