The Penn Calc Companion

Part I: Functions and Derivatives

About this Document

Contents

1 Functions

A *function* can be visualized as a machine that takes in an input x and returns an output $f(x)$. The collection of all possible inputs is called the *domain*, and the collection of all possible outputs is called the *range*.

This course deals with functions whose domains and ranges are $\mathbb{R}$ or subsets of $\mathbb{R}$ (this is the notation for the real numbers).

1.1 Examples

Example

1. Polynomials, e.g. $f(x) = x^3 - 5x^2 + x + 9$. Give the domain and range of f. (See Answer 1)
2. Trigonometric functions, e.g. sin, cos, tan. Give the domain and range for each of these. (See Answer 2)
3. The exponential function, e^x. Give the domain and range for the exponential. (See Answer 3)
4. The natural logarithm function, $\ln x$. Recall that this is the inverse of the exponential function. Give the domain and range for $\ln x$. (See Answer 4)
5. Is $\sin^{-1}$ a function? If so, why? If not, is there a way to make it into a function? (See Answer 5)

1.2 Operations on Functions

Composition

The *composition* of two functions, f and g, is defined to be the function that takes as its input x and returns as its output $g(x)$ fed into f.

$$f \circ g(x) = f(g(x))$$

Example

$$\sqrt{1 - x^2}$$

can be thought of as the composition of two functions, f and g. If $g = 1 - x^2$, f would be the function that takes an input $g(x)$ and returns its square root.

Example

Compute the composition $f \circ f$, i.e. the composition of f with itself, where $f(x) = \dfrac{1}{x+1}$. (See Answer 6)

Inverse

The *inverse* is the function that undoes f. If you plug $f(x)$ into f^{-1}, you will get x. Notice that this function works both ways. If you plug $f^{-1}(x)$ into $f(x)$, you will get back x again.

$$f^{-1}(f(x)) = x$$

$$f(f^{-1}(x)) = x$$

NOTE: f^{-1} denotes the inverse, not the reciprocal. $f^{-1}(x) \neq \frac{1}{f(x)}$.

Example

Let's consider $f(x) = x^3$. Its inverse is $f^{-1}(x) = x^{\frac{1}{3}}$.

$$f^{-1}(f(x)) = (x^3)^{\frac{1}{3}} = x$$

$$f(f^{-1}(x)) = (x^{\frac{1}{3}})^3 = x$$

1.3 Classes of Functions

Polynomials

A polynomial $P(x)$ is a function of the form

$$P(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

The top power n is called the degree of the polynomial. We can also write a polynomial using a summation notation.

$$P(x) = \sum_{k=0}^{n} c_k x^k$$

Rational functions

Rational functions are functions of the form $\dfrac{P(x)}{Q(x)}$ where each is a polynomial.

Example

$$\frac{3x-1}{x^2+x-6}$$

is a rational function. You have to be careful of the denominator. When the denominator takes a value of zero, the function may not be well-defined.

Powers

Power functions are functions of the form cx^n, where c and n are constant real numbers.

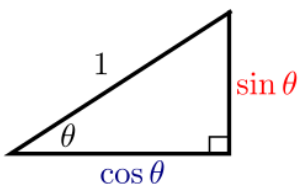
Other powers besides those of positive integers are useful.

Example

- What is x^0 ? (See Answer 7)
- What is $x^{-\frac{1}{2}}$? (See Answer 8)
- What is $x^{\frac{22}{7}}$? (See Answer 9)
- What is x^{π} ? We are not yet equipped to handle this, but we will come back to it later.

Trigonometrics

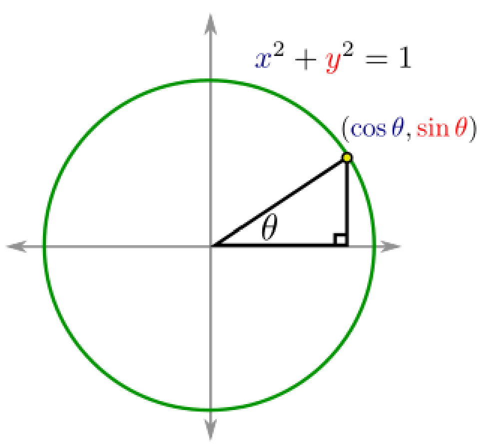
You should be familiar with the basic trigonometric functions sin, cos. One fact to keep in mind is $\cos^2\theta + \sin^2\theta = 1$ for any θ. This is known as a *Pythagorean identity*, which is so named because of one of the ways to prove it:



$$\cos^2\theta + \sin^2\theta = 1$$

By looking at a right triangle with hypotenuse 1 and angle θ, and labeling the adjacent and opposite sides accordingly, one finds by using Pythagoras' Theorem that $\cos^2\theta + \sin^2\theta = 1$.

Another way to think about it is to embed the above triangle into a diagram for the unit circle where we see that $\cos\theta$ and $\sin\theta$ returns the x and y coordinates, respectively, of a point on the unit circle with angle θ to the x-axis:



That explains the nature of the formula $\cos^2\theta + \sin^2\theta = 1$. It comes from the equation of the unit circle $x^2 + y^2 = 1$.

Other trigonometric functions:

- $\tan = \dfrac{\sin}{\cos}$
- $\cot = \dfrac{\cos}{\sin}$, the reciprocal of tan
- $\sec = \dfrac{1}{\cos}$, the reciprocal of cos
- $\csc = \dfrac{1}{\sin}$, the reciprocal of sin

All four of these have vertical asymptotes at the points where the denominator goes to zero.

Inverse Trigonometrics

We often write $\sin^{-1}$ to denote the inverse, but this can cause confusion. Be careful that $\sin^{-1} \neq \dfrac{1}{\sin}$. To avoid the confusion, the terminology arcsin is recommended for the inverse of the sin function.

The arcsin function takes on values $\left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$ and has a restricted domain $[-1, 1]$.

The arccos function likewise has a restricted domain $[-1, 1]$, but it takes values $[0, \pi]$.

The arctan function has an unbounded domain, it is well defined for all inputs. But it has a restricted range $\left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$.

Exponentials

Exponential functions are of the form c^x, where c is some positive constant. The most common such function, referred to as *the* exponential, is e^x. This is the most common because of its nice integral and differential properties (below).

Algebraic properties of the exponential function:

$$e^x e^y = e^{x+y}$$

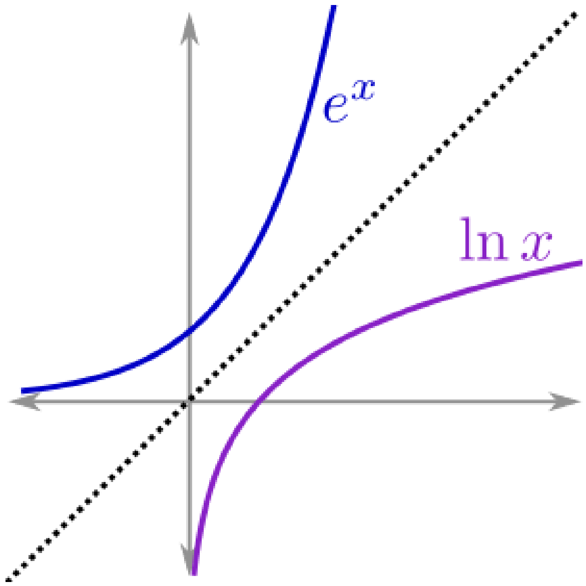
$$(e^x)^y = e^{xy}$$

Differential/integral properties:

$$\frac{d}{dx} e^x = e^x$$

$$\int e^x dx = e^x + C$$

Recall the graph of e^x, plotted here alongside its inverse, $\ln x$:



Note that the graphs are symmetric about the line $y = x$ (as is true of the graphs of a function and its inverse).

Before continuing, one might ask, what is e? There are several ways to define e, which will be revealed soon. For now, it is an irrational number which is approximately 2.718281828.

1.4 Euler's Formula

To close this lesson, we give a wonderful formula, which for now we will just take as a fact:

Euler's Formula

$$e^{ix} = \cos x + i \sin x$$

The i in the exponent is the imaginary number $\sqrt{-1}$. It has the properties $i^2 = -1$. i is not a real number. That doesn't mean that it doesn't exist. It just means it is not on a real number line.

Euler's formula concerns the exponentiation of an imaginary variable. What exactly does that mean? How is this related to trigonometric functions? This will be covered in our next lesson.

1.5 Additional Examples

Example

Find the domain of

$$f(x) = \frac{1}{\sqrt{x^2 - 3x + 2}}.$$

(See Answer 10)

Example

Find the domain of

$$f(x) = \ln(x^3 - 6x^2 + 8x).$$

(See Answer 11)

1.6 Answers to Selected Examples

1. The domain is $\mathbb{R}$, because we can plug in any real number into a polynomial. The range is $\mathbb{R}$, which we see by noting that this is a cubic function, so as $x \to -\infty$, $f(x) \to -\infty$, and as $x \to \infty$, $f(x) \to \infty$.

   (Return)

2. For sin and cos: domain is $\mathbb{R}$; range is $[-1, 1]$.

   For tan, the domain is $\{x \in \mathbb{R} : x \neq \frac{\pi}{2} + k\pi\}$; range is $\mathbb{R}$.

   (Return)

3. Domain is $\mathbb{R}$; range is $(0, \infty)$.

   (Return)

4. Domain is $(0, \infty)$; range is $\mathbb{R}$. Notice how the domain and range of the exponential relate to the domain and range of the natural logarithm.

   (Return)

5. $\sin^{-1}$ is not a function, because one input has many outputs. For example, $\sin^{-1}(0) = 0, \pi, 2\pi, \ldots$. By restricting the range of $\sin^{-1}$ to $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, one gets the function arcsin.

   (Return)

6. We find that

$$\begin{aligned} f \circ f(x) &= f(f(x)) \\ &= f\left(\frac{1}{x+1}\right) \\ &= \frac{1}{1/(x+1)+1} \\ &= \frac{x+1}{1+x+1} \\ &= \frac{x+1}{x+2}. \end{aligned}$$

   (Return)

7. $x^0 = 1$

   (Return)

8. Recall a fractional power denotes root. For example, $x^{\frac{1}{2}} = \sqrt{x}$. The negative sign in the exponent means that we take the reciprocal. So, $x^{-\frac{1}{2}} = \frac{1}{\sqrt{x}}$.

   (Return)

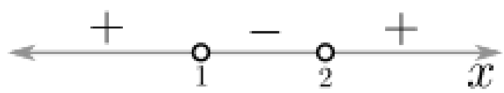
9. One can rewrite this as $\left(x^{22}\right)^{1/7}$. That means we take x to the 22nd power and then take the 7th root of the result. $x^{\frac{22}{7}} = \sqrt[7]{x^{22}}$

   (Return)

10. Note that the square root is only defined when its input is non-negative. Also, the denominator in a rational function cannot be 0. So we find that this function is well-defined if and only if $x^2 - 3x + 2 > 0$. Factoring gives

$$(x-2)(x-1) > 0.$$

   By plotting the points $x = 1$ and $x = 2$ (where the denominator equals 0) and testing points between them, one finds that $x^2 - 3x + 2 > 0$ when $x < 1$ or $x > 2$:



   So the domain of f is $x < 1$ or $2 < x$. In interval notation, this is $(-\infty, 1) \cup (2, \infty)$.

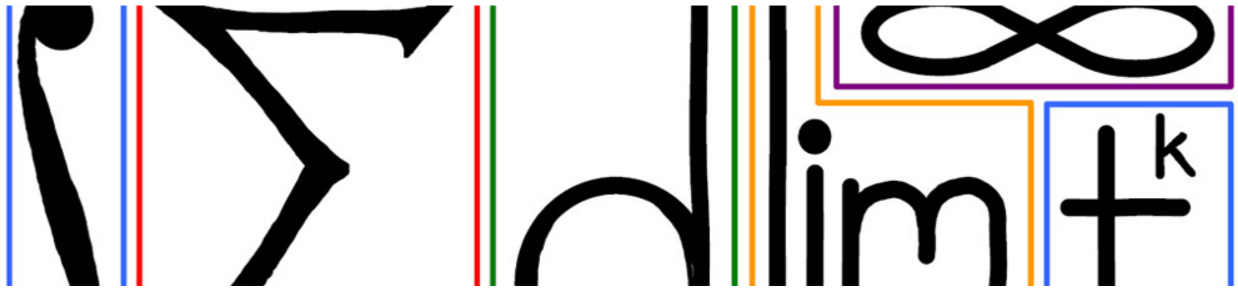
   (Return)

11. Since ln is only defined on the positive real numbers, we must have $x^3 - 6x^2 + 8x > 0$. Factoring gives

$$x(x^2 - 6x + 8) = x(x-2)(x-4) > 0$$

   As in the above example, plotting the points where this equals 0 and then testing points, we find that the domain is $0 < x < 2$ and $4 < x$. In interval notation, this is $(0, 2) \cup (4, \infty)$.

   (Return)



2 The Exponential

This module deals with a very important function: the exponential. The first question one might ask is: what *is* the exponential function e^x? We know certain values of the function such as $e^0 = 1$, but what about an irrational input such as e^π, or an imaginary input e^i? Is it possible to make sense of these values?

The following definition answers these questions.

The Exponential e^x

$$\begin{aligned} e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots \\ &= \sum_{k=0}^{\infty} \frac{x^k}{k!}, \end{aligned}$$

where

$$k! = k(k-1)(k-2)\cdots 3 \cdot 2 \cdot 1,$$

and $0! = 1$.

One can now plug values for x into the above sum to compute e^x. When $x = 0$, for instance, one finds that $e^0 = 1$, (since all the terms with x disappear) as expected. By plugging in $x = 1$, the true value of e is found to be $e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \cdots$.

2.1 A long polynomial

There are technical concerns when trying to add up an infinite number of things. These issues will be dealt with later in the modules on series. For now, treat the infinite sum above as a long polynomial (the actual term is the *Taylor series about* $x = 0$, which will be more formally dealt with in the next module). Polynomials are nice because they are easy to integrate and differentiate. Recall the power rule for differentiating and integrating a monomial x^k, where k is a constant:

$$\begin{aligned} \frac{d}{dx} x^k &= kx^{k-1} \\ \int x^k \, dx &= \frac{1}{k+1} x^{k+1} + C \quad (k \neq -1) \end{aligned}$$

2.2 Properties of e^x

Recall the following properties of the exponential function:

1. $e^{x+y} = e^x e^y$
2. $e^{x \cdot y} = (e^x)^y = (e^y)^x$
3. $\frac{d}{dx} e^x = e^x$
4. $\int e^x dx = e^x + C$.

Consider the last two properties in terms of the long polynomial.Taking the derivative of the long polynomial for e^x gives

$$\begin{aligned}\frac{d}{dx}\left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots\right) &= 0 + 1 + \frac{2x}{2!} + \frac{3x^2}{3!} + \frac{4x^3}{4!} + \cdots \\ &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots,\end{aligned}$$

which is the original long polynomial. Integrating also gives (up to the constant of integration) the original long polynomial. This agrees with facts about the derivative and integral of e^x. Thus, the long polynomial for e^x captures two of the key features of e^x; namely, e^x is its own derivative and its own integral.

2.3 Euler's formula

Recall that the imaginary number i is defined by $i = \sqrt{-1}$. So $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, and this continues cyclically (for a review of complex/imaginary numbers, see wikipedia). Assume the following fact, known as Euler's formula, mentioned in the last module.

Euler's formula

$$e^{ix} = \cos x + i \sin x.$$

Consider what happens when ix is plugged into the long polynomial for e^x. By simplifying the powers of i, and grouping the result into its real and imaginary parts, one finds

$$\begin{aligned}e^{ix} &= 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots \\ &= 1 + ix + \frac{i^2x^2}{2!} + \frac{i^3x^3}{3!} + \cdots \\ &= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} + \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots\right).\end{aligned}$$

If this is supposed to equal $\cos x + i \sin x$, then the real part must be $\cos x$, and the imaginary part must be $\sin x$. It follows that

$$\begin{aligned}\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \\ \sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.\end{aligned}$$

These formulas should be memorized, both in their long polynomial form and their more concise summation notation form.

Example

Use Euler's formula to show that $e^{i\pi} = -1$. (See Answer 1)

Example

Compute $1 - \frac{\pi^2}{2!} + \frac{\pi^4}{4!} - \cdots$. (See Answer 2)

Example

Check that taking the derivative of the long polynomial for $\sin x$ gives the long polynomial for $\cos x$ (hence, verify that $\frac{d}{dx} \sin x = \cos x$). (See Answer 3)

Example

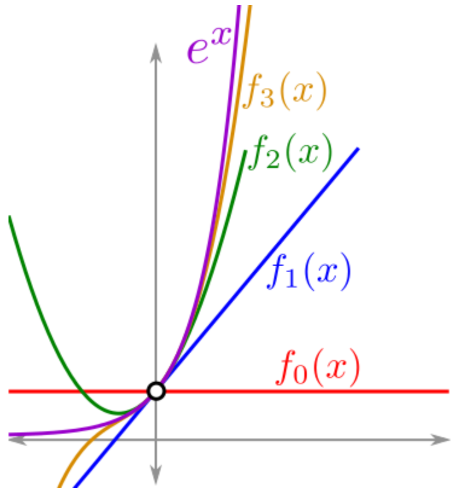
Show that the long polynomial for e^x satisfies the first property above, namely $e^{x+y} = e^x e^y$. Hint: start with the long polynomials for e^x and e^y and multiply these together, and carefully collect like terms to show it equals the long polynomial for e^{x+y}. (See Answer 4)

2.4 More on the long polynomial

The idea of a long polynomial is reasonable, because it actually comes from taking a sequence of polynomials with higher and higher degree:

$$\begin{aligned} f_0(x) &= 1 \\ f_1(x) &= 1 + x \\ f_2(x) &= 1 + x + \frac{x^2}{2} \\ f_3(x) &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} \\ &\vdots \end{aligned}$$

Each polynomial in the sequence is, in a sense, the best approximation possible of that degree. Put another way, taking the first several terms of the long polynomial gives a good polynomial approximation of the function. The more terms included, the better the approximation. This is how calculators compute the exponential function (without having to add up infinitely many things).



2.5 EXERCISES

- So, how good of an approximation is a polynomial truncation of e^x? Use a calculator to compare how close e is to the linear, quadratic, cubic, quartic, and quintic approximations. How many digits of accuracy do you seem to be gaining with each additional term in the series?
- Now, do the same thing with $1/e$ by plugging in $x = -1$ into the series. Do you have the same results? Are you surprised?
- Use the first three terms of the series for e^x to approximate $\sqrt[10]{e}$ and e^{10}. How accurate do you think these approximations are?
- Calculate the following sum using what you know:

$$\sum_{n=0}^{\infty}(-1)^n\frac{(\ln 3)^n}{n!}$$

- Write out the first four terms of the following series

$$\sum_{n=0}^{\infty}(-1)^n\frac{\pi^{2n}}{2^n n!}$$

- Write out the following series using summation notation:

$$1-\frac{2}{3!}+\frac{4}{5!}-\frac{8}{7!}+\cdots$$

- Estimate $\sin(1/2)$ to three digits of accuracy. How many terms in the series did this take?
- We've seen that $i = e^{i\pi/2}$ via Euler's formula. Using this and some algebra, tell me what is i^i. Isn't that nice? Now, tell me, what is $(i^i)^i$? Are you surprised? That's like, unreal!
- Practice your summation notation by rewriting the sum

$$\sum_{n=2}^{\infty}(-1)^n\frac{x^{n-2}}{n^3}$$

as a sum over an index that goes from zero to infinity.

- Use the first two nonzero terms of the Taylor series for $\cos(x)$ to approximate $\cos(\frac{1}{10})$.
- Use Euler's formula to derive the double angle formulas $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$ and $\sin(2\theta) = 2\sin(\theta)\cos(\theta)$.

2.6 Answers to Selected Examples

1. Setting $x = \pi$ in Euler's formula gives $e^{i\pi} = \cos\pi + i\sin\pi = -1$.

   (Return)

2. Note that this is the long polynomial for $\cos x$, evaluated at $x = \pi$. So the value is $\cos\pi = -1$.

   (Return)

3. Computing the derivative term by term gives

$$\begin{aligned}\frac{d}{dx}\sin(x) &= \frac{d}{dx}\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots\right)\\ &= 1 - 3\frac{x^2}{3!} + 5\frac{x^4}{5!} - \dots\\ &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots,\end{aligned}$$

   which is the long polynomial for $\cos x$, as desired.

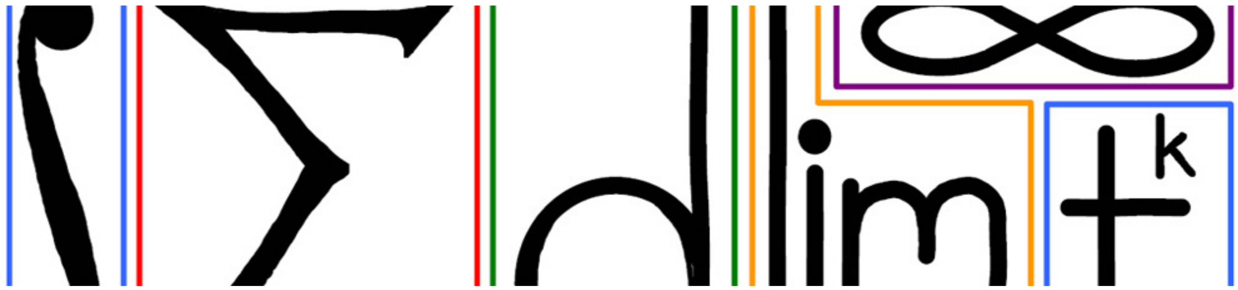
   (Return)

4. Beginning with $e^x \cdot e^y$, we find

$$\begin{aligned}e^x \cdot e^y &= \left(1 + x + \frac{x^2}{2!} + \cdots\right)\left(1 + y + \frac{y^2}{2!} + \cdots\right)\\ &= 1 + (x+y) + \left(\frac{x^2}{2!} + xy + \frac{y^2}{2!}\right) + \cdots\\ &= 1 + (x+y) + \frac{x^2 + 2xy + y^2}{2!} + \cdots\\ &= 1 + (x+y) + \frac{(x+y)^2}{2!} + \cdots,\end{aligned}$$

   which is the long polynomial for e^{x+y}, as desired.

   (Return)



3 Taylor Series

The long polynomial from the last module is actually called a Taylor series about $x = 0$ (this is referred to as a Maclaurin series in some textbooks, but this course will use the term Taylor series). The last module gave the Taylor series for e^x, $\sin x$, and $\cos x$. The logical next question is to ask whether every function has a Taylor series.

The answer is that most *reasonable* functions, and almost all of the functions encountered in this course, have a Taylor series. That is, every reasonable function f can be written as

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + \cdots .$$

This module describes how to compute the coefficients c_k for a given function f.

3.1 The definition of a Taylor series at x=0

The definition of the Taylor series of f at $x = 0$ is

Taylor series at $x = 0$

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!}x^k,$$

where $f^{(k)}(0)$ is the kth derivative of f evaluated at 0. In other words, the coefficient c_k mentioned above is given by

$$c_k = \frac{f^{(k)}(0)}{k!} = \frac{1}{k!} \cdot \left.\frac{d^k f}{dx^k}\right|_0$$

This seems circular, since the definition uses the function, and its derivatives, to write down the function. However, the definition only actually requires information about the function at a single point (in this case, 0). It is best to think of the Taylor series as a way of turning a function into a polynomial.

Example

Compute the Taylor series for e^x using the above definition to see that it matches the given series from the last module. (See Answer 1)

Example

Compute the Taylor series for $f(x) = \sin x$ using the above definition, and verify it matches the series found using Euler's formula. (See Answer 2)

Example

Compute the Taylor series for $f(x) = x^2 - 5x + 3$. (See Answer 3)

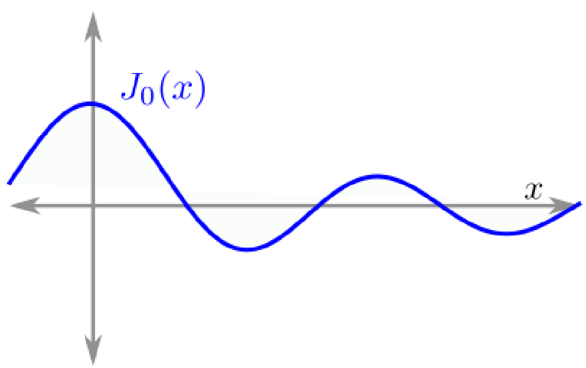
3.2 Why Taylor series matter

The big idea of this module is that the Taylor series can be thought of as an operator (a machine) which turns a function into a series. This is a useful operator because some functions are hard (or even impossible) to express using combinations of familiar functions. Nevertheless, these functions can often be understood by computing their Taylor series.

Example The *Bessel function*, denoted J_0, is best defined by its Taylor series:

$$\begin{aligned} J_0 &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2^{2k}(k!)^2} \\ &= 1 - \frac{1}{2^2}x^2 + \frac{1}{2^4(2!)^2}x^4 - \frac{1}{2^6(3!)^2}x^6 + \cdots \end{aligned}$$

This series has only the even powers of x, and it alternates, which is reminiscent of the series for cosine. One difference is that the denominator in the Bessel function grows more quickly than the denominator in the series for cosine. Thus, we might expect the graph to be a wave with a decreasing amplitude, which is exactly what we find:

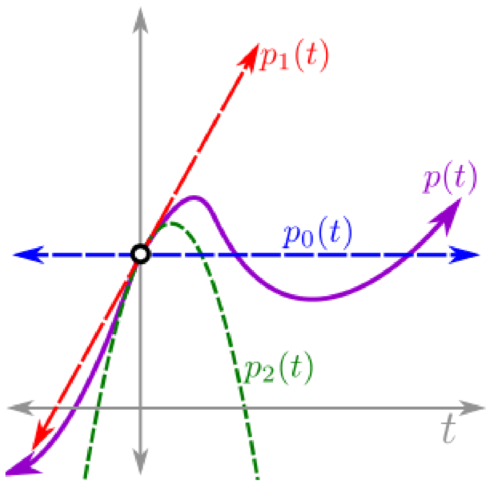


It turns out that the Bessel function describes many physical phenomena, including the shape of a hanging chain as it is rotated, and the shape of the waves formed after a stone is thrown into a pool of water.

3.3 Taylor series as polynomial approximants

The main reason Taylor series are useful is that they turn a potentially complicated function into something simple: a polynomial. Granted, this polynomial is infinitely long in general, but in practice it is only necessary to compute the first few terms to get a good, local approximation of the function. The more terms one includes, the better the polynomial approximates the function.

As an example, consider a particle on the number line with position function $p(t)$. At time 0, say its position is 5. Then one approximation of its position as a function of time is $p_0(t) = 5$. Given more information, say its velocity at time 0 is 3, the approximation becomes better. The next approximation as a function of time is $p_1(t) = 5 + 3t$. Now, suppose its acceleration at time 0 is -4. Then $p_2(t) = 5 + 3t - \frac{4}{2}t^2 = 5 + 3t - 2t^2$ is an even better polynomial approximation of the position function.



3.4 EXERCISES

- What is the Taylor series of $x^4 - 3x^3 + 2x^2 + 7x - 3$. This should be an easy one!
- What is the Taylor series of $(x-2)^2(x-3)$? This, also, should not be *too* hard...
- Compute a few derivatives and figure out the first few terms of the Taylor series of $\dfrac{1}{1-x}$. Have you seen this series before?
- What are the first two nonzero terms in the Taylor series of $\sqrt[3]{1-2x}$?
- What is the coefficient of the cubic term in the Taylor series of e^{-3x}?
- Use what you know about Taylor series to determine the third derivative of $\sin^3(2x)\cos^2(3x)$ at $x = 0$. That's a *lot* easier than computing the derivatives!
- The ERF function is defined in terms of a difficult integral:

$$ERF(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt$$

- Even if you don't remember integrals all that well, you know how to integrate a polynomial, right? So, Taylor expand the integrand and integrate term by term to get the Taylor series for ERF.
- What is the third derivative of ERF(x) at zero?
- Why does a Taylor series have all those $n!$ terms in the denominator? Let's see. Compute the Taylor series of $f(x) = (1+x)^5$ by (1) using the binomial theorem (or multiplication) to expand that power; then (2) by differentiating the function and using the Taylor series formula. What do you notice when you keep computing higher derivatives?

3.5 Answers to Selected Examples

1. Here, $f(x) = e^x$, and every derivative of e^x is e^x. Therefore, for all k we have

$$f^{(k)}(x) = e^x,$$

and so $f^{(k)}(0) = 1$ for all k. Plugging into the Taylor series formula gives

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \\ &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \\ &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots, \end{aligned}$$

as claimed.

(Return)

2. Computing the derivatives, and then evaluating at $x = 0$ gives the following table:

$$\begin{array}{rr} f(x) = \sin(x) & f(0) = 0 \\ f'(x) = \cos(x) & f'(0) = 1 \\ f''(x) = -\sin(x) & f''(0) = 0 \\ f'''(x) = -\cos(x) & f'''(0) = -1 \\ \vdots & \end{array}$$

Thus,

$$\begin{aligned} \sin(x) &= 0 + \frac{1}{1!}x + \frac{0}{2!}x^2 + \frac{-1}{3!}x^3 + \cdots \\ &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots, \end{aligned}$$

confirming what was found last time.

(Return)

3. Again, by directly using the definition:

$$\begin{array}{rr} f(x) = x^2 - 5x + 3 & f(0) = 3 \\ f'(x) = 2x - 5 & f'(0) = -5 \\ f''(x) = 2 & f''(0) = 2 \\ f'''(x) = 0 & f'''(0) = 0 \\ \vdots & \end{array}$$

So it follows that

$$f(x) = 3 - 5x + \frac{2}{2!}x^2 = 3 - 5x + x^2$$

(since all the subsequent derivatives are 0), which is the original function. This should not be a surprise, since the Taylor series represents a function as a long polynomial (henceforth called by its proper name: *series*). If f was a polynomial to begin with, it stands to reason that the Taylor series for f should just be f itself.

(Return)

4 Computing Taylor Series

The previous module gave the definition of the Taylor series for an arbitrary function. It turns out that this is not always the easiest way to compute a function's Taylor series. Just as functions can be added, subtracted, multiplied, and composed, so can their corresponding Taylor series.

Recall that the Taylor series for a function f is given by

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = f(0) + f'(0)x + \frac{f''(0)}{2!} x^2 + \cdots .$$

Using the definition of the Taylor series involves taking a lot of derivatives, which could be a lot of work, especially if the function involves compositions and products of functions, e.g. $f(x) = \sin(x^2)e^{x^3}$. This module will show how to compute the Taylor series of such functions more easily by using the Taylor series for functions we already know.

4.1 Substitution

Our first method, substitution, allows us to plug one function into the Taylor series of another. Consider the function

$$f(x) = \frac{1}{x}\sin(x^2).$$

Computing the Taylor series for f from the definition would involve the quotient rule, chain rule, and a lot of algebra. But by taking the series for $\sin x$ and substituting x^2 into this series, and then distributing the $\frac{1}{x}$, one finds

$$\begin{aligned} \frac{1}{x}\sin(x^2) &= \frac{1}{x}\left((x^2) - \frac{1}{3!}(x^2)^3 + \frac{1}{5!}(x^2)^5 - \cdots\right) \\ &= \frac{1}{x}\left(x^2 - \frac{1}{3!}x^6 + \frac{1}{5!}x^{10} - \cdots\right) \\ &= x - \frac{1}{3!}x^5 + \frac{1}{5!}x^9 - \cdots . \end{aligned}$$

Note that getting this many terms using the definition would involve taking nine derivatives of the original function, which would be a lot of work! To get a more complete description of the Taylor series, one can use the summation notation, and again substitute to find

$$\begin{aligned}
\frac{1}{x}\sin(x^2) &= \frac{1}{x}\sum_{k=0}^{\infty}(-1)^k\frac{(x^2)^{2k+1}}{(2k+1)!}\\
&= \frac{1}{x}\sum_{k=0}^{\infty}(-1)^k\frac{x^{4k+2}}{(2k+1)!}\\
&= \sum_{k=0}^{\infty}(-1)^k\frac{x^{4k+1}}{(2k+1)!}
\end{aligned}$$

Example

Find the Taylor series for e^{x^3} by substitution. (See Answer 1)

4.2 Combining like terms

Another way to use previous knowledge of one Taylor series to find another is by combining like terms. This requires some careful algebra, but it is no more difficult than multiplying two polynomials together. For example, consider the function

$$f(x) = \cos^2(x) = \cos(x)\cdot\cos(x).$$

Finding the series for a function multiplied by another function is the same as taking the series for each function and multiplying them together, and then collecting like terms. This is where some algebra is required.

$$\begin{aligned}
\cos(x)\cdot\cos(x) &= \left(1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\cdots\right)\left(1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\cdots\right)\\
&= 1+\left(-\frac{1}{2!}-\frac{1}{2!}\right)x^2+\left(\frac{1}{4!}+\frac{1}{2!}\frac{1}{2!}+\frac{1}{4!}\right)x^4+\cdots\\
&= 1-x^2+\frac{1}{3}x^4+\cdots.
\end{aligned}$$

To see where the coefficient of x^4 comes from, note that every x^4 term comes from some term from the left series multiplied together with some term from the right series whose powers add up to 4. There are three such pairs: 1 on the left paired with $\frac{1}{4!}x^4$ on the right; $-\frac{1}{2!}x^2$ on the left paired with $-\frac{1}{2!}x^2$ on the right; and $\frac{1}{4!}x^4$ on the left paired with 1 on the right. This is the same algebra one would do when multiplying two polynomials together; this is just a way of collecting like terms in a systematic way.

Example

Use the trigonometric identity

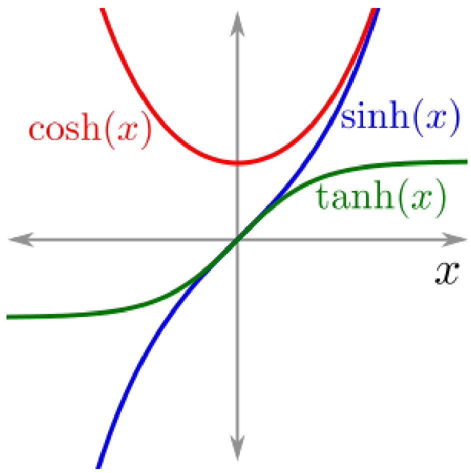
$$\cos^2 x = \frac{1+\cos(2x)}{2}$$

and substitution to find the series for $\cos^2 x$. Try to give the series in summation notation (other than the first term). (See Answer 2)

4.3 Hyperbolic trigonometric functions

The hyperbolic trigonometric functions $\sinh(x)$, $\cosh(x)$, and $\tanh(x)$ are defined by

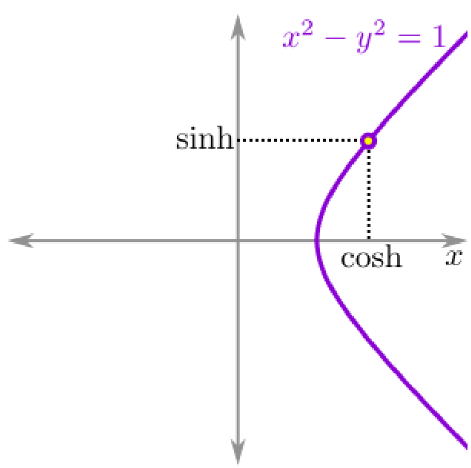
$$\sinh(x) = \frac{e^x - e^{-x}}{2}$$
$$\cosh(x) = \frac{e^x + e^{-x}}{2}$$
$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{\sinh(x)}{\cosh(x)}.$$



These hyperbolic trig functions, although graphically quite different from their traditional counterparts, have several similar algebraic properties, which is why they are so named. For example, the Pythagorean identity for cosine and sine has a version for hyperbolic cosine and sine:

$$\cosh^2(x) - \sinh^2(x) = 1.$$

One can verify this using the definitions and some algebra. But there is a geometric intuition for this relationship. Recall that cosine and sine give the x and y coordinates, respectively, for a point on the unit circle $x^2 + y^2 = 1$. The hyperbolic cosine and hyperbolic sine give the x and y coordinates, respectively, for points on the hyperbola $x^2 - y^2 = 1$:



Example

Using the Taylor series for e^x and substitution, show that the Taylor series for cosh and sinh are

$$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$$
$$\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.$$

Note that these are almost the same as the series for cosine and sine, respectively, except they do not alternate. This gives another reason for the names of these functions. (See Answer 3)

4.4 Manipulating Taylor series

Another way of using one Taylor series to find another is through differentiation and integration. For instance, to find the Taylor series for the derivative of f, one can differentiate the Taylor series for f term by term.

Example

By differentiating the Taylor series for sinh and cosh, show that

$$\frac{d}{dx}\sinh x = \cosh x$$
$$\frac{d}{dx}\cosh x = \sinh x.$$

This is yet another relationship which is similar (though not identical) to the relationship between sine and cosine. (See Answer 4)

4.5 Higher Order Terms in Taylor Series

In some situations, it will be convenient only to write the first few terms of a Taylor series. This is particularly true when combining or composing more than one Taylor series. Up until now, an ellipsis has been used to indicate that there are more terms in the series that are being omitted.

There is another way, sometimes used in this course, of notating the omitted terms in a Taylor series. That is by referring to them as Higher Order Terms (or H.O.T. for short). Having the extra HOT in a series means that all the remaining terms in the series have a higher degree than the previous terms.

Example

The function e^x can be written as

$$e^x = 1 + x + \frac{1}{2!}x^2 + \text{ HOT},$$

or it could also be written as

$$e^x = 1 + x + \text{ HOT}.$$

The point at which the higher order terms are cut-off is somewhat arbitrary and depends on the situation. There is a more formal way of keeping track of the higher order terms, called Big-O notation, which is presented in orders of growth.

Example

Find the first two non-zero terms of the Taylor series for

$$f(x) = 1 - 2xe^{\sin x^2}.$$

(See Answer 5)

4.6 Extra examples

Example

Compute the Taylor series (at 0) for $\sin^2 x$ up to and including terms of order 6. Try to give the full Taylor series in summation notation. (See Answer 6)

Example

Find the first three terms of the Taylor series for $\sqrt{f(x)}$, where

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots.$$

(See Answer 7)

4.7 EXERCISES

- Compute the Taylor series of $\cos(2x)\sin(3x)$ up to and including terms of degree 5. Don't try computing derivatives for this!
- Use a Taylor polynomial to give a cubic approximation to $2xe^{3x}$
- Compute the Taylor series of $e^{1-\cos t}$ in summation notation.
- Compute the Taylor series of $\cos(\sin(x))$ to fourth order.
- Compute the Taylor series of $\sin(\cos(x))$ to forth order. What happens that makes this different than the last problem? (Hint: $\cos(0) = 1$ but $\sin(0) = 0$...)
- Compute the first three nonvanishing terms in the Taylor series of $e^{2x}(\sinh 3x)/x$.
- Compute the Taylor series of $3x^2 e^{-x^2} \sin 2x^3$ up to and including terms of order eight (!) Wow, that means a lot of work, right? Think... which terms should you expand first?
- Compute the Taylor series of $\frac{1}{x}e^{-x^2}\sinh(2x)$ up to the fourth order term.
- What is the second derivative of the function $e^{x\cosh(x^2)}$ at $x = 0$?
- Compute the following limit $\lim_{x\to 0}(1 - e^x)\frac{\sin(x^2)}{x^3}$

4.8 Answers to Selected Exercises

1. Recall the series for e^x is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

Substituting x^3 into the series for e^x gives

$$\begin{aligned} e^{x^3} &= 1 + x^3 + \frac{(x^3)^2}{2!} + \frac{(x^3)^3}{3!} + \cdots \\ &= 1 + x^3 + \frac{x^6}{2!} + \frac{x^9}{3!} + \cdots \\ &= \sum_{k=0}^{\infty} \frac{(x^3)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{x^{3k}}{k!} \end{aligned}$$

(Return)

2. By the above identity,

$$\begin{aligned} \cos^2 x &= \frac{1}{2}\left(1 + \cos(2x)\right) \\ &= \frac{1}{2}\left(1 + \left(1 - \frac{(2x)^2}{2!} + \frac{(2x)^4}{4!} - \cdots\right)\right) \\ &= \frac{1}{2}\left(2 - \frac{4x^2}{2} + \frac{16x^4}{24} - \cdots\right) \\ &= 1 - x^2 + \frac{x^4}{3} - \cdots . \end{aligned}$$

In summation notation,

$$\begin{aligned} \cos^2 x &= \frac{1}{2}\left(1 + \sum_{k=0}^{\infty} (-1)^k \frac{(2x)^{2k}}{(2k)!}\right) \\ &= \frac{1}{2} + \frac{1}{2}\sum_{k=0}^{\infty} (-1)^k \frac{(2x)^{2k}}{(2k)!} \\ &= \frac{1}{2} + \sum_{k=0}^{\infty} (-1)^k \frac{2^{2k-1} x^{2k}}{(2k)!} \\ &= 1 + \sum_{k=1}^{\infty} (-1)^k \frac{2^{2k-1} x^{2k}}{(2k)!} . \end{aligned}$$

(Return)

3.

$$\begin{aligned}
\cosh(x) &= \frac{e^x + e^{-x}}{2} \\
&= \frac{1}{2}\left[(1 + x + \frac{x^2}{2!} + \cdots) + (1 - x + \frac{x^2}{2!} - \cdots)\right] \\
&= \frac{1}{2}\left[2 + 2\frac{x^2}{2!} + 2\frac{x^4}{4!} + \cdots\right] \\
&= 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots \\
&= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}.
\end{aligned}$$

$$\begin{aligned}
\sinh(x) &= \frac{e^x - e^{-x}}{2} \\
&= \frac{1}{2}\left[(1 + x + \frac{x^2}{2!} + \cdots) - (1 - x + \frac{x^2}{2!} - \cdots)\right] \\
&= \frac{1}{2}\left[2x + 2\frac{x^3}{3!} + 2\frac{x^5}{5!} + \cdots\right] \\
&= x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots \\
&= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.
\end{aligned}$$

(Return)

4. Differentiating hyperbolic sine gives

$$\begin{aligned}
\frac{d}{dx}\sinh x &= \frac{d}{dx}\sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{\infty}(2k+1)\frac{x^{2k}}{(2k+1)!} \\
&= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \\
&= \cosh x,
\end{aligned}$$

as desired. Similarly, differentiating hyperbolic cosine gives

$$\begin{aligned}
\frac{d}{dx}\cosh x &= \frac{d}{dx}\sum_{k=0}^{\infty}\frac{x^{2k}}{(2k)!} \\
&= \sum_{k=0}^{\infty}(2k)\frac{x^{2k-1}}{(2k)!} \\
&= \sum_{k=1}^{\infty}\frac{x^{2k-1}}{(2k-1)!} \\
&= \sum_{k=0}^{\infty}\frac{x^{2k+1}}{(2k+1)!}.
\end{aligned}$$

There was a little bit of reindexing there, but by writing out a few terms of each series, one can see that all of the above equalities are true.

(Return)

5. Beginning with the innermost function, in this case $\sin x^2$, we find that

$$\sin x^2 = x^2 - \frac{1}{3!}(x^2)^3 + \text{ HOT} = x^2 - \frac{1}{6}x^6 + \text{ HOT}.$$

Then plugging this into the series for e^x gives

$$\begin{aligned}
e^{\sin x^2} &= 1 + \left(x^2 - \frac{1}{6}x^6 + \text{ HOT}\right) + \frac{1}{2!}\left(x^2 + \text{ HOT}\right)^2 + \frac{1}{3!}\left(x^2 + \text{ HOT}\right)^3 + \text{ HOT} \\
&= 1 + x^2 + \frac{1}{2}x^4 + \left(-\frac{1}{6} + \frac{1}{6}\right)x^6 + \text{ HOT} \\
&= 1 + x^2 + \frac{1}{2}x^4 + \text{ HOT}
\end{aligned}$$

Then to complete the answer, plug this into the original function to find

$$\begin{aligned}
f(x) &= 1 - 2x\left(1 + x^2 + \frac{1}{2}x^4 + \text{ HOT}\right) \\
&= 1 - 2x - 2x^3 - x^5 + \text{ HOT}.
\end{aligned}$$

(Return)

6.

$$\begin{aligned}
\sin^2 x &= (x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots)(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots) \\
&= x^2 + (-\frac{1}{3!} - \frac{1}{3!})x^4 + (\frac{1}{5!} + \frac{1}{3!\cdot 3!} + \frac{1}{5!})x^6 + \cdots \\
&= x^2 - \frac{1}{3}x^4 + \frac{2}{45}x^6 - \cdots .
\end{aligned}$$

To get the full Taylor series, one can use the identity

$$\sin^2 x = \frac{1 - \cos(2x)}{2}$$

to find that

$$
\begin{aligned}
\sin^2 x &= \frac{1-\cos(2x)}{2} \\
&= \frac{1}{2}\left(1-\left(1-\frac{(2x)^2}{2!}+\frac{(2x)^4}{4!}-\cdots\right)\right) \\
&= \frac{1}{2}\left(\frac{(2x)^2}{2!}-\frac{(2x)^4}{4!}+\frac{(2x)^6}{6!}-\cdots\right) \\
&= \frac{1}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{(2x)^{2k}}{(2k)!}.
\end{aligned}
$$

(Return)

7. Let $g(x)=\sqrt{f(x)}$, where

$$g(x)=b_0+b_1x+b_2x^2+b_3x^3+\cdots .$$

Then $g(x)^2=f(x)$, and so the same holds for the Taylor series:

$$\left(b_0+b_1x+b_2x^2+b_3x^3+\cdots\right)^2=a_0+a_1x+a_2x^2+\cdots .$$

Multiplying out and collecting like terms gives

$$b_0^2+(b_0b_1+b_1b_0)x+(b_0b_2+b_1b_1+b_2b_0)x^2+\cdots=a_0+a_1x+a_2x^2+\cdots .$$

Now, equating coefficients of the monomials on the left and right gives the first few equations (of an infinite system of equations)

$$
\begin{aligned}
b_0^2 &= a_0 \\
2b_0b_1 &= a_1 \\
2b_0b_2+b_1^2 &= a_2.
\end{aligned}
$$

Solving these equations gives the first three coefficients of g:

$$
\begin{aligned}
b_0 &= \sqrt{a_0} \\
b_1 &= \frac{a_1}{2\sqrt{a_0}} \\
b_2 &= \frac{1}{2\sqrt{a_0}}\left(a_2-\frac{a_1^2}{4a_0}\right).
\end{aligned}
$$

Thus,

$$\sqrt{a_0+a_1x+a_2x^2+\cdots}=\sqrt{a_0}+\frac{a_1}{2\sqrt{a_0}}x+\frac{1}{2\sqrt{a_0}}\left(a_2-\frac{a_1^2}{4a_0}\right)x^2+\cdots .$$

(Return)

5 Convergence

A Taylor series can be thought of as an infinite polynomial. Up until now, we have not worried about the issues that come up when adding up infinitely many things. This module deals with two main issues:

1. A function may not have a Taylor series at all;
2. A function's Taylor series may not converge everywhere, even within the function's domain.

5.1 Functions without a Taylor series

The first problem is that some functions cannot be expressed in the form

$$f(x) = \sum_{k=0}^{\infty} c_k x^k = c_0 + c_1 x + c_2 x^2 + \cdots$$

Examples include tan, which has vertical asymptotes, and ln, which is not defined for $x \leq 0$. Polynomials are not able to capture these sorts of discontinuities and asymptotes.

The geometric series

The geometric series is an example of a Taylor series which is well behaved for some values of x and nonsensical for other values of x. The claim is that

$$1 + x + x^2 + x^3 + x^4 + \cdots = \frac{1}{1-x},$$

for $|x| < 1$. (See Justiification 1)

Example

Compute the Taylor series for $f(x) = \frac{1}{1-x}$ directly from the definition. (See Answer 2)

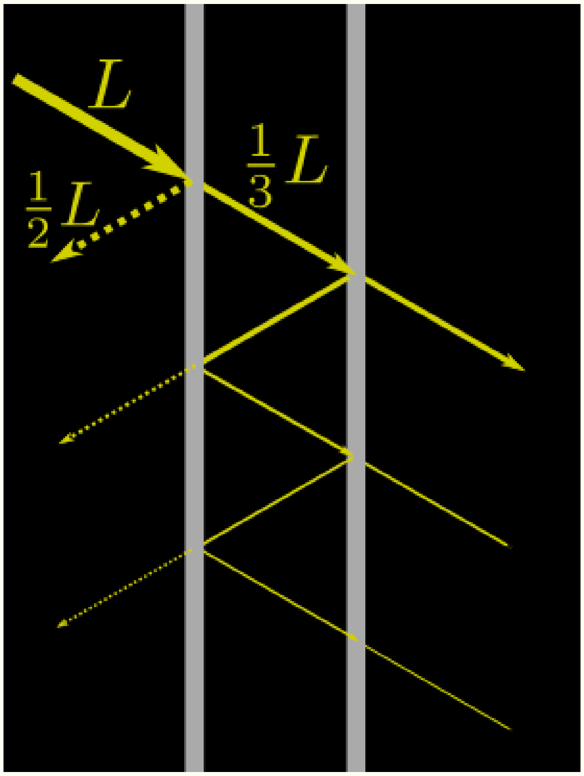
Note The geometric series only holds when $|x| < 1$. This makes sense, because if $|x| > 1$, the powers of x are getting bigger and bigger and so the series should not converge. If $x = 1$, then the series is adding 1 infinitely many times, which diverges. If $x = -1$, then the series oscillates between 1 and 0, and hence does not converge.

The takeaway is that every Taylor series has a convergence domain where the series is well-behaved, and outside that domain the series will not converge. For many functions, the domain is the whole real number line (e.g.

the series for e^x, sin, cos, cosh, and sinh all converge everywhere), but be aware that there are functions whose Taylor series do not converge everywhere. This will be covered more formally in Series Convergence And Divergence.

Example

A beam of light of intensity L hits a pane of glass. Half of the light is reflected, and a third of the light is transmitted; the rest is absorbed. When a beam of light of intensity L hits two parallel panes with an air gap between them, how much light is transmitted through both panes? (The following figure shows how the light gets reflected and rereflected. The first transmitted and reflected beams of light are labeled with their respective intensities. The question asks for the total of the beams of light emerging on the right side of the right pane of glass).



(See Answer 3)

Example

Use the Taylor series of $\frac{1}{1-x}$ to derive the Taylor series of $\ln(1+x)$. Hint: recall that $\ln(1+x) = \int \frac{1}{1+x}\,dx$. (See Answer 4)

Example

Use the fact that

$$\arctan x = \int \frac{1}{1+x^2}\,dx$$

to find the Taylor series for $\arctan x$. (See Answer 5)

Example

Another important function is the *binomial series* $(1+x)^\alpha$, where α is some constant. Show that

$$\begin{aligned}(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots \\ &= \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,\end{aligned}$$

where

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}.$$

This series also only holds for $|x| < 1$. (See Answer 6)

5.2 Summary

Here are all the series we have found so far. The following hold for all x:

$$\begin{aligned}e^x &= \sum_{k=0}^{\infty} \frac{x^k}{k!} \\ \cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \\ \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \\ \cosh x &= \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \\ \sinh x &= \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}.\end{aligned}$$

The following hold for $|x| < 1$:

$$\begin{aligned}\frac{1}{1-x} &= \sum_{k=0}^{\infty} x^k \\ \ln(1+x) &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \\ \arctan x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \\ (1+x)^\alpha &= \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.\end{aligned}$$

5.3 Electrostatics example

Here we use the geometric series and the binomial series from above in an example from electrostatics. An *electric dipole* is a pair of equally and oppositely charged particles separated by a short distance. One question of interest in electrostatics is the *electrostatic potential*, which is the sum of the point charge potentials from each pole.

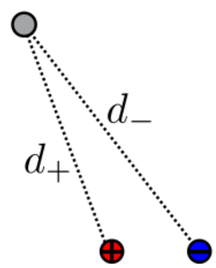
The point charge potential from a single particle with charge q, at a distance d from the particle, is

$$V = \frac{kq}{d},$$

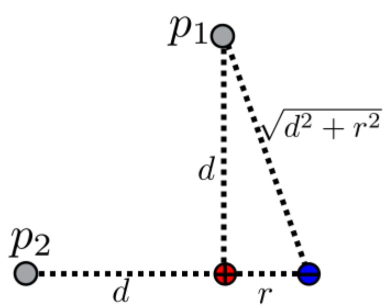
where k is a constant called the Coulomb constant. Then a dipole with particles of charge q and $-q$ has net electrostatic potential

$$V = \frac{kq}{d_+} - \frac{kq}{d_-},$$

where d_+ is the distance to the positively charged particle, and d_- is the distance to the negatively charged particle:



We will calculate the first order term for the electrostatic potential at two different locations: p_1 and p_2:



First consider p_1, located directly above and distance d from the positive particle. Let r be the distance between the charged particles. Then $d_+ = d$, and by the Pythagorean theorem, $d_- = \sqrt{d^2 + r^2}$. It follows that the electrostatic potential is

$$\begin{aligned} V &= \frac{kq}{d_+} - \frac{kq}{d_-} \\ &= \frac{kq}{d} - \frac{kq}{\sqrt{d^2 + r^2}}. \end{aligned}$$

Now, factoring out $\frac{kq}{d}$, and applying the binomial series with $\alpha = -\frac{1}{2}$, we find

$$\begin{aligned} V &= \frac{kq}{d}\left[1 - \frac{1}{\sqrt{1+(r/d)^2}}\right] \\ &= \frac{kq}{d}\left[1 - \left(1+(r/d)^2\right)^{-1/2}\right] \\ &= \frac{kq}{d}\left[1 - \left(1 - \frac{1}{2}(r/d)^2 + \text{ HOT}\right)\right] \\ &= \frac{1}{2}\frac{kqr^2}{d^3} + \text{ HOT}. \end{aligned}$$

At position p_2, which is directly left of and distance d from the positive particle, we have $d_+ = d$, and $d_- = d + r$, so we find that the electrostatic potential at p_2 is

$$\begin{aligned} V &= \frac{kq}{d_+} - \frac{kq}{d_-} \\ &= \frac{kq}{d} - \frac{kq}{d+r}. \end{aligned}$$

Again, factoring out $\frac{kq}{d}$ and expanding using the geometric series gives

$$\begin{aligned} V &= \frac{kq}{d}\left(1 - \frac{1}{1+\frac{r}{d}}\right) \\ &= \frac{kq}{d}\left(1 - \left(1 - \frac{r}{d} + \text{ HOT}\right)\right) \\ &= \frac{kqr}{d^2} + \text{ HOT}. \end{aligned}$$

5.4 EXERCISES

- Consider a snowman built from solid snowballs of radius 2^{-n}, for $n = 0, 1, 2, \ldots$, all stacked on top of one another. How many units tall is the snowman? How many cubic units of snow was required to build it?
- Compute the Taylor series about zero of

$$\ln \frac{1+3x}{1-3x}$$

- Compute the Taylor series about zero of

$$\frac{1}{\sqrt{1-x^2}}$$

- Using your answer to the previous problem, compute the Taylor series about zero of $\arcsin x$, using termwise integration and the fact that

$$\arcsin x = \int \frac{dx}{\sqrt{1-x^2}}$$

- For which values z is the Taylor series of $\sqrt[4]{3-2z^2}$ guaranteed to converge?

- Use the binomial series to give the Taylor expansion of $(1+x)^3$. Now, do it with your head: easier, right? Recall, we have said that the binomial series only converges when $|x < 1|$, but, clearly, that cannot be a *sharp* constraint, since $(1+x)^3$ is good for all x, right? Well, Horatio, there are more things... By the end of this course, we will learn when and how to bend some of these restrictions.
- Build a cylinder with radius 1 and height 3. Build a second cylinder with radius 1/2 and height 9, a third cylinder with radius 1/4, height 27, a fourth cylinder with radius 1/8 and height 81, and so on. What is the total volume of the cylinders?
- For which values of x does the Taylor series of $(\frac{1}{4} - 3x^2)^{1/4}$ converge?

5.5 Answers to Selected Examples

1. **Note** This is not a formal proof, which would require a few tools and definitions we have not yet learned. Let $y = 1 + x + x^2 + x^3 + \cdots$. Multiplying both sides by x gives

$$\begin{aligned} y &= 1 + x + x^2 + x^3 + \cdots \\ xy &= x + x^2 + x^3 + x^4 + \cdots \end{aligned}$$

Now, subtracting the second equation from the first, all the terms other than 1 cancel on the right, leaving us with

$$y(1-x) = 1.$$

Dividing by $1-x$ gives $y = \frac{1}{1-x}$.

(Return)

2.

$$\begin{aligned} f(x) &= \frac{1}{1-x} & f(0) &= 1 \\ f'(x) &= \frac{1}{(1-x)^2} & f'(0) &= 1 \\ f''(x) &= \frac{2}{(1-x)^3} & f''(0) &= 2 \\ f'''(x) &= \frac{6}{(1-x)^4} & f'''(0) &= 6. \end{aligned}$$

Notice the pattern that

$$f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}},$$

at least for the first few k. To see that the pattern continues, assume it holds for some k, and show that it holds for $k+1$ (this is a proof technique known as mathematical induction). If $f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}$, then

$$f^{(k+1)}(x) = \frac{(k+1)k!}{(1-x)^{k+2}} = \frac{(k+1)!}{(1-x)^{k+2}},$$

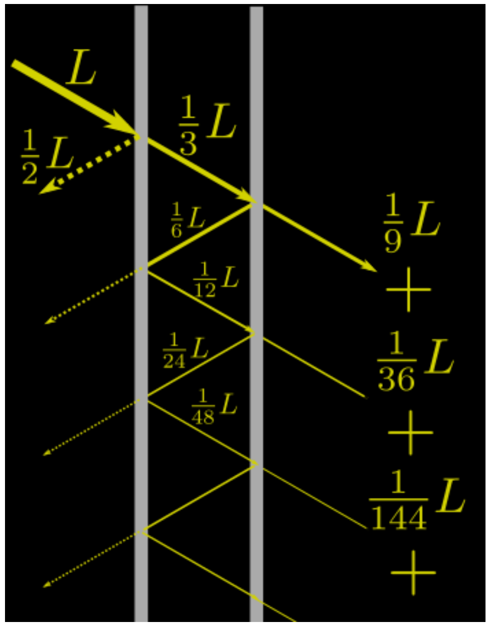
as desired. Then $f^{(k)}(0) = k!$, so according to the definition of Taylor series, it follows that

$$\begin{aligned} \frac{1}{1-x} &= 0! + 1!x + \frac{2!}{2!}x^2 + \frac{3!}{3!}x^3 + \cdots \\ &= 1 + x + x^2 + x^3 + \cdots, \end{aligned}$$

which agrees with the above.

(Return)

3. By labeling more of the transmitted and reflected beams of light, a pattern emerges among the beams of light on the right side of the right pane:



$\frac{1}{9}, \frac{1}{36}, \frac{1}{144}, \ldots$. Note that each beam is $\frac{1}{4}$ the previous beam. Thus, the total light emerging on the right side of the right pane of glass is

$$\begin{aligned}\frac{L}{9} + \frac{L}{36} + \frac{L}{144} + \cdots &= \frac{L}{9}\left(1 + \frac{1}{4} + \frac{1}{16} + \cdots\right) \\ &= \frac{L}{9}\left(\frac{1}{1 - 1/4}\right) \\ &= \frac{L}{9}\frac{4}{3} \\ &= \frac{4L}{27},\end{aligned}$$

by using the formula for the geometric series.

(Return)

4. Note that

$$\begin{aligned}\frac{1}{1+x} &= \frac{1}{1-(-x)} \\ &= 1 - x + x^2 - x^3 + x^4 - \cdots.\end{aligned}$$

Now, integrating gives $\int \frac{dx}{1+x} = \ln(1+x) + C$ on the one hand, and

$$\begin{aligned}\int (1 - x + x^2 - x^3 + x^4 - \cdots)dx &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \\ &= \sum_{k=1}^{\infty} (-1)^{k-1}\frac{x^k}{k},\end{aligned}$$

on the other hand. Plugging in $x = 0$ shows that $C = 0$, and so

$$\begin{aligned}\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots \\ &= \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}. \qquad (|x| < 1)\end{aligned}$$

Note that because this relied on the geometric series, which only holds for $|x| < 1$, the same restriction holds for the Taylor series for $\ln(1+x)$.

(Return)

5. Using the fact, and the geometric series, we find that

$$\begin{aligned}\arctan(x) &= \int \frac{1}{1+x^2}\, dx \\ &= \int \frac{1}{1-(-x^2)}\, dx \\ &= \int \left(1 - x^2 + x^4 - x^6 + \cdots\right)\, dx \qquad (|x| < 1) \\ &= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + C.\end{aligned}$$

Plugging in $x = 0$ gives that $C = 0$, since $\arctan 0 = 0$. Thus,

$$\begin{aligned}\arctan(x) &= x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \qquad (|x| < 1).\end{aligned}$$

So even though arctan is defined for all x, its Taylor series only converges for $|x| < 1$.

(Return)

6. For fixed α we have $f(x) = (1+x)^\alpha$. Then proceeding from the definition of the Taylor series, one computes

$$\begin{aligned}f(x) &= (1+x)^\alpha & f(0) &= 1 \\ f'(x) &= \alpha(1+x)^{\alpha-1} & f'(0) &= \alpha \\ f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2} & f''(0) &= \alpha(\alpha-1) \\ f'''(x) &= \alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3} & f'''(0) &= \alpha(\alpha-1)(\alpha-2) \\ &\vdots & &\vdots\end{aligned}$$

One finds that, in general, $f^{(k)}(0) = \alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)$. Thus, the Taylor expansion for $(1+x)^\alpha$ is

$$\begin{aligned}(1+x)^\alpha &= 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots \\ &= 1 + \binom{\alpha}{1}x + \binom{\alpha}{2}x^2 + \binom{\alpha}{3}x^3 + \cdots = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,\end{aligned}$$

as claimed.

(Return)

6 Expansion Points

Up until now, Taylor series expansions have all been at $x = 0$. The Taylor series at $x = 0$ gives a good approximation to the function near 0. But what if we want a good approximation to the function near a different point a? That is the topic of this module.

6.1 Expansion points

A function f has a Taylor series expansion about any point $x = a$ provided that f and all its derivatives exist at a. The definition of the Taylor series for f about $x = a$ is

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k. \end{aligned}$$

We say this is a series in $(x - a)$. A different way to view this series is by making the change of variables $x = a + h$. After cancellation, this yields

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}h^k. \end{aligned}$$

6.2 Taylor polynomial for approximation

Recall that the first few terms of the Taylor series for f about $x = 0$ gives a polynomial (the *Taylor polynomial*) which is a good approximation for f near 0. Similarly, the Taylor polynomial for f about $x = a$ gives a polynomial which is a good approximation of f near $x = a$. Note, however, that as the input gets further away from the expansion point a, the approximation gets worse.

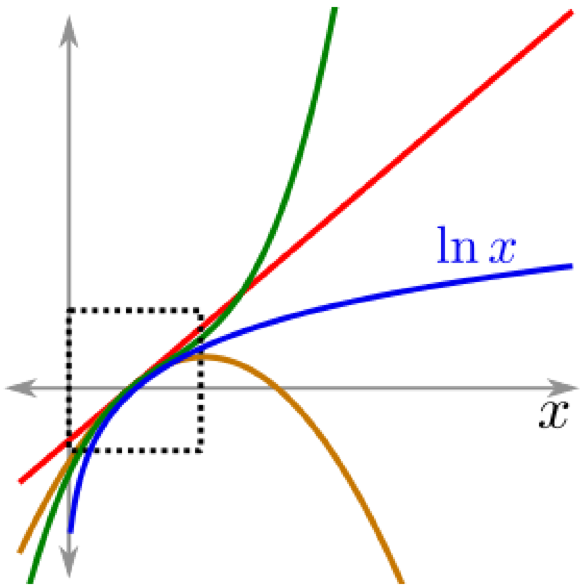
Example

Find the Taylor series for $f(x) = 3x^2 - x + 4$ about $x = 2$. (See Answer 1)

Example

Compute the Taylor series expansion for $\ln(x)$ about $x = 1$. (See Answer 2)

Note that the Taylor polynomial is only a good approximation to the function on the domain of convergence. For functions whose domain of convergence is the entire number line, this is not a concern. But for functions such as $\ln x$, the Taylor polynomials will only be a good approximation within the domain of convergence, which is $0 < x < 2$. Outside of that domain, the Taylor polynomials diverge wildly from $\ln x$, as shown here:



Even within a function's domain of convergence, a Taylor polynomial's approximation gets worse as the input gets further away from a. One way to improve an approximation is to include more and more terms of the Taylor series in the Taylor polynomial. However, this involves computing more and more derivatives. Another way to improve the approximation for $f(x)$ is to choose an expansion point a which is close to x.

Example

Use the Taylor polynomial of degree 2 for $f(x) = \sqrt{x}$ about $x = 1$ to approximate $\sqrt{10}$. Then repeat the process about $x = 9$ and compare the results. (See Answer 3)

6.3 Caveat for compositions

When computing the Taylor expansion for the composition $f \circ g$ about $x = a$, one must be careful of expansion points. In particular, one cannot simply take the series for g at $x = a$ and plug it into the series for f at $x = a$.

Example

Consider the expansion for $e^{\cos(x)}$ about $x = 0$. Although $\cos(x) = 1 - \frac{x^2}{2!} + \cdots$, and $e^x = 1 + x + \frac{x^2}{2!} + \cdots$, one will run into trouble trying to write

$$e^{\cos(x)} = 1 + (1 - \frac{x^2}{2!} + \cdots) + \frac{1}{2}(1 - \frac{x^2}{2!} + \cdots)^2 + \cdots .$$

The trouble is that collecting like terms requires adding up infinitely many things. For instance, the constant term above is $1+1+\frac{1}{2}+\cdots$. The reason this is a problem is that Taylor series are supposed to give a good polynomial approximation of a function without requiring too much computation or information about the function.

Remember that $e^x = 1 + x + \frac{x^2}{2} + \cdots$ is a good approximation when x is near 0. However, when x is near 0, $\cos(x)$ is near 1. So plugging the series for $\cos(x)$ into the series for e^x does not give a good approximation.

To avoid this problem when computing the Taylor series for the composition $f \circ g$ at $x = a$, one should plug the Taylor expansion of g about $x = a$ into the expansion of f about $x = g(a)$. In the above example, the expansion of e^x about $x = 1$ is

$$e^x = e + e(x-1) + \frac{e}{2!}(x-1)^2 + \cdots,$$

so

$$\begin{aligned} e^{\cos(x)} &= e + e\left[\left(1 - \frac{x^2}{2!} + \cdots\right) - 1\right] + \frac{e}{2!}\left[\left(1 - \frac{x^2}{2!} + \cdots\right) - 1\right]^2 + \cdots \\ &= e + e(-\frac{x^2}{2} + \cdots) + \frac{e}{2}(-\frac{x^2}{2} + \cdots)^2 + \cdots \\ &= e - \frac{e}{2}x^2 + \cdots. \end{aligned}$$

6.4 EXERCISES:

- Without using a calculator, find a decimal approximation to $\sqrt{83}$ by Taylor-expanding $\sqrt{x}$ about $a = 81$ and using the zero-th and first order terms.
- Without using a calculator, find a decimal approximation to $\sqrt[3]{124}$ using linear approximation. How close was your answer to truth?
- Without using a calculator, find a decimal approximation to $\cos(1)$ [in radians!] using linear approximation. How close was your answer to truth? (Hint: $\pi/3 \approx 1...$)
- Taylor expand $\sin x$ about $x = \pi$ and compute all the terms. Does what you get make sense?
- Use completing the square and the geometric series to get the Taylor expansion about $x = 2$ of $\frac{1}{x^2+4x+3}$
- Approximate $1004^{1/3}$ using the zeroth and first order terms of the Taylor series.

6.5 Answers to Selected Examples

1. Computing the derivatives, and evaluating at $x = 2$, one finds

$$\begin{aligned} f(x) &= 3x^2 - x + 4 & f(2) &= 14 \\ f'(x) &= 6x - 1 & f'(2) &= 11 \\ f''(x) &= 6 & f''(2) &= 6 \\ f'''(x) &= 0 & f'''(2) &= 0. \end{aligned}$$

And all the subsequent derivatives are 0. So from the definition, one finds that

$$\begin{aligned} f(x) &= 14 + 11(x-2) + \frac{6}{2!}(x-2)^2 \\ &= 14 + 11(x-2) + 3(x-2)^2. \end{aligned}$$

This appears to be different than the polynomial f with which we began. If one multiplies out this polynomial and collects like terms, however, the result is the original polynomial. This should not be surprising, since the best polynomial approximation to a polynomial is the polynomial itself, even factored into a slightly different form.

(Return)

2. Begin by computing the first few derivatives and evaluating at $x = 1$:

$$\begin{aligned} f(x) &= \ln(x) & f(1) &= 0 \\ f'(x) &= x^{-1} & f'(1) &= 1 \\ f''(x) &= -x^{-2} & f''(1) &= -1 \\ f'''(x) &= 2x^{-3} & f'''(1) &= 2. \end{aligned}$$

The pattern that emerges is $f^{(k)}(x) = (-1)^{k-1}(k-1)!x^{-k}$. To see that the pattern holds, check that

$$f^{(k+1)}(x) = (-1)^{k-1}(-k)(k-1)!x^{-k-1} = (-1)^k k! x^{-(k+1)},$$

as desired. So by induction, the pattern holds. It follows that $f^{(k)}(1) = (-1)^{k-1}(k-1)!$ for $k \geq 1$. Plugging in to the formula, one finds that

$$\begin{aligned} \ln(x) &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(k-1)!}{k!}(x-1)^k \\ &= \sum_{k=1}^{\infty} (-1)^{k-1} \frac{(x-1)^k}{k} \\ &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots . \end{aligned}$$

Note that with the change of variables $h = x - 1$ (and hence $x = h + 1$, we find that

$$\ln(1+h) = h - \frac{h^2}{2} + \frac{h^3}{3} - \frac{h^4}{4} + \cdots ,$$

which is the same series we found earlier for $\ln(1 + x)$.

(Return)

3. Using the definition, one finds

$$\begin{aligned} f(x) &= \sqrt{x} & f(1) &= 1 & f(9) &= 3 \\ f'(x) &= \frac{1}{2\sqrt{x}} & f'(1) &= \frac{1}{2} & f'(9) &= \frac{1}{6} \\ f''(x) &= -\frac{1}{4x^{3/2}} & f''(1) &= -\frac{1}{4} & f''(9) &= -\frac{1}{108} \end{aligned}$$

Thus, the Taylor polynomial about $x = 1$ is

$$\begin{aligned}\sqrt{x} &\approx 1 + \frac{1}{2}(x-1) - \frac{1/4}{2!}(x-1)^2 \\ &= 1 + \frac{1}{2}(x-1) - \frac{1}{8}(x-1)^2.\end{aligned}$$

And the corresponding approximation is

$$\begin{aligned}\sqrt{10} &\approx 1 + \frac{1}{2}\cdot 9 - \frac{1}{8}\cdot 9^2 \\ &\approx -4.6,\end{aligned}$$

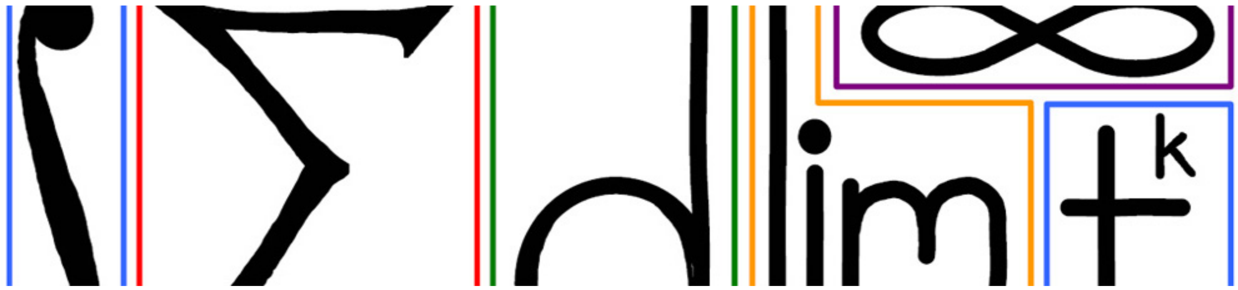
which is obviously quite far off the mark. On the other hand, the Taylor polynomial about $x = 9$ is

$$\begin{aligned}\sqrt{x} &\approx 3 + \frac{1}{6}(x-9) - \frac{1/108}{2!}(x-9)^2 \\ &= 3 + \frac{1}{6}(x-9) - \frac{1}{216}(x-9)^2.\end{aligned}$$

And the corresponding approximation is

$$\begin{aligned}\sqrt{10} &\approx 3 + \frac{1}{6}\cdot 1 - \frac{1}{216}\cdot 1^2 \\ &\approx 3.1620,\end{aligned}$$

which is quite a good approximation of $\sqrt{10} \approx 3.1623$.
(Return)



7 Limits

Having concluded our study of Taylor series, we now move on to limits. Some of the major topics of calculus (continuity, differentiation, and integration) can all be expressed using limits.

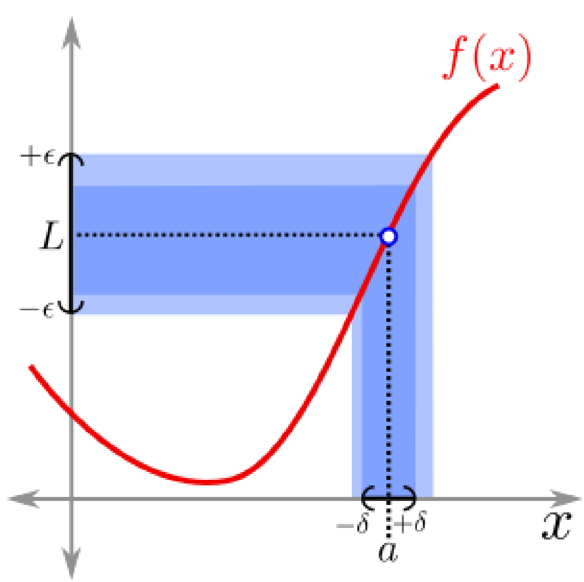
7.1 Definition of the limit

The limit formalizes the behavior of a function as its input approaches some value. The formal definition of the limit is

Limit

$\lim_{x \to a} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that $|f(x) - L| < \epsilon$ whenever $0 < |x - a| < \delta$. If there is no such L, then the limit does not exist.

In words, this says that the limit of a function exists if, when the input to f is very close to a (but not equal to a), the output from f is very close to L. This can also be thought of in terms of tolerances: given a certain ϵ tolerance for the output (seen as the band around L in the graph below), one can find a tolerance δ on the input (the band around a) so that for inputs within the tolerance, the corresponding outputs stays within ϵ of the desired output:



No matter how small ϵ is made, there must be some δ, which must depend on ϵ, generally. Actually finding δ often requires a little bit of work.

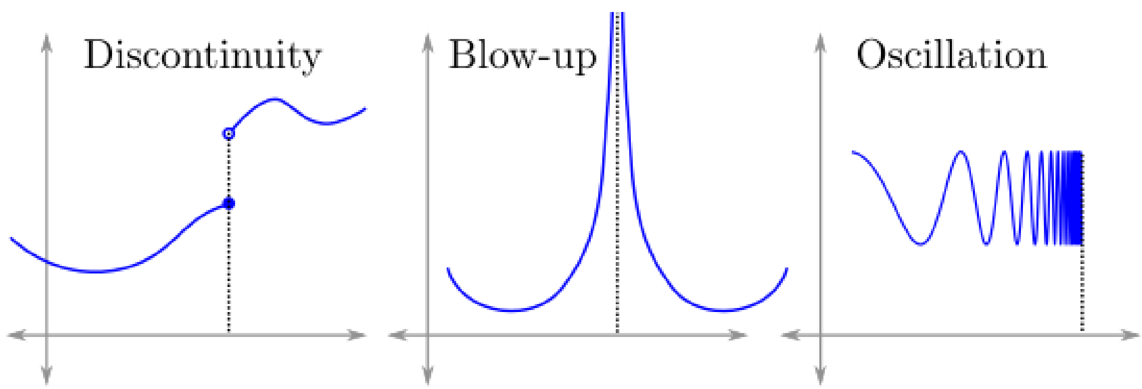
Example

Using the definition of the limit, show that $\lim_{x \to 3} x^2 = 9$. (See Answer 1)

7.2 When limits may not exist

There are a few ways a limit might not exist:

1. A *discontinuity*, or jump, in the graph of the function. In this case, the limit does not exist because the limit from the left and the limit from the right are not equal.
2. A *blow-up*, when the function has a vertical asymptote.
3. An *oscillation*, where the graph of the function oscillates infinitely up and down as the input approaches a certain value.



Most functions in this course will be well-behaved and will not have the above problems. The formal term for a well-behaved function is *continuous*.

7.3 Continuous functions

A function is *continuous at* the point a if the limit $\lim_{x \to a} f(x)$ exists and $\lim_{x \to a} f(x) = f(a)$. Intuitively, this says that there are no holes or jumps in the graph of f at a.

Finally, a function is *continuous* if it is continuous at every point in its domain.

7.4 Rules for limits

There are rules for adding, multiplying, dividing, and composing limits. Suppose that $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ exist. Then

1. (Sum) $\lim_{x \to a}(f+g)(x) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$.

2. (Product) $\lim_{x\to a}(f\cdot g)(x) = \left(\lim_{x\to a} f(x)\right)\left(\lim_{x\to a} g(x)\right)$.

3. (Quotient) $\lim_{x\to a}\left(\frac{f}{g}\right)(x) = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$, provided that $\lim_{x\to a} g(x) \neq 0$.

4. (Chain) $\lim_{x\to a}(f\circ g)(x) = f\left(\lim_{x\to a} g(x)\right)$, if f is continuous.

Almost all the functions encountered in this course are continuous, and so limits in most cases can be evaluated by simply plugging in the limiting input value into the function. The one case that sometimes gets complicated is the Quotient rule above when the limit of the denominator is 0.

Example

Show that $\lim_{x\to 0}\frac{\sin(x)}{x} = 1$. (See Answer 2)

Example

Find $\lim_{x\to 0}\frac{1-\cos x}{x}$. (See Answer 3)

Example

Compute $\lim_{x\to 0}\frac{\cos(x)-\sin(x)-1}{e^x-1}$. (See Answer 4)

Example

Compute $\lim_{x\to 0}\frac{\sqrt[3]{1+4x}-1}{\sqrt[5]{1+3x}-1}$. (See Answer 5)

There are other methods for computing these types of limits, including memorization, algebraic tricks, and l'Hopital's rule (more on that in the next module). However, in many cases, these different methods can all be replaced by a simple application of Taylor series.

7.5 EXERCISES

Compute the following limits:

$$\lim_{q\to 1}\frac{q^2+q+1}{q+3}$$

$$\lim_{x\to -2}\frac{x^2-4}{x+2}$$

$$\lim_{x\to 0}\frac{\sec x\tan x}{\sin x}$$

$$\lim_{x\to +\infty}\frac{6x^2-3x+1}{3x^2+4}$$

$$\lim_{x\to +\infty}\frac{x^2+x+1}{x^4-3x^2+2}$$

$$\lim_{y\to 0} \frac{\ln(1+2y)\sin y}{y^2\cos 2y}$$

$$\lim_{x\to 1} \frac{\ln x}{x^2}$$

$$\lim_{t\to 0} (3t^2+4t)\cot(t)$$

$$\lim_{z\to 0} \frac{z\cos(\sin(z))}{\sin(2z)}$$

$$\lim_{x\to 0} \frac{\ln(x+1)\arctan x}{x^2}$$

$$\lim_{x\to 0} \frac{\ln^2(\cos x)}{2x^4-x^5}$$

$$\lim_{s\to 0} \frac{e^s s\sin s}{1-\cos 2s}$$

$$\lim_{x\to 0^+} \frac{\sin(\arctan(\sin x))}{\sqrt{x}\sin 3x + x^2 + \arctan 5x}$$

$$\lim_{x\to 0} \frac{\sin x - \cos x - 1}{6xe^{2x}}$$

$$\lim_{x\to 0} \frac{\arctan x - 3\sin x + 2x}{3x^3}$$

$$\lim_{p\to 0} \frac{1-p-\cos 3p}{p^3}$$

$$\lim_{x\to\infty} x^{1/x}$$

7.6 Answers to Selected Examples

1. **Note** This is rather technical, and is only a demonstration of the process required to prove a limit exists from the definition. This course deals almost exclusively with continuous functions, where such proofs are not necessary.

   We must show that for any given $\epsilon > 0$, there exists δ (which depends on ϵ) such that $0 < |x-3| < \delta$ implies $|x^2-9| < \epsilon$.

   Let $\epsilon > 0$ be given. A little bit of algebra shows that

   $$|x^2-9| = |x-3|\cdot|x+3|.$$

   We get to control $|x-3|$ with δ. We also have (by using the triangle inequality) that

   $$|x+3| = |x-3+6| \le |x-3| + 6 < \delta + 6.$$

   Thus,

   $$|x^2-9| = |x-3|\cdot|x+3| < \delta\cdot(\delta+6).$$

   Now, if we pick δ to be the minimum of 1 and $\frac{\epsilon}{7}$, then we simultaneously guarantee that $\delta \le \frac{\epsilon}{7}$ and $\delta + 6 \le 7$, and so we find

   $$|x^2-9| < \delta\cdot(\delta+6) \le \frac{\epsilon}{7}\cdot 7 = \epsilon,$$

   as desired.

   (Return)

2. There are several proofs of this limit (e.g. memorization, l'Hospital's rule), but the simplest method is to use the Taylor series. Because x is near 0, the Taylor series expansion for $\sin x$ applies, and so

$$\begin{aligned}\lim_{x\to 0}\frac{\sin(x)}{x} &= \lim_{x\to 0}\frac{x-\frac{x^3}{3!}+\cdots}{x}\\ &= \lim_{x\to 0}\frac{x\left(1-\frac{x^2}{3!}+\cdots\right)}{x}\\ &= \lim_{x\to 0} 1-\frac{x^2}{3!}+\cdots\\ &= 1.\end{aligned}$$

This works because all the terms involving x go to 0 as x goes to 0.

(Return)

3. Replacing cos with its Taylor series (again, since x is near 0), we find

$$\begin{aligned}\lim_{x\to 0}\frac{1-\cos x}{x} &= \lim_{x\to 0}\frac{1-\left(1-\frac{1}{2!}x^2+\frac{1}{4!}x^4-\cdots\right)}{x}\\ &= \lim_{x\to 0}\frac{\frac{1}{2!}x^2-\frac{1}{4!}x^4+\cdots}{x}\\ &= \lim_{x\to 0}\frac{1}{2!}x-\frac{1}{4!}x^3+\cdots\\ &= 0.\end{aligned}$$

(Return)

4. Again, use the Taylor series (about $x = 0$) for each function:

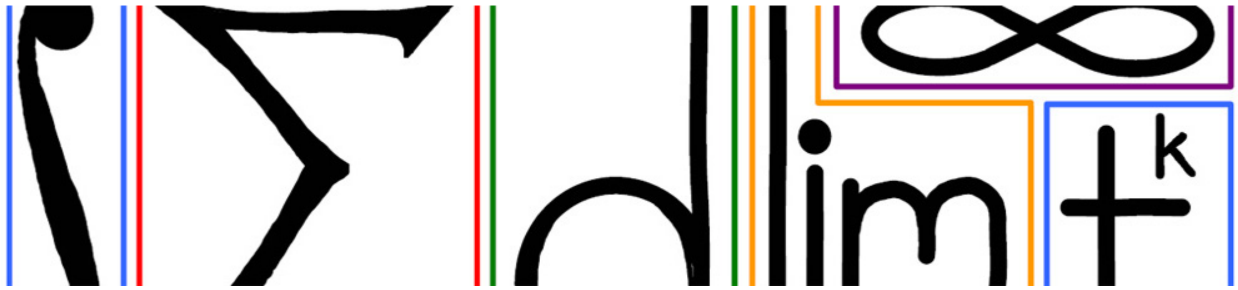
$$\begin{aligned}\lim_{x\to 0}\frac{\cos(x)-\sin(x)-1}{e^x-1} &= \lim_{x\to 0}\frac{(1-\frac{x^2}{2!}+\cdots)-(x-\frac{x^3}{3!}+\cdots)-1}{(1+x+\frac{x^2}{2!}+\cdots)-1}\\ &= \lim_{x\to 0}\frac{-x-\frac{x^2}{2!}+\cdots}{x+\cdots}\\ &= \lim_{x\to 0}\frac{x(-1-\cdots)}{x(1+\cdots)}\\ &= \lim_{x\to 0}\frac{-1-\cdots}{1+\cdots}\\ &= -1.\end{aligned}$$

(Return)

5. Here, we use the binomial series with $\alpha = \frac{1}{3}$ in the numerator, and $\alpha = \frac{1}{5}$ in the denominator. We find

$$\begin{aligned}
\lim_{x\to 0} \frac{(1+4x)^{1/3}-1}{(1+3x)^{1/5}-1} &= \lim_{x\to 0} \frac{\left(1+\frac{1}{3}(4x)+\text{ HOT}\right)-1}{\left(1+\frac{1}{5}(3x)+\text{ HOT}\right)-1} \\
&= \lim_{x\to 0} \frac{\frac{4}{3}x+\text{ HOT}}{\frac{3}{5}x+\text{ HOT}} \\
&= \lim_{x\to 0} \frac{\frac{4}{3}+\text{ HOT}}{\frac{3}{5}+\text{ HOT}} \\
&= \frac{\frac{4}{3}}{\frac{3}{5}} \\
&= \frac{20}{9}
\end{aligned}$$

(Return)



8 l'Hôpital's Rule

In previous modules, we saw that Taylor series are useful for computing certain limits of ratios. But sometimes, a fact known as l'Hôpital's rule is easier to use than Taylor series. While l'Hôpital's rule is commonly taught in a first calculus course, the justification for why it works is not usually taught. This module gives a justification for l'Hôpital's rule, using Taylor series.

8.1 l'Hôpital's rule

There are some limit situations where Taylor series are not particularly easy to use. For example, if the limit is being taken at a point about which the Taylor expansion is not already known, or the limit is at infinity, then using Taylor series is usually more work than it is worth. These are the situations where l'Hôpital's rule can be helpful.

l'Hôpital's Rule, $\frac{0}{0}$ case

If f and g are continuous functions such that $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$, provided this limit exists. If this is still of the form $\frac{0}{0}$, then derivatives may be taken again, and so on.

(See Justification 1)

Example

Using l'Hôpital's rule, compute two of the limits from the last module:

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

(See Answer 2)

Example

Compute $\lim_{x\to 0} \frac{\tan x}{\arcsin x}$. (See Answer 3)

Depending on the situation, it still might be easier to use Taylor series, especially if there are compositions and products of functions (assuming we know all the relevant Taylor series).

Example

Compute $\lim_{x\to 0} \frac{x^2 \ln(\cos x)}{\sin^2(3x^2)}$. (See Answer 4)

Example

Use l'Hopital's rule to compute $\lim_{x\to\pi} \frac{\sin(x)}{e^x \cos(x/2)}$. (See Answer 5)

l'Hôpital's Rule, $\frac{\infty}{\infty}$ case

If $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$, then $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$, again provided this limit exists.

Example

Compute $\lim_{x\to 0} \frac{\ln x}{\frac{1}{x^2}}$. (See Answer 6)

8.2 Other indeterminate forms

Some limits do not initially look like cases where l'Hôpital's rule applies, but with some algebra they can be rearranged into one of the applicable cases. These are called *indeterminate forms*.

Case: $\infty - \infty$

First, consider $\lim_{x\to a} f(x) - g(x)$, where $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$. Usually, one or both of f and g are ratios of other functions. In this case, getting a common denominator usually transforms the limit into one where l'Hospital's rule or a Taylor series approach applies.

Example

Compute $\lim_{x\to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2}$. (See Answer 7)

Case: $\infty \cdot 0$

Next, consider $\lim_{x\to a} f(x)g(x)$ where $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} g(x) = 0$. Since $\infty \cdot 0$ is not defined, it is not clear what this limit is. However, the product can be turned into one of the following ratios where l'Hôpital's rule applies:

$$\lim_{x\to a} f(x)g(x) = \lim_{x\to a} \frac{g(x)}{1/f(x)} \quad \lim_{x\to a} f(x)g(x) \qquad = \lim_{x\to a} \frac{f(x)}{1/g(x)}$$

since dividing by the reciprocal of a number is the same as multiplying. Now, note that $\lim_{x\to a} 1/f(x) = 0$, since $\lim_{x\to a} f(x) = \infty$. Thus, $\lim_{x\to a} \frac{g(x)}{1/f(x)}$ is now in the $\frac{0}{0}$ case of l'Hôpital's rule.

Similarly, $\lim_{x\to a} \frac{f(x)}{1/g(x)}$ is in the $\frac{\infty}{\infty}$ form of l'Hôpital's rule, and so it can be applied here too.

Deciding which of the above forms to use depends on the situation, but in many situations either form will work.

Example

Compute $\lim_{x\to 0^+} x \ln x$. (Here we use the one-sided limit $x \to 0^+$ becuase ln is only defined on the positive real numbers). (See Answer 8)

Case: ∞^0

Another indeterminate form arises when raising one function of x to a power which involves another function of x. Suppose $\lim_{x\to a} f(x) = \infty$ and $\lim_{x\to a} g(x) = 0$. Then what is $\lim_{x\to a} f(x)^{g(x)}$?

On the one hand, it seems that raising ∞ to any power should be ∞. On the other hand, raising anything to the 0th power should be 1. To find what the answer actually is, let

$$y = \lim_{x\to a} f(x)^{g(x)},$$

and take the ln of both sides. Now, recall that ln is a continuous function. Therefore, from the rules of limits in the last module, taking ln of a limit is the same as the limit of the ln:

$$\begin{aligned} \ln(y) &= \ln\left(\lim_{x\to a} f(x)^{g(x)}\right) \\ &= \lim_{x\to a} \ln\left(f(x)^{g(x)}\right) \\ &= \lim_{x\to a} g(x) \ln(f(x)) \end{aligned}$$

(the last step uses the fact that $\ln(a^b) = b \ln(a)$). Now, this is of the form $0\cdot\infty$, which was covered above. Note that when this limit is computed, it is $\ln(y)$ which has been found, and so the answer must be exponentiated to find y, the original limit.

Example

Compute $\lim_{x\to\infty} x^{1/x}$. (See Answer 9)

Case: 0^0

Consider the limit $f(x)^{g(x)}$, where $\lim_{x\to a} f(x) = 0$ and $\lim_{x\to a} g(x) = 0$. Because 0^0 is not defined, this is another indeterminate form. It can be dealt with as the ∞^0 case, by first taking ln, computing the resulting limit, and then exponentiating.

Example

Compute $\lim_{x \to 0^+} x^x$. (See Answer 10)

8.3 Limits going to infinity

I'Hôpital also works with limits going to infinity; the same hypothesis and conclusions hold. Before doing some examples, what does it mean for $\lim_{x \to \infty} f(x) = L$?

Limit at infinity

$\lim_{x \to \infty} f(x) = L$ if and only if for every $\epsilon > 0$ there exists $M > 0$ such that $|f(x) - L| < \epsilon$ whenever $x > M$. If there is no such L, then the limit does not exist.

Example

Compute $\lim_{x \to \infty} \frac{\ln x}{\sqrt{x}}$. (See Answer 11)

Example

Compute $\lim_{x \to \infty} \frac{e^x}{x^2}$. (See Answer 12)

While l'Hopital often works, there are situations where it fails to give an answer, and a little extra thought must be employed.

Example

Compute $\lim_{x \to \infty} \tanh x$. (See Answer 13)

Example

Compute $\lim_{x \to \infty} \frac{x \ln x}{\ln(\cosh x)}$. (See Answer 14)

It is also possible to deal with limits going to infinity using Taylor series, but it involves some algebra. The idea is to use a substitution to turn the limit going to infinity into a limit going to zero. Symbolically, if $x \to \infty$, then let $z = 1/x$. It follows that $z \to 0$ as $x \to \infty$, and by replacing x with $1/z$ throughout, the limit is transformed.

$$\lim_{x \to \infty} f(x) = \lim_{z \to 0} f(1/z).$$

This process works when the limit at 0 exists. A more general technique would only look at the one-sided limit from the right-hand side:

$$\lim_{x \to \infty} f(x) = \lim_{z \to 0^+} f(1/z).$$

Example

Compute $\lim_{x\to\infty} \dfrac{\sin(1/x)}{1/x}$. (See Answer 15)

8.4 EXERCISES

Compute the following limits. Should you use l'Hôpital's rule or Taylor expansion?

$$\lim_{x\to 2} \frac{x^3+2x^2-4x-8}{x-2}$$

$$\lim_{x\to \pi/3} \frac{1-2\cos x}{\pi-3x}$$

$$\lim_{x\to \pi} \frac{4\sin x\cos x}{\pi-x}$$

$$\lim_{x\to 9} \frac{2x-18}{\sqrt{x}-3}$$

$$\lim_{x\to 0} \frac{e^x-\sin x-1}{x^2-x^3}$$

$$\lim_{x\to 1} \frac{\cos(\pi x/2)}{1-\sqrt{x}}$$

$$\lim_{x\to 4} \frac{3-\sqrt{5+x}}{1-\sqrt{5-x}}$$

$$\lim_{x\to 0} \left(\frac{1}{x}-\frac{1}{\ln(x+1)}\right)$$

$$\lim_{x\to \pi/2} \frac{\sin x\cos x}{e^x\cos 3x}$$

$$\lim_{x\to +\infty} \frac{\ln x}{e^x}$$

$$\lim_{x\to +\infty} x\ln\left(1+\frac{3}{x}\right)$$

$$\lim_{x\to +\infty} \frac{(\ln x)(\sinh x)}{(x-1)e^x}$$

8.5 Answers to Selected Examples

1. The Taylor series for f and g about a are given by

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$$
$$g(x) = g(a) + g'(a)(x-a) + \frac{g(a)}{2!}(x-a)^2 + \cdots .$$

Since, by hypothesis, $f(a) = g(a) = 0$, it follows that

$$\begin{aligned}
\lim_{x\to a} \frac{f(x)}{g(x)} &= \lim_{x\to a} \frac{f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots}{g(a) + g'(a)(x-a) + \frac{1}{2}g(a)(x-a)^2 + \cdots} \\
&= \lim_{x\to a} \frac{f'(a)(x-a) + (1/2)f''(a)(x-a)^2 + \cdots}{g'(a)(x-a) + (1/2)g(a)(x-a)^2 + \cdots} \\
&= \lim_{x\to a} \frac{(x-a)\left[f'(a) + (1/2)f''(a)(x-a) + \cdots\right]}{(x-a)\left[g'(a) + (1/2)g(a)(x-a) + \cdots\right]} \\
&= \lim_{x\to a} \frac{f'(a) + (1/2)f''(a)(x-a) + \cdots}{g'(a) + (1/2)g(a)(x-a) + \cdots}.
\end{aligned}$$

Now, as $x \to a$, all the terms with $x - a$ go to 0, which leaves $\frac{f'(a)}{g'(a)}$. If this fraction is still 0/0, then l'Hôpital's rule says to take the derivative of the numerator and the denominator again. In terms of the Taylor series, this moves to the next leading terms in the numerator and denominator.

(Return)

2. These are both in the $\frac{0}{0}$ case, so differentiating numerator and denominator gives

$$\begin{aligned}
\lim_{x\to 0} \frac{\sin x}{x} &= \lim_{x\to 0} \frac{\cos x}{1} \\
&= 1.
\end{aligned}$$

$$\begin{aligned}
\lim_{x\to 0} \frac{1-\cos x}{x} &= \lim_{x\to 0} \frac{\sin x}{1} \\
&= 0.
\end{aligned}$$

(Return)

3. Since $\tan 0 = \arcsin 0 = 0$, we are in the $\frac{0}{0}$ case of l'Hôpital's rule. Recall that

$$\begin{aligned}
\frac{d}{dx} \tan x &= \sec^2 x \\
\frac{d}{dx} \arcsin x &= \frac{1}{\sqrt{1-x^2}}.
\end{aligned}$$

Thus, applying l'Hôpital's rule gives

$$\begin{aligned}
\lim_{x\to 0} \frac{\tan x}{\arcsin x} &= \lim_{x\to 0} \frac{\sec^2 x}{\frac{1}{\sqrt{1-x^2}}} \\
&= \frac{1}{1} \\
&= 1.
\end{aligned}$$

(Return)

4. We know all the relevant Taylor series for the functions in this problem, so that should be an easier

method. We find

$$\begin{aligned}
\lim_{x\to 0} \frac{x^2 \ln(\cos x)}{\sin^2(3x^2)} &= \lim_{x\to 0} \frac{x^2 \ln\left(1 - \frac{x^2}{2!} + \text{ HOT}\right)}{(3x^2 + \text{ HOT})^2} \\
&= \lim_{x\to 0} \frac{x^2\left(-\frac{x^2}{2} - \text{ HOT}\right)}{9x^4 + \text{ HOT}} \\
&= \lim_{x\to 0} \frac{-\frac{x^4}{2} + \text{ HOT}}{9x^4 + \text{ HOT}} \\
&= -\frac{1}{18}.
\end{aligned}$$

Using l'Hôpital here would be quite a lot of work. It turns out that we would have to apply the rule four times, which involves a lot of product and chain rule.

(Return)

5. Since $\sin, \cos, \exp$ are all continuous functions, and $\sin(\pi) = e^{\pi}\cos(\pi/2) = 0$, the hypotheses for l'Hôpital's rule are met. So it follows that

$$\begin{aligned}
\lim_{x\to\pi} \frac{\sin(x)}{e^x \cos(x/2)} &= \lim_{x\to\pi} \frac{[\sin(x)]'}{[e^x \cos(x/2)]'} \\
&= \lim_{x\to\pi} \frac{\cos(x)}{e^x\cos(x/2) - (1/2)e^x \sin(x/2)} \\
&= \frac{-1}{0 - (1/2)e^{\pi}\sin(\pi/2)} \\
&= 2e^{-\pi}.
\end{aligned}$$

Note that although we know the Taylor series for these functions at $x = 0$, the limit here is as $x \to \pi$. Thus, we cannot use the Taylor series approach, because a Taylor series about $x = 0$ does not give a good approximation when x is not close to 0.

(Return)

6. Note that $\ln x \to -\infty$ as $x \to 0$, and $\frac{1}{x^2} \to \infty$ as $x \to 0$. Therefore, the $\frac{\infty}{\infty}$ case of l'Hôpital's rule applies. Applying the rule,

$$\begin{aligned}
\lim_{x\to 0} \frac{\ln x}{x^{-2}} &= \lim_{x\to 0} \frac{\frac{1}{x}}{-2x^{-3}} \\
&= \lim_{x\to 0} \frac{x^2}{-2} \\
&= 0.
\end{aligned}$$

(Return)

7. Getting a common denominator, and then Taylor expanding gives

$$\begin{aligned}
\lim_{x\to 0} \frac{1}{\sin^2 x} - \frac{1}{x^2} &= \lim_{x\to 0} \frac{x^2 - \sin^2 x}{x^2 \sin^2 x} \\
&= \lim_{x\to 0} \frac{x^2 - \left(x - \frac{1}{3!}x^3 + \cdots\right)^2}{x^2\left(x - \frac{1}{3!}x^3 + \cdots\right)^2} \\
&= \lim_{x\to 0} \frac{x^2 - \left(x^2 - \frac{2}{3!}x^4 + \cdots\right)}{x^2\left(x^2 - \frac{2}{3!}x^3 + \cdots\right)} \\
&= \lim_{x\to 0} \frac{\frac{2}{6}x^4 + \cdots}{x^4 + \cdots} \\
&= \frac{1}{3}.
\end{aligned}$$

(Return)

8. This is of the form $0 \cdot (-\infty)$. In this case, it is easier to flip x into the denominator, because it is easier to take the derivative of x^{-1} than it is to take the derivative of $(\ln x)^{-1}$. So, we find

$$\lim_{x\to 0^+} x \ln x = \lim_{x\to 0^+} \frac{\ln x}{x^{-1}},$$

which is of the form $\frac{-\infty}{\infty}$, so applying l'Hôpital's rule gives

$$\begin{aligned}
\lim_{x\to 0^+} \frac{\ln x}{x^{-1}} &= \lim_{x\to 0^+} \frac{x^{-1}}{-x^{-2}} \\
&= \lim_{x\to 0^+} -x \\
&= 0.
\end{aligned}$$

(Return)

9. This is of the ∞^0 form. Let $y = \lim_{x\to\infty} x^{1/x}$. Then taking ln gives

$$\begin{aligned}
\ln(y) &= \lim_{x\to\infty} \ln\left(x^{1/x}\right) \\
&= \lim_{x\to\infty} \frac{1}{x}\ln(x) \\
&= \lim_{x\to\infty} \frac{\ln x}{x} \\
&= \lim_{x\to\infty} \frac{1/x}{1} \\
&= 0
\end{aligned}$$

(l'Hôpital's rule was used in the second to last step). So $\ln(y) = 0$, and so $y = 1$ is the answer.

(Return)

10. Letting $y = \lim_{x\to 0^+} x^x$, and taking logarithms gives

$$\begin{aligned}
\ln y &= \lim_{x\to 0^+} x \ln x \\
&= 0,
\end{aligned}$$

by an example above. Thus $y = e^0 = 1$.

(Return)

11. Both the numerator and denominator go to ∞ as $x \to \infty$, so applying l'Hôpital's rule gives

$$\begin{aligned} \lim_{x\to\infty} \frac{\ln x}{\sqrt{x}} &= \lim_{x\to\infty} \frac{x^{-1}}{\frac{1}{2}x^{-1/2}} \\ &= \lim_{x\to\infty} \frac{2}{\sqrt{x}} \\ &= 0. \end{aligned}$$

(Return)

12. The numerator and denominator both go to ∞ as $x \to \infty$, so this is the $\frac{\infty}{\infty}$ case of l'Hôpital's rule. It follows that

$$\begin{aligned} \lim_{x\to\infty} \frac{e^x}{x^2} &= \lim_{x\to\infty} \frac{[e^x]'}{[x^2]'} \\ &= \lim_{x\to\infty} \frac{e^x}{2x} \\ &= \lim_{x\to\infty} \frac{e^x}{2} \\ &= \infty. \end{aligned}$$

Note that l'Hôpital's rule was used twice here since $\lim_{x\to\infty} \frac{e^x}{2x}$ is still the $\frac{\infty}{\infty}$ case.

(Return)

13. Recall that $\tanh x = \frac{\sinh x}{\cosh x}$. Both $\sinh x \to \infty$ and $\cosh x \to \infty$ as $x \to \infty$. But applying l'Hôpital's rule (and then applying it again) gives

$$\begin{aligned} \lim_{x\to\infty} \tanh x &= \lim_{x\to\infty} \frac{\sinh x}{\cosh x} \\ &= \lim_{x\to\infty} \frac{\cosh x}{\sinh x} \\ &= \lim_{x\to\infty} \frac{\sinh x}{\cosh x}, \end{aligned}$$

so l'Hôpital's rule clearly will not give us an answer here. Instead, writing out the definition of $\tanh x$ and doing a little algebra, we find

$$\begin{aligned} \lim_{x\to\infty} \tanh x &= \lim_{x\to\infty} \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ &= \lim_{x\to\infty} \frac{e^x(1 - e^{-2x}}{e^x(1 + e^{-2x})} \\ &= 1. \end{aligned}$$

(Return)

14. Using l'Hôpital's rule, we find

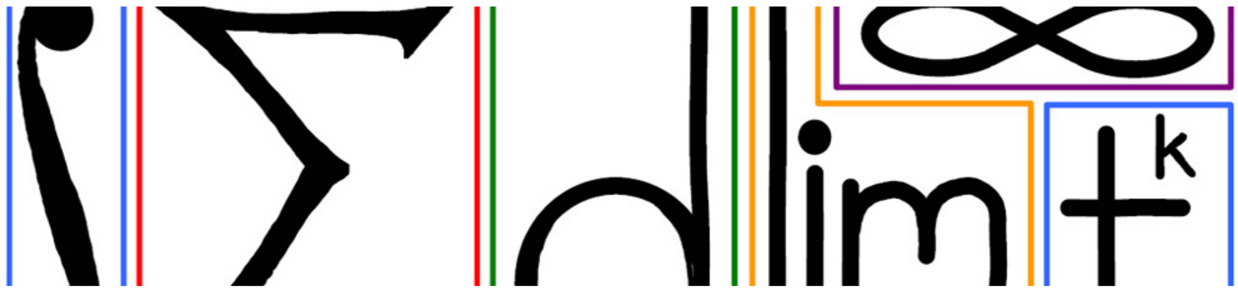
$$\begin{aligned}
\lim_{x\to\infty} \frac{x \ln x}{\ln(\cosh x)} &= \lim_{x\to\infty} \frac{\frac{x}{x} + 1 \cdot \ln x}{\frac{\sinh x}{\cosh x}} \\
&= \lim_{x\to\infty} \frac{1 + \ln x}{\tanh x} \\
&= \infty,
\end{aligned}$$

since the denominator goes to 1 (from the previous example) and the numerator goes to infinity.
(Return)

15. Using the substitution $z = 1/x$, the limit becomes

$$\begin{aligned}
\lim_{x\to\infty} \frac{\sin(1/x)}{1/x} &= \lim_{z\to 0} \frac{\sin(z)}{z} \\
&= 1,
\end{aligned}$$

as was shown in the last module.
(Return)

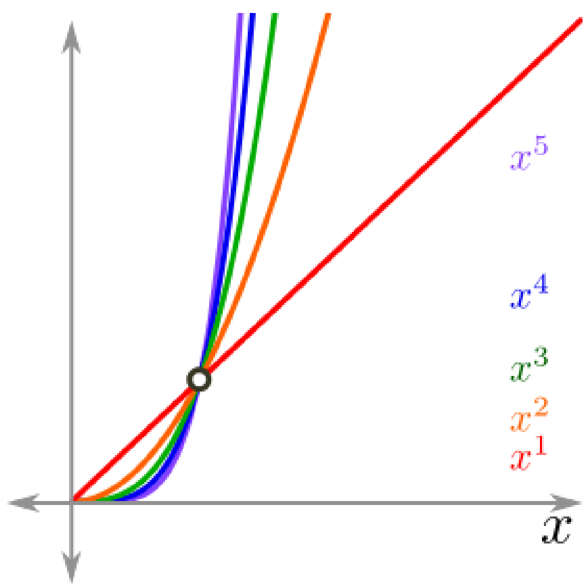


9 Orders of Growth

When dealing with limits as $x \to 0$, it is the lowest order term (i.e. the term with the smallest power) which matters the most, since higher powers of x are very small when x is close to 0. On the other hand, as $x \to \infty$, what is known as the *asymptotic growth* of a function. In this case, it is the highest order term which matters the most. This module deals with limits of both types and provides a more formal notion of how quickly a function grows or shrinks.

9.1 Hierarchy of functions going to infinity

First, consider the monomial x^n, where n is a fixed, positive integer. Looking at the graphs of these monomials, it becomes clear that as $x \to \infty$, $x^{n+1} > x^n$:



What happens when the functions involved are not polynomials? For example, how does the growth of the exponential compare to a polynomial? Or factorial and exponential?

In general, one can compare the asymptotic growth of the functions f and g by considering the limit $\lim_{x\to\infty} \frac{f(x)}{g(x)}$. If this limit is ∞, then f dominates. If the limit is 0, then g dominates. And if the limit is a constant, then f and g are considered equal (asymptotically).

Example

Compare exponential growth and polynomial growth. (See Answer 1)

Example

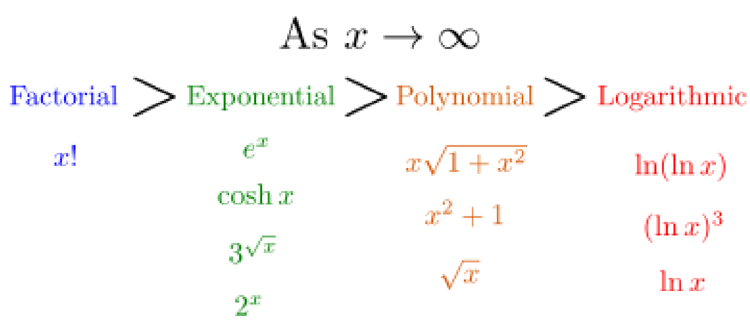
Compare the asymptotic behavior of $\ln^n x$ (for some fixed integer n) and x. (See Answer 2)

Example

Compare the asymptotic growth of the factorial and the exponential. (See Answer 3)

Thus we have the following hierarchy of growth, from greatest to smallest:

1. Factorial: $x! = x(x-1)(x-2)\cdots 3\cdot 2\cdot 1$.
2. Exponential: c^x for any $c > 1$ (usually $c = e$).
3. Polynomial: x^k for any $k > 0$.
4. Logarithmic: $\ln(x)$ and other related functions.



When a pair of functions are of a similar type, such as e^x and $3^{\sqrt{x}}$, one must compare these using the limit of the ratio of the functions, as above.

9.2 Hierarchy of functions going to 0

As above, first consider monomials x^n. As $x \to 0$, the inequality for monomials is the reverse of what it was for $x \to \infty$. That is, as $x \to 0$, we find that $x^n > x^{n+1}$. Intuitively, small numbers become even smaller when you raise them to higher powers.

It is important to keep track of whether a limit is going to 0 or ∞, since in the first case, the lowest order terms dominate, and in the second case the highest order terms dominate.

Example

Compute $\displaystyle\lim_{x\to 0} \frac{2x^3 - x^2 + x}{x^3 + 2x}$ and $\displaystyle\lim_{x\to\infty} \frac{2x^3 - x^2 + x + 1}{x^3 + 2x + 2}$. (See Answer 4)

9.3 Big-O notation

When dealing with limits as $x \to 0$ or $x \to \infty$, it is best to have a formal notation for the approximations which result from dropping higher or lower order terms. Big-O notation, pronounced "big oh", provides this formality.

Big-O notation, $x \to 0$

The function $f(x)$ is in $O(x^n)$, as $x \to 0$ if

$$|f(x)| < C|x|^n$$

for some constant C and all x sufficiently close to 0. Put another way, a function $f(x)$ is in $O(x^n)$, for x close to 0, if $f(x)$ approaches 0 at least as fast as a constant multiple of x^n.
More generally, a function $f(x)$ is in $O(g(x))$, as $x \to 0$ if

$$|f(x)| < C|g(x)|,$$

for some constant C and x sufficiently close to 0.

Big-O notation, as $x \to 0$, can be thought of as a more specific way of saying higher order terms. Just as Taylor series could include a different number of terms before indicating the rest is higher order terms (depending on the situation), the same is true for big-O.

Example

Express $\arctan(x)$ using big-O notation as $x \to 0$. (See Answer 5)

The definition for big-O as $x \to \infty$ is almost identical, except the bound needs to apply for all x sufficiently large.

Big-O notation, $x \to \infty$

The function $f(x)$ is in $O(x^n)$, as $x \to \infty$ if

$$|f(x)| < C|x|^n$$

for some constant C and all x sufficiently large. In other words, a function $f(x)$ is in $O(x^n)$, as $x \to \infty$, if $f(x)$ approaches infinity no faster than a constant multiple of x^n.
More generally, $f(x)$ is in $O(g(x))$, as $x \to \infty$ if

$$|f(x)| < C|g(x)|$$

for some constant C and all x sufficiently large.

Example

The monomial x^n is in $O(e^x)$ as $x \to \infty$ for any (fixed) n. This is a restatement of the above fact that the exponential dominates polynomials as $x \to \infty$.

Example

Show that $x\sqrt{x^2+3x+5} = x^2 + \frac{3}{2}x + O(1)$ as $x \to \infty$. Hint: use the binomial series. (See Answer 6)

Example

Justify the following statements as $x \to 0$:

1. $5x + 3x^2$ is in $O(x)$ but is not in $O(x^2)$.
2. $\sin x$ is in $O(x)$ but is not in $O(x^2)$.
3. $\ln(1+x) - x$ is in $O(x^2)$ but is not in $O(x^3)$.
4. $1 - \cos(x^2)$ is in $O(x^4)$ but is not in $O(x^5)$.
5. $\sqrt{x}$ is not in $O(x^n)$ for any $n \geq 1$.
6. e^{-1/x^2} is in $O(x^n)$ for all n.

(See Answer 7)

Example

Justify the following statements as $x \to \infty$:

1. $\arctan x$ is in $O(1)$ as well as $O(x^n)$ for any $n \geq 0$.
2. $x\sqrt{1+x^2}$ is in $O(x^2)$ but is not in $O(x^{3/2})$.
3. $\ln \sinh x$ is in $O(x)$ but is not in $O(\ln x)$.
4. $\cosh x$ is in $O(e^x)$ but is not in $O(x^n)$ for any $n \geq 0$.
5. $\ln(x^5)$ is in $O(\ln x)$ as well as $O(x^n)$ for all n.
6. x^x is in $O(e^{x^n})$ for all $n > 1$.

(See Answer 8)

9.4 Application: Error Analysis

When approximating a function by the first few terms of its Taylor series, there is a trade-off between convenience (the ease of the computation) and accuracy. Big-O notation can help keep track of this error.

Example

Consider the approximation, for x close to 0,

$$\begin{aligned}\sin(x^2)e^x &= \left(x^2 + O(x^6)\right)(1 + O(x)) \\ &= x^2 + x^2 \cdot O(x) + O(x^6) + O(x^6) \cdot O(x) \\ &= x^2 + O(x^3).\end{aligned}$$

So the error of approximating $\sin(x^2)e^x \approx x^2$ can be bounded by Cx^3, for some C when x is small. Determining a good C, in general, is tricky, and there is more on error bounds in Taylor Remainder Theorem.

9.5 Application: Computational Complexity

In computer science, an algorithm is a sequence of steps used to solve a problem. For example, there are algorithms to sort a list of numbers and algorithms to find the prime factorization of a number. Computational complexity is a measure of how efficient an algorithm is. Basically, a computer scientist wants to know roughly how the number of computations carried out by the computer will grow as the input (e.g. the length of the list to be sorted, or the size of the number to be factored) gets larger.

Example: Multiplication

Consider how much work is required to multiply two n-digit numbers using the usual grade-school method. There are two phases to working out the product: multiplication and addition.

First, multiply the first number by each of the digits of the second number. For each digit in the second number this requires n basic operations (multiplication of single digits) plus perhaps some "carries", so say a total of $2n$ operations for each digit in the second number. This means that the multiplication phase requires $n \cdot (2n)$ basic operations.

The addition phase requires repeatedly adding n digit numbers together a total of $n-1$ times. If each addition requires at most $2n$ operations (including the carries), and there are $n-1$ additions that must be made, it comes to a total of $(2n)(n-1)$ operations in the addition phase.

Adding these totals up gives about $4n^2$ total operations. Thus, the total number of basic operations that must be performed in multiplying two n digit numbers is in $O(n^2)$ (since the constant coefficient does not matter).

The reason the constant coefficient does not really matter when thinking about computational complexity, is that a faster computer can only improve the speed of a computation by a constant factor. The only way to significantly improve a computation is to somehow drastically cut the number of operations required to perform the operation. The next example shows an example of how important algorithmic improvements can be on computational complexity.

Example: Sorting

In sorting algorithms, the most basic operation is the comparison. For a sorting algorithm, one wants to know how many comparisons of two numbers will be made, on average.

One common sorting algorithm, which is used by most people who are sorting items by hand, is called Insertion Sort. It turns out that the number of comparisons for Insertion Sort, on average, is $O(n^2)$ as $n \to \infty$, where n is the length of the list of numbers to be sorted.

A more sophisticated sorting algorithm, called Mergesort, uses $O(n \ln(n))$ comparisons on average. This may not seem like a significant improvement over Insertion Sort, but consider the number of comparisons used to sort a list of 1000000 integers: Insertion Sort would use on the order of $1000000^2 = 10^{12}$ comparisons on average, where as Mergesort uses on the order of 13×10^6 comparisons. To put that in perspective, if Mergesort took a half second to complete the computation, Insertion Sort would take over ten hours!

9.6 Big O in Other Areas of Mathematics

Stirling's formula

Stirling's formula gives an asymptotic approximation for $x!$:

$$\ln(x!) = x \ln x - x + O(\ln x).$$

In a slightly more precise form, it can be written

$$x! = \sqrt{2\pi x}\left(\frac{x}{e}\right)^x \left(1 + O\left(\frac{1}{x}\right)\right).$$

Prime number theorem

In number theory, one very important function, $\pi(x)$, is defined to be the number of primes less than or equal to x. The Prime Number Theorem says that

$$\begin{aligned} \pi(x) &= \text{ number of primes } \leq x \\ &= \frac{x}{\ln x}\left(1 + O\left(\frac{1}{\ln x}\right)\right). \end{aligned}$$

9.7 EXERCISES

- Compute the following limits.

$$\lim_{x\to+\infty} \frac{e^{2x}}{x^3 + 3x^2 + 4}$$

$$\lim_{x\to+\infty} \frac{e^{3x}}{e^{x^2}}$$

$$\lim_{x\to+\infty} \frac{e^x (x-1)!}{x!}$$

$$\lim_{x\to+\infty} \frac{(^2}{x} - 3)(x^2+3)2x^4 - 2x^2 + 1$$

$$\lim_{x\to+\infty} \frac{2^x + 1}{(x+1)!}$$

$$\lim_{x\to+\infty} \frac{(3\ln x)^n}{(2x)^n}$$

- Simplify the following asymptotic expression as $x \to 0$

$$f(x) = \left(x - x^2 + O(x^3)\right) \cdot \left(1 + 2x + O(x^3)\right)$$

- Simplify the following asymptotic expression as $x \to \infty$

$$f(x) = \left(x^3 + 2x^2 + O(x)\right) \cdot \left(1 + \frac{1}{x} + O\left(\frac{1}{x^2}\right)\right)$$

- Here are some rules for positive functions with $x \to \infty$:

$$O(f(x)) + O(g(x)) = O(f(x) + g(x))$$
$$O(f(x)) \cdot O(g(x)) = O(f(x) \cdot g(x))$$

  Using these, show that

$$O\left(\frac{5}{x}\right) + O\left(\frac{\ln(x^2)}{4x}\right)$$

  simplifies to

$$O\left(\frac{\ln x}{x}\right)$$

- Which of the following are in $O(x^2)$ as $x \to 0$?

$$x \ln(1+x)$$
$$5x^2 + 6x + 1$$
$$1 - e^{-x}$$
$$x\sqrt{x^2 + 4x^3 + 5x^6}$$
$$x \sinh^2(3x)$$
$$\frac{x^2}{\ln(1+x)}$$

9.8 Answers to Selected Examples

1. Fix an integer n, and compute the limit (using l'Hopital repeatedly):

$$\begin{aligned}
\lim_{x\to\infty} \frac{x^n}{e^x} &= \lim_{x\to\infty} \frac{nx^{n-1}}{e^x} \\
&= \lim_{x\to\infty} \frac{n(n-1)x^{n-2}}{e^x} \\
&= \vdots \\
&= \lim_{x\to\infty} \frac{n!}{e^x} \\
&= 0
\end{aligned}$$

   (note that n is fixed, so $n!$ is a constant). Thus, the exponential dominates the monomial x^n (and thus any polynomial as well).

   (Return)

2. Again, using L'Hopital's rule repeatedly gives

$$\begin{aligned}\lim_{x\to\infty}\frac{\ln^n x}{x} &= \lim_{x\to\infty}\frac{n\ln^{n-1}x\frac{1}{x}}{1}\\ &= \lim_{x\to\infty}\frac{n\ln^{n-1}x}{x}\\ &= \vdots\\ &= \lim_{x\to\infty}\frac{n!}{x}\\ &= 0,\end{aligned}$$

since $n!$ is a constant here. This shows any polynomial beats any constant power of logarithm.

(Return)

3. First, when x is not an integer, the factorial is defined by

$$x! = \int_{t=0}^{\infty} t^x e^{-t}\,dt.$$

This definition shares properties with the traditional definition of factorial: $x! = x\cdot(x-1)!$, and $0! = 1$, and they coincide when x is an integer. It is not critical to know this integral definition, but know that $n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$ for an integer n.

When x is an integer, note that e^x is e multiplied with itself x times. On the other hand, $x!$ has x factors, most of which are bigger than e (at least when x is bigger than, say, 5). So as x increases to $x+1$, e^x only gains another factor of e, but $x!$ gains a factor of $x+1$. This explains why $x!$ grows faster than e^x (and similarly for any exponential function).

(Return)

4. As $x\to 0$, the higher powers of x go to 0 quickly, leaving the lowest order terms $\frac{x}{2x} = \frac{1}{2}$ in the limit.

As $x\to\infty$, it is the highest order terms (the x^3 terms) which dominate, so ignoring the lower order terms leaves $\frac{2x^3}{x^3} = 2$ in the limit.

(Return)

5. By taking the Taylor series for $\arctan(x)$, we find that

$$\arctan(x) = x - \frac{x^3}{3} + O(x^5)$$

as $x\to 0$.

We could also say

$$\arctan(x) = x - O(x^3)$$

as $x\to 0$.

(Return)

6. Recall that the binomial series

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \dots$$

requires that $|x| < 1$. So we must do a little algebra to get the square root to be of this form. Factoring out a x^2 will do the trick:

$$\begin{aligned} x\sqrt{x^2+3x+5} &= x\sqrt{x^2\left(1+\frac{3}{x}+\frac{5}{x^2}\right)} \\ &= x^2\left(1+\frac{3}{x}+\frac{5}{x^2}\right)^{1/2} \\ &= x^2\left(1+\frac{1}{2}\left(\frac{3}{x}+\frac{5}{x^2}\right)+O\left(\frac{1}{x^2}\right)\right) \\ &= x^2+\frac{3}{2}x+O(1), \end{aligned}$$

since all the terms involving $\frac{1}{x^2}$ become constants when multiplied by x^2.

(Return)

7. The first four of these can be justified by looking at the Taylor series and then finding the monomial of the lowest power (remember that as $x \to 0$, the dominant term is that of the lowest power).

   To see that $\sqrt{x}$ is not in $O(x^n)$ for any $n \geq 1$, we must show that given any constant C and $\epsilon > 0$, there exists some $x < \epsilon$ such that

   $$\sqrt{x} > Cx^n$$

   (this is the negation of the definition of big oh). Solving the above inequality for x, one finds that if

   $$x < \left(\frac{1}{C}\right)^{2/(2n-1)},$$

   then $\sqrt{x} > Cx^n$, and so $\sqrt{x}$ is not in $O(x^n)$ for any n.

   Finally, to see that e^{-1/x^2} is in $O(x^n)$, it suffices to show that

   $$e^{-1/x^2} < |x|^n$$

   for all x sufficiently small. Taking natural log of both sides (this preserves the inequality since log is an increasing function) gives

   $$-\frac{1}{x^2} - n\ln|x|.$$

   But recall from a previous module that $x \ln x \to 0$ as $x \to 0^+$. This ensures that no matter the (fixed) value of n, for sufficiently small x we will have

   $$\frac{1}{n} > -x^2 \ln|x|,$$

   and hence

   $$e^{-1/x^2} < |x|^n,$$

   as desired.

   (Return)

8. (a) Note that as $x \to \infty$, $\arctan x \to \frac{\pi}{2}$. Thus $\arctan x$ is bounded, and so it is in $O(1)$.

(b) Using the binomial series, as in a previous example, shows that

$$\begin{aligned}
x\sqrt{1+x^2} &= x\left(x^2\left(1+\frac{1}{x^2}\right)\right)^{1/2} \\
&= x^2\left(1+\frac{1}{x^2}\right)^{1/2} \\
&= x^2\left(1+\frac{1}{2}\frac{1}{x^2}+O\left(\frac{1}{x^4}\right)\right) \\
&= x^2+O(1).
\end{aligned}$$

Thus, $x\sqrt{1+x^2}$ is in $O(x^2)$, but no power smaller than x^2.

(a) For large x, e^{-x} is very small, and so $\sinh x \approx \frac{e^x}{2}$. Therefore,

$$\ln\sinh x \approx \ln\frac{e^x}{2} = x - \ln 2.$$

is in $O(x)$ but not in $O(\ln x)$, since logarithms are smaller than polynomials.

(a) Similarly, $\cosh x \approx \frac{e^x}{2}$ for large x. Hence cosh is in $O(e^x)$ but cannot be bounded by any polynomial.

(b) A handy property of logarithms tells us that

$$\ln(x^5) = 5\ln x,$$

which is in $O(\ln x)$ since it is itself a multiple of $\ln x$.

(a) Fix $n > 1$. It suffices to show that for large enough x, we have

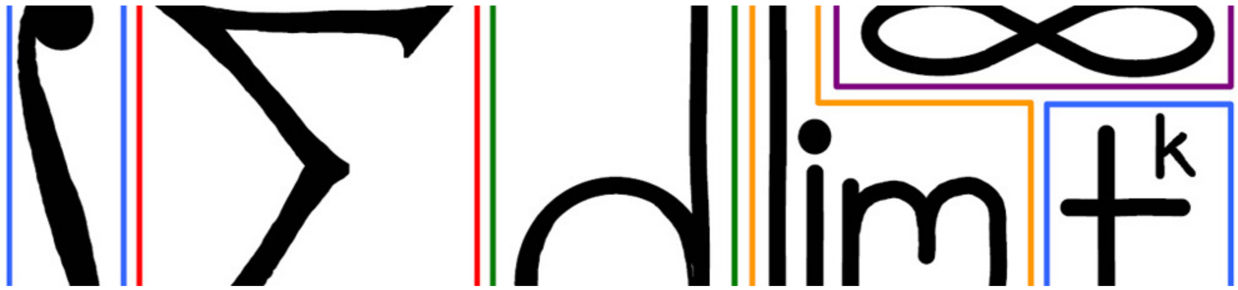
$$x^x < e^{x^n}.$$

Taking the logarithm of both sides gives

$$x\ln x < x^n \quad \Leftrightarrow \quad \ln x < x^{n-1}.$$

We saw earlier that any positive power of x beats the logarithm for large values of x. So (since $n > 1$) this inequality holds for large x, and so x^x is in $O(e^{x^n})$.

(Return)



10 Derivatives

There are several definitions of the derivative of a function $f(x)$ at $x = a$. These definitions are all equivalent, but they are all important because they emphasize different aspects of the derivative.

Derivative (first definition)

$$f'(a) = \left.\frac{df}{dx}\right|_{x=a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

If the limit does not exist, then the derivative is not defined at a.

This first definition emphasizes that the derivative is the rate of change of the output with respect to the input. The next definition is similar.

Derivative (second definition)

$$f'(a) = \left.\frac{df}{dx}\right|_{x=a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

If the limit does not exist, then the derivative is not defined at a.

This definition can be interpreted as the change in output divided by the change in input, as the change in input goes to 0. One can see this is equivalent to the first definition by making the substitution $h = x - a$. The third definition looks quite different from the first two.

Derivative (third definition)

The derivative of $f(x)$ at $x = a$, $f'(a)$, is the constant C such that for any variation to the input h, the following holds.

$$f(a+h) = f(a) + Ch + O(h^2).$$

That is, $f'(a)$ is the first-order variation of the output. If no such C exists, then the derivative does not exist.

To show the equivalence, one can do a little algebra to see that

$$\frac{f(a+h) - f(a)}{h} = C + O(h).$$

Then taking the limit on both sides as $h \to 0$ shows that $C = f'(a)$.

Example

Using the second and third definitions above, compute the derivative of $f(x) = x^n$, where n is a positive integer. (See Answer 1)

Example

Find the derivative of e^x using the third definition. (See Answer 2)

Example

Find the derivative of $\cos x$ using the third definition. Hint: use the identity

$$\cos(a + h) = \cos(a)\cos(h) - \sin(a)\sin(h).$$

(See Answer 3)

Example

Find the derivative of $f(x) = \sqrt{x}$ using the third definition. (See Answer 4)

10.1 Notation

There are several different notations for the derivative of $y = f(x)$. The best options are

$$\frac{df}{dx} \quad \text{or} \quad \frac{dy}{dx},$$

because they make it clear that the input is x and the output is $f(x)$ or y, respectively.

The next tier of options are fair, and have the advantage of requiring less writing, but they lose the benefit of knowing what the input variable is:

$$f' \quad \text{or} \quad \dot{y} \quad \text{or} \quad df.$$

The third option, df, is known as differential notation, which will be covered more in a later module.

Do not try to cancel the d's in the derivative. Do not write the d's in cursive, or replace the d's with Δ's (those notations have a different meaning).

10.2 Interpretations

The derivative is commonly interpreted as the *slope of the tangent line* to the graph of the function. This is fine when the function has one input and one output. But what happens in the (more realistic) situation of a function with more than one input and more than one output? How does one graph such a function? And if the units of the input and output are different, what is the unit of slope?

A better interpretation for the derivative is as the *rate of change of output with respect to input*. This interpretation makes sense with functions of many inputs and outputs. However, that will be covered in the multivariable sequel to this course.

10.3 Examples with respect to time

The most common use of the derivative is with respect to time. Here are several such examples from different areas.

Physics

Velocity $v(t)$ is the derivative of position $x(t)$, with respect to time. Similarly, acceleration $a(t)$ is the derivative of velocity with respect to time:

$$v = \frac{dx}{dt} \quad \text{and} \quad a = \frac{dv}{dt}$$

Electromagnetism

Electric current, I, in a circuit is the rate of change of charge, Q, with respect to time:

$$I = \frac{dQ}{dt}.$$

Chemistry

The reaction rate for the product P, denoted r_P, in a chemical reaction is proportional to the rate of change of the concentration of P, denoted $[P]$, with respect to time:

$$r_P = k\frac{d[P]}{dt}.$$

10.4 Examples with respect to other variables

Spring constant

The spring constant k for a spring is the derivative of force with respect to deflection:

$$k = \frac{d(\text{force})}{d(\text{deflection})}.$$

Elasticity

The elasticity modulus λ of a material is the rate of change of stress with respect to strain:

$$\lambda = \frac{d(\text{stress})}{d(\text{strain})}.$$

Viscosity

The viscosity of a fluid μ is related to the shear stress by the equation

$$\text{shear stress} = \mu\frac{d(\text{velocity})}{d(\text{height})}.$$

Tax rates

The marginal tax rate is the rate of change of tax with respect to income:

$$\text{marginal tax rate} = \frac{d(\text{tax})}{d(\text{income})}.$$

10.5 EXERCISES

- A rock is dropped from the top of a 320-foot building. The height of the rock at time t is given as $s(t) = -8t^2 + 320$, where t is measured in seconds. Find the speed (that is, the absolute value of the velocity) of the rock when it hits the ground in feet per second. Round your answer to one decimal place.
- A very rough model of population size P for an ant species is $P(t) = 2\ln(t+2)$, where t is time. What is the rate of change of the population at time $t = 2$?
- A particle's position, p, as a function of time, t, is represented by $p(t) = \frac{1}{3}t^3 - 3t^2 + 9t$. When is the particle at rest?
- Hooke's law states that the force F exerted by an "ideal" spring displaced a distance x from its equilibrium point is given by $F(x) = -kx$, where the constant k is called the "spring constant" and varies from one spring to another. In real life, many springs are nearly ideal for small displacements; however, for large displacements, they might deviate from what Hooke's law predicts. Much of the confusion between nearly-ideal and non-ideal springs is clarified by thinking in terms of series: for x near zero, $F(x) = -kx + O(x^2)$. Suppose you have a spring whose force follows the equation $F(x) = -2\tan 3x$. What is its spring constant?
- The profit, P, of a company that manufactures and sells N units of a certain product is modeled by the function $P(N) = R(N) - C(N)$. The revenue function, $R(N) = S \cdot N$, is the selling price S per unit times the number N of units sold. The company's cost, $C(N) = C_0 + C_{\text{op}}(N)$, is a sum of two terms. The first is a constant C_0 describing the initial investment needed to set up production. The other term, $C_{\text{op}}(N)$, varies depending on how many units the company produces, and represents the operating costs. Companies care not only about profit, but also "marginal profit," the rate of change of profit with respect to N. Assume that $S = \$50$, $C_0 = \$75,000$, $C_{\text{op}}(N) = \$50\sqrt{N}$, and that the company currently sells $N = 100$ units. Compute the marginal profit at this rate of production. Round your answer to one decimal place.
- In Economics, "physical capital" represents the buildings or machines used by a business to produce a product. The "marginal product of physical capital" represents the rate of change of output product with respect to physical capital (informally, if you increase the size of your factory a little, how much more product can you create?). A particular model tells us that the output product Y is given, as a function of capital K, by $Y = AK^{\alpha}L^{1-\alpha}$, where A is a constant, L is units of labor (assumed to be constant), and α is a constant between 0 and 1. Determine the marginal product of physical capital predicted by this model.

10.6 Answers to Selected Examples

1. Using the binomial expansion and the above definition, one finds

$$\begin{aligned} f'(a) &= \lim_{h\to 0} \frac{(a+h)^n - a^n}{h} \\ &= \lim_{h\to 0} \frac{a^n + na^{n-1}h + O(h^2) - a^n}{h} \\ &= \lim_{h\to 0} \frac{na^{n-1}h + O(h^2)}{h} \\ &= \lim_{h\to 0} na^{n-1} + O(h) \\ &= na^{n-1} \end{aligned}$$

   Using the third definition, and again the binomial expansion, one writes

$$\begin{aligned} f(a+h) &= (a+h)^n \\ &= a^n + na^{n-1}h + O(h^2), \end{aligned}$$

   so $f'(a) = na^{n-1}$.

   (Return)

2. Note that $e^{a+h} = e^a \cdot e^h$. Using our knowledge of the Taylor series for e^h, we have

$$\begin{aligned} e^{a+h} &= e^a e^h \\ &= e^a \left(1 + h + O(h^2)\right) \\ &= e^a + e^a h + O(h^2), \end{aligned}$$

   and so the derivative of e^x, evaluated at $x = a$, is e^a.

   (Return)

3. Using the above identity and our knowledge of Taylor series, we find

$$\begin{aligned} \cos(a+h) &= \cos(a)\cos(h) - \sin(a)\sin(h) \\ &= \cos(a)\left(1 + O(h^2)\right) - \sin(a)\left(h + O(h^3)\right) \\ &= \cos(a) - \sin(a)h + O(h^2), \end{aligned}$$

   so the derivative of $\cos x$, evaluated at $x = a$, is $-\sin(x)$.

   (Return)

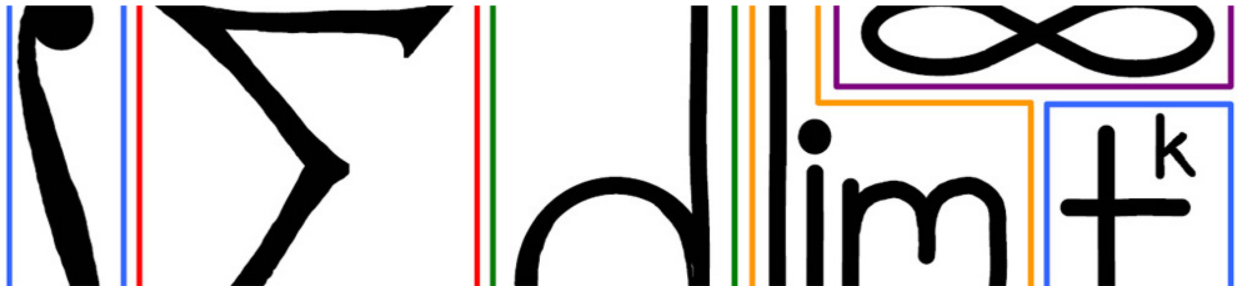
4. First, write

$$\begin{aligned} f(a+h) &= \sqrt{a+h} \\ &= \sqrt{a}\sqrt{1+\frac{h}{a}}. \end{aligned}$$

   Now, recalling the binomial series $(1+x)^\alpha = 1 + \alpha x + O(x^2)$, we find

$$\begin{aligned} \sqrt{a}\sqrt{1+\frac{h}{a}} &= \sqrt{a}\left(1 + \frac{1}{2}\frac{h}{a} + O(h^2)\right) \\ &= \sqrt{a} + \frac{1}{2\sqrt{a}}h + O(h^2), \end{aligned}$$

and so the derivative of $\sqrt{x}$, evaluated at $x = a$, is $\frac{1}{2\sqrt{a}}$.
(Return)



11 Differentiation Rules

Recall from the previous module the third definition given of the derivative, that

$$f(x+h) = f(x) + \frac{df}{dx}h + O(h^2).$$

The derivative can be thought of as a multiplier: a change of h to the input gets amplified by a factor of $\frac{df}{dx}$ to become a change of $\frac{df}{dx}h$ to the output. This interpretation helps make sense of the following rules.

11.1 Differentiation rules

Suppose u and v are differentiable functions of x. Then the following rules (written using the shorthand differential notation) hold:

Linearity

$$d(u+v) = du + dv \quad \text{and} \quad d(c \cdot u) = c \cdot du,$$

where c is a constant.

Product

$$d(u \cdot v) = u \cdot dv + v \cdot du.$$

Chain

$$d(u \circ v) = du \cdot dv.$$

(See Proof 1)

Another common way to express the Chain rule, using the more traditional derivative notation, is

$$\frac{du}{dx} = \frac{du}{dv} \cdot \frac{dv}{dx}.$$

Caveat

Note that in the chain rule, the output of v is being plugged in as an input to u. Therefore, in the above, if the derivative $d(u \circ v)$ is being evaluated at $x = a$, then du is evaluated at $x = v(a)$ and dv is evaluated at $x = a$. More explicitly,

$$d(u \circ v)\Big|_{x=a} = du\Big|_{x=v(a)} \cdot dv\Big|_{x=a}.$$

Example

Compute $\frac{d}{dx}\left(e^{\sin x}\right)$. (See Answer 2)

11.2 Other rules

There are a few other differentiation rules commonly taught in a first year calculus class, which can all be proven using the rules from above.

Reciprocal

$$d\left(\frac{1}{v}\right) = -\frac{1}{v^2}dv.$$

Quotient

$$d\left(\frac{u}{v}\right) = \frac{vdu - udv}{v^2}.$$

Inverse

$$d(u^{-1}) = \frac{1}{du}\Big|_{u^{-1}}$$

Note: u^{-1} is the *inverse* of u, *not* the reciprocal (which was covered above). (See Proof 3)

Example

Show that

$$\frac{d}{dx}\sec x = \sec x \tan x,$$

using the reciprocal rule. (See Answer 4)

Example

Show that

$$\frac{d}{dx}\tan x = \sec^2 x,$$

using the quotient rule. (See Answer 5)

Example

Show that

$$\frac{d}{dx} \ln x = \frac{1}{x},$$

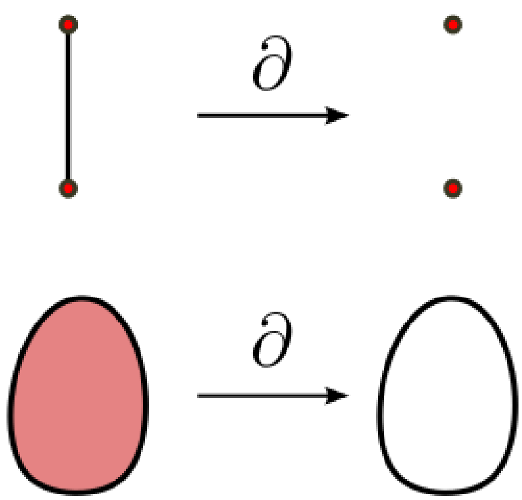
using the inverse derivative rule. (See Answer 6)

11.3 Bonus

There are operators in other areas of mathematics which act similarly to the derivative. Finding such similarities in disparate fields is useful, because theorems from one field can often be carried over to the other and proved using similar techniques. These connections give a deeper understanding of both fields.

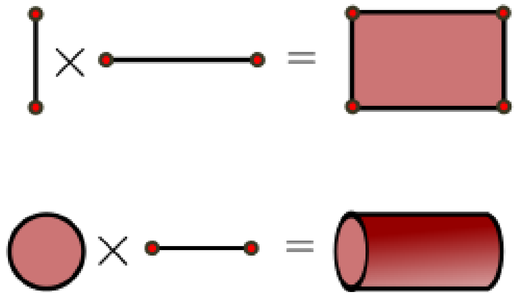
Boundary of spaces

Consider the *boundary* of a space. Loosely speaking, the boundary of a space is its outline, border, or edge. Think of the boundary as an operator, denoted ∂:



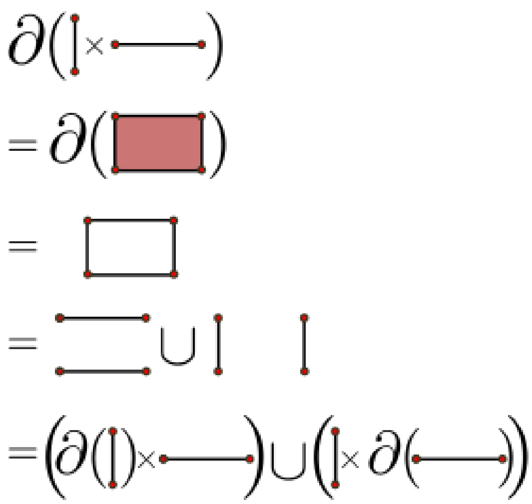
It turns out the the boundary operator ∂ acts similarly to a derivative. In particular, there is a product rule for ∂. But first, we need a notion of product for spaces.

The *Cartesian product* of two spaces X and Y, denoted $X \times Y$, is the set of ordered pairs (x, y) where $x \in X$ and $y \in Y$. This can be visualized by taking the first space and extruding it along the second space (easiest to visualize if the second space is a line segment), as in the following examples:

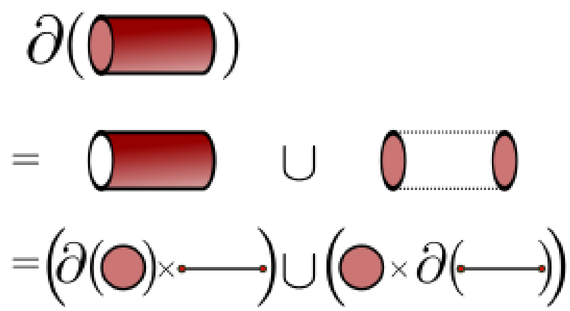


So the Cartesian product of two line segments is a filled in rectangle, and the Cartesian product of a disc and a line segment is a solid cylinder.

Now, consider the boundaries of these regions. The boundary of the filled in rectangle is its border, which can be thought of as the two vertical edges and the horizontal top and bottom edges. Note that this can be expressed as the boundary of the first segment times the second segment union the first segment times the boundary of the second segment:



Similarly, for the solid cylinder, the boundary is the union of the lateral area and the end caps. The lateral area is the Cartesian product of the circle with the line segment, which is the boundary of the disc times the line segment. The other piece (the end caps) is the disc times two points, which is the disc times the boundary of the line segment:



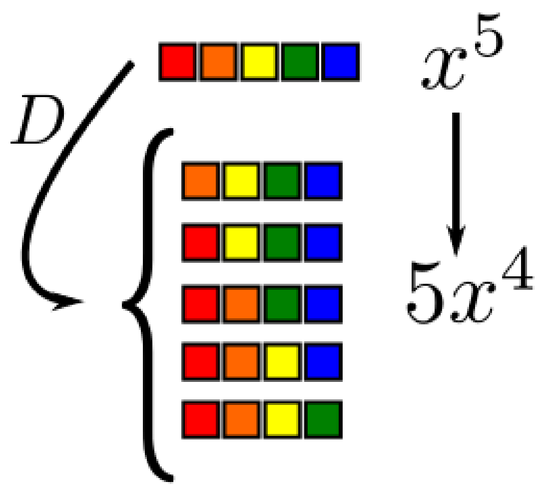
These examples suggest the (true) fact that for two spaces A and B, one has

$$\partial(A \times B) = \partial(A) \times B \cup A \times \partial(B).$$

Thinking of the boundary operator ∂ as the derivative, $\times$ as multiplication, and $\cup$ as addition, this is exactly like the product rule for functions given above. If you like this strange example, you may wish to take a course in Topology some day...

Lists

Consider a list of five distinct objects, all labeled with x. We might symbolize this by x^5. Now, consider the deletion operator D which deletes an object from the list. There are five different lists that might result (depending on which object was deleted), and this would logically be symbolized by $5x^4$, under the convention that a plus + stands for "logical OR":



There is an entire calculus for lists and other grammatical constructs in Computer Science. As one simple example, how would we algebraicize the empty list? It has zero elements, so logically it should be expressed as $x^0 = 1$. Now, let $\mathcal{L}$ denote the collection of all finite lists. The trivial observation that any list is either empty or has a first element can be translated into an algebraic equation:

$$\mathcal{L} = 1 + x\mathcal{L}.$$

Here's the cool part...solve for $\mathcal{L}$ and you get:

$$\mathcal{L} = \frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \cdots ,$$

the geometric series! Remember that plus mean "OR", so that this equation says any list is either the empty list, or the list of length 1, or the list of length 2, etc. If you find this sort of thing fun, you may wish to learn some Computer Science and some Analytic Combinatorics.

11.4 EXERCISES

- Find the derivative of $f(x) = \sqrt{x}(2x^2 - 4x)$.
- Find the derivative of $f(x) = 6x^4 - \dfrac{3}{x^2} - 2\pi$.
- Find the derivative of $f(x) = 7(x^3 + 4x)^5 \cos x$.
- Find the derivative of $f(x) = (e^x + \ln x)\sin x$.
- Find the derivative of $f(x) = \dfrac{\sqrt{x+3}}{x^2}$
- Find the derivative of $f(x) = \dfrac{\ln x}{\cos x}$.
- Find the derivative of $f(x) = \dfrac{\sqrt[3]{x} - 4}{x^3}$.
- Find the derivative of $f(x) = \sin^3(x^3)$.
- Find the derivative of $f(x) = e^{-1/x^2}$.
- Use the information about functions f and g from the following table to compute the value of $\left.\dfrac{d}{dx}\right|_{x=1} g(f(x))$.

x	$f(x)$	$f'(x)$	$g(x)$	$g'(x)$
-1	1	1	0	3
0	0	2	0	0
1	2	3	2	0
2	3	-1	-1	2
3	-1	0	3	-1

- Suppose that a certain quantity A is a function of another quantity B, which, in turn, depends on a third quantity C. We know that $B(C) = \sqrt{C}$. If the rate of change of A with respect to B is B^2, what is the rate of change of A with respect to C?

- This problem concerns the boundary operator ∂ from the bonus material. Denote by I the closed unit interval $[0,1]$. Then, as observed, $\partial I = \{0\} \cup \{1\}$ is the union of two points. Let's get a little "creative". Denote by I^n the "n-cube", that is, the Cartesian product of n intervals: $I^n = I \times I \times \cdots \times I$. This is a well-defined and perfectly reasonable n-dimensional cube. (Just because you can't visualize doesn't mean it can't exist!) Note that $I^1 = I$ and I^0 is a single point (a zero-dimensional cube!). As a step towards building a "calculus of spaces", let us write $\partial I^1 = 2I^0 = 2$ as a way of saying that the boundary of an interval consists of two points and that $I^0 = 1$. The boundary of an n-dimensional cube consists of a certain number of $(n-1)$-dimensional cubes (called "faces"). For example, a square I^2 has four faces. Using what you know about derivatives, answer this: how many faces does I^n have?

11.5 Answers to Selected Examples

1. **Linearity** Using the third definition of the derivative from the last module, we find

$$\begin{aligned}(u+v)(x+h) &= u(x+h) + v(x+h) \\ &= u(x) + \frac{du}{dx}h + O(h^2) + v(x) + \frac{dv}{dx}h + O(h^2) \\ &= (u+v)(x) + \left(\frac{du}{dx} + \frac{dv}{dx}\right)h + O(h^2),\end{aligned}$$

as desired. Similarly,

$$\begin{aligned}(c \cdot u)(x+h) &= c \cdot u(x+h) \\ &= c\left(u(x) + \frac{du}{dx}h + O(h^2)\right) \\ &= (c \cdot u)(x) + \left(c\frac{du}{dx}\right)h + O(h^2).\end{aligned}$$

Product Again using the third definition of the derivative, we find

$$\begin{aligned}(u \cdot v)(x+h) &= u(x+h) \cdot v(x+h) \\ &= \left(u(x) + \frac{du}{dx}h + O(h^2)\right) \cdot \left(v(x) + \frac{dv}{dx}h + O(h^2)\right) \\ &= u(x)v(x) + u(x)\frac{dv}{dx}h + v(x)\frac{du}{dx}h + O(h^2) \\ &= (u \cdot v)(x) + \left(u(x)\frac{dv}{dx} + v(x)\frac{du}{dx}\right)h + O(h^2),\end{aligned}$$

as desired.

Chain The chain rule is justified similarly, with a little bit more algebra:

$$\begin{aligned}(u \circ v)(x+h) &= u(v(x+h)) \\ &= u\left(v + \frac{dv}{dx}h + O(h^2)\right).\end{aligned}$$

To simplify the notation temporarily, let $\tilde{h} = \frac{dv}{dx}h + O(h^2)$. Then

$$\begin{aligned} u(v+\tilde{h}) &= u(v) + \frac{du}{dv}\tilde{h} + O(\tilde{h}^2) \\ &= u(v) + \frac{du}{dv}\left(\frac{dv}{dx}h + O(h^2)\right) + O((\frac{dv}{dx}h + O(h^2))^2) \\ &= u(v) + \frac{du}{dv}\cdot\frac{dv}{dx}h + O(h^2), \end{aligned}$$

as desired.

(Return)

2. In the above notation, $u(x) = e^x$ and $v(x) = \sin x$, and the question asks for the derivative $d(u \circ v)$. Again, remembering to evaluate du at $v(x)$, one finds that

$$\begin{aligned} d\left(e^{\sin x}\right) &= d(e^x)\bigg|_{\sin(x)} d(\sin x)\bigg|_x \\ &= e^{\sin x}\cos x. \end{aligned}$$

(Return)

3. **Reciprocal** One can think of this as an application of the chain rule, by writing

$$\frac{1}{v} = u \circ v$$

where $u(x) = \frac{1}{x}$. Or one can see it as a special case of the quotient rule.

Quotient Let $w = \frac{u}{v}$, so that $u = w \cdot v$. By the product rule,

$$du = w \cdot dv + v \cdot dw.$$

Solving for dw, replacing w with $\frac{u}{v}$, and clearing fractions gives

$$\begin{aligned} dw &= \frac{du - w \cdot dv}{v} \\ &= \frac{du - \frac{u}{v}dv}{v} \\ &= \frac{v \cdot du - u \cdot dv}{v^2}, \end{aligned}$$

as desired.

Inverse Note that $x = u \circ u^{-1}$ by the definition of the inverse of a function. Differentiating both sides of this equation, using the chain rule on the right, gives

$$1 = du\bigg|_{u^{-1}} d(u^{-1})\bigg|_x$$

Then solving for $d(u^{-1})$ gives

$$d(u^{-1}) = \frac{1}{du}\bigg|_{u^{-1}}$$

as desired.

(Return)

4. We find that

$$\begin{aligned}
\frac{d}{dx}\sec x &= \frac{d}{dx}\left(\frac{1}{\cos x}\right) \\
&= -\frac{1}{\cos^2 x}(-\sin x) \\
&= \frac{1}{\cos x}\cdot\frac{\sin x}{\cos x} \\
&= \sec x \tan x.
\end{aligned}$$

(Return)

5. We find that

$$\begin{aligned}
\frac{d}{dx}\tan x &= \frac{d}{dx}\frac{\sin x}{\cos x} \\
&= \frac{d(\sin x)\cos x - d(\cos x)\sin x}{\cos^2 x} \\
&= \frac{\sin^2 x + \cos^2 x}{\cos^2 x} \\
&= \frac{1}{\cos^2 x} \\
&= \sec^2 x.
\end{aligned}$$

(Return)

6. The inverse of the logarithm is the exponential, so $u = e^x$ in the inverse rule. Thus,

$$\begin{aligned}
\frac{d}{dx}\ln x &= \left.\frac{1}{d(e^x)}\right|_{\ln x} \\
&= \left.\frac{1}{e^x}\right|_{\ln x} \\
&= \frac{1}{x},
\end{aligned}$$

as desired.

(Return)

12 Linearization

One of the main uses of the derivative is *linearization*, which uses the first two terms (the constant and linear term) of the Taylor series as an approximation. In many applications, this gives a very good approximation, as we will see in some examples.

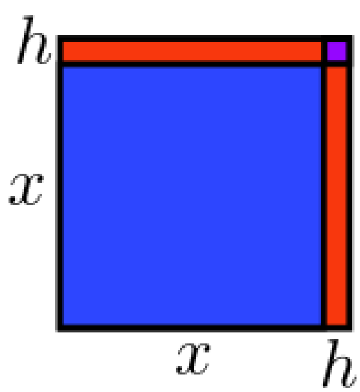
12.1 Linear variation visualized

There are several geometric examples where it is possible to see the linear variation as the change in area as a parameter is changed by a small amount.

Square

The area of a square of side length x is given by $A(x) = x^2$. When that is varied by a small amount h, the result is

$$A(x+h) = (x+h)^2 = x^2 + 2xh + h^2.$$



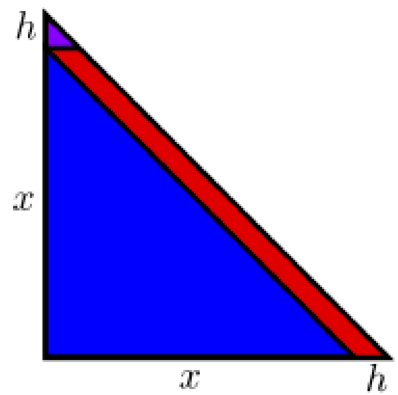
$$(x+h)^2 = x^2 + 2xh + h^2$$

The linear variation is $2xh$, which can be seen in the diagram as the rectangles along the right and top edges of the square. There are two of them, each with area xh. The final bit of area, the purple square in the diagram, has area h^2, which is higher order.

Triangle

The area of a right triangle with legs length x is $A(x) = \frac{1}{2}x^2$. When the leg is varied by h, the result is

$$A(x+h) = \frac{1}{2}x^2 + xh + \frac{1}{2}h^2.$$



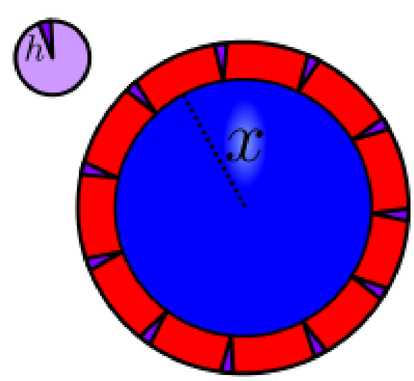
$$\frac{1}{2}(x+h)^2 = \frac{1}{2}x^2 + xh + \frac{1}{2}h^2$$

Visually, the linear variation xh comes from the red parallelogram, of base h and height x, running along the hypotenuse. The higher order term $\frac{1}{2}h^2$ comes from the small purple triangle at the tip of the triangle.

Disc

The area of a disc of radius x is $A(x) = \pi x^2$. If the radius is increased by h, the result is

$$A(x+h) = \pi(x+h)^2 = \pi x^2 + 2\pi xh + \pi h^2.$$



$$\pi(x+h)^2 = \pi x^2 + 2\pi xh + \pi h^2$$

Visually breaking this into the linear variation and higher order variation is a little bit harder. The best way is to imagine taking the ring formed by the increased radius and breaking it into rectangles and wedges. In the limit, the wedges can be rearranged into a disc of radius h, and the rectangles can be arranged to form a strip of length $2\pi x$ (the circumference of the inner circle) and width h.

12.2 Linear approximation

The equation underlying any linear approximation should be familiar, since it is just the first order Taylor series about $x = a$, after making the substitution $h = x - a$:

$$f(a + h) \approx f(a) + f'(a)h.$$

This will be a good linear approximation provided that h is small, i.e., the point a is close to the input we are trying to approximate. In general, one wants to pick a to be an input where it is easy to compute $f(a)$ and $f'(a)$ which is as close to the desired input as possible.

Example

Using a linear approximation, estimate $\sqrt{250}$. (See Answer 1)

Example

Using a linear approximation, estimate $\sqrt{104}$. Is this an over-approximation or an under-approximation? (See Answer 2)

Example

Approximate π^{20}. Hint: $\pi^2 \approx 9.86$. (See Answer 3)

Example

Approximate e^{30}. Hint: $e^3 \approx 20.1$, and $2^{10} \approx 1000$. (See Answer 4)

12.3 Newton's method

Another application of linearization gives a way of approximating the root of a function. This is called *Newton's method*.

Given a continuous, differentiable function f, the goal is to find a root (i.e. a value a such that $f(a) = 0$). Suppose x is an initial guess of a root. Then $x + h$ will be a root for some small value of h. Linearizing, and ignoring the higher order terms, gives

$$\begin{aligned} f(x + h) &= f(x) + f'(x)h + O(h^2) \\ &\approx f(x) + f'(x)h. \end{aligned}$$

Since we supposed that $x + h$ was a root of f, it follows that $f(x + h) = 0$. Therefore, setting the above equal to 0 and solving for h gives

$$f(x) + f'(x)h = 0 \quad \Rightarrow \quad h = -\frac{f(x)}{f'(x)}.$$

Thus, an even closer guess of a root is

$$x + h = x - \frac{f(x)}{f'(x)}.$$

By taking this new guess of a root, and repeating the above process, one (hopefully) gets a better and better approximation of a root.

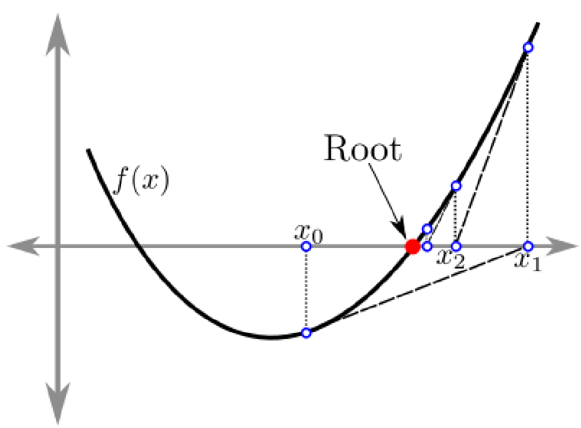
More formally, this is what is called a *difference equation*. Given an initial guess, called x_0, of a root of the function, one uses the update rule

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

to get x_1, and then x_2, and so on.

The resulting sequence hopefully converges to a root of f. Graphically, what is happening is as follows:

1. Pick a guess x_0.
2. Find the tangent line to f through the point $(x_0, f(x_0))$.
3. Let x_1 be the point where the tangent line intersects the x-axis.
4. Repeat steps 2 and 3 (see the figure).



Caveat

This sequence is only defined if $f'(x_n)$ exists and is non-zero for every x_n in the sequence. Even if the sequence is defined, it may not converge to anything. But if the sequence is defined and it does converge, say to L, then L is a root of f.

Example

Find the update rule for approximating $\frac{1}{a}$, the reciprocal of a number a. (See Answer 5)

Example

Find the update rule for using Newton's method to approximate $\sqrt{a}$. Use the update rule twice with initial guess $x_0 = 3$ to estimate $\sqrt{11}$. (See Answer 6)

Example

Find the update rule for finding $\sqrt[3]{a}$. Use the update rule once with the initial guess of $x_0 = 5$ to estimate $\sqrt[3]{100}$. (See Answer 7)

12.4 EXERCISES

- Use a linear approximation to estimate $\sqrt[3]{67}$. Round your answer to four decimal places. Hint: what is 4^3?
- Use a linear approximation to estimate the cosine of an angle of 66°. Round your answer to four decimal places. Hint: remember that $60^\circ = \dfrac{\pi}{3}$, and hence $6^\circ = \dfrac{\pi}{30}$.
- Use Newton's method to determine the intersection of e^{-x} and x.
- The golden ratio $\varphi = \dfrac{1+\sqrt{5}}{2}$ is a root of the polynomial $x^2 - x - 1$. If you use Newton's method to estimate its value, what is the appropriate update rule for the sequence x_n ?
- To approximate $\sqrt{10}$ using Newton's method, what is the appropriate update rule for the sequence x_n ?
- You want to build a square pen for your new chickens, with an area of $1200\,\text{ft}^2$. Not having a calculator handy, you decide to use Newton's method to approximate the length of one side of the fence. If your first guess is 30 ft, what is the next approximation you will get?
- You are in charge of designing packaging materials for your company's new product. The marketing department tells you that you must put them in a cube-shaped box. The engineering department says that you will need a box with a volume of $500\,\text{cm}^3$. What are the dimensions of the cubical box? Starting with a guess of 8 cm for the length of the side of the cube, what approximation does one iteration of Newton's method give you? Round your answer to two decimal places.
- Without using a calculator, approximate 9.98^{98}. Here are some hints. First, 9.98 is close to 10, and $10^{98} = 1\,\text{E}\,98$ in scientific notation. What does linear approximation give as an estimate when we decrease from 10^{98} to 9.98^{98}?
- A diving-board of length L bends under the weight of a diver standing on its edge. The free end of the board moves down a distance $D = PL^3/3EI$ where P is the weight of the diver, E is a constant of elasticity (that depends on the material from which the board is manufactured), and I is a moment of inertia. (These last two quantities will again make an appearance in Lectures 13 and 41, but do not worry about what exactly they mean now...) Suppose our board has a length $L = 2\,\text{m}$, and that it takes a deflection of $D = 20\,\text{cm}$ under the weight of the diver. Use a linear approximation to estimate the deflection that it would take if its length was increased by 20 cm.

12.5 Answers to Selected Examples

1. The function here is $f(x) = \sqrt{x}$. Possible choices for a are perfect squares, because it is easy to compute the square root of squares. The nearest perfect square is $256 = 16^2$, so we choose $a = 256$. Thus,

$h = x - a = -6$. Then

$$\begin{aligned}
f(x) &\approx f(a) + f'(a)h \\
\sqrt{250} &\approx \sqrt{256} + \frac{1}{2\sqrt{256}}(-6) \\
&= 16 - \frac{6}{32} \\
&= 15\frac{13}{16} \\
&\approx 15.8,
\end{aligned}$$

which is very close to the calculator's answer of $15.811\ldots$.

(Return)

2. As above, $a = 100$ is the closest point where it is easy to compute $\sqrt{x}$ and derivatives. Then $h = 4$, so the linear approximation is

$$\begin{aligned}
f(a+h) &\approx f(a) + f'(a)h \\
&= \sqrt{100} + \frac{1}{2\sqrt{100}} \cdot 4 \\
&= 10 + \frac{4}{20} \\
&= 10.2.
\end{aligned}$$

This is an over-approximation. One way to see why is to consider the graph of $\sqrt{x}$, which is concave down, so the linear approximation is above the true value. Another argument is to consider the next term of the Taylor series, which is negative (since the second derivative of $\sqrt{x}$ is negative).

For comparison, the value according to a calculator is $\sqrt{104} \approx 10.198$.

(Return)

3. From the hint, $\pi^{20} = (\pi^2)^{10} \approx 9.86^{10}$. Thus, we are trying to approximate $f(x) = x^{10}$ at $x = 9.86$. The nearest easy input is $a = 10$, so we find

$$\begin{aligned}
f(9.86) &\approx f(10) + f'(10)(-.14) \\
&= 10^{10} + 10(10)^9(-.14) \\
&= 10^{10}(1 - .14) \\
&= 8.6 \cdot 10^9.
\end{aligned}$$

The true answer is approximately $8.77 \cdot 10^9$, so this estimate is within 2%.

(Return)

4. From the first hint, $e^{30} = (e^3)^{10} \approx (20.1)^{10}$, so consider the linear approximation for $f(x) = x^{10}$ near $a = 20$:

$$\begin{aligned}
f(x) &\approx f(20) + f'(20)(x - 20) \\
x^{10} &\approx 20^{10} + 10 \cdot 20^9(x - 20)
\end{aligned}$$

So it follows that

$$\begin{aligned}
e^{30} &\approx (20.1)^{10} \\
&\approx 20^{10} + 10 \cdot 20^9 \cdot (.1) \\
&\approx 20^{10}\,[1 + (.5)(.1)] \\
&\approx 2^{10} \cdot 10^{10} \cdot 1.05 \\
&\approx 1.05 \cdot 10^{13}
\end{aligned}$$

(the last step used the second hint that $2^{10} \approx 10^3$). The true answer is approximately $e^{30} \approx 1.068 \cdot 10^{13}$, so the error is less than 2%.

(Return)

5. First, we must find a function which has $\frac{1}{a}$ as a root. We might try $f(x) = x - \frac{1}{a}$, but unfortunately (after a little algebra) this leads to the update rule

$$x_{n+1} = \frac{1}{a},$$

which is not particularly helpful, since that is the quantity we are trying to approximate.

Another try would be $f(x) = a - \frac{1}{x}$. This will work. Note that $f'(x) = \frac{1}{x^2}$, so the update rule is

$$\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{a - \frac{1}{x_n}}{\frac{1}{x_n^2}} \\
&= x_n - \left(x_n^2 a - x_n\right) \\
&= 2x_n - a x_n^2.
\end{aligned}$$

Note that this rule does not involve division, but only multiplication and subtraction.

(Return)

6. The first step is to find a function whose root is $\sqrt{a}$. A good choice is $f(x) = x^2 - a$. Then according to Newton's method, the update rule is

$$\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{x_n^2 - a}{2x_n} \\
&= x_n - \frac{x_n}{2} + \frac{a}{2x_n} \\
&= \frac{x_n}{2} + \frac{a}{2x_n}.
\end{aligned}$$

Using the update rule with $x_0 = 3$ and $a = 11$ gives

$$\begin{aligned}
x_1 &= \frac{x_0}{2} + \frac{11}{2x_0} \\
&= \frac{3}{2} + \frac{11}{6} \\
&= \frac{20}{6} \\
&= \frac{10}{3}
\end{aligned}$$

Updating one more time gives

$$
\begin{aligned}
x_2 &= \frac{x_1}{2} + \frac{11}{2x_1} \\
&= \frac{5}{3} + \frac{11}{\frac{20}{3}} \\
&= \frac{199}{60} \\
&\approx 3.3166
\end{aligned}
$$

The calculator gives that $\sqrt{11} \approx 3.3166$, so we get a good approximation with only a little bit of work. One could repeat the update rule several more times to get an even better approximation.

(Return)

7. The function in this case is $f(x) = x^3 - a$. Then $f'(x) = 3x^2$, and so the update rule is

$$
\begin{aligned}
x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\
&= x_n - \frac{x_n^3 - a}{3x_n^2} \\
&= x_n - \frac{1}{3}x_n + \frac{a}{3x_n^2} \\
&= \frac{2}{3}x_n + \frac{a}{3x_n^2}.
\end{aligned}
$$

Using this to estimate $\sqrt[3]{100}$ with the initial guess $x_0 = 5$, one finds

$$
\begin{aligned}
x_1 &= \frac{2}{3} \cdot x_0 + \frac{100}{3x_0^2} \\
&= \frac{10}{3} + \frac{4}{3} \\
&= \frac{14}{3} \\
&\approx 4.666
\end{aligned}
$$

If we compare this to the answer from a calculator, we find that $\sqrt[3]{100} \approx 4.64$, and so even after just one step, we are within 1% of the true answer.

(Return)

13 Higher Derivatives

The nth derivative of a function $f(x)$, denoted $f^{(n)}(x)$ or $\frac{d^n}{dx^n}(f)$, is defined recursively by

$$f^{(n)}(x) = \frac{d}{dx} f^{(n-1)}(x).$$

In other words, the nth derivative is what one gets by taking the derivative n times. Note that in the $\frac{d}{dx}$ notation, the power n goes to the right of dx, to emphasize the fact that the nth derivative of f is achieved by iterating n times the operator $\frac{d}{dx}$. So

$$\left(\frac{d}{dx}\right)^n f = \frac{d^n}{dx^n} f.$$

13.1 Interpretations

Let $x(t)$ denote the position of a moving body as a function of time. Then the velocity $v(t)$ of the body is

$$v(t) = \frac{dx}{dt}.$$

The *acceleration* of an object is the second derivative of its position function (i.e. the derivative of its velocity):

$$a(t) = \frac{dv}{dt} = \frac{d}{dt}\left(\frac{dx}{dt}\right) = \frac{d^2x}{dt^2}.$$

The *jerk* of an object is the third derivative of its position function (i.e. the derivative of its acceleration):

$$j(t) = \frac{d^3x}{dt^3}.$$

The *snap* (or *jounce*) of an object is the fourth derivative of its position:

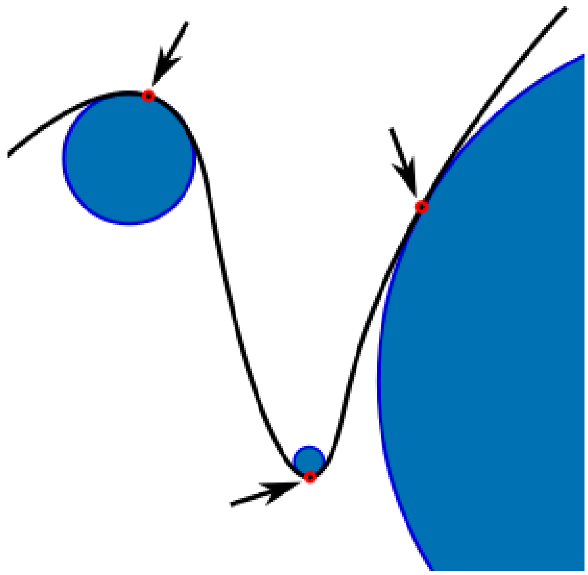
$$s(t) = \frac{d^4x}{dt^4}.$$

Quadrotors

The maneuverability of nano quadrotors depends on controlling both the jerk and the snap (in addition to velocity and acceleration).

Curvature

In geometry, curvature can (informally) be thought of as how quickly the graph of the function curves. Consider the largest circle that can comfortably sit tangent to the graph of f at the point a. Intuitively, the larger the radius R of the circle that fits, the smaller the curvature. Here is a curve with several of these circles (called *osculating circles*) drawn it at different indicated points.



In fact, the curvature of a curve f, denoted κ, is defined by $\kappa = \frac{1}{R}$, where R is the radius of the largest circle which fits the curve to second order. With some algebra, one finds the following expression for curvature in terms of the first and second derivatives of f:

Curvature of a function f

$$\kappa = \frac{|f''|}{(1+(f')^2)^{3/2}}.$$

(See Justification 1)

Note that for a straight line, the second derivative $f'' = 0$, and so $\kappa = 0$, which matches intuition. In this case, the osculating circle is infinite. Similarly, $\kappa = 0$ at inflection points of f.

Elasticity

Consider an elastic beam with uniform cross section and static load $q(x)$, where x is the location of the load along the beam. Then the deflection $u(x)$ (the amount the beam sags at location x) satisfies the equation

$$EI\frac{d^4u}{dx^4} = q(x),$$

where E and I are constants: E is the constant of elasticity (depends on the material), and I is the moment of inertia (depends on the shape of the beam).

Taylor series

As seen in previous modules, information about the derivatives of a function evaluated at a single point gives information about the function for inputs near that point via the Taylor series:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

Taking a few terms of this series gives a polynomial which is a good approximation of f near a. The more derivatives one knows, the more terms one can include in the series, and the better the approximation.

13.2 Bonus: another look at Taylor series

Consider the alternative way to express the Taylor series, in terms of the distance h from the base point a:

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{f''(a)}{2!}h^2 + \frac{f'''(a)}{3!}h^3 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}h^k \\ &= \sum_{k=0}^{\infty} \frac{h^k}{k!}\left(\left.\frac{d}{dx}\right|_a\right)^k f \\ &= \sum_{k=0}^{\infty} \frac{1}{k!}\left(h\left.\frac{d}{dx}\right|_a\right)^k f. \end{aligned}$$

Note that this resembles the Taylor series for the exponential e^x, where

$$x = h\left.\frac{d}{dx}\right|_a.$$

This may seem a little unusual. But the idea of exponentiating an operator to get another operator is a useful tool, which comes up in other areas of mathematics. In this notation, we can write

$$f(a+h) = e^{\left(h\left.\frac{d}{dx}\right|_a\right)} f.$$

Another way to think of this is that $e^{h\frac{d}{dx}}$ is the *shift operator*, which takes in the function $f(x)$ and gives back the function $f(x+h)$.

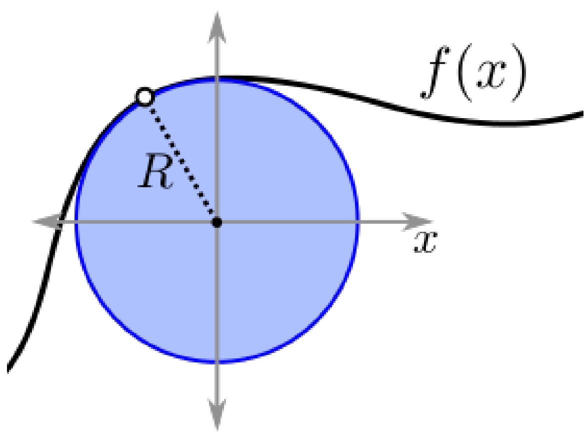
13.3 EXERCISES

- You are given the position, velocity and acceleration of a particle at time $t = 0$. The position is $p(0) = 2$, the velocity $v(0) = 4$, and the acceleration $a(0) = 3$. Using this information, which Taylor series should they use to approximate $p(t)$, and what is the estimated value of $p(4)$ using this approximation?
- If a particle moves according to the position function $s(t) = t^3 - 6t$, what are its position, velocity and acceleration at $t = 3$?
- If the position of a car at time t is given by the formula $p(t) = t^4 - 24t^2$, for which times t is its velocity decreasing?

- What is a formula for the second derivative of $f(t) = t^2 \sin 2t$? Use this formula to compute $f''(\pi/2)$.
- Use a Taylor series expansion to compute the third derivative of $f(x) = \sin^3(\ln(1+x))$ at zero.
- What is the curvature of the graph of the function $f(x) = -2\sin(x^2)$ at the point $(0,0)$?

13.4 Answers to Selected Examples

1. For a given point on the curve, draw its osculating circle, say of radius R (right now, R is unknown to us;we will eventually find R). Then place the coordinate axes so that the origin is at the center of the circle (note that the curvature at a given point only depends on the radius of the osculating circle, which is independent of where the axes are placed):



The equation of the osculating circle is $x^2 + y^2 = R^2$. Solving for y gives

$$y = \sqrt{R^2 - x^2} = (R^2 - x^2)^{1/2},$$

whose first and second derivatives should respectively match the first and second derivatives of the function f at the point (that is what it means for the circle to match the function up to second order). Remember that the first derivative and second derivative of f at the given point are just constants. We will set the derivative and second derivative of the equation of the circle equal to these constants, respectively, and then solve for R.

The first derivative of the equation of the circle is

$$\begin{aligned}\frac{d}{dx}(R^2 - x^2)^{1/2} &= \frac{1}{2}(R^2 - x^2)^{-1/2}(-2x)\\ &= -x(R^2 - x^2)^{-1/2}.\end{aligned}$$

The second derivative of the equation of the circle (using the product rule) is

$$\begin{aligned}
\frac{d}{dx}\left(-x(R^2-x^2)^{-1/2}\right) &= -(R^2-x^2)^{-1/2} - x\left(-\frac{1}{2}\right)(R^2-x^2)^{-3/2}(-2x) \\
&= -(R^2-x^2)^{-1/2} - x^2(R^2-x^2)^{-3/2} \\
&= -\frac{1}{(R^2-x^2)^{1/2}} - \frac{x^2}{(R^2-x^2)^{3/2}} \\
&= -\frac{\left(R^2-x^2\right)}{(R^2-x^2)^{3/2}} - \frac{x^2}{(R^2-x^2)^{3/2}} \\
&= \frac{-R^2}{(R^2-x^2)^{3/2}} \\
&= -R^2(R^2-x^2)^{-3/2}.
\end{aligned}$$

So setting the corresponding derivatives of f equal to the derivatives of the circle gives

$$\begin{aligned}
f' &= -x(R^2-x^2)^{-1/2} \\
f'' &= -R^2(R^2-x^2)^{-3/2}.
\end{aligned}$$

Now we do some algebra to solve for R in terms of f' and f''. Squaring the first equation gives

$$(f')^2 = x^2(R^2-x^2)^{-1}.$$

Solving this equation for x^2 gives

$$x^2 = \frac{(f')^2R^2}{1+(f')^2}.$$

Plugging this into the second equation, and doing some algebra gives

$$\begin{aligned}
f'' &= -R^2(R^2-x^2)^{-3/2} \\
&= -R^2\left(R^2 - \frac{(f')^2R^2}{1+(f')^2}\right)^{-3/2} \\
&= -R^2\left[R^2\left(1 - \frac{(f')^2}{1+(f')^2}\right)\right]^{-3/2} \\
&= \frac{-R^2}{R^3}\left(\frac{1+(f')^2-(f')^2}{1+(f')^2}\right)^{-3/2} \\
&= -\frac{1}{R}\left(1+(f')^2\right)^{3/2}.
\end{aligned}$$

Taking absolute values (since the radius of a circle should not be negative) gives

$$|f''| = \frac{1}{R}\left(1+(f')^2\right)^{3/2}.$$

Now, solving for $\kappa = \frac{1}{R}$ gives

$$\kappa = \frac{1}{R} = \frac{|f''|}{(1+(f')^2)^{3/2}},$$

as desired.

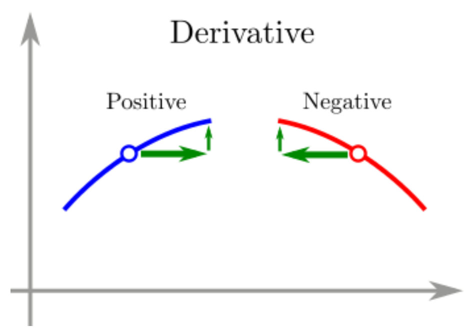
(Return)

14 Optimization

One of the most important applications of derivatives is *optimization*. In some introductory calculus classes these types of problems are called *max/min problems*: given a function, what is the maximum or minimum output subject to some constraints. This module will review how derivatives can be used in these problems and give some of the reasons why these methods work.

14.1 Critical points

First, observe that for a differentiable function f, if the derivative is not zero at a point, then that point cannot be a maximum or a minimum. For instance, if the derivative is positive, then the output is increasing with respect to the input, so by increasing the input, one can increase the output. Hence, the point is not a maximum. If the derivative is negative, then decreasing the input will increase the output, so that point cannot be a maximum.

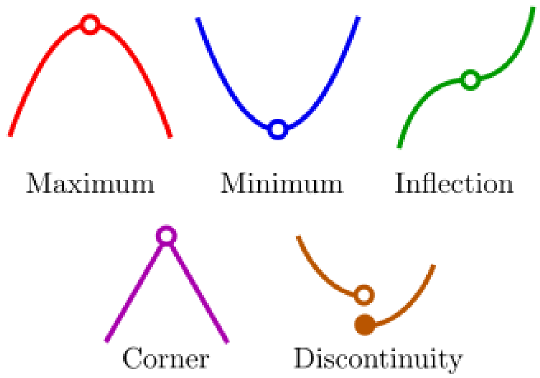


Similarly, a point cannot be a minimum if the derivative is not zero. Thus, the only possible inputs where a maximum or minimum can occur are those where the derivative is zero. This motivates the following definition

Critical point

A *critical point* of a function f is an input $x = a$ where either $f'(a) = 0$ or where the derivative is undefined.

Critical points include maximum and minimum points (called *extrema*) as well as inflection points; these are the points where the derivative is 0. Other critical points occur at corner points or discontinuities, where the derivative is undefined. The reason for including points where the derivative is not defined is that such a point could be a maximum or minimum:



Example

Compute the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$. (See Answer 1)

14.2 Classifying critical points

Once one has computed a critical points $x = a$, one can classify whether it is a maximum or minimum using the *second derivative test:*

Second Derivative Test

Suppose $x = a$ is a critical point of f where $f'(a) = 0$.

1. If $f''(a) > 0$, then f has a local minimum at a.
2. If $f''(a) < 0$, then f has a local maximum at a.
3. If $f''(a) = 0$, then the test fails.

In the third case, one can use the Taylor expansion about $x = a$ to determine the behavior of the function. In this case, $x = a$ could still be a local maximum, minimum, or inflection point.

(See Justification 2)

Example

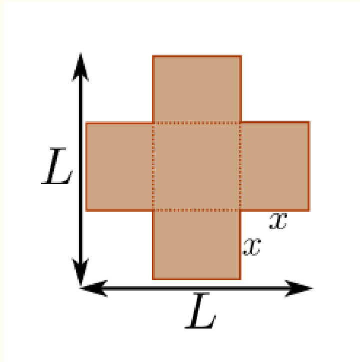
Use the Taylor series about $x = 0$ for

$$\sin^2 x \ln(\cos x)$$

to determine whether the function has a local maximum, local minimum, or inflection point at $x = 0$. (Take as a given that $x = 0$ is a critical point). (See Answer 3)

Example

Consider a square sheet of cardboard of side length L. By cutting equal sized squares of side length x from each corner of the sheet and folding up the flaps which are formed, one gets an open box:



Note that as x gets bigger, the box gets taller but the area of the base of the box shrinks. As x gets smaller, the area of the base grows, but the height shrinks. Find the value of x which maximizes the volume of the resulting box. (See Answer 4)

Example

Classify the critical points of $f(x) = x^3 - 6x^2 + 9x - 5$. (See Answer 5)

Example

Suppose a firm producing widgets expects to sell $3000 - 10p^2$ units (where p is the price of the unit). What price p should the firm set to maximize revenue (note that revenue here is just price times quantity sold)? (See Answer 6)

14.3 Global Extrema

While a local maximum or minimum is sometimes useful information, what is usually more important is the global maximum and minimum values of a function on a closed interval $[a, b]$ (or subject to some other constraint such as $x \geq 0$). These are called the *global extrema*, or *absolute extrema*, of a function.

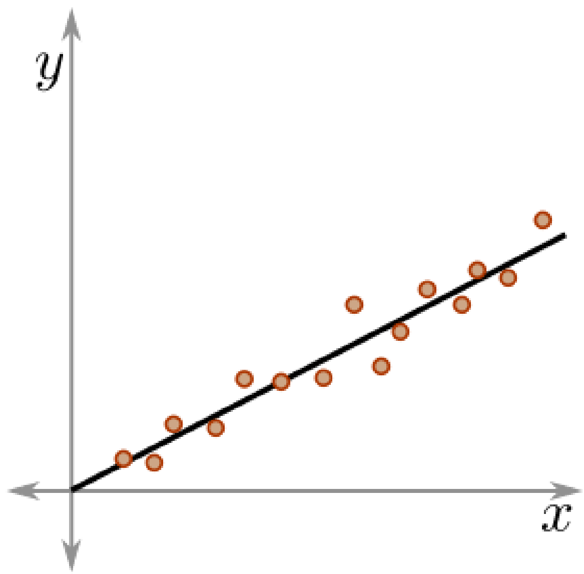
Global extrema on the interval $[a, b]$ either occur at critical points of f or at the endpoints of the interval. So in addition to finding the critical points of f in the interval and checking their values, one must also evaluate f at the endpoints of the interval to find the global extrema.

Example

Find the global extrema of $f(x) = x^3 - 6x^2 + 9x - 5$ on the interval $[2, 4]$. (See Answer 7)

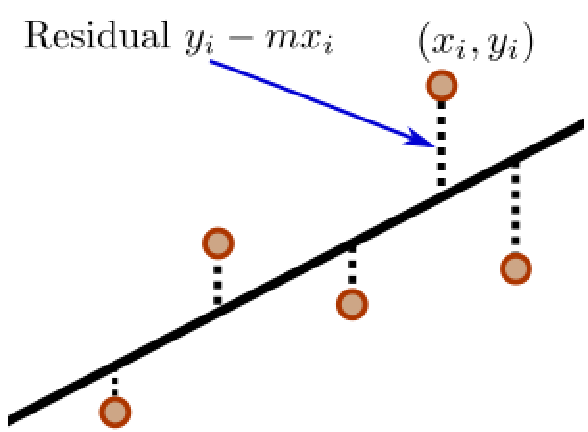
14.4 Application: Statistics

In statistics, one often takes experimental data points of the form (x_i, y_i) and looks for a relationship. A very simple relationship is the linear relationship $y = mx$. The data may not follow this relationship perfectly, and there may be some slight experimental error or other noise, so one tries to find the value of m which best fits the data:



This process, called a *linear regression*, can be framed as an optimization problem. But what is the quantity being optimized?

There are several different linear regression models, depending on the quantity being minimized. These different quantities yield different best fit lines. One of the most common models is called *ordinary least squares*. This method seeks to minimize the sum of the squares of the *residuals*, which are the vertical distances from the points to the line:



As shown above, the residual for a given point (x_i, y_i) is $y_i - mx_i$. Thus, the quantity being minimized is

$$S(m) = \sum_i (y_i - mx_i)^2.$$

Taking the derivative with respect to m gives

$$\begin{aligned}\frac{dS}{dm} &= \sum_i 2(y_i - mx_i)(-x_i)\\ &= \sum_i \left(-2x_i y_i + 2mx_i^2\right)\\ &= -2\sum_i x_i y_i + 2m\sum_i x_i^2.\end{aligned}$$

Setting this equal to 0 and solving for m gives

$$m = \frac{\displaystyle\sum_i x_i y_i}{\displaystyle\sum_i x_i^2}.$$

Applying the second derivative test, we compute

$$\frac{d^2S}{dm^2} = 2\sum_i x_i^2 > 0,$$

so the above value of m minimizes the sum of squares (hence the *least squares* name).

Note

To find the line of best fit of the form $y = mx + b$ requires methods of multivariable calculus (because there are two variables, m and b, which need to be optimized). Optimization with multiple variables is not much more difficult than for a single variable, but these methods are beyond the scope of this course.

14.5 EXERCISES

- Find all the local maxima and minima of the function $y = xe^{-x^2}$.
- Which type of critical point does the function $f(x) = e^{\sin(x^4)}\cos(x^2)$ have at zero ?
- Use a Taylor series about $x = 0$ to determine whether the function $f(x) = \sin^3(x^3)$ has a local maximum or local minimum at the origin.
- Find the location of the global maximum and minimum of $f(x) = x^3 - 6x^2 + 1$ on the interval $[-1, 7]$.
- Consider a stretch of highway in which cars are traveling at an average speed v. The "traffic density" u is the total amount of cars on our stretch of road divided by its length. These two quantities are related: the less cars on the road, the faster drivers are able to go. On the other hand, if traffic becomes heavy, drivers will naturally decrease their speed. The so-called "parabolic model" assumes that this relationship is dictated by the equation:
$$u = u_{\max}\left(1 - \frac{v}{v_{\max}}\right)$$
where $u_{\max}$ represents the capacity of the road, and $v_{\max}$ the speed limit on it. The amount of cars passing through our road is called the "traffic flux" or "throughput," and is given by the product of the traffic density and the average speed: $F = uv$. Using the parabolic model, find out at what average speed v_* the flux through our road is maximized.

- A manufacturing company wants to know how many workers it should hire. If it employs too many people, the machines in the factory will be overutilized and the workers will have to wait until they are free, thus reducing the number of units each one will produce in a day's work. On the other hand, too few workers would leave the machines idle for long periods of time. A rough model for the relationship between the number n of workers and their productivity p is given by the equation

$$p = p_{\text{max}} \left(1 - \frac{n}{n_{\text{max}}}\right)$$

where $p_{\text{max}} = 10$ is the maximum number of units a worker can produce in a day and $n_{\text{max}} = 100$ is the maximum number of workers the factory can accommodate. The amount of units U manufactured in the whole factory in one day is equal to the product of the number of workers and the number of units each one produces: $U = np$. How many workers should the company hire in order to maximize its production?

- A technology company has just invented a new gadget. In order to maximize the profit derived from its sale, the company must make a critical decision: at what price should it be sold? A market study suggests that the number N of units sold would approximately follow the equation $N = N_{\text{max}} e^{-P/\lambda}$, where P is the sale price, $N_{\text{max}} = 10,000,000$ is the number of units that would saturate the market, and $\lambda = \$50$. If it costs \$250 to manufacture one of these gadgets, at what price P_* would be profit of the company be maximized?

- The manufacturing process of a certain chemical substance is exothermic, that is, it releases heat. The amount of heat released, Q, depends on the temperature T at which the process is carried out, and it is given by the equation $Q = \alpha(T - T_0)^{-2} e^{(T-T_0)/\lambda}$, where $T_0 = 70^oF$ is the room temperature of the manufacturing plant, and $\alpha = 3000\, J\, (^oF)^2$ and $\lambda = 50^oF$. If the temperature T must be maintained above 100^oF, at what temperature T_* would be the heat loss be minimized?

- Classify the critical point $x = 0$ of the function $f(x) = \frac{\sin^2(3x^2)\cos(x)}{x}$ using Taylor series.

- Construct a box without a top whose base is a square. The material cost for the bottom is \$10 per square feet, the cost for the side is \$5 per square feet. The box must have volume 8 cubic feet. Determine the dimension of the box that will minimize the cost.

14.6 Answers to Selected Examples

1. The derivative $f'(x) = 3x^2 - 12x + 9$ is defined everywhere, so the critical points are where $f'(x) = 0$. Since

$$\begin{aligned} f'(x) &= 3x^2 - 12x + 9 \\ &= 3(x^2 - 4x + 3) \\ &= 3(x-3)(x-1), \end{aligned}$$

the critical points are $x = 3$ and $x = 1$.

(Return)

2. The second derivative test is justified by considering the Taylor series for f about $x = a$:

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2 + \cdots \\ &= f(a) + \frac{1}{2} f''(a)(x-a)^2 + \cdots, \end{aligned}$$

since $f'(a) = 0$. Thus, when x is close to a, $f(x)$ behaves like a parabola centered at $x = a$. Recall that the sign of the coefficient of the square term in a parabola determines if the parabola opens up or down. A positive coefficient means the parabola opens upward, and a negative coefficient means the parabola opens downward.

Here, the coefficient of the $(x-a)^2$ is $\frac{1}{2}f''(a)$. So if $f''(a) > 0$, then the parabola opens upward, meaning $f(a)$ is a local minimum of f. If $f''(a) < 0$, then the parabola opens downward, meaning $f(a)$ is a local maximum of f. If $f''(a) = 0$, then one has to look at more terms of the Taylor series to determine f's behavior at a.

(Return)

3. Expanding and multiplying the Taylor series gives

$$\begin{aligned}\sin^2 x \ln(\cos x) &= \left(x + O(x^3)\right)^2 \ln\left(1 - \frac{1}{2!}x^2 + O(x^4)\right)\\ &= (x + O(x^3))(x + O(x^3))\left(-\frac{1}{2}x^2 + O(x^4)\right)\\ &= -\frac{1}{2}x^4 + O(x^6).\end{aligned}$$

Thus, near $x = 0$ the function behaves like $-\frac{1}{2}x^4$, which is downward opening (because of the negative coefficient) and U shaped (because it is an even power). Therefore, the function has a local maximum at $x = 0$.

Note that the second derivative test, besides being tricky to apply with all of the product rules and chain rules, would ultimately be inconclusive in this example.

(Return)

4. The volume of the box is the area of the base times the height. The base is a square of side length $L-2x$ (since x has been cut from both sides). The height of the box is x. Thus

$$V = (L-2x)^2 \cdot x = 4x^3 - 4Lx^2 + L^2x$$

Finding the critical points means taking the derivative with respect to x and setting equal to 0:

$$\frac{dV}{dx} = 12x^2 - 8Lx + L^2 = 0.$$

This factors as

$$(6x - L)(2x - L) = 0,$$

so the critical points are $x = \frac{L}{6}$ and $x = \frac{L}{2}$. To apply the second derivative test, we compute

$$\frac{d^2V}{dx^2} = 24x - 8L$$

and evaluate at each critical point:

$$\left.\frac{d^2V}{dx^2}\right|_{x=L/2} = 4L > 0 \quad \text{and} \quad \left.\frac{d^2V}{dx^2}\right|_{x=L/6} = -4L < 0.$$

Thus, $x = \frac{L}{2}$ is a local minimum and $x = \frac{L}{6}$ is a local maximum. (Note also that for $x = \frac{L}{2}$ there is no cardboard left, since the removed corners have consumed the entire square!).

The volume that results from $x = \frac{L}{6}$ is

$$V = \left(\frac{2L}{3}\right)^2 \frac{L}{6} = \frac{2L^3}{27}.$$

(Return)

5. As found in a previous example, the critical points of f are $x = 3$ and $x = 1$. The second derivative of f is $f''(x) = 6x - 12$. Thus, $f''(3) = 6 > 0$ and $f''(1) = -6 < 0$, and it follows from the second derivative test that 3 is a local minimum of f, and 1 is a local maximum of f.

   (Return)

6. Revenue is $R(p) = (3000 - 10p^2)p = 3000p - 10p^3$. Taking the derivative gives $R'(p) = 3000 - 30p^2$, and setting equal to 0 gives

$$\begin{aligned} 3000 - 30p^2 &= 0 \\ 3000 &= 30p^2 \\ 100 &= p^2 \\ p &= \pm 10. \end{aligned}$$

   So R has critical point $p = 10$ (ignore $p < 0$ since price should be positive). The second derivative is *R(p) = -60p.*$ThusR(10) = -600 < 0$, and $p = 10$ is a local maximum.

   At this price, the revenue is

$$R(10) = 2000 \cdot 10 = 20000.$$

   (Return)

7. From the prior examples, $x = 1$ and $x = 3$ are the critical points of f. But $x = 1$ is not in the interval $[2, 4]$, so disregard it. Then evaluate f at 2,3,4 to find the extreme values:

$$\begin{aligned} f(2) &= 8 - 24 + 18 - 5 = -3 \\ f(3) &= 27 - 54 + 27 - 5 = -5 \\ f(4) &= 64 - 96 + 36 - 5 = -1. \end{aligned}$$

   Thus the absolute maximum is -1 and occurs when $x = 4$. The absolute minimum is -5 and occurs at $x = 3$.

   (Return)

15 Differentials

This module deals with differentials, e.g. dx or du. A formal treatment of differential forms is beyond the scope of this course. For now, the best way to think about the differential dx or du is to think of them as rates of change, and relate them with the chain rule:

$$du = \frac{du}{dx} dx.$$

In words, the rate of change of u equals the rate of change of u with respect to x times the rate of change of x. This is not a perfect interpretation, but it will serve our purposes for this course.

15.1 The differential as an operator

Think of the differential d as an operator which can be applied to an equation $f = g$ to give back $df = dg$. This process allows one to find the derivative of functions which are defined *implicitly,* i.e., functions which cannot be "solved for y" as $y = f(x)$. This method is called *implicit differentiation.*

Example

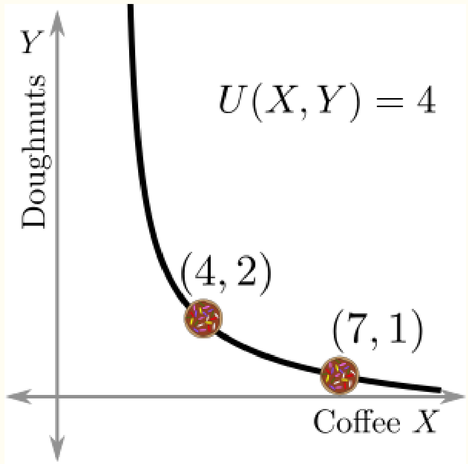
Find $\frac{dy}{dx}$ of the circle $x^2 + y^2 = r^2$. (See Answer 1)

Example

In economics, the *marginal rate of substitution* (MRS) of X for Y is the rate at which a consumer is willing to exchange good Y for good X to maintain an equal level of satisfaction (called *utility* in economics jargon). Let $U(X, Y)$ denote a particular consumer's utility function. As a particular example, let X be coffee, in ounces, and Y be doughnuts. Then the curve $U(X, Y) = C$ represents all the different combinations of coffee and doughnuts where the consumer is equally happy. For example, if the utility function is

$$U(X, Y) = Y^2(X - 3),$$

then the following shows the graph of the curve $U(X, Y) = 4$. The two plotted points show that this consumer is equally satisfied with 4 ounces of coffee and 2 doughnuts as with 7 ounces of coffee and 1 doughnut.



The MRS, then, is the rate of exchange of Y for X (doughnuts for coffee) so that utility stays the same (i.e. we stay on the curve). Mathematically, this is

$$\text{MRS} = -\frac{dY}{dX}.$$

Find MRS for the utility function given above, $U(X, Y) = Y^2(X - 3)$. Then calculate the MRS at the points $(4, 2)$ and $(7, 1)$ and interpret the results. (See Answer 2)

Example

Find the derivative of the function $y = f(x)$ defined implicitly by the equation

$$ye^x + x\ln(y) = e.$$

(See Answer 3)

15.2 Related rates

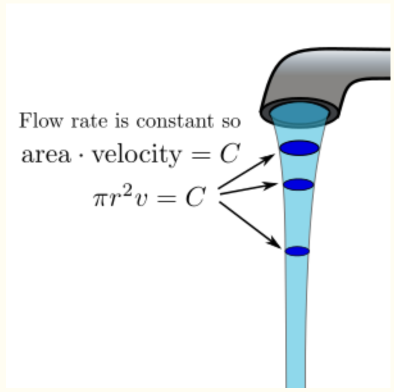
The differential is often used in *related rates* problems. A related rates problem typically has a physical description and asks for the rate at which some quantity is changing. The description must be translated into an implicit relation between the variables involved, and then implicit differentiation is used to find the desired derivative.

Example

Suppose a 10 foot ladder is leaning against a wall. The base of the ladder starts sliding away from the wall at a rate of 4 feet per second. At the moment when the base of the ladder is 6 feet away from the wall, at what rate is the top of the ladder sliding down the wall? (See Answer 4)

Example

Consider the shape of a stream of water as it flows from a faucet. The stream has a circular cross-section which gets narrower lower in the stream, and the goal is to find how the radius of that cross-section is changing with respect to time.
Assume on the one hand that the water is flowing at a constant rate C. On the other hand, the area of the cross section times the velocity through that cross section equals the flow:



Dividing by π and taking the differential gives

$$\begin{aligned} r^2 v &= C \\ 2r\,dr\,v + r^2\,dv &= 0 \\ dr &= -\frac{r^2}{2rv}\,dv = -\frac{r}{2v}\,dv. \end{aligned}$$

So

$$\frac{dr}{dv} = -\frac{r}{2v}.$$

Now, using the chain rule and a few facts from physics gives that

$$\begin{aligned} \frac{dr}{dt} &= \frac{dr}{dv}\frac{dv}{dt} \\ &= -\frac{r}{2v}\frac{dv}{dt} \\ &= -\frac{rg}{2v} \\ &= -\frac{rg}{2(v_0+gt)}. \end{aligned}$$

This is known as a *differential equation* (in particular, a *separable* differential equation). One can solve this equation to get an explicit expression for r in terms of t; see the module on separable differential equations.

15.3 Relative rates of change

We can normalize the differential du by dividing by u. This gives $\frac{du}{u}$. This is known as the *relative rate of change*. Note that

$$\frac{du}{u} = d(\ln u).$$

Example

For a given resistor in an electrical circuit, Ohm's law says that

$$V = IR,$$

where V is voltage across the resistor, I is the current, and R is the resistance of the resistor. If the voltage across a variable resistor is fixed, find the relationship between the relative rates change of resistance and current. (See Answer 5)

Example

In a geometric solid (say, a sphere or a cube), how does the relative rate of change in volume compare to that of surface area? (See Answer 6)

Example

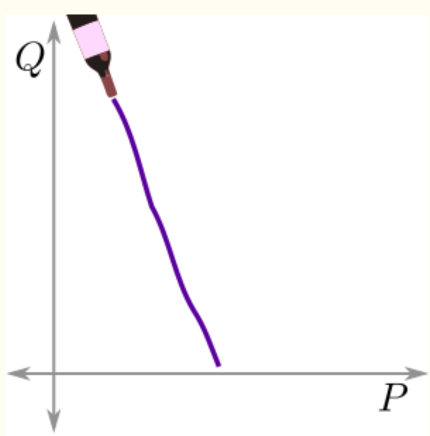
In economics, the *demand curve* for a good is the quantity Q of the good that a consumer would purchase as a function of the price P of the good. The demand curve slopes downward since a consumer will typically buy less of a good if it is more expensive (the exception being a Giffen good).

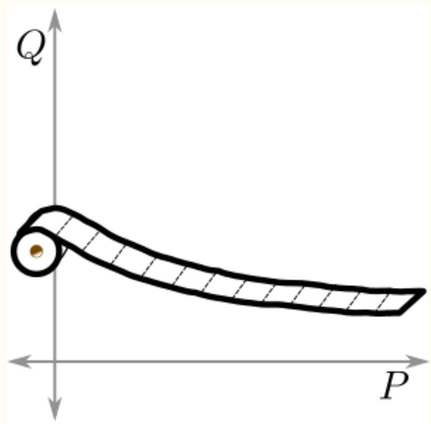
The *price elasticity of demand*, E, for a good is the rate of change of relative quantity fluctuation with respect to relative price fluctuation. Informally, it can be thought of as the percent change in quantity resulting from a percent change in price. One can also think of E as a measure of how price sensitive a consumer is for that good at that price. Mathematically,

$$E = -\frac{dQ/Q}{dP/P}.$$

The negative sign is there to force elasticity to be positive (without it, $\frac{dQ/Q}{dP/P}$ would always be negative due to the downward slope of the demand curve).

A good is said to be *elastic* at a certain price if $E > 1$ (that is, a consumer is highly sensitive to price changes). A good is *inelastic* at a certain price if $E < 1$. An example of an elastic good is wine, since a small increase in the relative price of wine can result in a consumer substituting a different alcohol for it. An example of an inelastic good is toilet paper, since regardless of price changes, a consumer is likely to require about the same amount of toilet paper:





Revenue, R, is given by $R = P \cdot Q$. How does one maximize the revenue with respect to price? Express the criterion in terms of E.

15.4 EXERCISES

- Use implicit differentiation to find $\frac{dy}{dx}$ from the equation $y^2 - y = \sin 2x$.
- Find the derivative $\frac{dy}{dx}$ if x and y are related through $xy = e^y$.
- Use implicit differentiation to find $\frac{dy}{dx}$ if $\sin x = e^{-y\cos x}$.
- Find the derivative $\frac{dy}{dx}$ from the equation $x \tan y - y^2 \ln x = 4$.
- Model a hailstone as a round ball of radius R. As the hailstone falls from the sky, its radius increases at a constant rate C. At what rate does the volume V of the hailstone change?
- The volume of a cubic box of side-length L is $V = L^3$. How are the relative rates of change of L and V related?
- Consider a box of height h with a square base of side length L. Assume that L is increasing at a rate of 10% per day, but h is decreasing at a rate of 10% per day. Use a linear approximation to find at what (approximate) rate the volume of the box changing. "Hint:" consider the relative rate of change of the volume of the box.
- A large tank of oil is slowly leaking oil into a containment tank surrounding it. The oil tank is a vertical cylinder with a diameter of 10 meters. The containment tank has a square base with side length of 15 meters and tall vertical walls. The bottom of the oil tank and the bottom of the containment tank are concentric (the round base inside the square base). Denote by h_o the height of the oil inside of the oil tank, and by h_c the height of the oil in the containment tank. How are the rates of change of these two quantities related?
- The "stopping distance" D_{stop} is the distance traveled by a vehicle from the moment the driver becomes aware of an obstacle in the road until the car stops completely. This occurs in two phases.

(1) The first one, the "reaction phase," spans from the moment the driver sees the obstacle until he or she has completely depressed the brake pedal. This entails taking the decision to stop the vehicle, lifting the foot from the gas pedal and onto the brake pedal, and pressing the latter down its full distance to obtain maximum braking power. The amount of time necessary to do all this is called the "reaction time" t_{react}, and is independent of the speed at which the vehicle was traveling. Although this quantity varies from driver to driver, it is typically between 1.5 s and 2.5 s. For the purposes of this problem, we will use an average value of 2 s. The distance traversed by the vehicle in this time is unsurprisingly called "reaction distance" D_{react} and is given by the formula $D_{\text{react}} = v t_{\text{react}}$, where v is the initial speed of the vehicle.

(2) In the "braking phase," the vehicle decelerates and comes to a complete stop. The "braking distance" D_{brake} that the vehicle covers in this phase is proportional to the square of the initial speed of the vehicle: $D_{\text{brake}} = \alpha v^2$. The constant of proportionality α depends on the vehicle type and condition, as well as on the road conditions. Consider a typical value of $10^{-2}\,\text{s}^2/\text{m}$.

If the initial speed of the vehicle is $108\,\text{km/h} = 30\,\text{m/s}$, what is the ratio between the relative rate of change of the stopping distance and the relative rate of change of the initial speed?

- Assume that you possess equal amounts of a product X and Y, but you value them differently. Specifically, your "utility function" is of the form $U(X, Y) = CX^{\alpha}Y^{\beta}$ for α, β, and C positive constants. What is your marginal rate of substitution (MRS) of Y for X?

15.5 Answers to Selected Examples

1. Taking the differential (remembering the chain rule) and doing some algebra gives

$$\begin{aligned}
2x(dx) + 2y(dy) &= 0 \\
2y(dy) &= -2x(dx) \\
\frac{dy}{dx} &= \frac{-2x}{2y} \\
&= -\frac{x}{y}.
\end{aligned}$$

So $\frac{dy}{dx} = -\frac{x}{y}$.

(Return)

2. Using implicit differentiation, we find

$$\begin{aligned}
Y^2(X-3) &= C \\
2Y\,dY\,(X-3) + Y^2\,dX &= 0.
\end{aligned}$$

Now solving for $-\frac{dY}{dX}$ we find

$$\text{MRS} = -\frac{dY}{dX} = \frac{Y^2}{2Y(X-3)} = \frac{Y}{2(X-3)}.$$

Evaluating at $(4, 2)$ gives a MRS of 1, which means that at that point the consumer is willing to substitute 1 doughnut for an ounce of coffee. At $(7, 1)$ the MRS is $\frac{1}{8}$, which means that the consumer is only willing to give up $\frac{1}{8}$ of a doughnut for an additional ounce of coffee.

In a sense the MRS is a measure of how a consumer reacts to the scarcity of one good relative to another and how that affects her willingness to exchange the goods.

(Return)

3. Taking the differential (remembering the product rule) gives

$$(dy)e^x + y(e^x dx) + (dx)\ln(y) + x\left(\frac{1}{y}dy\right) = 0.$$

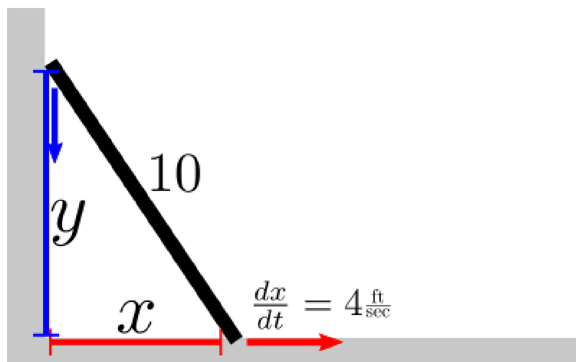
Ultimately, the goal is to solve for $\frac{dy}{dx}$, so collecting all the terms with dy on the left and the terms with dx on the right gives

$$\left(e^x + \frac{x}{y}\right) dy = (-ye^x - \ln(y))\, dx.$$

Finally, dividing through gives

$$\frac{dy}{dx} = \frac{-ye^x - \ln(y)}{e^x + x/y}.$$

(Return)

4. 

Let x be the distance from the base of the ladder to the wall and y be the distance from the top of the ladder to the floor. Then by the Pythagorean theorem, $x^2 + y^2 = 10^2$. At the moment in question, $x = 6$ so $y = 8$.

The differential of this equation gives

$$2xdx + 2ydy = 0,$$

and solving for $\frac{dy}{dx}$ gives $\frac{dy}{dx} = -\frac{x}{y} = -\frac{6}{8} = -\frac{3}{4}$. But the question asked for $\frac{dy}{dt}$, not $\frac{dy}{dx}$. But by the Chain rule,

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

Since $\frac{dy}{dx} = -\frac{3}{4}$ and $\frac{dx}{dt} = 4$ (given in the problem), it follows that

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = \left(-\frac{3}{4}\right) \cdot 4 = -3.$$

So the top of the ladder is sliding down the wall at a rate of 3 feet per second.

(Return)

5. Beginning with Ohm's law and taking the differential gives

$$V = IR$$
$$dV = R\,dI + I\,dR.$$

Now, note that $dV = 0$ since voltage was assumed to be constant. Then dividing through by IR gives

$$\begin{aligned} R\,dI + I\,dR &= 0 \\ \frac{R\,dI}{IR} + \frac{I\,dR}{IR} &= 0 \\ \frac{dI}{I} + \frac{dR}{R} &= 0 \\ \frac{dI}{I} &= -\frac{dR}{R}. \end{aligned}$$

Thus, the relative rates of change of resistance and current are equal and opposite.

(Return)

6. For a sphere of radius r, the volume and area (and their differentials) are

$$\begin{aligned} V &= \frac{4}{3}\pi r^3 & dV &= 4\pi r^2\,dr \\ A &= 4\pi r^2 & dA &= 8\pi r\,dr. \end{aligned}$$

Then the relative rates of change of volume and area are

$$\begin{aligned} \frac{dV}{V} &= \frac{4\pi r^2\,dr}{\frac{4}{3}\pi r^3} \\ &= 3\frac{dr}{r}. \end{aligned}$$

and

$$\begin{aligned} \frac{dA}{A} &= \frac{8\pi r\,dr}{4\pi r^2} \\ &= 2\frac{dr}{r}. \end{aligned}$$

Thus, the relative rate of change of volume is $\frac{3}{2}$ that of area. Written another way,

$$\frac{dV/V}{dA/A} = \frac{3}{2}.$$

For a cube of side length s, it turns out the same holds. The volume and area are

$$\begin{aligned} V &= s^3 & dV &= 3s^2\,ds \\ A &= 6s^2 & dA &= 12s\,ds. \end{aligned}$$

The relative rates are $\frac{dV}{V} = 3\frac{ds}{s}$ and $\frac{dA}{A} = 2\frac{ds}{s}$, so once again

$$\frac{dV/V}{dA/A} = \frac{3}{2}.$$

(Return)

16 Differentiation As An Operator

The sum, product, quotient, and chain rules make it possible to differentiate many functions. However, there are some more exotic functions which cannot be differentiated using these tools alone. For example, what is the derivative of 2^x or x^x or x^{x^x}?

The derivative should be interpreted as a rate of change, but what about the act of differentiation? Differentiation is an *operator:* it takes in a function and gives out another function. Other examples of operators include the logarithm, exponentiation, and integration. These (and other) operators can be applied to an entire equation to transform a hard problem to an easy problem and (once the solution is found) back again. This idea will allow us to compute the derivatives of the exotic functions above, and more.

16.1 Logarithmic differentiation

A common combination is called *logarithmic differentiation,* which consists of applying the logarithm operator followed by the differentiation operator. It is best demonstrated by example.

Example

Find the derivative of e^x using logarthmic differentiation. (See Answer 1)

Example

Use logarithmic differentiation to show that $\frac{d}{dx}(a^x) = a^x \ln a$. (See Answer 2)

Example

Find the derivative of x^x. (See Answer 3)

Example

Find the derivative of $f(x)^{g(x)}$. (See Answer 4)

16.2 Other operators

There are other operators which can be used prior to differentiation. Consider the following examples.

Example

Compute the derivative of $\ln x$ by using exponentiation followed by differentiation. (See Answer 5)

Note that $\ln x$ is the inverse of e^x, and so when we exponentiated the equation, it could be thought of as applying the inverse of $\ln x$. This same method works for many other inverse functions. In particular, applying a trigonometric function can be thought of as an operator as well. This method can be used to find the derivatives of the various inverse trigonometric functions.

Example

Find the derivative of $\arcsin(x)$. (See Answer 6)

Example

Find the derivative of $\arctan(x)$. (See Answer 7)

16.3 Operators in other contexts

Besides being useful in computing derivatives of exotic functions, operators (especially the logarithm) can also be useful in computing limits. The method is similar to the above method for derivatives.

Example

Show that

$$\lim_{x\to\infty}\left(1+\frac{a}{x}\right)^x = e^a.$$

This is a common limit which will come up again in the course. (See Answer 8)

16.4 Infinite Power Tower

Consider the infinite power tower

$$y = x^{x^{x^{x^{\cdots}}}}.$$

That is, x raised to the x raised to the x etc. This is certainly an unusual function. A better way to define this function is implicitly:

$$y = x^y.$$

To see that this makes intuitive sense, note that the exponent of the first x in the infinite tower is *itself* an infinite tower, so replacing the exponent of x with y is sensible.

Use logarithmic differentiation to find the derivative of this function (that is, $\frac{dy}{dx}$).

(See Answer 9)

It turns out that this function is well-defined and differentiable on

$$e^{-e} < x < e^{1/e}$$

One can check that the (x, y) pairs $(e^{-e}, e^{-1}), (1, 1), (\sqrt{2}, 2), (e^{1/e}, e)$ all satisfy the above implicit equation.

16.5 EXERCISES

- Find the derivative of $f(x) = (\ln x)^x$
- Find the derivative of $f(x) = x^{\ln x}$
- Compute $\lim_{x\to+\infty} \left(\frac{x+2}{x+3}\right)^{2x}$
- Compute $\lim_{x\to 0^+} [\ln(1+x)]^x$
- Compute $\lim_{x\to 0} \left(1 + \arctan\frac{x}{2}\right)^{2/x}$
- Compute $\lim_{x\to 0^+} \left(\frac{2}{x}\right)^{\sin x}$
- Let

$$\alpha = 1 + \frac{2}{2 + \frac{2}{2+\frac{2}{2+\cdots}}}$$

What is the value of α?

- Let

$$\varphi = \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}}}}$$

What is the value of φ?

16.6 Answers to Selected Examples

1. We already know the derivative of e^x, but suppose we did not. Let $y = e^x$. Then applying the logarithm operator to this equation gives
$$\ln y = \ln(e^x) = x.$$
Now applying the differentiation operator to the equation (and remembering the chain rule) gives
$$\frac{dy}{y} = dx,$$
so $\frac{dy}{dx} = y$, and since $y = e^x$ from our original definition, we have
$$\frac{dy}{dx} = e^x,$$
as expected.
(Return)

2. Similarly to the above example, let $y = a^x$. Then taking the logarithm gives

$$\ln y = \ln(a^x) = x \ln a.$$

Differentiating gives

$$\frac{dy}{y} = (\ln a)\, dx$$

And so

$$\frac{dy}{dx} = y(\ln a) = a^x \ln a,$$

as desired.

(Return)

3. Let $y = x^x$. Taking the logarithm gives

$$\ln y = x \ln x.$$

Differentiating (using the product rule on the right) gives that

$$\frac{dy}{y} = x\frac{1}{x}\, dx + \ln x\, dx.$$

Next, factoring and solving for $\frac{dy}{dx}$ gives

$$\frac{dy}{dx} = y(1 + \ln x) = x^x(1 + \ln x)$$

(Return)

4. Let $y = f(x)^{g(x)}$. Applying the logarithm gives $\ln y = g(x) \ln f(x)$, and differentiating gives

$$\frac{dy}{y} = g(x)\frac{1}{f(x)}f'(x)\, dx + g'(x) \ln f(x)\, dx.$$

Factoring and solving for $\frac{dy}{dx}$ shows

$$\begin{aligned}\frac{dy}{dx} &= y\left[g(x)\frac{1}{f(x)}f'(x) + g'(x) \ln f(x)\right] \\ &= f(x)^{g(x)}\left[\frac{g(x)}{f(x)}f'(x) + g'(x) \ln f(x)\right]\end{aligned}$$

(Return)

5. Letting $y = \ln x$, we exponentiate the equation to find

$$e^y = x.$$

Now, differentiating gives

$$e^y\, dy = dx,$$

and solving for $\frac{dy}{dx}$ gives

$$\frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{e^{\ln x}} = \frac{1}{x},$$

as desired.

(Return)

6. Letting $y = \arcsin x$, take the sine of the equation to find

$$\sin y = x.$$

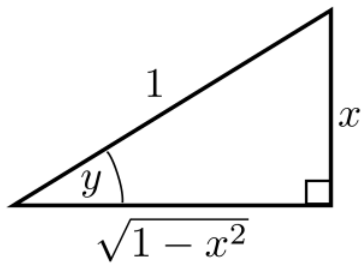
Now, differentiating gives

$$\cos y \, dy = dx.$$

Thus, we find that

$$\frac{dy}{dx} = \frac{1}{\cos y}.$$

Now, a little bit of trigonometry helps rewrite $\sec y$ in terms of x. Our original equation had $y = \arcsin x$. That means that y is the angle such that $\sin y = x$. Since sine is the opposite over the hypotenuse, we can express this relationship with the following right triangle:



where the adjacent leg comes from the Pythagorean theorem. It follows that $\cos y = \sqrt{1-x^2}$, and so we find

$$\frac{dy}{dx} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-x^2}}.$$

(Return)

7. Let $y = \arctan(x)$. Applying tan to the equation gives

$$\tan y = x,$$

and differentiating gives

$$\sec^2 y \, dy = dx.$$

Therefore,

$$\frac{dy}{dx} = \frac{1}{\sec^2 y}$$

We can now do similar right triangle trig as in the previous example. Or we can recall that by the Pythagorean identity for tangent and secant, we have

$$\tan^2 y + 1 = \sec^2 y$$

Making this substitution gives

$$\begin{aligned}\frac{dy}{dx} &= \frac{1}{\sec^2 y} \\ &= \frac{1}{\tan^2 y + 1} \\ &= \frac{1}{x^2+1}.\end{aligned}$$

(Return)

8. Let y be the function given by

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

Take the logarithm of both sides, and use the power property of logarithms to see

$$\begin{aligned}\ln y &= \ln\left(1 + \frac{a}{x}\right)^x \\ &= x\ln\left(1 + \frac{a}{x}\right).\end{aligned}$$

Now, taking the limit of both sides gives

$$\lim_{x\to\infty} \ln y = \lim_{x\to\infty} x\ln\left(1 + \frac{a}{x}\right).$$

Now, since $x \to \infty$, we have that $\frac{a}{x}$ is small, and so we can use our Taylor series for $\ln(1+x)$, which gives us that

$$\begin{aligned}\lim_{x\to\infty} \ln y &= \lim_{x\to\infty} x\ln\left(1 + \frac{a}{x}\right) \\ &= \lim_{x\to\infty} x\left(\frac{a}{x} + O\left(\frac{1}{x^2}\right)\right) \\ &= \lim_{x\to\infty} a + O\left(\frac{1}{x}\right) \\ &= a.\end{aligned}$$

Recall from our limit rules that the order of the logarithm and the limit can be switched since the logarithm is a continuous function. Thus

$$\ln\left(\lim_{x\to\infty} y\right) = a.$$

So, exponentiating both sides, we have that

$$\lim_{x\to\infty} y = e^a,$$

as desired.

(Return)

9. First, taking the logarithm gives

$$\ln y = y\ln x.$$

Now, differentiating the equation (implicitly) gives

$$\frac{dy}{y} = \ln x\, dy + y\frac{dx}{x}$$

Factoring and solving for $\frac{dy}{dx}$ gives

$$\begin{aligned}\left(\frac{1}{y} - \ln x\right) dy &= y\frac{dx}{x} \\ \frac{dy}{dx} &= \frac{y}{x\,(1/y - \ln x)} \\ &= \frac{y^2}{x(1 - y\ln x)}.\end{aligned}$$

This shows that although a function may be difficult to understand, it can nevertheless be fairly easy to find its derivative.

(Return)

The Penn Calc Companion

Part II: Integration and Applications

About this Document

This material is taken from the wiki-format Penn Calc Wiki, created to accompany Robert Ghrist's Single-Variable Calculus class, as presented on Coursera beginning January 2013. All material is copyright 2012-2013 Robert Ghrist. Past editors and contributors include: Prof. Robert Ghrist, Mr. David Lonoff, Dr. Subhrajit Bhattachayra, Dr. Alberto Garcia-Raboso, Dr. Vidit Nanda, Ms. Lee Ling Tan, Mr. Brett Bernstein, Prof. Antonella Grassi, and Prof. Dennis DeTurck.

Contents

17 Antidifferentiation

This module begins our study of integration. Integration, or *anti-differentiation*, can be thought of as running differentiation in reverse, or undoing the derivative.

This motivates the following definition:

The Indefinite Integral

The *indefinite integral of* $f(t)$, denoted $\int f(t)\,dt$, is the class of functions whose derivative is $f(t)$. $\int f(t)\,dt$ is also referred to as the *anti-derivative of* f. The act of taking the indefinite integral is an operator which is referred to as *anti-differentiation* or *integration*.

Note

The indefinite integral of a function is only defined up to an added constant, called the constant of integration. In other words, if $F(x)$ is an anti-derivative of $f(x)$, then $G(x) = F(x) + C$, C a constant, is also an anti-derivative of f, because C disappears when differentiated. Conversely, any two indefinite integrals of $f(x)$ differ only by some constant.

Any of the known derivatives from the previous chapter can be rephrased as an integral. For example, just as there was a power rule for differentiating monomials, there is a corresponding power rule for integrating monomials. And any anti-derivative can easily be checked by taking the derivative and seeing that the result gives back the original function.

Example

Give the integral of each of the following functions: x^n, $\frac{1}{x}$, $\sin x$, $\cos x$, e^x. (See Answer 1)

There are other functions which are harder to integrate by merely using one of the derivatives we already know. Some of these can be integrated using other techniques from upcoming modules, but there are also functions whose anti-derivative cannot be expressed in terms of simple functions.

17.1 Differential equations

The motivating problem for the study of anti-differentiation is solving a *differential equation*. A differential equation is an equation involving a function and its derivative. In this course, we deal with *ordinary* differential equations, ODEs, which are differential equations involving only functions of one variable and the derivative with respect to that variable (future courses deal with *partial differential equations*, which involve functions of several variables and partial derivatives).

Solving a differential equation means finding the function (or class of functions, usually) which satisfy the differential equation.

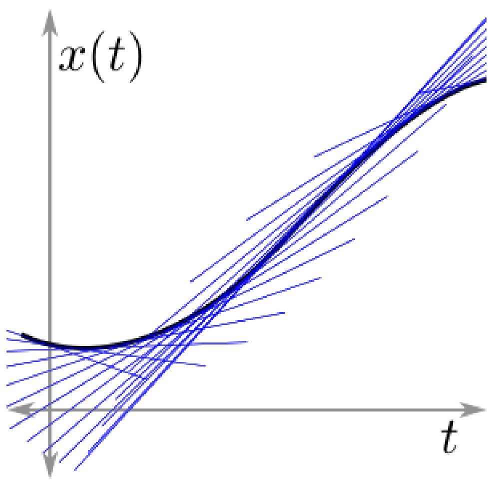
A Simple ODE

The simplest differential equation is of the form

$$\frac{dx}{dt} = f(t).$$

Here, the goal is to find the function $x(t)$ whose derivative with respect to t is $f(t)$. But this is precisely what the integral is. And so, the solution of the differential equation $\frac{dx}{dt} = f(t)$ is given by $x(t) = \int f(t)dt$.

Using the interpretation of the derivative as slope, one can think of the function $f(t)$ as describing the slope of the function $x(t)$:



Thus, $x(t)$ is a function which fits the slopes prescribed by $f(t)$. Note that any constant vertical shift of a solution $x(t)$ will still have the same slope at each point. This is one interpretation of the integration constant: it represents the potential vertical shifts to a solution of the differential equation.

Example

Consider a falling object. Let $x(t)$ be the height of the object at time t, $v(t)$ be the velocity of the object, and assume that acceleration is the constant $-g$ (negative because gravity pulls down). Express the height of the object as a function of t, v_0, and x_0; here, v_0 and x_0 are the velocity and height of the object, respectively, at time $t = 0$. (See Answer 2)

The Next Simplest ODE

Another slightly more complex ODE is of the form

$$\frac{dx}{dt} = f(x).$$

Before we discuss how to solve this in general, we consider a specific example, which is one of the most famous differential equations:

$$\frac{dx}{dt} = ax,$$

where a is a constant. We solve this differential equation in three different ways:

1. (Guess) Solve this differential equation by first observing that $x = Ce^t$ satisfies $\frac{dx}{dt} = x$ and then adjusting the exponent so that an extra factor of a comes out when differentiating. Hint: remember the chain rule. (See Answer 3)

2. (Series) Solve the differential equation by assuming

$$x(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots$$

   and then determining what the constants c_i must be to satisfy the differential equation. (See Answer 4)

3. (Integration) Rearrange the differential equation into the form

$$\frac{dx}{x} = a\,dt$$

   and integrate both sides to solve the differential equation. (See Answer 5)

The differential equation from this example is sometimes used as a simple model of population growth. In words, the differential equation says that the growth of a population is proportional to the size of the population. As the solution above slows, this model implies the population has exponential growth. This is not a very good model for most populations because of competition for resources and overcrowding. But under certain conditions and for short periods of time, some populations (for instance, bacteria with an abundant food supply) do exhibit exponential growth. For more examples of exponential growth, see the next module.

17.2 Initial value problems

Although a general differential equation's solution often depends on a constant (sometimes several), an additional condition called an *initial value* or *initial condition* can specify a specific solution. This condition is usually of the form $y(t_0) = y_0$. A differential equation with such an initial condition is called an *initial value problem*. To solve such a problem, first find the general solution and then use the initial value to find the specific constant of integration which satisfies the initial condition.

In the context of population growth, the initial value is typically the size of the population at time 0. This is particularly nice in the exponential growth model, because the solution is of the form $P(t) = De^{At}$. So if $P(0) = P_0$ is given, then plugging this in gives $P(t) = P_0 e^{At}$.

17.3 EXERCISES

- $\int (4x^3 + 3x^2 + 2x + 1)\,dx =$
- $\frac{d}{dx} \int \ln \tan x\,dx =$
- $\int \left(\frac{d}{dx} e^{-x}\right) dx =$

- Find the general solution of the differential equation

$$\frac{dx}{dt} = t^2$$

- Find the general solution of the differential equation

$$\frac{dx}{dt} = x^2$$

- There is a large class of differential equations – the so-called "linear" ones – for which we can find solutions using the Taylor series method discussed in the Lecture. One such differential equation is

$$t\frac{d^2x}{dt^2} + \frac{dx}{dt} + tx = 0$$

  This is a particular case of the more general "Bessel differential equation," and one solution of it is given by the Bessel function $J_0(t)$ that we saw earlier. Notice that this involves not only the first derivative but also the second derivative. For this reason, it is said to be a "second order" differential equation.

  In this problem we will content ourselves with finding a relationship (specifically, a "recurrence relation") on the coefficients of a Taylor series expansion about $t = 0$ of a solution to our equation. Hence consider the Taylor series

$$x(t) = \sum_{k=0}^{\infty} c_k t^k$$

  Substituting this into the differential equation above will give you two conditions. The first one is $c_1 = 0$. What is the other one?

  "Note:" this problem involves some nontrivial manipulation of indices in summation notation. Do not get discouraged if it feels more difficult than other problems: it is!

- Find the general solution of the differential equation

$$\frac{dx}{dt} = t^3 + x^2t^3$$

17.4 Answers to Selected Examples

1.

$$\int x^n dx = \frac{1}{n+1}x^{n+1} + C$$
$$\int \frac{1}{x} dx = \ln|x| + C$$
$$\int \sin x dx = -\cos x + C$$
$$\int \cos x dx = \sin x + C$$
$$\int e^x dx = e^x + C$$

(Don't forget the constant!)

(Return)

2. We know from an earlier module that
$$\frac{dv}{dt} = a = -g.$$
Beginning with the second of these equations, we find that
$$v(t) = \int (-g)\, dt = -gt + C.$$
We can determine C by plugging in $t = 0$. This cancels that $-gt$ and leaves us with $C = v(0) = v_0$, the initial velocity. Thus,
$$v(t) = -gt + v_0$$
Now, using the fact that
$$\frac{dx}{dt} = v,$$
we find that
$$\begin{aligned} x(t) &= \int v(t)\, dt \\ &= \int (-gt + v_0)\, dt \\ &= -\frac{1}{2}gt^2 + v_0 t + C. \end{aligned}$$
Again, we can find C by plugging in $t = 0$. This leaves us with $x(0) = C$, and so $C = x_0$, the initial height. Thus,
$$x(t) = -\frac{1}{2}gt^2 + v_0 t + x_0.$$
(Return)

3. Observe that $x = Ce^{at}$ will get an extra factor of a when differentiated by the chain rule. That is,
$$\frac{d}{dt}\left(Ce^{at}\right) = aCe^{at}.$$
And so $x(t) = Ce^{at}$ is a solution of the differential equation.

(Return)

4. Assuming that
$$x(t) = c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots,$$
and then taking the derivative of this series, term by term, we find
$$\frac{dx}{dt} = 0 + c_1 + 2c_2 t + 3c_3 t^2 + \cdots.$$
On the other hand, from the original differential equation we have
$$\begin{aligned} \frac{dx}{dt} &= ax \\ &= a\left(c_0 + c_1 t + c_2 t^2 + c_3 t^3 + \cdots\right) \\ &= ac_0 + ac_1 t + ac_2 t^2 + ac_3 t^3 + \cdots. \end{aligned}$$

Because these two series both equal $\frac{dx}{dt}$, they must be equal to each other. But two series are equal if and only if their corresponding coefficients are equal. Therefore,

$$\begin{aligned} c_1 &= ac_0 \\ 2c_2 &= ac_1 \\ 3c_3 &= ac_2, \end{aligned}$$

and so on. Solving these equations one by one gives

$$\begin{aligned} c_1 &= ac_0 \\ c_2 &= \frac{1}{2}ac_1 = \frac{1}{2}a^2c_0 \\ c_3 &= \frac{1}{3}ac_2 = \frac{1}{6}a^3c_0 \end{aligned}$$

And, generally, $c_n = \frac{1}{n!}a^nc_0$ (this can be proven using a method called induction). Doing a little bit of factoring and grouping of factors, we find

$$\begin{aligned} x(t) &= c_0 + ac_0t + \frac{1}{2!}a^2c_0t^2 + \frac{1}{3!}a^3c_0t^3 + \cdots \\ &= c_0\left(1 + (at) + \frac{1}{2!}(at)^2 + \frac{1}{3!}(at)^3 + \cdots\right) \\ &= c_0e^{at}, \end{aligned}$$

which is, again, of the form Ce^{at}.

(Return)

5. Using the chain rule and substituting according to the differential equation, we have

$$\begin{aligned} dx &= \frac{dx}{dt}\,dt \\ dx &= ax\,dt \\ \frac{dx}{x} &= a\,dt. \end{aligned}$$

Integrating both sides of the equation gives

$$\ln x = at + C$$

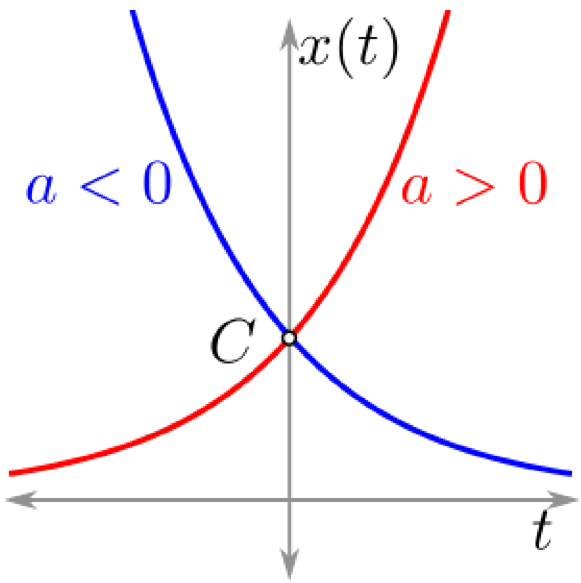
(only one constant of integration is necessary here, because a constant on the left side could be subtracted from both sides and absorbed into C). Now, exponentiating the equation gives $x = e^{at+C}$. By exponential rules, $e^{at+C} = e^{at}e^C$, and the e^C is often rewritten as a new constant, often written C again.

Thus, the solution to $\frac{dx}{dt} = ax$ is $x(t) = Ce^{at}$, where C is any constant.

(Return)

18 Exponential Growth Examples

Recall from the last module that the differential equation $\frac{dx}{dt} = ax$ has solution $x = Ce^{at}$, where C is some constant. The constant C can be thought of as an initial condition, the value of the function at time $t = 0$. When $a > 0$, the function has *exponential growth*. When $a < 0$, the function has *exponential decay*:



This module is devoted to several examples of exponential growth and decay.

18.1 Radioactive decay

Carbon-14 is a radioactive isotope of carbon which exists in organic materials. It is known that the rate at which carbon-14 atoms decay is proportional to the number of carbon-14 atoms present. If I represents the number of atoms, then the differential equation is

$$\frac{dI}{dt} = -\lambda I,$$

where λ is positive (and so the number of atoms is decreasing).

18.2 Population growth

For bacteria with an abundant food supply, the population P satisfies

$$\frac{dP}{dt} = bP,$$

for some positive b. But as the food supply dwindles, or overcrowding occurs, the population growth will necessarily slow (or else the bacteria would eventually consume the earth). Thus, this is not usually an accurate population model.

18.3 Interest accumulation

Consider a bank account with initial deposit (also called the *principal*) P, annual interest rate r, and which is compounded n times a year (so $n = 1$ gives simple interest, $n = 4$ is quarterly interest, etc.). Then the value of the account at the end of k years is $P\left(1 + \frac{r}{n}\right)^{nk}$. What happens as n gets bigger and bigger? Recall that $\lim_{n\to\infty}\left(1 + \frac{\alpha}{n}\right)^n = e^\alpha$. Then it follows that

$$\begin{aligned}\lim_{n\to\infty} P\left(1 + \frac{r}{n}\right)^{nk} &= \lim_{n\to\infty} P\left[\left(1 + \frac{r}{n}\right)^n\right]^k \\ &= Pe^{rk}.\end{aligned}$$

This is called *continuous compounding*. So an account with continuous compounding is worth Pe^{rk} after k years, where P is the initial investment, and r is the annual interest rate.

Rule of 70

The *Rule of 70* is a mental math trick that approximates the number of years it takes for a continuously compounded account to double in value. The rule says that

$$\#\text{ years for the account to double } = \frac{70}{100r}.$$

In other words, the number of years is 70 divided by the annual percentage. Verify the Rule of 70. (See Answer 1)

18.4 Linguistics

Historical linguists study (among other things) word usage and the rate at which words fall out of use. Let $W(t)$ be the number of words which are in active use in English after t years. One model predicts that $\frac{dW}{dt} = -\lambda W$, hence $W = Ce^{-\lambda t}$, and C is the number of common words at time 0.

Example

Suppose the writing of John Milton (from around 1667) consists of 20% words which are unfamiliar to us today. Use the above model to estimate the fraction of words that Shakespeare's audience (from around 1600) recognized of Chaucer's writing (from around 1400). (See Answer 2)

This model ignores the creation of new words as well as the fact that definitions of existing words can evolve over time. So it is probably not the most accurate model. It is worth noting that even if the mathematics are correct, if the underlying model is not very good, then the resulting answer will not be very good either.

18.5 Zombies

Suppose in the zombie apocalypse that the rate of change of the infected population $Z(t)$ is proportional to the uninfected population $U(t)$. Let P be the total population (assumed to be constant). Then

$$\frac{dZ}{dt} = rU = r(P - Z),$$

since $U = P - Z$ (the number of uninfected is the total population less the infected). Taking the derivative of $U = P - Z$ gives

$$\begin{aligned} \frac{dU}{dt} &= \frac{dP}{dt} - \frac{dZ}{dt} \\ &= -\frac{dZ}{dt} \\ &= -rU, \end{aligned}$$

since P was assumed to be constant. Thus, $U(t) = U_0 e^{-rt}$, where U_0 is the initial uninfected population. And so the zombie population is given by

$$Z(t) = P - U_0 e^{-rt}.$$

Note the population is doomed, as $Z(t) \to P$ as $t \to \infty$.

Although a zombie apocalypse is unlikely, this model is still useful for other phenomena involving how quickly something spreads. This includes the spread of disease, propaganda, and technology. Another example is heat transfer, discussed in detail below.

18.6 Newton's law of cooling

Newton's law of cooling says that the rate of change of the temperature of a body is proportional to the difference of the temperatures of the body and ambient environment. Let T be the temperature of the body and A be the ambient temperature. Then

$$\frac{dT}{dt} = k(A - T),$$

for some positive constant k. Separating gives

$$\frac{dT}{T - A} = -k\,dt,$$

and integrating and exponentiating gives $T - A = Ce^{-kt}$. Thus the solution is $T = A + Ce^{-kt}$. Note that k must be positive for this to model to make sense. Indeed, as $t \to \infty$, the temperature of the body should approach the ambient temperature, which will only happen if $k > 0$. Also, note that $C = T_0 - A$ is the initial difference in temperature.

18.7 EXERCISES

- After drinking a cup of coffee, the amount C of caffeine in a person's body obeys the differential equation

  $$\frac{dC}{dt} = -\alpha C$$

  where the constant α has an approximate value of 0.14 hours^{-1}

  How many hours will it take a human body to metabolize half of the initial amount of caffeine?

- The amount I of a radioactive substance in a given sample will decay in time according to the following equation:

$$\frac{dI}{dt} = -\lambda I$$

Nuclear engineers and scientists tend to be concerned with the "half-life" of a substance, that is, the time it takes for the amount of radioactive material to be halved.

Find the half-life of a substance in terms of its decay constant λ.

- In a highly viscous fluid, a falling spherical object of radius r decelerates right before reaching the bottom of the container. A simple model for this behavior is provided by the equation

$$\frac{dh}{dt} = -\frac{\alpha}{r}h,$$

where h is the height of the object measured from the bottom, and α is a constant that depends on the viscosity of the fluid.

Find the time it would take the object to drop from $h = 6r$ to $h = 2r$ in terms of α and r.

- On a cold day you want to brew a nice hot cup of tea. You pour boiling water (at a temperature of 212° F) into a mug and drop a tea bag in it. The water cools down in contact with the cold air according to Newton's law of cooling:

$$\frac{dT}{dt} = \kappa(A - T)$$

where T is the temperature of the water, $A = 32°$F the ambient temperature, and $\kappa = 0.36$ min^{-1}.

The threshold for human beings to feel pain when entering in contact with something hot is around 107° F. How many seconds do you have to wait until you can safely take a sip?

- On the night of April 14, 1912, the British passenger liner RMS Titanic collided with an iceberg and sank in the North Atlantic Ocean. The ship lacked enough lifeboats to accommodate all of the passengers, and many of them died from hypothermia in the cold sea waters. Hypothermia is the condition in which the temperature of a human body drops below normal operating levels (around 36°C). When the core body temperature drops below 28°C, the hypothermia is said to have become severe: major organs shut down and eventually the heart stops.

If the water temperature that night was -2°C, how long did it take for passengers of the Titanic to enter severe hypothermia? Recall from lecture that heat transfer is described by Newton's law of cooling:

$$\frac{dT}{dt} = \kappa(A - T)$$

where T is the body temperature of a passenger, A the water temperature, and $\kappa = 0.016$ min^{-1}.

- The birthrate of a population (number of births per year $\times$ 100 / number of population) is 20%, and the mortality rate (number of deaths per year $\times$ 100/ number of population) is 5%. If the initial population is 10,000, find the function $P(t)$, the population as a function of time. How long does it take for the population to double?

- Let $y(t)$ denote the number of atoms of a particular radioactive isotope of carbon at year t. We know that the rate at which $y(t)$ decreases is proportional to y itself. (a) What differential equation does $y(t)$ satisfy? (b) If it takes 5 years for the number to decrease to half of its initial number, what is the constant involved in your answer of part (a)?

18.8 Answers to Selected Examples

1. The problem asks for k when the balance has doubled. That is, find k such that

$$2P = Pe^{rk}.$$

Dividing by P and taking logarithms gives $rk = \ln(2)$, so $k = \frac{\ln(2)}{r}$. Since $\ln(2) \approx .7$, this shows that

$$\begin{aligned} k &= \frac{\ln(2)}{r} \\ &\approx \frac{.7}{r} \\ &\approx \frac{70}{100r}, \end{aligned}$$

as desired.

(Return)

2. Let t denote time measured in years. Let $WM(t)$ denote the number of words from Milton's time which are in common use at time t. Then

$$WM(t) = WM(1667)e^{-\lambda(t-1667)}$$

(the $t - 1667$ is used in the exponent because we are interested in the number of years since Milton). Let $WC(t)$ denote the number of words from Chaucer's time which are in common use at time t. Then

$$WC(t) = WC(1400)e^{-\lambda(t-1400)}.$$

From the given information, we have that

$$\frac{WM(2013)}{WM(1667)} = \frac{4}{5}.$$

And according to the above formula, we have

$$\frac{WM(1667)e^{-\lambda\cdot 346}}{WM(1667)} = \frac{4}{5}.$$

The factors of $WM(1667)$ cancel, leaving

$$e^{-\lambda\cdot 346} = \frac{4}{5}.$$

Taking the logarithm of both sides and dividing by -346 gives

$$\lambda = \frac{-\ln(4/5)}{346} \approx 6.5 \times 10^{-4}.$$

Knowing λ allows us to compute the fraction of words from Chaucer's time that would be recognized in Shakespeare's time:

$$\begin{aligned} \frac{WC(1600)}{WC(1400)} &= \frac{WC(1400)e^{-\lambda(200)}}{WC(1400)} \\ &= e^{-\lambda(200)} \\ &\approx .88 \end{aligned}$$

So, according to this model, approximately 88% of words from Chaucer's work would have been understood in Shakespeare's time.

(Return)

19 More Differential Equations

Recall that an ordinary differential equation is an equation involving a function and its derivatives. The solution to a differential equation is a function which satisfies the equation. An earlier module introduced a few basic differential equations. This module deals with a few different families of differential equations and the methods of solving them.

19.1 Autonomous differential equations

A differential equation is called *autonomous* if the derivative of the function $x(t)$ is independent of t, i.e. the equation is of the form

$$\frac{dx}{dt} = f(x).$$

Logically, a *nonautonomous* differential equation is one where the derivative equals a function of both x and t:

$$\frac{dx}{dt} = f(x, t).$$

In general, nonautonomous differential equations can be very difficult, but certain types yield to a little algebra and integration. These include separable differential equations and linear first order differential equations, which are covered here.

19.2 Separable differential equations

A *separable differential equation* is one where, with a little algebra, we are able to express the differential equation in the form

$$\frac{dx}{dt} = f(x)g(t).$$

This may involve some algebra. Note in particular that any autonomous equation is separable (think of $g(t) = 1$). Once a differential equation is factored this way, it can be solved by using the chain rule and some algebra:

$$\begin{aligned} dx &= \frac{dx}{dt}\,dt \\ dx &= f(x)g(t)\,dt \\ \frac{dx}{f(x)} &= g(t)\,dt. \end{aligned}$$

(This is why these equations are called separable–the variables can be separated to opposite sides of the equation). To solve this equation, one must find the functions whose derivatives are $\frac{1}{f(x)}$ and $g(t)$, respectively. In other words, one integrates both sides.

Example

Solve the differential equation

$$y' = 3yx^2.$$

(See Answer 1)

Example

Solve the differential equation

$$\frac{dx}{dt} = e^{t-x}.$$

(See Answer 2)

Example

Solve the initial value problem $\frac{dy}{dx} = xy + x$, with $y(0) = 3$. (See Answer 3)

Example

Assume that a falling body with mass m has a drag force proportional to velocity $v(t)$. Then the downward acceleration mg is being counteracted by the upward acceleration κv, for some constant κ. Thus,

$$\begin{aligned} m\frac{dv}{dt} &= mg - \kappa v \\ &= -\kappa\left(v - \frac{mg}{\kappa}\right). \end{aligned}$$

which is separable. Solve this differential equation. (See Answer 4)

19.3 Linear 1st order differential equations

The product rule gives a technique (integration by parts) for seemingly difficult integrals; the product rule also gives a technique for solving a certain class of non-separable differential equations called *linear 1st order differential equations*. This is a differential equation which can be written in the form

$$\frac{dx}{dt} = A(t)x + B(t)$$

(as for separable differential equations, this may involve a little algebra). This form gives the reason for calling these equations *linear*, since dropping the t's gives

$$\frac{dx}{dt} = Ax + B,$$

which is reminiscent of the equation of a line. *1st order* means that the equation only involves the function x and its derivative $\frac{dx}{dt}$ (and no higher derivatives), along with functions of t.

The *standard form* of a linear 1st order differential equation is achieved by bringing all the terms involving x to the left side, which gives

$$\frac{dx}{dt} - A(t)x = B.$$

Example

Identify which of these is a linear 1st order differential equation, and put it in standard form if it is. In the cases that are not, identify which condition is violated. Are all separable equations also linear 1st order?

1. $tx' + x = 0$
2. $x' - e^t x^2 = 0$
3. $x' = x\sin(t)$
4. $x'' - t^2 \frac{dx}{dt} = 0$

(See Answer 5)

Integrating factors

The method for solving linear 1st order differential equations is to use the product rule to factor the sum of two derivatives into the derivative of a product. It is best explained by example.

Example

In part 1 from the previous example, note that $tx' + x = (tx)'$ by the product rule. So that differential equation can be written as $(tx)' = 0$. Integrating both sides gives $tx = C$ for some constant C. Thus the solution is $x = \frac{C}{t}$.
However, not all linear 1st order differential equations are expressed so nicely. For instance, in example 3 above, one cannot rewrite $x' - \sin(t)x$ as the derivative of a product of functions. This is where an *integrating factor* is used.
The integrating factor, denoted by $I(t)$ in this course, is a function which is multiplied through the entire differential equation, giving

$$I\frac{dx}{dt} - IAx = IB.$$

$I(t)$ is chosen so that the left side of this equation can be factored as a derivative of a product using the product rule. Symbolically, the goal is to choose $I(t)$ so that

$$I\frac{dx}{dt} - IAx = \frac{d}{dt}(Ix).$$

To find I, expand the product $\frac{d}{dt}(Ix) = I\frac{dx}{dt} + \frac{dI}{dt}x$. For this to equal the left side of the above equation, it must be that $-IA = \frac{dI}{dt}$. This differential equation is separable, and one finds that $\frac{dI}{I} = -A\,dt$. Integrating and exponentiating gives that

$$I(t) = e^{\int -A(t)\,dt}.$$

One need not work through all this algebra every time but can jump straight to writing down the integrating factor. Multiplying through by the integrating factor allows the left side to be rewritten by the product rule, and integrating both sides finishes the problem.

To summarize the method:

1. Get the differential equation into standard form $\frac{dx}{dt} - A(t)x = B(t)$.
2. Compute the integrating factor $I(t) = e^{-\int A(t)\,dt}$.
3. Multiply the entire equation by $I(t)$, which gives

$$I\frac{dx}{dt} - IAx = IB.$$

1. Rewrite the left side as the derivative of a product (this works because of the way $I(t)$ was chosen): $\frac{d}{dt}(Ix) = IB$.
2. Integrate both sides and then divide by I.
3. The final answer, then, is given by

$$x(t) = e^{\int A} \cdot \int Be^{-\int A}.$$

Example

Solve the differential equation

$$tx' + tx = t^2.$$

Hint: $\int te^t\,dt = te^t - e^t + C$. (See Answer 6)

Example

Suppose a 1000 gallon tank is 90% full. An additive is is pumped into the tank at a rate of 10 gallons per minute. The mixture is well stirred and drained at a rate of 5 gallons per minute.
What is the concentration of the additive when the tank is full? (See Answer 7)

19.4 EXERCISES

- Solve the differential equation $\dfrac{dx}{dt} = \dfrac{x}{t}$.
- Solve the differential equation $\dfrac{dx}{dt} = \dfrac{\sqrt{1-x^2}}{\sqrt{1-t^2}}$.
- Given that $x(0) = 0$ and $\dfrac{dx}{dt} = te^x$, compute $x(1)$.
- What integrating factor should be used to solve the linear differential equation

$$t^2\frac{dx}{dt} = 4t - t^5x$$

- Solve the differential equation $\dfrac{dx}{dt} - 5x = 3$.
- Solve the differential equation $\dfrac{dx}{dt} = \dfrac{x}{1+t} + 2$.

- Suppose that, in order to buy a house, you obtain a mortgage. If the lender advertises an annual interest rate r, your debt $D(t)$ will increase exponentially according to the simple O.D.E.

$$\frac{dD}{dt} = rD$$

If you pay your debt at a rate of P (annual rate, paid continuously), the evolution of your debt will then (under assumptions of continual compounding and payment) obey the linear differential equation

$$\frac{dD}{dt} = rD - P$$

Using this model, answer the following question: if initial amount of the mortgage is for \$400,000, the annual interest rate is 5%, and you pay at a rate of \$40,000 every year, how many years will it take you to pay off the debt?

- German physician Ernst Heinrich Weber (1795-1878) is considered one of the fathers of experimental psychology. In his study of perception, he noticed that the perceived difference between two almost-equal stimuli is proportional to the *percentual* difference between them. In terms of differentials, we can express Weber's law as

$$dp = k\frac{dS}{S},$$

where p is the perceived intensity of a stimulus and S its actual strength. Observe the relative rate of change on the right hand side. In what way must the magnitude of a stimulus change in time for a human being to perceive a linear growth? Linearly? Logarithmically? Polynomially?

- Some nonlinear differential equations can be reduced to linear ones by a clever change of variables. Bernouilli equations

$$\frac{dx}{dt} + p(t)x = q(t)x^{\alpha}, \qquad \alpha \in \mathbb{R}$$

constitute the most important case. Notice that for $\alpha = 0$ or $\alpha = 1$ the above equation is already linear. For other values of α, the substitution $u = x^{1-\alpha}$ yields a linear differential equation in the variable u.

Apply the above change of variables in the case

$$\frac{dx}{dt} + 2tx = x^3$$

What differential equation on u do you get?

19.5 Answers to Selected Examples

1. Separating gives

$$\frac{dy}{y} = 3x^2\,dx.$$

Integrating both sides, we have

$$\int \frac{dy}{y} = \int 3x^2\,dx$$
$$\ln y = x^3 + C.$$

Now exponentiating both sides gives

$$\begin{aligned} y &= e^{x^3+C} \\ &= Ce^{x^3}, \end{aligned}$$

for some constant C (remember that the C is not the same in the first and second line above, but we just rewrite it for convenience).

(Return)

2. First, using a law of exponents on the right side, we have

$$e^{t-x} = e^t e^{-x}.$$

Now, separating gives

$$e^x \, dx = e^t \, dt.$$

We might be tempted at this point to say $x = t$ because of the symmetry of this equation. But we must integrate both sides, which introduces an integration constant:

$$e^x = e^t + C.$$

Now taking the logarithm gives

$$x = \ln(e^t + C).$$

Note We must have the integration constant contained within the natural logarithm. In general, it is best to introduce the integration constant as soon as the integration occurs. A common mistake is to forget the constant and then at the very end of the problem add it. This is frequently incorrect, as in this case.

(Return)

3. First, this differential equation does not look like it is of the form of a separable differential equation. However, with a little factoring, one finds that $\frac{dy}{dx} = x(y+1)$. Thus,

$$\frac{dy}{y+1} = x \, dx.$$

Anti-differentiating gives $\ln|y+1| = \frac{1}{2}x^2 + C$. Then, exponentiating gives

$$|y+1| = e^{x^2/2+C} = De^{x^2/2}.$$

Apply the initial condition by plugging in $x = 0$ and $y = 3$, which gives that $D = 4$. Thus, the solution to the initial value problem is $y = 4e^{x^2/2} - 1$. One should double check that this satisfies the differential equation.

(Return)

4. Separating gives

$$\frac{dv}{v - mg/\kappa} = -\frac{\kappa}{m} dt.$$

Integrating both sides gives $\ln(v - mg/\kappa) = -\kappa t/m + C$, and so exponentiating and solving for v gives

$$v = Ce^{-\kappa t/m} + \frac{mg}{\kappa}$$

(here C is replacing the constant e^C from exponentiating). Note that as $t \to \infty$, the exponential term goes to 0, and so $v(t) \to \frac{mg}{\kappa}$, which is the terminal velocity of the falling body (when the force of gravity and drag cancel each other).

(Return)

5. 1. Linear 1st order. Standard form is $x' + \frac{1}{t}x = 0$.
   2. Not linear because of the x^2. Note that this is separable though.
   3. Linear 1st order. Standard form is $x' - \sin(t)x = 0$.
   4. Not 1st order because of the presence of x''.

   Number 2 shows that a differential equation can be separable even though it is not linear 1st order.

   (Return)

6. Divide through by t to get the equation in standard form: $x' + x = t$. Compute the integrating factor

$$I = e^{\int dt} = e^t.$$

   Multiplying through gives

$$e^t x' + e^t x = e^t t.$$

   Rewriting this using the product rule gives $(e^t x)' = te^t$. Integrating both sides and using the hint gives

$$e^t x = te^t - e^t + C.$$

   Finally, dividing by e^t gives

$$x = t - 1 + \frac{C}{e^t}$$

   as a final answer.

   (Return)

7. Begin by setting up a few variables to help make sense of what is happening. Let $V(t)$ be the volume of the total mixture at time t. Let $Q(t)$ be the total amount of additive in the mixture at time t. Let $C(t)$ be the concentration of the mixture, i.e.

$$C(t) = \frac{Q(t)}{V(t)}.$$

   The volume is not too difficult to compute. Since there are 10 gallons per minute entering the tank, and 5 gallons per minute leaving the tank, the net amount of fluid entering the tank is 5 gallons per minute. The tank begins at 90% full, which is 900 gallons. So

$$V(t) = 900 + 5t,$$

   and is full at $t = 20$. Next, consider the rate at which the quantity $Q(t)$ of additive in the tank is changing. There is 10 gallons of pure additive entering per minute and 5 gallons of mixture leaving. Therefore the amount of additive leaving is $5C$. Putting this together gives

$$\frac{d}{dt}Q = 10 - 5C = 10 - \frac{5Q}{900 + 5t}.$$

   Rearranging this gives

$$\frac{dQ}{dt} + \frac{1}{180 + t}Q = 10.$$

This is a 1st order linear differential equation. Computing the integrating factor we find

$$\begin{aligned} I &= \exp\left(\int \frac{1}{180+t}\,dt\right) \\ &= \exp(\ln(180+t)) \\ &= 180+t. \end{aligned}$$

Multiplying through gives

$$(180+t)\frac{dQ}{dt} + Q = 10(180+t).$$

Now, as always, the left side can be rewritten using the product rule to give

$$\frac{d}{dt}\left[Q(180+t)\right] = 1800 + 10t.$$

Integrating both sides gives

$$Q \cdot (180+t) = 1800t + 5t^2 + K$$

(using K here to avoid confusion with concentration C). We can find K by setting $t = 0$ in this equation. Since $Q(0) = 0$ (there is no additive in the tank initially), we find $K = 0$.

Now, solving for Q and evaluating at $t = 20$ gives

$$\begin{aligned} Q &= \frac{1800t + 5t^2}{180+t} \\ &= 190. \end{aligned}$$

So the final concentration of the full tank is

$$\frac{190}{1000} = 19\%.$$

(Return)

20 ODE Linearization

We have seen techniques for solving two types of differential equations: separable and linear. Unfortunately, there are a lot of differential equations which do not fit into these categories. In some of these cases, we can use linearization to determine the behavior of such differential equations.

20.1 Oscillation

How does one model oscillation? It turns out that a first order differential equation will not work, but a second order (i.e. involving the second derivative) equation will:

$$\frac{d^2x}{dt^2} = -a^2x.$$

Solving such an equation is beyond the scope of this course, but in a course on differential equations one finds the pair of solutions

$$\begin{aligned} x &= C_1 \cos(at) \\ x &= C_2 \sin(at). \end{aligned}$$

For this course we will look at a simpler way to model oscillation.

20.2 Simple Oscillator

Consider a spinner where $\theta(t)$ represents the angle the arrow makes with the positive x-axis at time t. Then θ increases linearly with t and whenever θ gets to 2π, it goes back to 0: (Click Here: Simple Oscillator Animated GIF)

This can be modeled by

$$\begin{aligned} \frac{d\theta}{dt} &= a \\ \theta &= at + \theta_0 \mod 2\pi \end{aligned}$$

where $\mod 2\pi$ means "take the remainder when divided by 2π". Here a can be thought of as the frequency of the spinner (e.g. how many revolutions per minute it makes).

20.3 Coupled Oscillators

Now consider two simple oscillators, θ_1 and θ_2, with the same frequency a, but which are slightly out of phase with each other (i.e. one arrow is slightly ahead of the other): (Click Here: Two Oscillators Animated GIF)

Now suppose these oscillators are *coupled* so that each exerts a small influence on the other (e.g. by connecting their axles with a rod). One way to represent this mathematically is to adjust the rates of change of the oscillators so that they are affected by the difference in angles:

$$\begin{aligned}\frac{d\theta_1}{dt} &= a + \epsilon \sin(\theta_2 - \theta_1)\\ \frac{d\theta_2}{dt} &= a - \epsilon \sin(\theta_2 - \theta_1).\end{aligned}$$

Here, ϵ is some small constant which represents the strength of the effect of the coupling. When θ_2 is bigger than θ_1, the above differential equations speed up θ_1 slightly and slow down θ_2 slightly. One can find by simulation that this coupling effect causes the oscillators to synchronize relatively quickly, depending on how big the phase is between them and how big ϵ is: (Click Here: Coupled Oscillators Animated GIF)

Synchronization

To analyze the synchronization effect mathematically, consider the *phase* φ between the two oscillators:

$$\varphi = \theta_2 - \theta_1.$$

Looking at how the phase φ changes with respect to time gives

$$\begin{aligned}\frac{d\varphi}{dt} &= \frac{d}{dt}(\theta_2 - \theta_1)\\ &= \frac{d\theta_2}{dt} - \frac{d\theta_1}{dt}\\ &= (a - \epsilon\sin(\theta_2-\theta_1)) - (a + \epsilon\sin(\theta_2-\theta_1))\\ &= -2\epsilon\sin(\theta_2-\theta_1)\\ &= -2\epsilon\sin(\varphi).\end{aligned}$$

This is a separable differential equation, but solving it to find φ as an explicit function of t is not so easy, and does not really help us understanding the synchronization phenomenon. But linearization will help us understand the synchronization effect and how quickly it occurs.

Linearization

Going back to the differential equation for the phase, suppose we replace $\sin\varphi$ with its linearization:

$$\begin{aligned}\frac{d\varphi}{dt} &= -2\epsilon\sin\varphi\\ &= -2\epsilon\left(\varphi + O(\varphi^3)\right)\\ &\approx -2\epsilon\varphi.\end{aligned}$$

This will be a good approximation assuming the phase is small (the oscillators are not too far out of sync). This is a familiar differential equation, which gives us the approximate solution

$$\varphi(t) \approx \varphi_0 e^{-2\epsilon t},$$

where φ_0 is the initial phase. This is called the *linearized solution* to the original differential equation. Here, the linearized solution predicts that the phase decays exponentially, which is consistent with the above simulation.

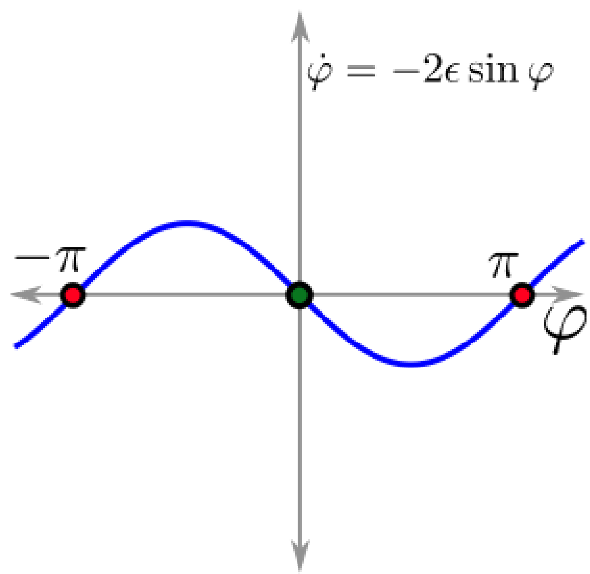
20.4 Equilibria

Another way to study differential equations and predict their behavior, is to study the equilibria of the equation. An *equilibrium* of the equation

$$\dot{x} = f(x)$$

(here, $\dot{x} = \frac{dx}{dt}$), is a solution $x(t) = C$, C a constant, such that $\dot{x} = 0$. In other words, an equilibrium is a root of f. In terms of the differential equation, an equilibrium is a steady state where the quantity x does not change.

One way to find the equilibria of a differential equation is to plot the derivative of a function versus the function itself. From the phase differential equation above, we plot $\dot{\varphi}$ on the y-axis and φ on the x-axis and look for roots:



The roots of this equation are the values of φ for which $\sin \varphi = 0$. For the range of values in which we are interested, the roots are $\varphi = -\pi, 0, \pi$. The equilibrium at 0 is familiar, because that is the state of synchronization to which the above simulation converged. The other two correspond to a phase of π, which means the oscillators are completely opposite one another (it is the same for $-\pi$ since these angles are co-terminal).

Stable and Unstable

A logical question at this point is why did the above coupled oscillator simulation eventually synchronize rather than ending up in opposite directions?

In general, some equilibria are attractive in the sense that if the quantity x gets near such an equilibrium, it will be drawn towards it and stay at it. Some equilibria are repellent in the sense that even if x is very close to such an equilibrium, it will be pushed away from it. Formally,

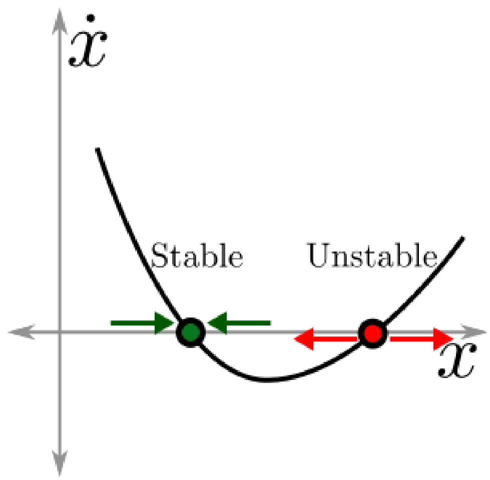
Stable and Unstable Equilibria

An equilibrium C of the differential equation

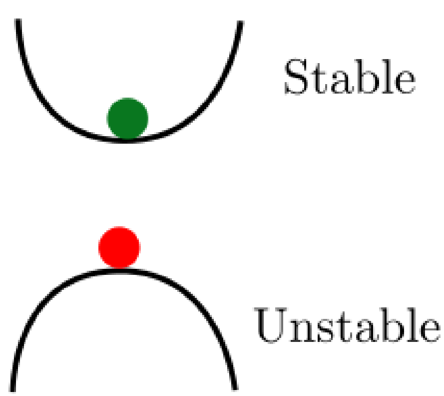
$$\frac{dx}{dt} = f(x)$$

is *stable* if $f'(C) < 0$ and is *unstable* if $f'(C) > 0$.

It is best to make sense of these definitions visually. Plot $\dot{x}$ versus x. Then each root of this equation is an equilibrium. If the graph crosses from positive to negative (going from left to right), then the equilibrium is stable. If the graph crosses from negative to positive (again, from left to right), then the equilibrium is unstable:



Another way to think of stable and unstable equilibria is to visualize one ball sitting in a bowl, and another ball sitting on top of an inverted bowl:



Each of these balls is in equilibrium (it will stay where it is as long as it is not disturbed). But the ball in the bowl is stable because if we nudge it in either direction, it will return to its equilibrium. However, the ball on the inverted bowl is unstable because if it is nudged in either direction it will roll off the bowl.

Example

Find and classify the equilibria of the equation

$$\frac{dx}{dt} = x^2 - 4x + 3.$$

(See Answer 1)

20.5 EXERCISES

- The differential equation

$$\frac{dx}{dt} = (e^x - 1)(x - 1)$$

has an equilibrium at $x = 0$. What is the linearized equation at this equilibrium? Hint: Taylor-expand the right hand side about zero.

- There is also an equilibrium at $x = 1$ in the equation above. What is the linearized equation at this equilibirum? Hint: let $h = x - 1$ be a local coordinate and compute $\dot{h} = \dot{x} = \cdots$ by Taylor expanding the right hand side at $x = 1$.

- Recall from Lecture 18, Newton's Law of Heat Transfer, which states that

$$\frac{dT}{dt} = \kappa(A - T),$$

where $\kappa > 0$ is a thermal conductivity constant and A is the (constant) ambient temperature. Find and classify the equilibria in this system (using the derivative of the right hand side at the equilibria, recall...). Wasn't that easy?

- Find and classify all the equilibria of the ODE

$$\frac{dy}{dt} = -2y + y^2 + y^3$$

- Recall from Lecture 19 how we computed the terminal velocity of a falling body with linear drag given by

$$m\frac{dv}{dt} = mg - \kappa v,$$

where, of course, m is mass, g is gravitation, v is velocity, and $\kappa > 0$ is the drag coefficient. Can you see how easily one can solve for the equilibrium $v_\infty = mg/\kappa$? Do it!

- Very good. Now, let's use a more realistic model of drag that is quadratic as opposed to linear:

$$m\frac{dv}{dt} = mg - \lambda v^2,$$

where $\lambda > 0$ is a constant drag coefficient. This differential equation is not as easy to solve (but soon you will learn how). Is there is terminal velocity? What is it?

- Recall that with continuous compounding at an interest rate of $r > 0$, an investment $I(t)$ with initial investment $I_0 = I(0)$ is $I(t) = I_0 e^{rt}$. What happens if you wish to withdraw funds from the investment at a rate of spending S, where $S > 0$ is constant? The differential equation is:

$$\frac{dI}{dt} = rI - S.$$

Your goals are as follows. You have an initial investment I_0, and you cannot change it or the rate r. You want to be able to spend as much as possible but you also don't want to ever spend all your money. What amount of spending rate S can you bear? Hint: if you're not sure what to do, find and classify the equilibria in this model and think about which initial conditions lead to which long-term behaviors.

- In our lesson, we looked at two oscillators with "sinusoidal" coupling. Other types of coupling are possible as well. Consider the system of two oscillators modeled by

$$\frac{d\theta_1}{dt} = 2 + \epsilon(e^{\theta_1 - \theta_2} - 1) \quad ; \quad \frac{d\theta_2}{dt} = 2 + \epsilon(1 - e^{\theta_1 - \theta_2})$$

Consider the phase difference $\varphi = \theta_2 - \theta_1$. Note that $\varphi = 0$ (where the oscillators are coupled) is an equilibrium. What is the linearized equation for φ about 0?

This looks intimidating, but is very straightforward. If you're not sure how to start, compute $\frac{d\varphi}{dt}$. Then linearize this about $\varphi = 0$.

20.6 Answers to Selected Examples

1. Here, $f(x) = x^2 - 4x + 3$. Factoring, one finds

$$\frac{dx}{dt} = (x-1)(x-3).$$

So the roots (and hence the equilibria) are $x = 1$ and $x = 3$. Taking the derivative, we find

$$\begin{aligned} f'(x) &= 2x - 4 \\ f'(1) &= -2 < 0 \\ f'(3) &= 2 > 0. \end{aligned}$$

Thus $x = 1$ is a stable equilibrium, and $x = 3$ is an unstable equilibrium.

(Return)

21 Integration By Substitution

The previous modules gave some of the motivation for integration as a method of solving differential equations. In this and the next few modules, we turn to techniques of integration.

21.1 Integration rules

Since integration is the inverse of differentiation, one can turn differentiation rules into integration rules. For example, by the linearity of the derivative, we have linearity of the integral:

$$\int (u+v)\,dx = \int u\,dx + \int v\,dx$$
$$\int (cu)\,dx = c\int u\,dx$$

where c is a constant. In other words, integration is a linear operator.

The rest of this module deals with turning the chain rule for differentiation into a rule for integration. This rule is called *substitution*, or *u-substitution* traditionally.

21.2 Substitution: the chain rule in reverse

Recall the chain rule, which says that if $u = u(x)$ is a function of x, then

$$du = \frac{du}{dx}dx.$$

Now if $f = f(u)$ is a function of u, then we find

$$\int f(u)\,du = \int f(u(x))\frac{du}{dx}\,dx$$

(To get from the left side to the right, all we have done is replace u and du by $u(x)$ and $\frac{du}{dx}dx$, respectively). This is the formula for substitution, or u-substitution.

Substitution is a useful technique but is not always easy to apply. In a typical problem, one encounters the right side of the above equation, but without knowing what f and u are. If one can find the correct f and u so that the integral can be expressed as above, then one can switch over to the left side of the above equation, which is usually easier to evaluate.

Example

Compute

$$\int e^{\sin x} \cos x \, dx.$$

(See Answer 1)

Example

Compute

$$\int 2xe^{x^2} \, dx.$$

(See Answer 2)

It is not always so easy to see the ideal choice of u, and sometimes it might take a few tries to find the right substitution. Usually, a good strategy is to look for the *inner function* of a composition of functions and let that be u. Another idea is to look for a function whose derivative is also a factor of the integrand.

Example

Compute

$$\int x\sqrt{x-1} \, dx$$

(See Answer 3)

Another general tip for integration by substitution is to try to simplify the integrand as much as possible before integrating.

Example

Compute

$$\int \cot\theta \csc\theta \, d\theta.$$

(See Answer 4)

Example

The Gompertz model for the size $N(t)$ of a tumor at time t is

$$\frac{dN}{dt} = -aN \ln(bN),$$

where $a > 0$ and $0 < b < 1$ are constants. Solve this differential equation. Hint: it is separable. Then find the limit behavior

$$\lim_{t\to\infty} N(t).$$

Finally, find the equilibria of the original differential equation and classify them as stable or unstable. (See Answer 5)

Example

Compute

$$\int 4(2x+5)^4\,dx.$$

(See Answer 6)

21.3 Perspective

The big idea of this module is that a change of variables (a substitution of one variable for a function of another) can change a difficult integral into an easier one. After computing the easier integral, we can change the variables back again. This idea will come up again in this course and in multivariable calculus.

21.4 Additional Examples

Example

Compute

$$\int \frac{(\ln x)^2}{x}\,dx.$$

(See Answer 7)

Example

Compute

$$\int \tan\theta\,d\theta.$$

(See Answer 8)

Example

Compute

$$\int x^5\sqrt{1+x^3}\,dx.$$

(See Answer 9)

21.5 EXERCISES

- Compute the integral $\displaystyle\int 3\cos x\,dx$
- Compute the integral $\displaystyle\int x\sec^2 x^2\,dx$
- Compute the integral $\displaystyle\int \frac{4x}{(x^2-1)^3}\,dx$

- Compute the integral $\int \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx$
- Compute the integral $\int \frac{\ln(15x^5)}{x}\,dx$
- Compute the integral $\frac{x\,dx}{\sqrt{x+3}}\,dx$
- Compute the integral $\frac{dx}{x\sqrt{x^2-1}}\,dx$ using the substitution $u = \sqrt{x^2-1}$
- Now do the same integral using the substitution $u = x^{-1}$ What is going on here?
- Compute the integral $\frac{dx}{x\sqrt{x^2+1}}\,dx$

21.6 Answers to Selected Examples

1. Let $u = \sin x$ and $f(u) = e^u$. Then
$$du = \frac{du}{dx}\,dx = \cos x\,dx,$$
and
$$\begin{aligned}\int e^{\sin x}\cos x\,dx &= \int f(\sin x)d(\sin x)\\ &= \int f(u)\,du\\ &= \int e^u\,du\\ &= e^u + C\\ &= e^{\sin x} + C.\end{aligned}$$
Note that after evaluating the integral in terms of u, we usually replace u with its function of x, since the original integral was with respect to x.
We can check that this is the correct antiderivative by differentiating (remembering to apply the chain rule) and seeing that we get back the function which we were integrating originally.
Typically, we need not write out all the details of what f is. It is sufficient to identify u and du and then make the necessary substitutions.
(Return)

2. The inner function here looks like $u = x^2$. Then $du = 2x\,dx$, which is the remaining factor in the integrand. Thus,
$$\begin{aligned}\int 2xe^{x^2}\,dx &= \int e^u\,du\\ &= e^u + C\\ &= e^{x^2} + C.\end{aligned}$$
(Return)

3. Here, the inner function seems to be $x - 1$. So let $u = x - 1$. Then $du = dx$. This takes care of what is under the square root, and the differential, but what about the extra factor of x? We can take care of this factor by noting that $u = x - 1$ implies $x = u + 1$. So the integral becomes

$$\begin{aligned}
\int x\sqrt{x-1}\,dx &= \int (u+1)\sqrt{u}\,du \\
&= \int (u+1)u^{1/2}\,du \\
&= \int u^{3/2} + u^{1/2}\,du \\
&= \frac{2}{5}u^{5/2} + \frac{2}{3}u^{3/2} + C \\
&= \frac{2}{5}(x-1)^{5/2} + \frac{2}{3}(x-1)^{3/2} + C.
\end{aligned}$$

From the third to the fourth line above, we used the power rule on each term of the integrand. Again, we can differentiate to make sure we get back to our original integrand (though it might require a little algebra to show that they are in fact equal).

(Return)

4. Since cotangent and cosecant are not very familiar functions, it is helpful to rewrite them in terms of sine and cosine. This gives

$$\begin{aligned}
\int \cot\theta\csc\theta\,d\theta &= \int \frac{\cos\theta}{\sin\theta}\cdot\frac{1}{\sin\theta}\,d\theta \\
&= \int \frac{\cos\theta}{\sin^2\theta}\,d\theta.
\end{aligned}$$

It is easier now to see that $u = \sin\theta$ is a good choice, since its derivative $du = \cos\theta\,d\theta$ is in the numerator. Making this substitution shows

$$\begin{aligned}
\int \frac{\cos\theta}{\sin^2\theta}\,d\theta &= \int \frac{1}{u^2}\,du \\
&= \int u^{-2}\,du \\
&= -u^{-1} + C \\
&= -\frac{1}{\sin\theta} + C \\
&= -\csc\theta + C.
\end{aligned}$$

Of course, if one happened to remember the fact that

$$\frac{d}{d\theta}\csc\theta = -\csc\theta\cot\theta,$$

then we would not require a substitution. But substitution allows us to do these integrals (and harder ones) without needing to memorize a lot of information.

(Return)

5. Separating variables and integrating both sides gives

$$\int \frac{dN}{N\ln(bN)} = \int -a\,dt.$$

The right side is easy, but the left side requires some work. Looking for a function whose derivative is a factor in the integrand, we see that $u = \ln(bN)$ is a good choice. In this case

$$du = \frac{1}{bN} \cdot b\,dN = \frac{dN}{N}.$$

And so, the integral on the left above becomes

$$\begin{aligned}\int \frac{dN}{N\ln(bN)} &= \int \frac{1}{u}\,du \\ &= \ln u + C \\ &= \ln\ln(bN) + C.\end{aligned}$$

Putting this together with the integral on the right above (and combining the integration constants to one integration constant on the right), we have

$$\ln\ln(bN) = -at + C$$

Exponentiating twice and then dividing by b gives

$$N = \frac{1}{b}e^{e^{-at+C}} = \frac{1}{b}e^{Ce^{-at}}.$$

By plugging in $t = 0$, we find that $C = \ln(bN_0)$, where N_0 is the initial size of the tumor.

In the long run, the exponential $e^{-at} \to 0$, since $a > 0$ by assumption. Therefore, the entire exponent is going to 0, and so

$$\lim_{t\to\infty} N(t) = \frac{1}{b}.$$

Note that $N(t) = \frac{1}{b}$ is an equilibrium solution to the original differential equation, since it gives $\frac{dN}{dt} = 0$. It is a stable equilibrium since the graph of of $-aN\ln(bN)$ goes from positive to negative as N goes from less than $\frac{1}{b}$ to greater than $\frac{1}{b}$.

Another equilibrium is $N = 0$. This is unstable, since $N > 0$ means $\frac{dN}{dt} > 0$. Intuitively, even if the tumor is very tiny, it will grow according to this model.

(Return)

6. In this case, the inner function is $u = 2x + 5$, and one finds that $du = 2dx$. Thus $dx = \frac{du}{2}$, which gives

$$\begin{aligned}\int 4(2x+5)^4 dx &= \int 4u^4 \frac{du}{2} \\ &= \int 2u^4 du \\ &= \frac{2}{5}u^5 + C \\ &= \frac{2}{5}(2x+5)^5 + C.\end{aligned}$$

(Return)

7. Here, a good inner function is $u = \ln x$, because the derivative $du = \frac{1}{x}\,dx$. Thus

$$\begin{aligned}\int \frac{(\ln x)^2}{x}\,dx &= \int u^2\,du \\ &= \frac{u^3}{3} + C \\ &= \frac{(\ln x)^3}{3} + C.\end{aligned}$$

(Return)

8. First, rewrite tangent in terms of sine and cosine:

$$\int \tan\theta\,d\theta = \int \frac{\sin\theta}{\cos\theta}\,d\theta$$

Now, note that $u = \sin\theta$ would not work, because its derivative, $\cos\theta$ is in the denominator. On the other hand, $u = \cos\theta$ is a good substitution because its derivative (up to a constant) is in in the numerator. That is, $du = -\sin\theta\,d\theta$. Therefore,

$$\begin{aligned}\int \frac{\sin\theta}{\cos\theta}\,d\theta &= \int -\frac{1}{u}\,du \\ &= -\ln(u) + C \\ &= -\ln(\cos\theta) + C.\end{aligned}$$

(Return)

9. The logical choice of inner function is $u = 1 + x^3$, which gives $du = 3x^2\,dx$ and so

$$dx = \frac{du}{3x^2}$$

Substituting in, we find

$$\begin{aligned}\int x^5\sqrt{1+x^3}\,dx &= \int x^5\sqrt{u}\frac{du}{3x^2} \\ &= \frac{1}{3}\int x^3\sqrt{u}\,du.\end{aligned}$$

This seems problematic, because we haven't been able to get everything in terms of u. But we can use our original substitution to help. Since $u = 1 + x^3$, we have that $x^3 = u - 1$, and so

$$\begin{aligned}\frac{1}{3}\int x^3\sqrt{u}\,du &= \frac{1}{3}\int (u-1)\sqrt{u}\,du \\ &= \frac{1}{3}\int u^{3/2} - u^{1/2}\,du \\ &= \frac{1}{3}\left(\frac{2}{5}u^{5/2} - \frac{2}{3}u^{3/2}\right) + C \\ &= \frac{2}{15}(1+x^3)^{5/2} - \frac{2}{9}(1+x^3)^{3/2} + C.\end{aligned}$$

(Return)

22 Integration By Parts

This module uses the product rule to derive another useful integration technique: integration by parts. Recall the product rule:

$$d(u \cdot v) = u \cdot dv + v \cdot du.$$

Integrating both sides gives

$$\int d(u \cdot v) = \int u\,dv + \int v\,du.$$

Solving for $\int u\,dv$ gives

Integration by parts

If $u = u(x)$ and $v = v(x)$ are two functions of x, then

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

Intuitively, we are given a difficult integral $\int u\,dv$. By breaking the integrand into u and dv and applying the above formula, we are hopefully able to wind up with an easier integral $\int v\,du$. Like with the substitution technique, it requires a little bit of thought to choose suitable u and dv. Once u and dv are picked, it is a fairly mechanical process to apply the formula (assuming a good choice of u and dv).

Note that the selection is constrained by the fact that $u\,dv$ must be the entire integrand. So whatever choice is made for u, whatever factors are left over become dv. Note also that the formula involves finding v, and so dv must be integrable. Ideally, dv should be easy to integrate, which can help guide the selection.

Example

Compute

$$\int xe^x\,dx.$$

(See Answer 1)

Example

Compute

$$\int \ln(x)dx.$$

(See Answer 2)

Example

Try to compute

$$\int \frac{\sin x}{x}\, dx$$

Hint: it cannot be done using integration by parts. (See Answer 3)

22.1 LIPET: A tip for choosing u and dv

It is not always obvious how to choose u and dv. The mnemonic LIPET gives a suggestion for how to select u, and then whatever is left over becomes dv.

1. **L**ogarithm
2. **I**nverse function
3. **P**olynomial
4. **E**xponential
5. **T**rigonometric.

When picking u, go down the list until some factor of the integrand first matches something from the list. So in the first example above, there was no logarithm, no inverse function, but there was a polynomial, x, which was chosen for u. In the second example, there was a logarithm, so that became u.

This will not always work perfectly, because (as the above example showed) some integrals simply cannot be computed using integration by parts. But in most examples where integration by parts works, the above mnemonic will help give the correct selection of u and dv.

Example

Compute

$$\int \frac{\ln x}{x^2}\, dx.$$

(See Answer 4)

22.2 Repeated use

Sometimes integration by parts requires repeated use, if the integral $\int v\, du$ is not easy to compute. It is not always easy to tell when repeating integration by parts will help, but with practice it becomes easier.

Example

Compute

$$\int e^x \cos(x) dx.$$

(See Answer 5)

Example

Compute

$$\int e^{2x} \sin(3x)\, dx.$$

(See Answer 6)

There are integrals that require several applications of integration by parts before they are finished. Unfortunately, it is not always clear when it will work and when it will not. Doing a lot of practice can help develop the intuition to tell the difference.

As the next example shows, sometimes an integral that looks like a perfect candidate for integration by parts does not yield to this method.

Example

Compute

$$\int e^{x} \cosh x\, dx.$$

(See Answer 7)

22.3 Reduction formulae

A final application of integration by parts is to prove what are known as *reduction formulae*. These formulae express one integral in terms of another slightly simpler integral. One can use a reduction formula to repeatedly simplify an integral, eventually reaching a known integral. These formulae are invariably derived by using integration by parts and some algebra.

Example

For a fixed integer $n \geq 0$, show that

$$\int x^n \cos x\, dx = x^n \sin x + nx^{n-1} \cos x - n(n-1) \int x^{n-2} \cos x\, dx.$$

Use this formula to find

$$\int x^2 \cos x dx.$$

(See Answer 8)

Example

Similar algebra as in the above example shows that for $n \geq 0$

$$\int x^n \sin x\, dx = -x^n \cos x + nx^{n-1} \sin x - n(n-1) \int x^{n-2} \sin x\, dx.$$

Example

Find a reduction formula for

$$\int x^n e^x \, dx.$$

Use it to evaluate

$$\int x^2 e^x \, dx.$$

(See Answer 9)

Example

Show that for $n \geq 2$,

$$\int \sec^n(x) \, dx = \frac{1}{n-1} \sec^{n-2}(x) \tan(x) + \frac{n-2}{n-1} \int \sec^{n-2}(x) \, dx.$$

(See Answer 10)

22.4 Additional examples

Example

Compute

$$\int x \sin(x) \, dx.$$

(See Answer 11)

Example

Compute

$$\int \arctan x \, dx.$$

Hint: recall that

$$\frac{d}{dx} \arctan x = \frac{1}{1+x^2}.$$

(See Answer 12)

22.5 EXERCISES

Compute the following integrals:

- $\displaystyle\int x e^{x/2} \, dx$
- $\displaystyle\int x^2 e^{x/2} \, dx$

- $\displaystyle\int 3x \ln x \, dx$
- $\displaystyle\int 3x^2 \ln x \, dx$
- $\displaystyle\int x^2 \cos \frac{x}{2} \, dx$
- $\displaystyle\int e^{2x} \sin 3x \, dx$
- $\displaystyle\int \ln x \, dx$
- $\displaystyle\int \ln^2 x \, dx$
- $\displaystyle\int \sin(\ln x) \, dx$
- $\displaystyle\int \arcsin(2x) \, dx$

To solve the integral $\displaystyle\int e^x \cos x \, dx$, we used the method of integration by parts twice. Based on how we solved the integral of $e^x \cosh x$, we can try the same with the cosine version, using the fact that $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$. Try! The integration is the easy part...the hard part is getting the algebra to work out (hello again, Euler's formula...)

- Compute $\displaystyle\int \sin(2x) \cos(3x) \, dx$

22.6 Answers to Selected Examples

1. Letting $u = x$ and $dv = e^x dx$ (we see that these factors together make up our integrand), one finds by differentiating u and integrating dv that $du = dx$ and $v = e^x$. Many students find it helps to organize this information in a grid:

$$\begin{aligned} u &= x & du &= dx \\ dv &= e^x \, dx & v &= e^x. \end{aligned}$$

   Then, from the formula it follows that

$$\begin{aligned} \int xe^x dx &= xe^x - \int e^x dx \\ &= xe^x - e^x + C \\ &= (x-1)e^x + C. \end{aligned}$$

   One can check that the derivative of this gives back the original integrand, as desired.

   (Return)

2. The selection of u and dv that works is

$$\begin{aligned} u &= \ln(x) & du &= \frac{1}{x}\,dx \\ dv &= dx & v &= x. \end{aligned}$$

Note that the only other possibility would be $dv = \ln(x)dx$. But this choice would mean that to find v we would need to find the integral of $\ln(x)$, which is the problem at hand.

Applying the formula, we find

$$\begin{aligned} \int \ln(x)dx &= x\ln(x) - \int x \cdot \frac{1}{x}\,dx \\ &= x\ln(x) - \int dx \\ &= x\ln(x) - x + C \\ &= x(\ln(x) - 1) + C. \end{aligned}$$

Again, we can check that the derivative of this function gives back $\ln x$, our original integrand.

(Return)

3. One choice we might try is

$$\begin{aligned} u &= \sin x & du &= \cos x\,dx \\ dv &= \frac{1}{x}\,dx & v &= \ln x. \end{aligned}$$

However, this requires us to compute the integral

$$\int v\,du = \int \ln x \cos x\,dx,$$

which is no better than the original integral. Another possible choice is

$$\begin{aligned} u &= \frac{1}{x} & du &= -\frac{1}{x^2}\,dx \\ dv &= \sin x\,dx & v &= -\cos x. \end{aligned}$$

This leads to the integral

$$\int v\,du = \int \frac{\cos x}{x^2}\,dx,$$

which is again no better than the original integral.

It turns out no selection will work. Some integrals cannot be computed using integration by parts. And some integrals (like this one) have no elementary answer (i.e. some combination of trigonometric functions, polynomials, exponentials, etc.).

That said, we could expand $\sin(x)$ as a Taylor series, divide by x and integrate term by term, which gives a series solution. This gives a perfectly suitable solution provided that x is not too far from 0.

(Return)

4. A good choice is

$$u = \ln x \qquad du = \frac{1}{x}\,dx$$
$$dv = \frac{1}{x^2}\,dx \qquad v = -\frac{1}{x}.$$

Then

$$\begin{aligned}\int \frac{\ln x}{x^2}\,dx &= -\frac{1}{x}\ln x - \int -\frac{1}{x^2}\,dx \\ &= -\frac{1}{x}\ln x - \frac{1}{x} + C.\end{aligned}$$

(Return)

5. We can take

$$u = e^x \qquad du = e^x\,dx$$
$$dv = \cos x\,dx \qquad v = \sin x.$$

(it turns out that this would work equally well if we reversed these). Then the formula says that

$$\int e^x \cos(x) dx = e^x \sin(x) - \int \sin(x) e^x dx.$$

This does not seem much better than the original problem. However, with some persistence and algebra, this will work. Let $I = \int e^x \cos(x) dx$ be the original integral, and let $J = \int \sin(x) e^x dx$ be the new integral. So the above calculation shows

$$I = e^x \sin(x) - J$$

Using integration by parts on J, we pick

$$u = e^x \qquad du = e^x\,dx$$
$$dv = \sin x\,dx \qquad v = -\cos x.$$

Then it follows that

$$\begin{aligned}J &= \int \sin(x) e^x dx \\ &= e^x(-\cos(x)) - \int (-\cos(x)) e^x dx \\ &= -e^x \cos(x) + \int e^x \cos(x) dx \\ &= -e^x \cos(x) + I.\end{aligned}$$

So the problem has come back to the original integral I. This might seem like cause for despair, but putting together the previous calculations shows

$$\begin{aligned}I &= e^x \sin(x) - J \\ &= e^x \sin(x) - (-e^x \cos(x) + I) \\ &= e^x \sin(x) + e^x \cos(x) - I.\end{aligned}$$

Now, solving for I gives

$$\begin{aligned} 2I &= e^x \sin(x) + e^x \cos x + C \\ I &= \frac{1}{2}(e^x \sin(x) + e^x \cos(x)) + C. \end{aligned}$$

(Return)

6. The algebra is similar to the above example, but care must be taken with the constants. Let

$$I = \int e^{2x} \sin(3x)\, dx.$$

Let

$$\begin{aligned} u &= e^{2x} & du &= 2e^{2x}\, dx \\ dv &= \sin(3x)\, dx & v &= -\frac{1}{3}\cos(3x). \end{aligned}$$

Then

$$\begin{aligned} I &= -\frac{1}{3}e^{2x}\cos(3x) - \int \left(-\frac{1}{3}\cos(3x)\right) 2e^{2x}\, dx \\ &= -\frac{1}{3}e^{2x}\cos(3x) + \frac{2}{3}\int e^{2x}\cos(3x)\, dx. \end{aligned}$$

Now, let

$$J = \int e^{2x} \cos(3x)\, dx.$$

Selecting

$$\begin{aligned} u &= e^{2x} & du &= 2e^{2x}\, dx \\ dv &= \cos(3x)\, dx & v &= \frac{1}{3}\sin(3x), \end{aligned}$$

we have

$$\begin{aligned} J &= \frac{1}{3}e^{2x}\sin(3x) - \frac{2}{3}\int e^{2x}\sin(3x)\, dx \\ &= \frac{1}{3}e^{2x}\sin(3x) - \frac{2}{3}I. \end{aligned}$$

Putting this all together, we have

$$\begin{aligned} I &= -\frac{1}{3}e^{2x}\cos(3x) + \frac{2}{3}J \\ &= -\frac{1}{3}e^{2x}\cos(3x) + \frac{2}{3}\left(\frac{1}{3}e^{2x}\sin(3x) - \frac{2}{3}I\right) \\ &= -\frac{1}{3}e^{2x}\cos(3x) + \frac{2}{9}e^{2x}\sin(3x) - \frac{4}{9}I. \end{aligned}$$

Solving for I gives

$$\begin{aligned}
\frac{13}{9} I &= -\frac{1}{3} e^{2x} \cos(3x) + \frac{2}{9} e^{2x} \sin(3x) \\
I &= \frac{9}{13} \left(-\frac{1}{3} e^{2x} \cos(3x) + \frac{2}{9} e^{2x} \sin(3x) \right) \\
&= \frac{1}{13} e^{2x} \left(-3\cos(3x) + 2\sin(3x) \right).
\end{aligned}$$

Remember that any indefinite integral has an integration constant, so the final answer is

$$\int e^{2x} \sin(3x)\, dx = \frac{1}{13} e^{2x} \left(-3\cos(3x) + 2\sin(3x) \right) + C.$$

(Return)

7. This looks so similar to the above examples, that it is reasonable to expect that two applications of integration by parts will allow us to algebraically find this integral. Unfortunately, there is a problem that will soon present itself. Let

$$I = \int e^x \cosh x\, dx.$$

If we set

$$\begin{aligned}
u &= e^x & du &= e^x\, dx \\
dv &= \cosh x\, dx & v &= \sinh x,
\end{aligned}$$

we find

$$\begin{aligned}
I &= e^x \sinh x - \int e^x \sinh x \\
&= e^x \sinh x - J,
\end{aligned}$$

where we have set

$$J = \int e^x \sinh x\, dx.$$

Letting

$$\begin{aligned}
u &= e^x & du &= e^x\, dx \\
dv &= \sinh x\, dx & v &= \cosh x,
\end{aligned}$$

we find

$$\begin{aligned}
J &= e^x \cosh x - \int e^x \cosh x\, dx \\
&= e^x \cosh x - I.
\end{aligned}$$

Putting it all together,

$$\begin{aligned}
I &= e^x \sinh x - J \\
&= e^x \sinh x - (e^x \cosh x - I) \\
&= e^x \sinh x - e^x \cosh x + I.
\end{aligned}$$

Here is where our problem arises. We cannot solve for I because there is a positive I on both sides. This problem is due to the fact that (unlike sine and cosine), the hyperbolic sine and cosine do not introduce negative signs when integrated or differentiated, respectively.

So what do we do? Rewrite our integral using the definition of $\cosh x$ and it becomes easy:

$$\begin{aligned}\int e^x \cosh x\,dx &= \int e^x \cdot \left(\frac{e^x + e^{-x}}{2}\right) \\ &= \frac{1}{2}\int \left(e^{2x} + 1\right)\,dx \\ &= \frac{1}{2}\left(\frac{1}{2}e^{2x} + x\right) + C \\ &= \frac{1}{4}e^{2x} + \frac{1}{2}x + C.\end{aligned}$$

(Return)

8. Let

$$\begin{aligned} u &= x^n & du &= nx^{n-1}\,dx \\ dv &= \cos x\,dx & v &= \sin x. \end{aligned}$$

Then according to the formula,

$$\int x^n \cos x\,dx = x^n \sin x - \int nx^{n-1} \sin x\,dx.$$

Now, since we want to get our integral in terms of an integral involving $\cos x$ and a power of x, we can apply integration by parts to

$$\int x^n \sin x\,dx.$$

Here, we let

$$\begin{aligned} u &= x^n & du &= nx^{n-1}\,dx \\ dv &= \sin x\,dx & v &= -\cos x. \end{aligned}$$

This gives

$$\int x^n \sin x\,dx = -x^n \cos x + \int nx^{n-1} \cos x\,dx.$$

Now, using this in our earlier equation (though with n replaced by $n-1$), we find

$$\begin{aligned}\int x^n \cos x\,dx &= x^n \sin x - \int nx^{n-1} \sin x\,dx \\ &= x^n \sin x - n\int x^{n-1} \sin x\,dx \\ &= x^n \sin x - n\left(-x^{n-1}\cos x + \int (n-1)x^{n-2}\cos x\,dx\right) \\ &= x^n \sin x + nx^{n-1}\cos x - n(n-1)\int x^{n-2}\cos x\,dx.\end{aligned}$$

The formula says that

$$\begin{aligned}\int x^2 \cos x\, dx &= x^2 \sin x + 2x \cos x - 2 \int \cos x\, dx \\ &= x^2 \sin x + 2x \cos x - 2 \sin x + C \\ &= x^2 \sin x + 2x \cos x - 2 \sin x + C.\end{aligned}$$

(Return)

9. Letting

$$\begin{aligned} u &= x^n & du &= nx^{n-1}\, dx \\ dv &= e^x\, dx & v &= e^x, \end{aligned}$$

we find that

$$\int x^n e^x = x^n e^x - n \int x^{n-1} e^x\, dx.$$

Applying this when $n = 2$ (then applying it again) gives

$$\begin{aligned}\int x^2 e^x &= x^2 e^x - 2 \int x e^x\, dx \\ &= x^2 e^x - 2 \left(x e^x - \int e^x\, dx \right) \\ &= x^2 e^x - 2x e^x + 2e^x + C.\end{aligned}$$

(Return)

10. In integrating a power of a trigonometric function, it can be hard to pick how many factors become u and how many become dv. The fact that dv is supposed to be easy to integrate can guide this selection. Since $\int \sec^2 x dx = \tan x$, letting $dv = \sec^2 x dx$ should work well.

Thus, $u = \sec^{n-2}(x)$ and $dv = \sec^2 x dx$, which means $du = (n-2)\sec^{n-3}(x)\sec(x)\tan(x)dx$ (by the chain rule), and $v = \tan x$. Recalling the Pythagorean identity $\tan^2 x = \sec^2 x - 1$, one finds that

$$\begin{aligned}\int \sec^n(x)dx &= \sec^{n-2}(x)\tan(x) - \int (n-2)\sec^{n-2}(x)\tan^2(x)dx \\ &= \sec^{n-2}(x)\tan(x) - (n-2)\int \sec^{n-2}(x)(\sec^2(x) - 1)dx \\ &= \sec^{n-2}(x)\tan(x) - (n-2)\int (\sec^n(x) - \sec^{n-2}(x))dx \\ &= \sec^{n-2}(x)\tan(x) - (n-2)\int \sec^n(x)dx + (n-2)\int \sec^{n-2}(x)dx\end{aligned}$$

Now, solving for $\int \sec^n(x)dx$ gives

$$\int \sec^n(x)dx = \frac{1}{n-1}\sec^{n-2}(x)\tan(x) + \frac{n-2}{n-1}\int \sec^{n-2}(x)dx,$$

as desired.

(Return)

11. The logical choice (either by the LIPET mnemonic, or by picking u to be something which gets simpler when differentiated) for parts is

$$\begin{aligned} u &= x & du &= dx \\ dv &= \sin(x)\,dx & v &= -\cos(x) \end{aligned}$$

Therefore,

$$\begin{aligned} \int x\sin(x)\,dx &= -x\cos(x) - \int -\cos(x)\,dx \\ &= -x\cos(x) + \int \cos(x)\,dx \\ &= -x\cos(x) + \sin(x) + C. \end{aligned}$$

(Return)

12. As in some earlier examples, the only choice we have is to set

$$\begin{aligned} u &= \arctan x & du &= \frac{1}{1+x^2}\,dx \\ dv &= dx & v &= x. \end{aligned}$$

Therefore,

$$\int \arctan x\,dx = x\arctan x - \int \frac{x}{1+x^2}\,dx.$$

This second integral can be solved with a substitution of

$$\begin{aligned} u &= 1+x^2 \\ du &= 2x\,dx. \end{aligned}$$

So $dx = \frac{du}{2x}$. Making the substitution gives

$$\begin{aligned} \int \frac{x}{1+x^2}\,dx &= \int \frac{x}{u}\cdot\frac{du}{2x} \\ &= \frac{1}{2}\int\frac{du}{u} \\ &= \frac{1}{2}\ln u + C \\ &= \frac{1}{2}\ln(1+x^2) + C. \end{aligned}$$

Putting it all together, we find

$$\begin{aligned} \int \arctan x\,dx &= x\arctan x - \int \frac{x}{1+x^2}\,dx \\ &= x\arctan x - \frac{1}{2}\ln(1+x^2) + C. \end{aligned}$$

(Return)

23 Trigonometric Substitution

There is another class of integrals which usually do not involve trigonometric functions, but which can be solved by substituting the variable with a trigonometric function. This can be thought of as using the substitution formula, from Integration By Substitution, in the other direction. That is, going from the left side to right side in the equality

$$\int f(x)\,dx = \int f(x(\theta))\frac{dx}{d\theta}\,d\theta,$$

where we have made the substitution $x = x(\theta)$. We often use θ when making a trigonometric substitution.

Example

Compute

$$\int \frac{dx}{1+x^2}.$$

(See Answer 1)

23.1 Typical substitutions

Trigonometric substitution makes use of the Pythagorean identities. In general, the basic trigonometric substitutions are:

Form	Substitution	Identity used
$1+x^2$	$x = \tan\theta$	$1+\tan^2\theta = \sec^2\theta$
$1-x^2$	$x = \sin\theta$	$1-\sin^2\theta = \cos^2\theta$
x^2-1	$x = \sec\theta$	$\sec^2\theta - 1 = \tan^2\theta$

Caveat

The form x^2-1 often leads to a messy integral involving $\sec(\theta)$. This can often be avoided using a hyperbolic trigonometric substitution (see below).

After a substitution has been made, the resulting integral will often involve a product of trigonometric functions, possibly raised to powers. These types of integrals are covered in more detail in Trigonometric Integrals. For now, here are a few of the useful identities in evaluating these integrals:

Power reduction
$\sin^2(\theta) = \frac{1-\cos(2\theta)}{2}$
$\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}$

Double angle
$\sin(2\theta) = 2\sin\theta\cos\theta$
$\cos(2\theta) = \cos^2\theta - \sin^2\theta$

Example

Compute

$$\int \sqrt{1-x^2}\,dx.$$

(See Answer 2)

Example

Compute

$$\int \frac{dx}{\sqrt{1-x^2}}.$$

(See Answer 3)

23.2 Forms with other constants

There are other forms which are similar to the above forms but have different constants involved. These are dealt with using similar substitutions which make the constants cancel and factor so that the same identities can be used.

Example

Compute

$$\int \frac{dx}{x^2\sqrt{x^2+4}}$$

(See Answer 4)

Example

Compute

$$\int \frac{dx}{4+9x^2}.$$

(See Answer 5)

The following table summarizes the substitutions to be made when other constants are involved. The identities used are the same Pythagorean identities given in the above table.

Form	Substitution
$a^2x^2+b^2$	$x = \frac{b}{a}\tan\theta$
$b^2-a^2x^2$	$x = \frac{b}{a}\sin\theta$
$a^2x^2-b^2$	$x = \frac{b}{a}\sec\theta$

23.3 Completing the square

Sometimes it is not obvious at first that an integral is of the form where a trigonometric substitution is helpful. It may take a little bit of algebra to see what the right substitution is. This common algebraic tool is known as *completing the square*, which simply rewrites a quadratic expression as the square of a binomial plus a constant. To review the algebra involved in this process, check Wikipedia:Completing the square.

Example

Compute

$$\int \frac{dx}{\sqrt{3+2x-x^2}}$$

(See Answer 6)

23.4 Hyperbolic trigonometric substitutions

Recall that the hyperbolic trigonometric functions $\sinh(x)$ and $\cosh(x)$ are defined by

$$\sinh(\theta) = \frac{e^{\theta} - e^{-\theta}}{2}$$
$$\cosh(\theta) = \frac{e^{\theta} + e^{-\theta}}{2}.$$

These functions satisfy the Pythagorean identity $\cosh^2(\theta) - \sinh^2(\theta) = 1$. Also, note that $\frac{d}{d\theta}\cosh(\theta) = \sinh(\theta)$, and $\frac{d}{d\theta}\sinh(\theta) = \cosh(\theta)$. This means hyperbolic substitutions are another option for dealing with the following forms:

Form	Substitution	Identity used
$1+x^2$	$x = \sinh\theta$	$1+\sinh^2\theta = \cosh^2\theta$
x^2-1	$x = \cosh\theta$	$\cosh^2\theta - 1 = \sinh^2\theta$

This often gives a simpler answer than the $x = \sec\theta$ substitution suggested above, but the trade-off is that the answer will involve hyperbolic functions. Here are some of the other identities for the hyperbolic functions, which are similar to those for regular trigonometric functions:

Double Angle
$\sinh(2\theta) = 2\sinh(\theta)\cosh(\theta)$
$\cosh(2\theta) = \cosh^2(\theta) + \sinh^2(\theta)$
$\cosh(2\theta) = 2\cosh^2(\theta) - 1$
$\cosh(2\theta) = 2\sinh^2(\theta) + 1$

Power reduction
$\sinh^2(\theta) = \frac{\cosh(2\theta)-1}{2}$
$\cosh^2(\theta) = \frac{\cosh(2\theta)+1}{2}$

Example

Compute

$$\int \frac{dx}{\sqrt{1+x^2}}.$$

(See Answer 7)

Example

Compute

$$\int \sqrt{1+x^2}\,dx.$$

(See Answer 8)

23.5 Blow-ups

Sometimes a differential equation can be solved by using a trigonometric substitution. But this can sometimes lead to an unreasonable solution due to blow-ups or singularities, which exist for many trigonometric functions.

Example

Consider a financial model which predicts that marginal profits equal some positive constant plus something which is proportional to the square of net profits. Mathematically,

$$\frac{dP}{dt} = b^2 + a^2P^2,$$

for constants a and b (we square them to ensure that they are positive). Solve this differential equation and find where it has a blow-up. (See Answer 9)

23.6 EXERCISES

Compute the following integrals:

- $\displaystyle\int \frac{x^2}{\sqrt{4-x^2}}\,dx$
- $\displaystyle\int \frac{dx}{\sqrt{x^2-2x}}$
- $\displaystyle\int \frac{\sqrt{1-x^2}}{x^2}\,dx$
- $\displaystyle\int (1-x^2)^{-3/2}\,dx$
- $\displaystyle\int \frac{x}{\sqrt{1+x^2}}\,dx$
- $\displaystyle\int \frac{dx}{x\sqrt{x^2-1}}$

- $\displaystyle\int \frac{dx}{\sqrt{x^2 - 6x + 10}}$
- $\displaystyle\int \frac{dx}{\sqrt{x^2 - 2x - 8}}$

23.7 Answers to Selected Examples

1. Consider the substitution $x = \tan(\theta)$. Then one finds that $dx = \sec^2\theta\, d\theta$. Making these substitutions and recalling the Pythagorean identity $1 + \tan^2\theta = \sec^2\theta$, the integral becomes

$$\begin{aligned}\int \frac{dx}{1+x^2} &= \int \frac{\sec^2\theta\, d\theta}{1+\tan^2\theta} \\ &= \int \frac{\sec^2\theta\, d\theta}{\sec^2\theta} \\ &= \int d\theta \\ &= \theta + C \\ &= \arctan(x) + C.\end{aligned}$$

The last line comes from our original substitution:

$$x = \tan\theta \quad \Leftrightarrow \quad \arctan x = \theta.$$

(Return)

2. According to the above guide, the substitution to make is $x = \sin\theta$. Then $dx = \cos\theta\, d\theta$, and it follows that

$$\begin{aligned}\int \sqrt{1-x^2}\, dx &= \int \sqrt{1-\sin^2\theta}\cos\theta\, d\theta \\ &= \int \sqrt{\cos^2\theta}\cos\theta\, d\theta \\ &= \int \cos^2\theta\, d\theta.\end{aligned}$$

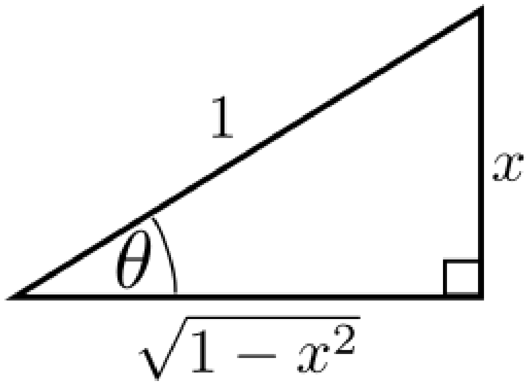
Now using the power reduction identity for cosine, we have

$$\begin{aligned}\int \cos^2\theta\, d\theta &= \int \frac{1}{2}(1+\cos(2\theta))\, d\theta \\ &= \frac{\theta}{2} + \frac{1}{4}\sin(2\theta) + C.\end{aligned}$$

Finally, we must get this back in terms of x. We know that $\theta = \arcsin x$. But to take care of $\sin 2\theta$, we must use the double angle formula from above. This gives

$$\begin{aligned}\sin 2\theta &= 2\sin\theta\cos\theta \\ &= 2x\sqrt{1-x^2}\end{aligned}$$

In the last line above, we knew $\sin\theta = x$ from the original substitution. We found $\cos\theta$ by drawing a right triangle which relates x and θ according to the substitution $\sin\theta = x$:



Putting this all together and doing a little simplification, we find

$$\int \sqrt{1-x^2} = \frac{1}{2}\arcsin x + \frac{1}{2}x\sqrt{1-x^2} + C.$$

(Return)

3. By the above table, the substitution $x = \sin\theta$ should be used (hence $\theta = \arcsin(x)$). Then $dx = \cos\theta\, d\theta$, so the integral becomes

$$\begin{aligned}\int \frac{dx}{\sqrt{1-x^2}} &= \int \frac{\cos\theta\, d\theta}{\sqrt{1-\sin^2\theta}}\\ &= \int \frac{\cos\theta\, d\theta}{\sqrt{\cos^2\theta}}\\ &= \int d\theta\\ &= \theta + C\\ &= \arcsin(x) + C.\end{aligned}$$

(Return)

4. The form $x^2 + 4$ in the denominator reminds us of the substitution we made earlier for $x^2 + 1$, which was the substitution $x = \tan\theta$. This is the correct impulse, but unfortunately it does not work quite right here since there is no nice simplification for $\tan^2\theta + 4$.

We can fix this by adjusting the coefficients. The idea is that we could factor out a 4 if we had

$$4\tan^2\theta + 4 = 4(\tan^2\theta + 1).$$

To get that extra factor of 4, we can make the substitution $x = 2\tan\theta$. Then $dx = 2\sec^2\theta\, d\theta$, and the integral becomes

$$\begin{aligned}\int \frac{dx}{x^2\sqrt{x^2+4}} &= \int \frac{2\sec^2\theta\, d\theta}{4\tan^2\theta\sqrt{4\tan^2\theta+4}}\\ &= \int \frac{2\sec^2\theta\, d\theta}{4\tan^2\theta\sqrt{4(\tan^2\theta+1)}}\\ &= \frac{1}{4}\int \frac{\sec^2\theta\, d\theta}{\tan^2\theta\sec\theta}.\end{aligned}$$

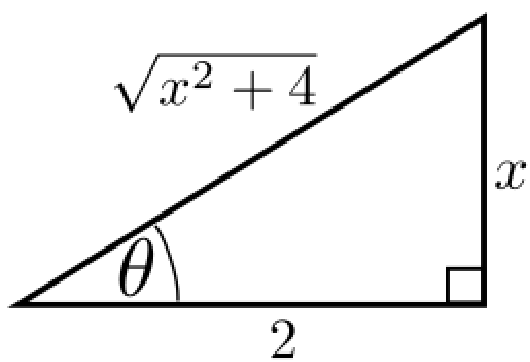
The last equality above comes from again using the identity $\tan^2\theta + 1 = \sec^2\theta$. Doing a little simplification and rewriting in terms of sine and cosine gives

$$\begin{aligned}\frac{1}{4}\int \frac{\sec\theta\, d\theta}{\tan^2\theta} &= \frac{1}{4}\int \frac{1}{\cos\theta}\cdot\frac{\cos^2\theta}{\sin^2\theta}\, d\theta \\ &= \frac{1}{4}\int \frac{\cos\theta\, d\theta}{\sin^2\theta}.\end{aligned}$$

This we can handle with a substitution of $u = \sin\theta$ and $du = \cos\theta\, d\theta$, which gives

$$\begin{aligned}\frac{1}{4}\int \frac{\cos\theta\, d\theta}{\sin^2\theta} &= \frac{1}{4}\int \frac{du}{u^2} \\ &= \frac{1}{4}\left(-\frac{1}{u}\right) + C \\ &= -\frac{1}{4u} + C \\ &= -\frac{1}{4\sin\theta} + C\end{aligned}$$

Now, we must do one final bit of right triangle trigonometry to get $\sin\theta$ in terms of x. By the original substitution we have $\tan\theta = \frac{x}{2}$, and this can be expressed by the following triangle:



It follows that $\sin\theta = \frac{x}{\sqrt{x^2+4}}$. Putting it all together, we have

$$\begin{aligned}\int \frac{dx}{x^2\sqrt{x^2+4}} &= -\frac{1}{4\sin\theta} + C \\ &= -\frac{\sqrt{x^2+4}}{4x} + C.\end{aligned}$$

(Return)

5. This is another example which looks like $x = \tan\theta$ is the right type of substitution to make. However, again we need to adjust the coefficient since $4 + 9\tan^2\theta$ does not simplify nicely.

The key is to get the constants to cancel and factor. The substitution $x = \frac{2}{3}\tan\theta$ will work, and in this case $\theta = \arctan(\frac{3}{2}x)$. Then $dx = \frac{2}{3}\sec^2\theta d\theta$, and the integral becomes

$$\begin{aligned}
\int \frac{dx}{4+9x^2} &= \frac{2}{3}\int \frac{\sec^2\theta\, d\theta}{4+9(4/9)\tan^2\theta} \\
&= \frac{2}{3}\int \frac{\sec^2\theta\, d\theta}{4(1+\tan^2\theta)} \\
&= \frac{2}{3}\int \frac{\sec^2\theta\, d\theta}{4\sec^2\theta} \\
&= \frac{2}{3}\cdot\frac{1}{4}\int d\theta \\
&= \frac{1}{6}\theta + C \\
&= \frac{1}{6}\arctan(\frac{3}{2}x) + C.
\end{aligned}$$

(Return)

6. Start by completing the square for the quadratic:

$$\begin{aligned}
3+2x-x^2 &= -x^2+2x+3 \\
&= -(x^2-2x)+3 \\
&= -(x^2-2x+1)+4 \\
&= -(x-1)^2+4 \\
&= 4-(x-1)^2.
\end{aligned}$$

So we can rewrite the integral as

$$\begin{aligned}
\int \frac{dx}{\sqrt{3+2x-x^2}} &= \int \frac{dx}{\sqrt{4-(x-1)^2}} \\
&= \int \frac{du}{\sqrt{4-u^2}},
\end{aligned}$$

where we substituted $u = x-1$ and $du = dx$. This can now be dealt with using a trigonometric substitution of $u = 2\sin\theta$ (remember, the extra factor of 2 is there so that the 4 will factor out). So $du = 2\cos\theta\, d\theta$, and the integral becomes

$$\begin{aligned}
\int \frac{du}{\sqrt{4-u^2}} &= \int \frac{2\cos\theta\, d\theta}{\sqrt{4-4\sin^2\theta}} \\
&= \int \frac{2\cos\theta\, d\theta}{\sqrt{4}\sqrt{1-\sin^2\theta}} \\
&= \int \frac{2\cos\theta\, d\theta}{2\cos\theta} \\
&= \int d\theta \\
&= \theta + C.
\end{aligned}$$

Solving our original substitution for θ, we see that

$$\begin{aligned}\theta &= \arcsin\left(\frac{u}{2}\right) \\ &= \arcsin\left(\frac{x-1}{2}\right).\end{aligned}$$

So the final answer is

$$\int \frac{dx}{\sqrt{3+2x-x^2}} = \arcsin\left(\frac{x-1}{2}\right) + C.$$

(Return)

7. Using a regular trigonometric substitution, we would set $x = \tan\theta$, and $dx = \sec^2\theta\, d\theta$, which, after the usual algebra, gives

$$\begin{aligned}\int \frac{dx}{\sqrt{1+x^2}} &= \int \frac{\sec^2\theta\, d\theta}{\sqrt{1+\tan^2\theta}} \\ &= \int \sec\theta\, d\theta.\end{aligned}$$

But the integral of secant is not easy to remember, nor easy to rederive. If instead, we make the hyperbolic trigonometric substitution $x = \sinh u$, so $dx = \cosh u\, du$, then we have

$$\begin{aligned}\int \frac{dx}{\sqrt{1+x^2}} &= \int \frac{\cosh u\, du}{\sqrt{1+\sinh^2 u}} \\ &= \int \frac{\cosh u\, du}{\cosh u} \\ &= \int du \\ &= u + C \\ &= \operatorname{arcsinh} x + C.\end{aligned}$$

So the hyperbolic trigonometric substitution led to a much easier integral to evaluate. The trade-off is that the final result involves the inverse hyperbolic trigonometric functions, as opposed to more familiar functions.

(Return)

8. Using the hyperbolic trigonometric substitution $x = \sinh(\theta)$ gives

$$\begin{aligned}\int \sqrt{1+x^2}\, dx &= \int \sqrt{1+\sinh^2\theta}\cosh\theta\, d\theta \\ &= \int \sqrt{\cosh^2\theta}\cosh\theta\, d\theta \\ &= \int \cosh^2\theta\, d\theta \\ &= \frac{1}{2}\int (\cosh(2\theta)+1)\, d\theta \\ &= \frac{1}{2}\left(\theta + \frac{1}{2}\sinh(2\theta)\right) + C \\ &= \frac{1}{2}\theta + \frac{1}{4}2\sinh(\theta)\cosh(\theta) + C \\ &= \frac{1}{2}\sinh^{-1} x + \frac{1}{2}x\sqrt{1+x^2} + C.\end{aligned}$$

(Return)

9. This is a separable equation. Separating the variables and integrating both sides gives

$$\int \frac{dP}{b^2 + a^2P^2} = \int dt.$$

On the left, we can use the trigonometric substitution

$$\begin{aligned} P &= \frac{b}{a}\tan\theta \\ dP &= \frac{b}{a}\sec^2\theta\, d\theta. \end{aligned}$$

Note then that $\theta = \arctan \frac{a}{b}P$. This gives

$$\begin{aligned} \int \frac{dP}{b^2 + a^2P^2} &= \int \frac{\frac{b}{a}\sec^2\theta\, d\theta}{b^2(1+\tan^2\theta)} \\ &= \frac{b}{a}\int \frac{\sec^2\theta\, d\theta}{b^2\sec^2\theta} \\ &= \frac{1}{ab}\int d\theta \\ &= \frac{1}{ab}\theta \\ &= \frac{1}{ab}\arctan\frac{a}{b}P. \end{aligned}$$

(leaving off the constant for now). On the right side we get $t + C$, so

$$\frac{1}{ab}\arctan\frac{a}{b}P = t + C.$$

Solving this for P gives

$$P(t) = \frac{b}{a}\tan(abt + C).$$

If initial profits, at $t = 0$, are 0, then $C = 0$, so the final answer is

$$P(t) = \frac{b}{a}\tan(abt).$$

Since tangent blows up at $\frac{\pi}{2}$, this model implies profit goes to infinity at $t = \frac{\pi}{2ab}$, which is a sign that this model is not perfect.

(Return)

24 Partial Fractions

So far, the techniques of integration covered in this course have all been derived from differentiation rules run in reverse. This module gives an algebraic method for integrating *rational functions*. Recall that a rational function is a function of the form

$$f(x) = \frac{P(x)}{Q(x)},$$

where $P(x)$ and $Q(x)$ are polynomials. It turns out that with a little bit of algebraic manipulation, many of these integrals are not too difficult to compute.

Example

Compute

$$\int \frac{3x^2 - 5}{x - 2}\, dx.$$

(See Answer 1)

The rest of this module expands on this method (in particular, when the denominator is of a higher degree), which is known as the method of *partial fractions*.

24.1 Partial fractions

Given a rational function $\frac{P(x)}{Q(x)}$, and P has a lower power than Q, the method of partial fractions uses algebra to rewrite the function as a sum of simpler terms which are easy to integrate. While there are some cases to deal with, the basic outline of the method is:

1. Given the integral $\int \frac{P(x)}{Q(x)} dx$ where P and Q are polynomials.
2. Factor $Q(x) = (x - r_1)(x - r_2) \dots (x - r_n)$. Assume for now that each of these factors is distinct.
3. We use the following fact, that the rational function can be expressed as
$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - r_1} + \frac{A_2}{x - r_2} + \dots + \frac{A_n}{x - r_n}.$$
4. Use algebra to find what each of the constants A_i is. This step requires the most work.
5. Then

$$\int \frac{P(x)}{Q(x)}dx = \int \left(\frac{A_1}{x-r_1} + \ldots + \frac{A_n}{x-r_n}\right)dx$$
$$= A_1 \ln|x-r_1| + \ldots + A_n \ln|x-r_n| + C.$$

Example

Compute
$$\int \frac{3x-1}{x^2-2x-3}dx.$$
(See Answer 2)

Example

Compute
$$\int \frac{2x^2-6x-2}{x^3-x^2-2x}dx.$$
(See Answer 3)

Example

Compute
$$\int \frac{x^2+2x-1}{2x^3+3x^2-2x}\,dx.$$
(See Answer 4)

Example

A simple model for the deflection $x(t)$ of a thin beam under a load proportional to λ^2 is
$$\frac{dx}{dt} = \lambda^2 x - x^3 = x(\lambda - x)(\lambda + x).$$
Solve this differential equation (but do not solve for $x(t)$ explicitly). Then find the equilibria of the differential equation and classify them as stable or unstable. (See Answer 5)

Example

The *logistic model* for population dynamics says that the rate of change of a population P with respect to time is
$$\frac{dP}{dt} = rP - bP^2$$
where r and b are positive constants which can be thought of as the reproduction rate and death rate, respectively. Factoring and letting $K = \frac{r}{b}$, we have
$$\frac{dP}{dt} = bP(K-P).$$
Solve this differential equation. What is the long run population behavior? (See Answer 6)

24.2 Other technicalities

Higher degree numerator

For the algebra to work out above, the degree of the numerator, $P(x)$, must be lower than that of the denominator, $Q(x)$. However, it is easy to deal with the case when the numerator has equal or higher degree. One can use long division to rewrite the quotient as a divisor plus a remainder, just like writing an improper fraction as a mixed number in middle school.

Repeated factors

If the denominator has one or more repeated factors, i.e.

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x-r_1)^{m_1}\dots(x-r_k)^{m_k}},$$

where one or more of the m_i is greater than 1. Then the way to express the function is

$$\begin{aligned}\frac{P(x)}{(x-r_1)^{m_1}\dots(x-r_n)^{m_n}} &= \frac{A_1}{x-r_1} + \frac{A_2}{(x-r_1)^2} + \dots + \frac{A_{m_1}}{(x-r_1)^{m_1}} \\ &+ \frac{B_1}{x-r_2} + \frac{B_2}{(x-r_2)^2} + \dots + \frac{B_{m_2}}{(x-r_2)^{m_2}} \\ &+ \dots .\end{aligned}$$

Now, the algebra proceeds as before to find the constants in the numerators. It is easiest to see this through an example.

Example

Compute

$$\int \frac{2x^2-4x-2}{(x+1)(x-1)^2}dx.$$

(See Answer 7)

Quadratic factors

Suppose one of the factors of the denominator is a quadratic which cannot be factored (e.g. x^2+1). Then the numerator of this factor in the expansion should be of the form $Ax+B$. Then the algebra proceeds as before.

Example

Compute

$$\int \frac{3x^2-2x+1}{(x-1)(x^2+1)}dx.$$

(See Answer 8)

24.3 EXERCISES

Compute the following integrals:

- $\int \frac{5+x}{x^2+x-6}\,dx$
- $\int \frac{2x+3}{6x^2+5x+1}\,dx$
- $\int \frac{x}{(x+1)(x+2)}\,dx$
- $\int \frac{x^2-x+5}{(x-2)(x-1)(x+3)}\,dx$
- $\int \frac{2x-1}{x^3-x}\,dx$
- $\int \frac{x^2-3}{x^2-4}\,dx$
- $\int \frac{x^3+10x^2+33x+36}{x^2+4x+3}\,dx$
- $\int \frac{x+2}{(x-1)^2}\,dx$
- $\int \frac{dx}{x^4-6x^3+12x^2}$

24.4 Answers to Selected Examples

1. By doing polynomial long division on this ratio, we find

$$\begin{array}{r l}
 & 3x+6 \\
x-2 & \overline{\big)\, 3x^2-5} \\
 & -\underline{(3x^2-6x)} \\
 & \qquad 6x-5 \\
 & \qquad -\underline{(6x-12)} \\
 & \qquad\qquad 7
\end{array}$$

$$\tfrac{3x^2-5}{x-2} = 3x+6+\tfrac{7}{x-2}$$

(For more on polynomial long division, see wikipedia).

The above observation, which is entirely based on algebra, allows us to evaluate the integral as

$$\begin{aligned}
\int \frac{3x^2-5}{x-2}\,dx &= \int \left(3x+6+\frac{7}{x-2}\right)\,dx \\
&= \frac{3}{2}x^2+6x+7\ln|x-2|+C.
\end{aligned}$$

(Return)

2. Factoring the denominator gives $x^2 - 2x - 3 = (x+1)(x-3)$. Thus, the goal is to find constants A and B such that
$$\frac{3x-1}{(x+1)(x-3)} = \frac{A}{x+1} + \frac{B}{x-3}.$$
There are several methods for finding the constants, but one of the simplest is to clear denominators, which gives
$$3x - 1 = A(x-3) + B(x+1).$$
This equation must hold for every value of x. In particular, one can pick convenient values of x which make the algebra easy. In this case, by plugging in $x = 3$, the first term on the right disappears. Thus, the equation becomes $8 = B \cdot 4$, and so $B = 2$. Similarly, picking $x = -1$ makes the second term on the right disappear. Thus, $-4 = A \cdot (-4)$ so $A = 1$. It follows that
$$\begin{aligned}\int \frac{3x-1}{x^2-2x-3}\,dx &= \int \frac{3x-1}{(x+1)(x-3)}\,dx \\ &= \int \left(\frac{1}{x+1} + \frac{2}{x-3}\right) dx \\ &= \ln|x+1| + 2\ln|x-3| + C.\end{aligned}$$
(Return)

3. Factoring the denominator gives
$$x^3 - x^2 - 2x = x(x+1)(x-2)$$
So we are looking for constants A, B, and C such that
$$\frac{2x^2-6x-2}{x(x+1)(x-2)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-2}.$$
Clearing fractions gives
$$2x^2 - 6x - 2 = A(x+1)(x-2) + Bx(x-2) + Cx(x+1).$$
Now, picking the following convenient values of x allows us to find each constant:
$$\begin{array}{lll} x = 0 & -2 = -2A & A = 1 \\ x = -1 & 6 = 3B & B = 2 \\ x = 2 & -6 = 6C & C = -1. \end{array}$$
So we have that
$$\begin{aligned}\int \frac{2x^2-6x-2}{x(x+1)(x-2)} &= \int \left(\frac{1}{x} + \frac{2}{x+1} - \frac{1}{x-2}\right) dx \\ &= \ln|x| + 2\ln|x+1| - \ln|x-2| + C.\end{aligned}$$
(Return)

4. Factoring gives

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

So we are looking for constants A, B, C such that

$$\frac{x^2+2x-1}{2x^3+3x^2-2x} = \frac{A}{x} + \frac{B}{2x-1} + \frac{C}{x+2}.$$

As before, we clear fractions which gives

$$x^2 + 2x - 1 = A(2x-1)(x+2) + Bx(x+2) + Cx(2x-1).$$

Now we pick convenient values of x to make the factors cancel and solve for the constants:

$$\begin{array}{lll} x = 0 & -1 = -2A & A = \dfrac{1}{2} \\ x = \dfrac{1}{2} & \dfrac{1}{4} = \dfrac{5}{4}B & B = \dfrac{1}{5} \\ x = -2 & -1 = 10C & C = -\dfrac{1}{10}. \end{array}$$

So we find

$$\begin{aligned} \int \frac{x^2+2x-1}{x(2x-1)(x+2)}\,dx &= \int \left(\frac{1/2}{x} + \frac{1/5}{2x-1} + \frac{-1/10}{x+2}\right) dx \\ &= \frac{1}{2}\ln|x| + \frac{1}{5}\cdot\frac{1}{2}\ln|2x-1| - \frac{1}{10}\ln|x+2| + C \\ &= \frac{1}{2}\ln|x| + \frac{1}{10}\ln|2x-1| - \frac{1}{10}\ln|x+2| + C. \end{aligned}$$

Note the extra factor of $\frac{1}{2}$ for the middle term comes from doing a substitution of $u = 2x - 1$, which implies $dx = \frac{1}{2}\,du$.

(Return)

5. Factoring gives

$$\frac{dx}{dt} = x(\lambda - x)(\lambda + x).$$

Separating and integrating gives

$$\int \frac{dx}{x(\lambda-x)(\lambda+x)} = \int dt.$$

Now, we use partial fractions on the left side:

$$\frac{1}{x(\lambda-x)(\lambda+x)} = \frac{A}{x} + \frac{B}{\lambda-x} + \frac{C}{\lambda+x}$$

Clearing fractions gives

$$1 = A(\lambda-x)(\lambda+x) + Bx(\lambda+x) + Cx(\lambda-x).$$

Picking convenient values of x gives

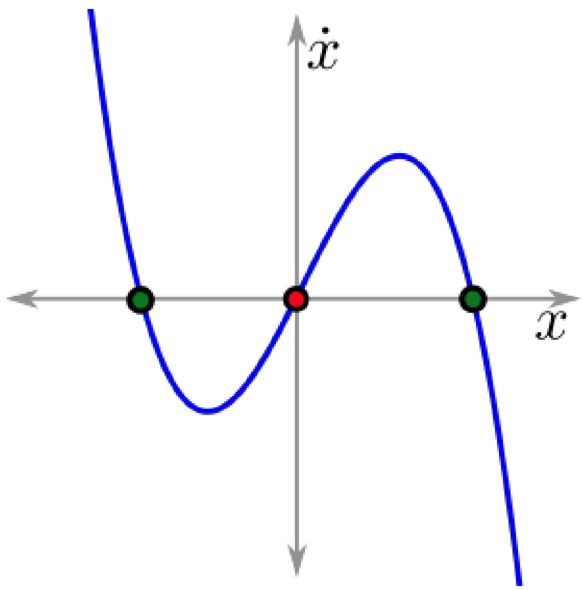
$$\begin{aligned} x &= 0 & 1 &= \lambda^2 A & A &= \frac{1}{\lambda^2} \\ x &= \lambda & 1 &= 2\lambda^2 B & B &= \frac{1}{2\lambda^2} \\ x &= -\lambda & 1 &= -2\lambda^2 C & C &= -\frac{1}{2\lambda^2}. \end{aligned}$$

So we have

$$\begin{aligned} \int \frac{dx}{x(\lambda-x)(\lambda+x)} &= \int \left(\frac{1/\lambda^2}{x} + \frac{1/2\lambda^2}{\lambda-x} - \frac{1/2\lambda^2}{\lambda+x} \right) dx \\ &= \frac{1}{\lambda^2} \ln|x| - \frac{1}{2\lambda^2} \ln|\lambda-x| - \frac{1}{2\lambda^2} \ln|\lambda+x|. \end{aligned}$$

All of this equals $t + C$ on the right.

The equilibria of the differential equation are $x = 0$, $x = \lambda$, and $x = -\lambda$. The equilibrium at 0 is unstable and the other two are stable, as the graph shows:



(Return)

6. Separating gives

$$\frac{dP}{P(K-P)} = bdt. \tag{1}$$

Integrating the left side is done using partial fractions, and the denominator is already factored. So the next step is to find A and B such that

$$\frac{1}{P(K-P)} = \frac{A}{P} + \frac{B}{K-P}.$$

Clearing denominators gives $1 = A(K-P) + BP$. Remember, K and A are constants, and this equation must hold for every value of P. Setting $P = K$ cancels the first term and gives $1 = BK$, so $B = \frac{1}{K}$.

Setting $P = 0$ cancels the second term and gives $A = \frac{1}{K}$. Thus,

$$\begin{aligned}\int \frac{dP}{P(K-P)} &= \int \frac{1}{K}\left(\frac{1}{P} + \frac{1}{K-P}\right) dP \\ &= \frac{1}{K}\left(\ln P - \ln(K-P)\right) \\ &= \frac{1}{K}\ln\frac{P}{K-P},\end{aligned}$$

by a property of logarithms. Multiplying through by K gives

$$\begin{aligned}\ln\left(\frac{P}{K-P}\right) &= \int Kb\,dt \\ &= \int r\,dt \\ &= rt + C\end{aligned}$$

(recall that $Kb = r$ by the definition of K). Now, exponentiating gives

$$\frac{P}{K-P} = \tilde{C}e^{rt},$$

for a new constant $\tilde{C}$. By plugging in $t = 0$, we find that

$$\tilde{C} = \frac{P_0}{K-P_0},$$

where P_0 is the initial population. Multiplying through by $K - P$ and doing a little algebra gives

$$\begin{aligned}P &= \tilde{C}e^{rt}(K-P) \\ P + P\tilde{C}e^{rt} &= \tilde{C}e^{rt}K \\ P(1+\tilde{C}e^{rt}) &= \tilde{C}e^{rt}K \\ P &= \frac{\tilde{C}e^{rt}K}{1+\tilde{C}e^{rt}}.\end{aligned}$$

Replacing $\tilde{C} = \frac{P_0}{K-P_0}$ gives

$$\begin{aligned}P &= \frac{P_0}{K-P_0}e^{rt}K \cdot \frac{1}{1+\frac{P_0}{K-P_0}e^{rt}} \\ &= \frac{KP_0e^{rt}}{K-P_0+P_0e^{rt}} \\ &= \frac{KP_0}{(K-P_0)e^{-rt}+P_0}\end{aligned}$$

(From the first to the second line, we distributed $(K - P_0)$ in the denominator. From the second to third line, we multiplied the top and bottom by e^{-rt}.).

Note that if $P_0 = 0$ (i.e. there was no population to begin with), then the population will stay at 0. This is consistent with the above equation. On the other hand, if $P_0 > 0$, then as $t \to \infty$, the e^{-rt} in the denominator goes to 0, and so

$$\lim_{t\to\infty} P(t) = \frac{KP_0}{P_0} = K.$$

We can think of K as the carrying capacity for the population, a sort of ideal size for the population. Alternatively, by looking at the original differential equation, we see that K is an equilibrium. It is stable, since populations above K have a negative derivative (hence are decreasing), and populations below K have a positive derivative (hence are increasing).

On the other hand 0 is an unstable equilibrium. The model implies that as long as the population is not extinct to begin with, it will grow and eventually equal K.

(Return)

7. The denominator is already factored, so write

$$\frac{2x^2-4x-2}{(x+1)(x-1)^2} = \frac{A}{x+1} + \frac{B}{x-1} + \frac{C}{(x-1)^2}.$$

Clearing the denominators gives

$$2x^2-4x-2 = A(x-1)^2 + B(x+1)(x-1) + C(x+1).$$

Plugging in $x = 1$ cancels the first two terms on the right, leaving $-4 = 2C$, so $C = -2$. Plugging in $x = -1$ cancels the second two terms and leaves $4 = 4A$, so $A = 1$.

Now, it seems that there are no more nice values of x to help solve for B. But remember that the equation must hold for any value of x. Picking $x = 0$ (which is an easy value to use), gives $-2 = A - B + C$. Knowing $A = 1$ and $C = -2$ gives $B = 1$.

Thus,

$$\begin{aligned}\int \frac{2x^2-4x-2}{(x+1)(x-1)^2}dx &= \int \left(\frac{1}{x+1} + \frac{1}{x-1} + \frac{-2}{(x-1)^2}\right)dx \\ &= \ln|x+1| + \ln|x-1| + \frac{2}{x-1} + C\end{aligned}$$

(Return)

8. Write

$$\frac{3x^2-2x+1}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}.$$

Clearing fractions gives $3x^2 - 2x + 1 = A(x^2 + 1) + (Bx + C)(x - 1)$. Picking $x = 1$ gives $2 = 2A$, so $A = 1$. Now, picking any other two values for x will allow finding B and C. For instance, $x = 0$ gives $1 = A - C$, and so $C = 0$. Finally, picking $x = -1$ gives $6 = 2A + (-B)(-2)$, so $B = 2$.

Thus,

$$\begin{aligned}\int \frac{3x^2-2x+1}{(x-1)(x^2+1)}dx &= \int \left(\frac{1}{x-1} + \frac{2x}{x^2+1}\right)dx \\ &= \ln|x-1| + \ln(x^2+1) + C.\end{aligned}$$

(Return)

25 Definite Integrals

This module moves from the indefinite integral, which is a class of functions, to the definite integral, which is a number. The relationship between these seemingly unrelated topics will be revealed in the next module.

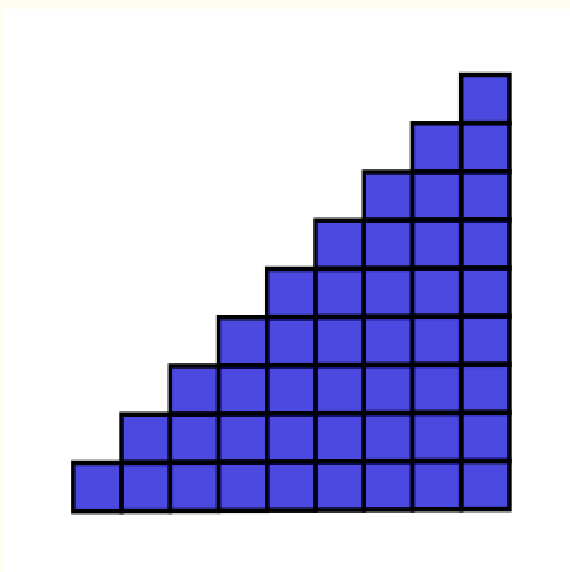
The idea underlying the definite integral is that adding up local increments leads to a global total. Before getting into the details of what this means, consider a simple example.

Example

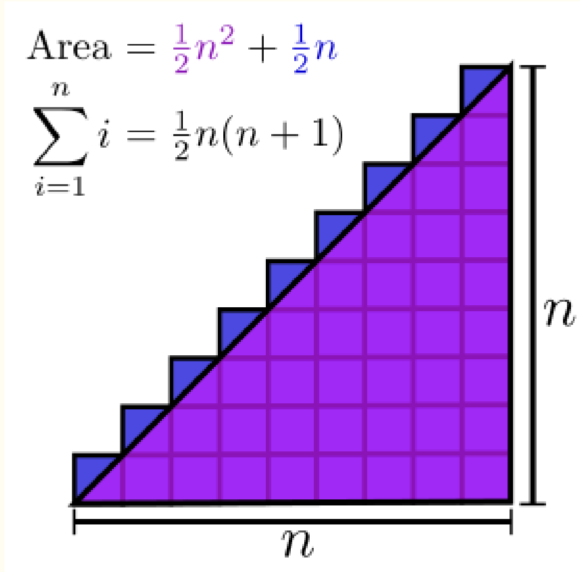
Consider

$$\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n.$$

One can visualize this sum as the area of a triangular stack of 1×1 boxes. The first column has 1 box, the second column has 2 boxes, and so on through the nth column with n boxes:



The area of this roughly triangular region can be found by splitting it into two regions: a right triangle of base and height n, and the half boxes left over:



The total area is therefore $\frac{1}{2}n(n+1)$, and so we find that

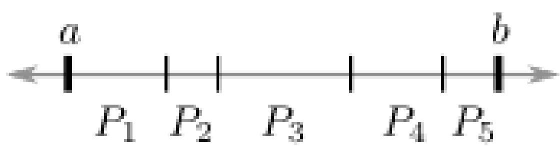
$$\sum_{i=1}^{n} i = \frac{1}{2}n(n+1).$$

The point of this example is to compare the amount of computation (e.g. the number of additions) required to do the sum using local information (adding up the terms one by one), verses the global information (evaluating the product on the right above). It is much easier to simply evaluate the product.

The definite integral takes this type of idea and generalizes it to more difficult sums. Before we can define it, we need a few definitions.

25.1 Partitions and Riemann sums

Given an interval $[a, b]$, a *partition* P of $[a, b]$ is a division of the interval $[a, b]$ into subintervals P_i. Visually, think of placing hash marks along the interval $[a, b]$ and then labeling the subintervals $P_1, P_2, \ldots$ from left to right:



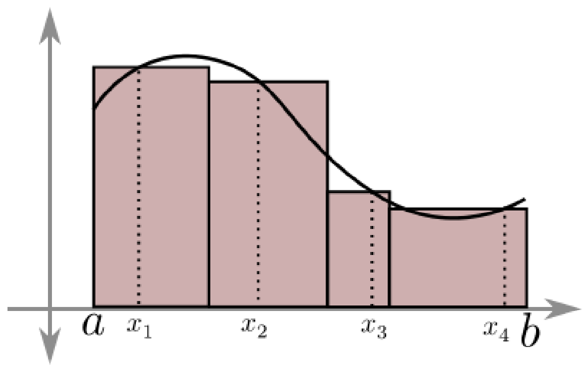
Let $(\Delta x)_i$ be the width of the ith subinterval, P_i.

Choose a *sample point* x_i from the ith subinterval (this can be a point chosen at random from the subinterval or systematically; it does not matter).

Given a function f, a partition P for an interval $[a, b]$, and sample points x_i, the *Riemann sum* of f on P is given by

$$\sum_{i=1}^{N} f(x_i)(\Delta x)_i.$$

The Riemann sum can be interpreted as an approximation of the area under the curve of f from a to b using rectangles. The width and height of the ith rectangle are $(\Delta x)_i$ and $f(x_i)$, respectively. Note that in this area interpretation, a rectangle which is below the x-axis has negative area (since $f(x_i) < 0$ in this case). For an example with $N = 4$ rectangles, consider the following figure:



25.2 The definite integral

The definite integral

The *definite integral* of a function f from a to b, denoted

$$\int_{x=a}^{b} f(x)\,dx,$$

is defined by

$$\int_{x=a}^{b} f(x)\,dx = \lim_{\Delta x \to 0} \sum_{i=1}^{N} f(x_i)(\Delta x)_i.$$

The function f being integrated is called the *integrand*.

In other words, the definite integral is the limit of the Riemann sums as the lengths of the subintervals approach 0. In the area interpretation, the widths of all the rectangles are getting arbitrarily small, which ultimately gives the area under the curve:

(Link to Riemann Sum Limit Animated GIF)

Remember that when interpreting the definite integral as the area under the curve, any region which is below the x-axis contributes negative area to the total.

Example

Using the definition of the definite integral, compute

$$\int_{x=0}^{1} x\,dx.$$

(See Answer 1)

Notation

Sums The integral sign $\int$ and the summation sign $\sum$ are both short for sum. The integral sign $\int$ looks like a stylized S, and the summation sign is the Greek sigma, short for sum.

Limits Including the variable in the limits of integration is not strictly necessary, but is a useful habit to develop for future courses where integration will be happening with respect to several variables. It is also fine to suppress the notation and just have $\int_a^b f(x)\,dx$:

$$\int_a^b f(x)\,dx = \int_{x=a}^{b} f(x)\,dx.$$

Variables The variable used in the integrand does not matter; it is sometimes referred to as a dummy variable:

$$\int_{x=a}^{b} f(x)\,dx = \int_{t=a}^{b} f(t)\,dt = \int_{z=a}^{b} f(z)\,dz.$$

However, if there is a variable used in one of the limits of integration (as will happen from time to time), it is important to avoid using that as the dummy variable too. For example,

$$\int_a^x f(t)\,dt \quad \text{instead of} \quad \int_a^x f(x)\,dx.$$

Caveat

Note that, although their notation is similar, definite integrals are not the same as indefinite integrals! The indefinite integral of a function is a class of functions, whereas the definite integral of a function over an interval is a number.

That said, it is no accident that they have similar notations, because of their relationship, which is given by the Fundamental Theorem of Integral Calculus in the next module.

25.3 Properties of definite integrals

Linearity

The definite integral is linear, i.e.

$$\int_{x=a}^{b} (f(x) + g(x))\ dx = \int_{x=a}^{b} f(x)\,dx + \int_{x=a}^{b} g(x)\,dx.$$

$$\int_{x=a}^{b} c \cdot f(x)\,dx = c\int_{x=a}^{b} f(x)\,dx.$$

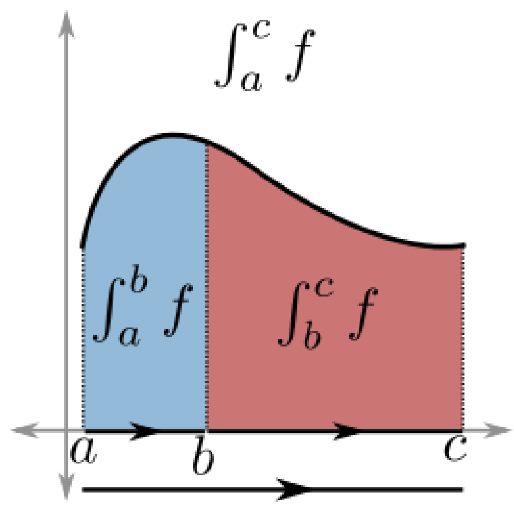
(See Justification 2)

Additivity

When integrating the same function over two adjacent intervals, we have additivity:

$$\int_a^b f(x)\,dx + \int_b^c f(x)\,dx = \int_a^c f(x)\,dx.$$

In the area interpretation, this can be thought of as taking the area under the curve from a to b and adding the area under the curve from b to c, which gives the area under the curve from a to c:



Another way of thinking about it is adding the intervals $[a, b]$ and $[b, c]$ together to get $[a, c]$. It is important to note that the orientation of the interval matters, as discussed in the next subsection.

Orientation

The orientation of the interval over which we integrate matters. Integrating from left to right is positive, and integrating from right to left is negative:

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

(See Justification 3)

Dominance

This is another intuitive property. If $f(x) \geq 0$ for all x in the interval $[a, b]$, then

$$\int_a^b f(x)\,dx \geq 0.$$

Also, if $f(x) \geq g(x)$ for all x in the interval, then

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

(See Justification 4)

25.4 More examples

There are a few definite integrals that we can compute directly from the definition. But for most functions, it is not easy to work directly with the definition.

Example

Compute

$$\int_a^b c\,dx$$

(See Answer 5)

Example

Compute

$$\int_a^b x\,dx$$

(See Answer 6)

25.5 Odd and even functions

There are a few final cases where certain definite integrals can be simplified by using properties of the integrand.

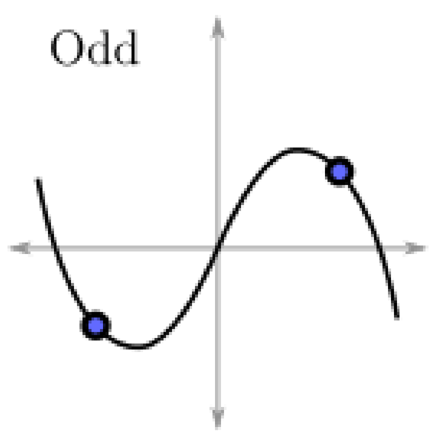
Odd and even functions

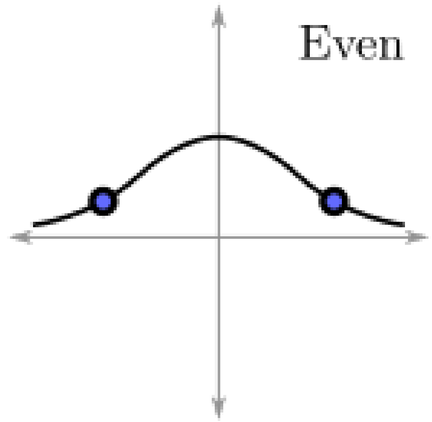
A function $f(x)$ is called *odd* if

$$f(-x) = -f(x).$$

A function $g(x)$ is called *even* if

$$g(-x) = g(x).$$





The reason for the terminology comes from Taylor series. A function is odd if and only if every term in its Taylor series has *odd* power. Similarly, a function is even if and only if every term in its Taylor series has *even* power. (See Justification 7)

Example

Sine and hyperbolic sine are both odd functions because they only have odd powers in their Taylor series. Cosine and hyperbolic cosine are both even functions because they only have even powers in their Taylor series.

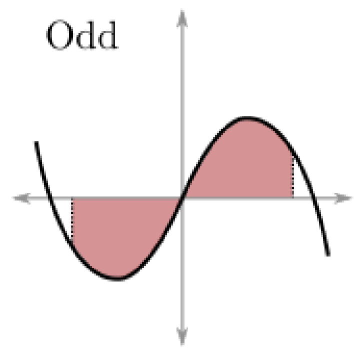
Odd function over a symmetric domain

If an odd function f is integrated over a domain that is symmetric about the origin (i.e., an interval of the form $[-L, L]$, then

$$\int_{x=-L}^{L} f(x)\,dx = 0.$$

Formally, any subinterval's on the left half of the interval will make a contribution to the Riemann sum which is equal and opposite to the contribution of the corresponding subinterval on the right half of the interval. These equal and opposite sums cancel, and so the definite integral over the entire interval is 0.

In terms of the area interpretation, the net area under the curve over the left half of the interval will be equal and opposite in sign to the net area under the curve over the right half of the interval. Therefore, the total area will be 0:



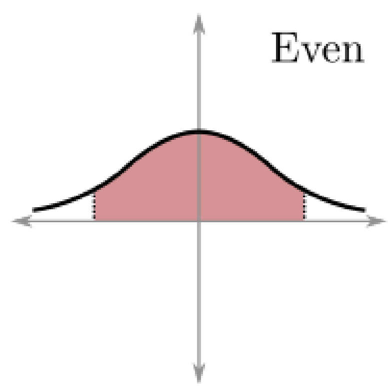
Even function over a symmetric domain.

If an even function g is integrated over a domain that is symmetric about the origin (i.e., an interval of the form $[-L, L]$), then

$$\int_{x=-L}^{L} g(x)\,dx = 2\int_{x=0}^{L} g(x)\,dx.$$

Formally, each subinterval on the left half of the interval has a corresponding subinterval on the right with an equal contribution to the Riemann sum. So one can just take the Riemann sum on the right and double it.

Using the area interpretation, one can see that the region under the curve on the left will be the mirror image of the region under the curve on the right, so the total area is just twice the area on the right:



25.6 EXERCISES

- One particular choice of partition and sampling that can be used to numerically evaluate definite integrals is the following. With n fixed, divide the interval $[a, b]$ into n subintervals P_i of common length $(\Delta x)_i = (b-a)/n$. For the sampling, choose the right endpoint of each P_i; this gives you the formula:

  $$x_i = a + i\frac{b-a}{n}$$

  With these choices of partition and sampling, compute the Riemann sums for the integral

$$\int_{x=1}^{2} \frac{dx}{x}$$

for $n = 1, 2, 3$ subdivisions. Note: in the next Lecture we will learn that

$$\int_{x=1}^{2} \frac{dx}{x} = \ln 2 \simeq 0.693$$

How does this compare to the values you obtained from the Riemann sums?

- With the same choices of partition and sampling as in the previous problem, evaluate the Riemann sum for the integral

$$\int_{x=0}^{3} x^2 \, dx$$

for an arbitrary number n of subdivisions. You may need to use the following:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \quad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}, \quad \sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$$

- The line $y = x$, the x-axis and the vertical line $x = 2$ bound a triangle of area 2. Thus,

$$I = \int_{x=0}^{2} x \, dx = 2$$

Evaluating the Riemann sum for n subdivisions for the above integral with the same choices of partition and sampling as in the previous problem yields an approximation $RS(n)$ for its value I. The error $E(n)$ we commit by using this approximation is defined to be the difference

$$E(n) = RS(n) - I$$

Show that $E(n)$ is in $O(n^{-k})$ for some $k > 0$. What's the best value of k?

- What is the following integral? Think!

$$\int_{x=-\pi/4}^{\pi/4} \left(x^2 + \ln|\cos x|\right) \sin\frac{x}{2} \, dx$$

- Using the definition of definite integrals, compute $\int_0^1 x^3 \, dx$. Use a uniform partition and the fact that $\sum_{i=1}^{n} i^3 = \frac{n^2(n+1)^2}{4}$.

25.7 Answers to Selected Examples

1. Let the partition P divide the interval $[0, 1]$ into N equally sized subintervals. Then the ith subinterval of P is given by $[(i-1)\frac{1}{N}, i\frac{1}{N}]$, and $(\Delta x)_i = \frac{1}{N}$. Choose the right endpoint of each subinterval to be its sample point, i.e. $x_i = \frac{i}{N}$. Finally, note that as $N \to \infty$, $\Delta x \to 0$. It follows that

$$\begin{aligned}
\int_0^1 x\,dx &= \lim_{\Delta x \to 0} \sum_{i=1}^{N} f(x_i)(\Delta x)_i \\
&= \lim_{N\to\infty} \sum_{i=1}^{N} \frac{i}{N} \cdot \frac{1}{N} \\
&= \lim_{N\to\infty} \frac{1}{N^2} \sum_{i=1}^{N} i \\
&= \lim_{N\to\infty} \frac{1}{N^2} \frac{N(N+1)}{2} \\
&= \lim_{N\to\infty} \frac{N^2+N}{2N^2} \\
&= \frac{1}{2}.
\end{aligned}$$

   We used the fact from earlier that

$$\sum_{i=1}^{n} i = \frac{1}{2} n(n+1)$$

   (Return)

2. The definite integral is defined as the limit of Riemann sums. Note that for any partition P of the interval,

$$\begin{aligned}
\sum_{i=1}^{N} (f+g)(x_i)(\Delta x)_i &= \sum_{i=1}^{N} \left[f(x_i) + g(x_i)\right](\Delta x)_i \\
&= \sum_{i=1}^{N} f(x_i)(\Delta x)_i + g(x_i)(\Delta x)_i \\
&= \sum_{i=1}^{N} f(x_i)(\Delta x)_i + \sum_{i=1}^{N} g(x_i)(\Delta x)_i,
\end{aligned}$$

   because of linearity of finite sums. Therefore, as one takes the limit as $\Delta x \to 0$, one finds (by the linearity of limits) that

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

   The argument for a constant multiple is almost identical: we can pull a constant out from a sum, and pull a constant out from a limit.

   (Return)

3. Consider what happens if one computes

$$\int_a^b f(x)\,dx + \int_b^a f(x)\,dx.$$

By the additivity property (where c has been replaced by a), this is

$$\int_a^a f(x)\, dx.$$

But this equals 0, which is intuitive in the area interpretation. (More formally, any partition of an interval with 0 width has subintervals of 0 width, so the Riemann sums equal 0). Therefore,

$$\int_a^b f(x)\, dx + \int_b^a f(x)\, dx = 0,$$

and rearranging gives

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

as desired.

(Return)

4. For the first part, note that regardless of the partition of $[a, b]$, the Riemann sum

$$\sum_{i=1}^{N} f(x_i)(\Delta x)_i \geq 0,$$

because $f(x_i) \geq 0$ by the above assumption. Since each Riemann sum is non-negative, the limit is non-negative.

For the second part, note that

$$f(x) \geq g(x) \Longrightarrow f(x) - g(x) \geq 0.$$

So applying the first part, we have

$$\int_a^b \left(f(x) - g(x)\right)\, dx \geq 0.$$

Then by linearity of the definite integral (above),

$$\int_a^b f(x)\, dx - \int_a^b g(x)\, dx \geq 0,$$

and rearranging gives

$$\int_a^b f(x)\, dx \geq \int_a^b g(x)\, dx.$$

(Return)

5. If we use the partition of $[a, b]$ into n equal intervals, then

$$(\Delta x)_i = \frac{b-a}{n}$$

Also, note that $f(x_i) = c$ for all i. So

$$\begin{aligned}\int_a^b c\,dx &= \lim_{\Delta x \to 0} \sum_{i=1}^{n} c\frac{b-a}{n} \\ &= \lim_{\Delta x \to 0} c \cdot n \cdot \frac{b-a}{n} \\ &= c \cdot (b-a).\end{aligned}$$

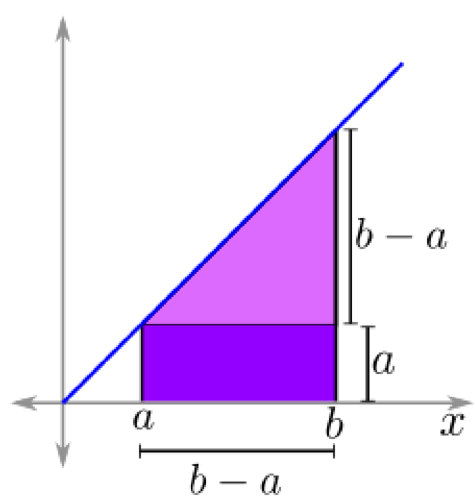
We could also see this by interpreting this definite integral as the area under the curve $y = c$ between $x = a$ and $x = b$, which is simply a rectangle of base $b - a$ and height c.

(Return)

6. Again using a partition into n equal sized subintervals, we have that $(\Delta x)_i = \frac{b-a}{n}$. If we take our sample point to be the right endpoint of each subinterval, then we have $x_i = a + \frac{b-a}{n}i$. So

$$\begin{aligned}\int_{x=a}^b x\,dx &= \lim_{n \to \infty} \sum_{i=1}^{n} \left(a + \frac{b-a}{n}i\right)\frac{b-a}{n} \\ &= \lim_{n \to \infty} \sum_{i=1}^{n} a\frac{b-a}{n} + \sum_{i=1}^{n} i\left(\frac{b-a}{n}\right)^2 \\ &= \lim_{n \to \infty} n \cdot a\frac{b-a}{n} + \left(\frac{b-a}{n}\right)^2 \frac{n(n+1)}{2} \\ &= a(b-a) + \frac{(b-a)^2}{2} \\ &= \frac{2ab - 2a^2 + b^2 - 2ab + a^2}{2} \\ &= \frac{1}{2}(b^2 - a^2).\end{aligned}$$

This can also be found by interpreting the definite integral as the area under the curve $y = x$, which can be broken into a rectangle with base $b - a$ and height a and a triangle with base and height $b - a$:



(Return)

7. If f only has odd powers in its Taylor series, then

$$f(x) = a_1 x + a_3 x^3 + a_5 x^5 + \cdots$$

for some constants $a_1, a_3, \cdots$. So evaluating $f(-x)$ and doing a little algebra, we find

$$\begin{aligned} f(-x) &= a_1(-x) + a_3(-x)^3 + a_5(-x)^5 + \cdots \\ &= -a_1 x - a_3 x^3 - a_5 x^5 - \cdots \\ &= -\left(a_1 x + a_3 x^3 + a_5 x^5 + \cdots\right) \\ &= -f(x), \end{aligned}$$

as desired. Similarly, if $g(x)$ has even powers, then

$$g(x) = a_0 + a_2 x^2 + a_4 x^4 + \cdots$$

and it follows that

$$\begin{aligned} g(-x) &= a_0 + a_2(-x)^2 + a_4(-x)^4 + \cdots \\ &= a_0 + a_2 x^2 + a_4 x^4 + \cdots \\ &= g(x), \end{aligned}$$

as desired.

(Return)

26 Fundamental Theorem Of Integral Calculus

Computing definite integrals from the definition is difficult, even for fairly simple functions. Fortunately, there is a powerful tool—the *Fundamental Theorem of Integral Calculus*—which connects the definite integral with the indefinite integral and makes most definite integrals easy to compute.

The Fundamental Theorem of Integral Calculus (FTIC)

Given a continuous function f, it follows that

1. $\frac{d}{dx}\left(\int_a^x f(t)dt\right) = f(x)$ and

2. $\int_a^b f(x)dx = \left(\int f(x)dx\right)\Big|_a^b$

(where $F(x)\Big|_a^b = F(b) - F(a)$).

In other words, Part 1 says that the function

$$F(x) = \int_{t=a}^{x} f(t)dt$$

is an anti-derivative of f.

Part 2 says that the definite integral can be computed by finding the indefinite integral of f and subtracting the evaluation at the bottom bound from the evaluation at the top bound. Note that even though the indefinite integral is actually a class of functions that differ by a constant, $F(b) - F(a)$ has the same value for any function F in such a class, so when computing the antiderivative for the purpose of computing a definite integral, it is allowable (and convenient) to forego the constant of integration.

Part 2 can be expressed in a slightly different way which is illustrative. For a differentiable function F we have

$$F\Big|_{x=a}^{b} = \int_{x=a}^{b} dF.$$

This says that the net change in quantity (given on the left side) equals the integral of the rate of change (given on the right side). This interpretation will be used in many applications in the next chapter. (See Rough Proof 1)

Example

Compute

$$\int_{x=1}^{T} \frac{1}{x}\, dx.$$

(See Answer 2)

Example

Compute

$$\int_{1}^{3} x^2 dx.$$

(See Answer 3)

Example

Suppose a publisher prints 12000 books per month with expected revenue of \$60 per book. The marginal cost of each book is given by

$$MC(x) = 10 + \frac{1}{2000}x.$$

What would be the change in profit from a 25% increase in production? (See Answer 4)

Example

Find

$$\frac{d}{dx}\left(\int_{0}^{x} \sin(t)dt\right).$$

(See Answer 5)

Example

Find

$$\frac{d}{dx}\left(\int_{x}^{x^3} \sin(t)dt\right).$$

(See Answer 6)

Caveat

Note that if the integrand $f(t)$ fails to be defined or continuous at a point in the interval $[a, b]$, then the FTIC does not hold. The following example shows this using the singularities of a rational function.

Example

Compute

$$\int_{1}^{4} \frac{dx}{x^2 - 5x + 6}.$$

(See Answer 7)

26.1 Limits of integration and substitution

One must be careful when using Part 2 of the Fundamental Theorem of Integral Calculus along with the method of substitution. The reason this can cause problems is that when a substitution is made, the old limits of integration are still in terms of the original variable. Therefore, one must either get the antiderivative in terms of the original variable before evaluating (this is what we usually did at the end of the substitution anyway), or one can change the limits of integration to reflect the new variable.

Consider the following example which demonstrates both techniques.

Example

Compute
$$\int_{x=0}^{1} x(x-1)^n \, dx,$$
where n is some fixed positive constant. (See Answer 8)

Another case where one must be careful of the limits of integration is with Integration By Parts. One can compute the indefinite integral completely and then apply the limits of integration, or one can apply them as one goes, as in the following:
$$\int_a^b u \, dv = uv\Big|_a^b - \int_a^b v \, du.$$

Example

Again compute the definite integral
$$\int_{x=0}^{1} x(x-1)^n \, dx,$$
but this time using integration by parts. (See Answer 9)

26.2 Additional examples

Example

Compute
$$\int_{x=e^3}^{e^5} \frac{\ln x}{x} \, dx.$$
(See Answer 10)

Example

Compute
$$\int_{x=-1}^{1} \frac{1}{1+3x^2} \, dx.$$
(See Answer 11)

26.3 EXERCISES

- Evaluate the following integrals:

$$\int_{x=-1}^{1} \frac{dx}{1+x^2}$$

$$\int_{x=0}^{3} 5x\sqrt{x+1}\,dx$$

$$\int_{x=-\pi}^{\pi} \frac{d}{dx}(x\cos x)\,dx$$

- Compute the following derivatives:

$$\frac{d}{dx}\int_{x=-\pi}^{\pi} x\cos x\,dx$$

$$\frac{d}{dx}\int_{t=0}^{x} \cos t\,dt$$

$$\frac{d}{dx}\int_{t=x^2}^{x^4} e^{-t^2}\,dt$$

$$\frac{d}{dx}\int_{t=1}^{x} \frac{1}{\sqrt{t}}\,dt$$

$$\frac{d}{dx}\int_{t=0}^{\arcsin x} \ln\left|\sin t + \cos t\right|\,dt$$

$$\frac{d}{dx}\int_{t=\sin x}^{\tan x} e^{-t^2}\,dt$$

- What is the leading-order term in the Taylor series about $x = 0$ of

$$f(x) = \int_0^x \ln(\cosh(t))dt$$

- We usually use Riemann sums to approximate integrals, but we can go the other way, too, using an antiderivative to approximate a sum. Using only your head (no paper, no calculator), compute an approximation for

$$\sum_{n=0}^{100} n^3$$

Hint: what integral does this resemble?

- Compute $\displaystyle\frac{d}{dx}\int_{3x}^{x^4} e^{-t^2/2}dt$

- Find the critical points of the function $f(x) = \int_e^{x^2} \ln(1+t^2)\cos(\sqrt{t})dt$

26.4 Answers to Selected Examples

1. Part 1: By the definition of the derivative, and the definition of the definite integral,

$$\begin{aligned}
\frac{d}{dx}\left(\int_a^x f(t)dt\right) &= \lim_{h\to 0}\frac{\int_a^{x+h} f(t)dt - \int_a^x f(t)dt}{h} \\
&= \lim_{h\to 0}\frac{\int_x^{x+h} f(t)dt}{h} \\
&= \lim_{h\to 0}\frac{f(x^*)\Delta x}{h} \\
&= \lim_{h\to 0}\frac{f(x^*)h}{h} \\
&= \lim_{h\to 0} f(x^*) \\
&= f(x),
\end{aligned}$$

   since $x \leq x^* \leq x + h$.

   Part 2: By Part 1, $F(x) = \int_a^x f(t)dt$ is an anti-derivative of f. Furthermore, we have

$$F(b) - F(a) = \int_a^b f(t)dt - \int_a^a f(t)dt = \int_a^b f(t)dt - 0 = \int_a^b f(t)dt.$$

   Let $G(x)$ be some anti-derivative of f. Since anti-derivatives of the same function differ only by a constant, $G(x) = F(x) - C$ for some constant C. Then we have

$$G(b) - G(a) = (F(b) - C) - (F(a) - C) = F(b) - F(a)$$

   as desired.

   (Return)

2. Using Part 2 of the Fundamental Theorem, we find that

$$\begin{aligned}
\int_{x=1}^{T} \frac{1}{x}\,dx &= \ln x\Big|_{x=1}^{T} \\
&= \ln T - \ln 1 \\
&= \ln T.
\end{aligned}$$

   Note that the above definite integral is sometimes used as the definition of the natural logarithm.

   (Return)

3. By Part 2 of FTIC,

$$\begin{aligned}
\int_1^3 x^2 dx &= \frac{1}{3}x^3\Big|_1^3 \\
&= \frac{1}{3}(3^3 - 1^3) \\
&= \frac{26}{3}.
\end{aligned}$$

   (Return)

4. The additional 25% means an extra 3000 books. So the goal is to find the change in profit P as x goes from 12000 to 15000. That is,
$$P\Big|_{x=12000}^{15000} = \int_{x=12000}^{15000} dP,$$
according to Part 2 of the Fundamental Theorem. Now, since profit is revenue minus cost, it follows that marginal profit is given by
$$dP = dR - dC.$$
Here,
$$dR = MR(x)\,dx = 60\,dx$$
since the marginal revenue from each book is \$60. And
$$dC = MC(x)\,dx = 10 + \frac{1}{2000}x\,dx.$$
Putting it all together, we find that
$$\begin{aligned} P\Big|_{x=12000}^{15000} &= \int_{x=12000}^{15000} dP \\ &= \int_{x=12000}^{15000} \left(60 - \left(10 + \frac{1}{2000}x\right)\right)\,dx \\ &= \int_{x=12000}^{15000} \left(50 - \frac{1}{2000}x\right)\,dx \\ &= 50x - \frac{1}{4000}x^2\Big|_{x=12000}^{15000} \\ &= \left(50\cdot 15000 - \frac{15000^2}{4000}\right) - \left(50\cdot 12000 - \frac{12000^2}{4000}\right) \\ &= \$129750 \end{aligned}$$
(Return)

5. By Part 1 of FTIC, $\frac{d}{dx}\left(\int_0^x \sin(t)dt\right) = \sin(x)$.
(Return)

6. One must be careful with a function of x in one or both bounds. A good way to break the problem down is to write $F(x) = \int_0^x \sin(t)dt$. By Part 1 of FTIC, $F'(x) = \sin(x)$. Now, note that
$$\begin{aligned} \int_x^{x^3} \sin(t)dt &= \int_0^{x^3} \sin(t)dt - \int_0^x \sin(t)dt \\ &= F(x^3) - F(x). \end{aligned}$$
Next, taking the derivative (and remembering the chain rule) gives
$$\begin{aligned} \frac{d}{dx}\left(\int_x^{x^3} \sin(t)dt\right) &= \frac{d}{dx}\left(F(x^3) - F(x)\right) \\ &= F'(x^3)(3x^2) - F'(x)(1) \\ &= \sin(x^3)(3x^2) - \sin(x). \end{aligned}$$
(Return)

7. This is a rational function, and the denominator factors as $(x-3)(x-2)$, so use partial fractions to express

$$\frac{1}{(x-3)(x-2)} = \frac{A}{x-3} + \frac{B}{x-2}.$$

Clearing denominators gives $1 = A(x-2) + B(x-3)$. Setting $x = 3$ gives $A = 1$. Setting $x = 2$ gives $B = -1$. Thus,

$$\begin{aligned} \int \frac{dx}{x^2-5x+6} &= \int \left(\frac{1}{x-3} - \frac{1}{x-2}\right) dx \\ &= \ln|x-3| - \ln|x-2|. \end{aligned}$$

Then trying to apply FTIC would give

$$\begin{aligned} \int_1^4 \frac{dx}{x^2-5x+6} &= \ln|x-3| - \ln|x-2| \Big|_1^4 \\ &= \ln(1) - \ln(2) - (\ln(2) - \ln(1)) \\ &= -2\ln(2). \end{aligned}$$

However, because $\frac{1}{(x-3)(x-2)}$ has singularities at $x = 2$ and $x = 3$, one must evaluate the improper integral as follows:

$$\int_1^4 \frac{dx}{x^2-5x+6} = \int_1^2 \frac{dx}{x^2-5x+6} + \int_2^3 \frac{dx}{x^2-5x+6} + \int_3^4 \frac{dx}{x^2-5x+6},$$

and none of these integrals exist, as will be shown in the next section on improper integrals, so the entire integral does not exist.

(Return)

8. First, we will try the method of substituting back in before evaluating. We make the substitution

$$\begin{aligned} u &= x - 1 \\ du &= dx. \end{aligned}$$

Then the integral becomes

$$\begin{aligned} \int_{x=0}^{1} x(x-1)^n\, dx &= \int_{x=0}^{1} (u+1)u^n\, du \\ &= \int_{x=0}^{1} u^{n+1} + u^n\, du \\ &= \frac{u^{n+2}}{n+2} + \frac{u^{n+1}}{n+1}\Big|_{x=0}^{1}. \end{aligned}$$

This is where a lot of students might make the mistake of plugging in the limits of x when trying to evaluate the integral. This is a reason to include the $x = a$ at the bottom limit of integration: it helps remind us that those limits are in terms of x, even if we have made one or more substitution.

To avoid this pitfall, we now finish getting the antiderivative in terms of x so that we can evaluate:

$$\begin{aligned}\left.\frac{u^{n+2}}{n+2}+\frac{u^{n+1}}{n+1}\right|_{x=0}^{1} &= \left.\frac{(x-1)^{n+2}}{n+2}+\frac{(x-1)^{n+1}}{n+1}\right|_{x=0}^{1} \\ &= 0+0-\frac{(-1)^{n+2}}{n+2}-\frac{(-1)^{n+1}}{n+1} \\ &= (-1)^{n+2}\left(\frac{1}{n+1}-\frac{1}{n+2}\right) \\ &= \frac{(-1)^{n}}{(n+1)\,(n+2)}.\end{aligned}$$

On the other hand, as soon as we made the substitution $u = x - 1$, we could have changed the limits of integration to reflect our new variable.

Namely, at the lower limit $x = 0$, what is the corresponding value of u? Well, $u = x - 1$, so when $x = 0$, we have $u = -1$. Similarly, when $x = 1$, the corresponding value of u is $u = 0$. So as we were making our substitution, we could make a corresponding change in the limits of integration so that the new definite integral is entirely in terms of u:

$$\int_{x=0}^{1} x(x-1)^n\,dx = \int_{u=-1}^{0}(u+1)u^n\,du.$$

Now the calculation proceeds as before, and gives the same answer. Changing the limits can sometimes be easier, especially in a complicated integral which may involve (for example) a u-substitution and a trigonometric substitution. Otherwise, the algebra can get messy as one substitutes back in and then substitutes back in again.

(Return)

9. The logical choice of parts is

$$\begin{aligned} u &= x & du &= dx \\ dv &= (x-1)^n\,dx & v &= \frac{(x-1)^{n+1}}{n+1}.\end{aligned}$$

Then by the formula for parts, we have

$$\begin{aligned}\int_{x=0}^{1} x(x-1)^n\,dx &= \left. x\frac{(x-1)^{n+1}}{n+1}\right|_{x=0}^{1} - \int_{x=0}^{1}\frac{(x-1)^{n+1}}{n+1}\,dx \\ &= 0-\left.\frac{(x-1)^{n+2}}{(n+1)\,(n+2)}\right|_{x=0}^{1} \\ &= \frac{(-1)^{n+2}}{(n+1)\,(n+2)} \\ &= \frac{(-1)^{n}}{(n+1)\,(n+2)}.\end{aligned}$$

Note that from the first to second line above, we have $x\frac{(x-1)^{n+1}}{n+1}$ is 0 at both $x = 1$ and at $x = 0$, so that entire term disappears.

(Return)

10. This integral is best computed with a substitution of

$$u = \ln x$$
$$du = \frac{1}{x}\,dx$$

(Integration by parts works too, but it involves a little bit of algebra). Here it is convenient to change the limits of integration as we go, so note that when $x = e^3$, we have $u = 3$. When $x = e^5$ we have $u = 5$. Thus,

$$\begin{aligned}
\int_{x=e^3}^{e^5} \frac{\ln x}{x}\,dx &= \int_{u=3}^{5} u\,du \\
&= \frac{1}{2}u^2\Big|_{u=3}^{5} \\
&= \frac{1}{2}(25-9) \\
&= \frac{1}{2}\cdot 16 \\
&= 8.
\end{aligned}$$

(Return)

11. This looks like a good candidate for a trigonometric substitution. Because of the extra factor of 3, the logical substitution is

$$x = \frac{1}{\sqrt{3}}\tan\theta$$
$$dx = \frac{1}{\sqrt{3}}\sec^2\theta\,d\theta.$$

Remember the constant $\frac{1}{\sqrt{3}}$ is there so that when it is squared it will cancel with the factor of 3 and allow us to use the identity $1+\tan^2 = \sec^2$. Again, we will change the bounds as we go. Note that

$$\begin{aligned}
x = -1 \quad &\Rightarrow \quad \frac{1}{\sqrt{3}}\tan\theta = -1 \\
&\Rightarrow \quad \tan\theta = -\sqrt{3}.
\end{aligned}$$

The value of θ for which this holds is $\theta = -\frac{\pi}{3}$. Similarly, when $x = 1$ we have $\theta = \frac{\pi}{3}$. Proceeding, we have

$$\begin{aligned}
\int_{x=-1}^{1} \frac{1}{1+3x^2} &= \frac{1}{\sqrt{3}} \int_{\theta=-\pi/3}^{\pi/3} \frac{\sec^2\theta\, d\theta}{1+3(\frac{1}{\sqrt{3}}\tan\theta)^2} \\
&= \frac{1}{\sqrt{3}} \int_{\theta=-\pi/3}^{\pi/3} \frac{\sec^2\theta\, d\theta}{1+\tan^2\theta} \\
&= \frac{1}{\sqrt{3}} \int_{\theta=-\pi/3}^{\pi/3} \frac{\sec^2\theta}{\sec^2\theta}\, d\theta \\
&= \frac{1}{\sqrt{3}} \int_{\theta=-\pi/3}^{\pi/3} d\theta \\
&= \frac{1}{\sqrt{3}} \theta \Big|_{-\pi/3}^{\pi/3} \\
&= \frac{1}{\sqrt{3}} \left(\frac{\pi}{3} - \frac{-\pi}{3} \right) \\
&= \frac{2\pi}{3\sqrt{3}}
\end{aligned}$$

(Return)

27 Improper Integrals

An *improper integral* is a definite integral which cannot be evaluated using the Fundamental Theorem of Integral Calculus (FTIC). This situation arises because the integral either

1. has a point in its interval of integration which is not in the domain of the integrand (the function being integrated) or

2. has ∞ or $-\infty$ as a bound of integration.

As an example of the first type, consider

$$\int_0^2 \frac{dx}{x}.$$

This integral is improper because the left endpoint, 0, is not in the domain of $\frac{1}{x}$.

Another example of the first type is

$$\int_0^4 \frac{dx}{\sqrt[3]{x-2}}.$$

This is improper because the point 2 is in the interval of integration but is not in the domain of $\frac{1}{\sqrt[3]{x-2}}$.

For an example of the second type, consider

$$\int_0^\infty \frac{dx}{1+x^2}.$$

This is improper because the upper bound is ∞.

For an example of the danger of trying to apply the Fundamental Theorem of Integral Calculus when it does not apply, consider

$$\int_{x=-1}^1 \frac{1}{x^2}\,dx.$$

This is an improper integral (hence FTIC does not apply) because the point $x = 0$ is not in the domain of $\frac{1}{x^2}$. If we were to try to apply FTIC anyway, we would find

$$\begin{aligned}\int_{x=-1}^1 \frac{1}{x^2}\,dx &= -\frac{1}{x}\bigg|_{x=-1}^1 \\ &= -1-1 \\ &= -2.\end{aligned}$$

This is problematic since $\frac{1}{x^2} > 0$ for all x (hence should have a positive integral by the dominance property).

27.1 Dealing with improper integrals

To deal with an improper integral of the first type, first consider the integral $\int_a^b f(x)\,dx$, where a is not in the domain of $f(x)$, but f is continuous on the rest of the interval $(a, b]$. In this situation, one replaces the lower bound with a variable T which is slightly larger than a, and then takes the limit as T approaches a from the right (this is denoted by $T \to a^+$).

$$\int_a^b f(x)\,dx = \lim_{T\to a^+} \int_T^b f(x)\,dx.$$

This replacement allows the integral $\int_T^b f(x)\,dx$ to be computed using FTIC, since f is continuous on that interval. After that integral is computed (in terms of T), the limit is computed. If the limit exists, then the original integral exists and equals the result. If the limit does not exist or is infinite, then the original integral does not exist either.

Example

Determine whether the integral

$$\int_0^2 \frac{dx}{x}$$

exists, and if it exists, what its value is. (See Answer 1)

If the right endpoint, b, of the integral $\int_a^b f(x)\,dx$ is not in the domain of f, then b gets replaced with a variable T slightly smaller than b, and again the integral is replaced with a limit, this time as $T \to b^-$ (that is, the limit as T approaches b from the left).

$$\int_a^b f(x)\,dx = \lim_{T\to b^-} \int_a^T f(x)\,dx.$$

Again, the integral equals this limit, if it exists. If the limit does not exist, then the integral does not exist.

Example

Determine whether the integral

$$\int_1^3 \frac{dx}{(x-3)^2}$$

exists. If it exists, find its value. (See Answer 2)

For integrals where a point inside the interval of integration is not in the domain of the integrand, the integral is first split at the bad point, and then each integral is evaluated separately using the above techniques. So, consider $\int_a^b f(x)\,dx$ where the point c is not in the domain of f and $a < c < b$. Then

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx \\ &= \lim_{T\to c^-} \int_a^T f(x)\,dx + \lim_{U\to c^+} \int_U^b f(x)\,dx.\end{aligned}$$

Both of the resulting limits must exist for the original integral to exist.

Example

Determine if the integral

$$\int_{-1}^{1} \frac{dx}{x^{4/5}}$$

exists. (See Answer 3)

27.2 Bounds at infinity

For integrals with one bound at infinity, the integral is defined as follows.

$$\int_{a}^{\infty} f(x)\,dx = \lim_{T\to\infty} \int_{a}^{T} f(x)\,dx.$$

Similarly,

$$\int_{-\infty}^{a} f(x)\,dx = \lim_{T\to-\infty} \int_{T}^{a} f(x)\,dx.$$

In the case of bounds of ∞ and $-\infty$, one can first split the integral at any real number c, and then compute each integral as above:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{\infty} f(x)\,dx \\ &= \lim_{T\to-\infty} \int_{T}^{c} f(x)\,dx + \lim_{U\to\infty} \int_{c}^{U} f(x)\,dx. \end{aligned}$$

As before, both of these limits must exist for the original integral to exist. It is *not* equivalent to computing a single limit such as

$$\lim_{T\to\infty} \int_{-T}^{T} f(x)\,dx.$$

Example

Determine if the integral

$$\int_{0}^{\infty} \frac{dx}{1+x^2}$$

exists. If it exists, find its value. (See Answer 4)

27.3 The p-integral

One class of improper integrals is common enough to get its own name: *the p-integral*. This is the name given to integrals of the form

$$\int \frac{1}{x^p}\,dx = \int x^{-p}\,dx.$$

There are two versions of this integral that are of interest to us. First, with a limit at infinity, and second with a limit at 0.

Example

Show that

$$\int_{x=1}^{\infty} x^{-p}\,dx = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p \leq 1. \end{cases}$$

(See Answer 5)

Example

Consider the p-integral with a limit at 0:

$$\int_{x=0}^{1} x^{-p}\,dx.$$

This integral is improper because 0 is not in the domain of x^{-p}. Show that

$$\int_{x=0}^{1} x^{-p}\,dx = \begin{cases} \infty & \text{if } p \geq 1 \\ \frac{1}{1-p} & \text{if } p < 1. \end{cases}$$

(See Answer 6)

27.4 Converge or diverge

In some contexts, it is enough to know whether an integral *converges* (has a finite answer) or *diverges* (goes to infinity or does not exist). Using the knowledge of the above p-integrals, and some asymptotic tools from earlier in the course, can help quickly determine whether certain improper integrals converge or diverge.

Example

Determine if

$$\int_{x=0}^{1} \frac{dx}{\sqrt{x^2+x}}$$

converges or diverges. (See Answer 7)

Example

Determine if

$$\int_{x=1}^{\infty} \frac{dx}{\sqrt{x^2+x}}$$

converges or diverges. (See Answer 8)

Example

Determine if

$$\int_{x=-\infty}^{\infty} \frac{2x}{1+x^2}\,dx$$

converges or diverges. (See Answer 9)

27.5 EXERCISES

- Use a Talyor expansion of the integrand at $x = 0$ to determine whether the following integrals converge or diverge:

$$\int_{x=0}^{1} \frac{e^{-x}}{x}\,dx$$

$$\int_{x=0}^{1} \frac{\cos^2 x}{\sqrt{x}}\,dx$$

- Use the asymptotics of the integrand as $x \to \infty$ to determine whether the following integrals converge or diverge:

$$\int_{x=1}^{+\infty} \frac{\sqrt[3]{x+3}}{x^3}\,dx$$

$$\int_{x=1}^{+\infty} \frac{1-5^{-x}}{x}\,dx$$

- Compute the following integrals, if they converge, by evaluating a limit.

$$\int_{x=0}^{+\infty} e^{-x}\sin x\,dx$$

$$\int_{x=1}^{2} \frac{dx}{\sqrt{x-1}}$$

$$\int_{x=0}^{4} \frac{2\,dx}{\sqrt{16-x^2}}$$

- The following integral is improper both at $x = 1$ and at $x \to \infty$:

$$\int_{x=1}^{+\infty} \frac{1}{\sqrt{x^3-1}}\,dx$$

  First, as $x \to \infty$, show that the integrand is $x^{-3/2} + O(x^{-9/2})$, so that it converges at this limit.

  Next, as $x \to 1^+$, show that the integrand is $(3(x-1))^{-1/2} + O((x-1)^{1/2})$, so that it converges at this limit also.

- Consider the following two integrals:

$$I_1 = \int_{x=2}^{+\infty} \frac{dx}{\sqrt{x^3-8}}, \qquad I_2 = \int_{x=2}^{+\infty} \frac{1}{\sqrt{(x-2)^3}}\,dx$$

  One converges and one does not. Which is which and why?

- Until now we have used asymptotic analysis to relate an improper integral to a p-integral. But sometimes the leading order term is not a power. Identify the leading order term as $x \to +\infty$ of the integrand of

$$\int_{x=1}^{+\infty} \frac{1}{\sinh x}\, dx$$

and determine whether the integral converges or diverges.

- For $p \geq 0$ an integer, consider the following integral:

$$I_p = \int_{x=1}^{+\infty} \frac{dx}{\ln^p x}$$

Show that this diverges for any value of p.

Hint 1: think about the growth of $\ln^p x$ as $x \to +\infty$ as compared to polynomial growth.

Hint 2: recall from Lecture 25 that if $g(x) \geq f(x)$ for every $x \in [a, b]$, then

$$\int_{x=a}^{b} g(x)\, dx \geq \int_{x=a}^{b} f(x)\, dx$$

This is also true if the domain of integration is unbounded and the integrals are defined...

- Determine whether the following integral converges or diverges.

$$\int_{2}^{\infty} \frac{1}{(x^5 - 4x^3)^{1/4}}\, dx$$

27.6 Answers to Selected Examples

1. Since the left endpoint is not in the domain of $\frac{1}{x}$, the integral becomes

$$\begin{aligned}\int_0^2 \frac{dx}{x} &= \lim_{T\to 0^+} \int_T^2 \frac{dx}{x} \\ &= \lim_{T\to 0^+} \ln x \Big|_T^2 \\ &= \lim_{T\to 0^+} (\ln(2) - \ln(T)).\end{aligned}$$

This limit does not exist because $\lim_{T\to 0^+} \ln(T)$ diverges. Hence, the original integral does not exist.

(Return)

2. The integral becomes

$$\begin{aligned}\int_1^3 \frac{dx}{(x-3)^2} &= \lim_{T\to 3^-} \int_1^T \frac{dx}{(x-3)^2} \\ &= \lim_{T\to 3^-} -\frac{1}{x-3}\Big|_1^T \\ &= \lim_{T\to 3^-} -\frac{1}{T-3} - (-\frac{1}{-2}).\end{aligned}$$

This limit does not exist since $\lim_{T\to 3^-} \frac{1}{T-3}$ diverges. Hence, the integral does not exist.

(Return)

3. The only point not in the domain of the function $\frac{1}{x^{4/5}}$ is 0. Thus, the integral becomes

$$
\begin{aligned}
\int_{-1}^{1} \frac{dx}{x^{4/5}} &= \int_{-1}^{0} \frac{dx}{x^{4/5}} + \int_{0}^{1} \frac{dx}{x^{4/5}} \\
&= \lim_{T\to 0^-} \int_{-1}^{T} \frac{dx}{x^{4/5}} + \lim_{U\to 0^+} \int_{u}^{1} \frac{dx}{x^{4/5}} \\
&= \lim_{T\to 0^-} 5x^{1/5}\Big|_{-1}^{T} + \lim_{U\to 0^+} 5x^{1/5}\Big|_{u}^{1} \\
&= \lim_{T\to 0^-} \left(5T^{1/5} - 5(-1)^{1/5}\right) + \lim_{U\to 0^+} \left(5 - 5u^{1/5}\right) \\
&= 0 + 5 + 5 - 0 \\
&= 10.
\end{aligned}
$$

So the integral exists and equals 10.

(Return)

4. Because the top bound is ∞, the integral becomes

$$
\begin{aligned}
\int_{0}^{\infty} \frac{dx}{1+x^2} &= \lim_{T\to\infty} \int_{0}^{T} \frac{dx}{1+x^2} \\
&= \lim_{T\to\infty} \arctan(x)\Big|_{0}^{T} \\
&= \lim_{T\to\infty} \arctan(T) - \arctan(0) \\
&= \lim_{T\to\infty} \arctan(T) \\
&= \frac{\pi}{2}.
\end{aligned}
$$

So the integral exists and equals $\frac{\pi}{2}$.

(Return)

5. First, if $p \neq 1$, then we can use the power rule on x^{-p}. Here we find

$$
\begin{aligned}
\int_{x=1}^{\infty} x^{-p}\, dx &= \lim_{T\to\infty} \frac{x^{-p+1}}{-p+1}\Big|_{x=1}^{T} \\
&= \lim_{T\to\infty} \frac{T^{-p+1}}{1-p} - \frac{1}{1-p}.
\end{aligned}
$$

Note that T^{-p+1} diverges to infinity if $p < 1$, since in this case we have a positive power of T, which goes to infinity as T goes to infinity. On the other hand, T^{-p+1} goes to 0 if $p > 1$, since in that case it is a negative power of T. Putting this together with the above equations, we have

$$
\int_{x=1}^{\infty} x^{-p}\, dx = \begin{cases} \frac{1}{p-1} & \text{if } p > 1 \\ \infty & \text{if } p < 1. \end{cases}
$$

Finally, consider the case $p = 1$. In this case,

$$
\begin{aligned}
\int_{x=1}^{\infty} \frac{1}{x}\, dx &= \lim_{T\to\infty} \ln x\Big|_{1}^{T} \\
&= \lim_{T\to\infty} \ln T \\
&= \infty.
\end{aligned}
$$

And so the cases are as claimed above.

(Return)

6. As in the previous example, we first consider when $p \neq 1$ and use the power rule:

$$\int_{x=0}^{1} x^{-p}\,dx = \lim_{T\to 0^+} \left.\frac{x^{1-p}}{1-p}\right|_{x=T}^{1}$$
$$= \lim_{T\to 0^+} \frac{1}{1-p} - \frac{T^{1-p}}{1-p}.$$

Now, note that the limit is as $T \to 0^+$. So we have T^{1-p} diverges if $p > 1$ (since a negative power of T diverges as $T \to 0^+$), and converges to 0 if $p < 1$. Putting this together with the previous computation gives

$$\int_{x=0}^{1} x^{-p}\,dx = \begin{cases} \frac{1}{1-p} & \text{if } p1. \end{cases}$$

Finally, if $p = 1$, then

$$\int_{x=0}^{1} \frac{1}{x}\,dx = \lim_{T\to 0^+} \ln x\Big|_{x=T}^{1}$$
$$= \lim_{T\to 0^+} 0 - \ln T$$

which diverges to infinity. Thus, the original integral diverges for all $p \geq 1$, as claimed.

(Return)

7. This is not a function for which we can easily find an antiderivative. However, since we are interested in the behavior of the function near 0 (that is where the blow-up occurs), we can do a little bit of algebra to see that

$$\frac{1}{\sqrt{x^2+x}} = \frac{1}{\sqrt{x}} \cdot \frac{1}{\sqrt{x+1}}$$
$$= \frac{1}{\sqrt{x}} \cdot (1+x)^{-1/2}$$
$$= \frac{1}{\sqrt{x}} \left(1 + O(x)\right),$$

by the binomial expansion. Therefore, the leading order term of this function as $x \to 0$ is $\frac{1}{\sqrt{x}}$. This is a convergent p-integral, because $p = \frac{1}{2}$ and from the above example

$$\int_{x=0}^{1} \frac{dx}{x^p}$$

converges when $p < 1$.

(Return)

8. We must do a slightly different analysis for this integral, since we are interested in the behavior as $x \to \infty$. Because we want to take advantage of the binomial series again (which requires its argument to be less

than 1), we do the following algebra:

$$\begin{aligned}\frac{1}{\sqrt{x^2+x}} &= \frac{1}{\sqrt{x^2}} \cdot \frac{1}{\sqrt{1+x^{-1}}} \\ &= \frac{1}{x}\left(1+\frac{1}{x}\right)^{-1/2} \\ &= \frac{1}{x}\left(1+O\left(\frac{1}{x}\right)\right).\end{aligned}$$

So for this integral, the leading order term is $\frac{1}{x}$, which is the p-integral with $p = 1$. This diverges, as the above example shows, and so the original integral diverges.

(Return)

9. As mentioned above, it is not valid to do this with a single limit such as

$$\lim_{T\to\infty} \int_{x=-T}^{T} \frac{2x}{1+x^2}\, dx.$$

The integral must be split and treated as two separate integrals with limits at infinity:

$$\int_{x=-\infty}^{\infty} \frac{2x}{1+x^2}\, dx = \lim_{S\to-\infty} \int_{x=S}^{0} \frac{2x}{1+x^2}\, dx + \lim_{T\to\infty} \int_{x=0}^{T} \frac{2x}{1+x^2}\, dx.$$

Again, a little bit of asymptotic analysis with the help of the geometric series helps determine the behavior of this function:

$$\begin{aligned}\frac{2x}{1+x^2} &= \frac{2x}{1+x^2} \cdot \frac{x^{-2}}{x^{-2}} \\ &= \frac{2}{x} \cdot \frac{1}{1+x^{-2}} \\ &= \frac{2}{x}\left(1+O(x^{-2})\right).\end{aligned}$$

Here, the geometric series was used to write

$$\frac{1}{1+x^{-2}} = 1 - x^{-2} + x^{-4} - \cdots = 1 + O(x^{-2}),$$

which is justified since $x \to \infty$ and so x^{-2} is very small. So the original integrand behaves like $\frac{2}{x}$, which diverges when integrated to infinity. Therefore

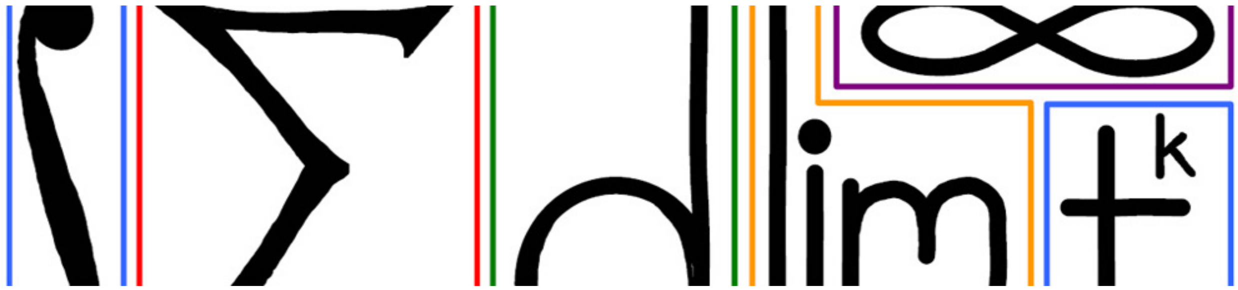
$$\lim_{T\to\infty} \int_{x=0}^{T} \frac{2x}{1+x^2}\, dx$$

diverges, and (but for a sign change) the same reasoning shows

$$\lim_{S\to-\infty} \int_{x=S}^{0} \frac{2x}{1+x^2}\, dx$$

diverges too, so the original integral diverges.

(Return)



28 Trigonometric Integrals

A trigonometric integral is an integral involving products and powers of trigonometric functions: cosine, sine, tangent, secant, cosecant, and cotangent. Many of these integrals can be handled with u-substitution, but there are other methods which are outlined in this module. The three families of integrals discussed in this module are

$$\int \sin^m\theta\cos^n\theta\,d\theta$$
$$\int \tan^m\theta\sec^n\theta\,d\theta$$
$$\int \sin(m\theta)\cos(n\theta)\,d\theta.$$

28.1 Product of sines and cosines

Consider the integral

$$\int \sin^m\theta\cos^n\theta\,d\theta.$$

There are several cases to consider based on whether m and n are odd and even.

m is odd

If m is odd, then one factor of $\sin\theta$ can be set aside. This leaves behind an even power of $\sin\theta$, which can be expressed in terms of $\cos\theta$ using the Pythagorean identity. Then the substitution $u = \cos\theta$ can be made.

Example

Find

$$\int \sin^3(x)\cos(x)\,dx.$$

(See Answer 1)

n is odd

If n is odd, the procedure is very similar. This time, we set aside a factor of $\cos\theta$. This leaves an even power of $\cos\theta$ which can be expressed in terms of $\sin\theta$ using the Pythagorean identities.

Example

Find
$$\int \sin^2(x)\cos^3(x)\,dx.$$
(See Answer 2)

Both m and n are even

If neither m nor n is odd, then both are even. This is a bit more difficult and requires using the power reduction formulas:

Power reduction
$\sin^2(\theta) = \frac{1-\cos(2\theta)}{2}$
$\cos^2(\theta) = \frac{1+\cos(2\theta)}{2}$

Example

Find
$$\int \sin^2 x\,dx.$$
(See Answer 3)

Example

Find
$$\int \sin^2\theta\cos^2\theta\,d\theta.$$
(See Answer 4)

Example

Find
$$\int \cos^4(x)\,dx.$$
(See Answer 5)

28.2 Product of tangents and secants

Next consider the integral
$$\int \tan^m\theta\sec^n\theta\,d\theta.$$
As with the product of sines and cosines, the method will depend on whether m and n are odd or even.

m is odd

If m is odd, we will set aside a factor of $\tan\theta\sec\theta$. Note that this is the derivative of $\sec\theta$ and so this sets up a substitution of $u = \sec\theta$. After setting aside these factors, we are left with an even power of $\tan\theta$, which can be expressed in terms of $\sec\theta$ using the Pythagorean identity

$$\tan^2\theta = \sec^2\theta - 1.$$

Now, the integral can be computed using the substitution $u = \sec\theta$.

Example

Compute

$$\int \tan^3\theta \sec\theta \, d\theta.$$

(See Answer 6)

n is even

If n is even, then we can set aside a factor of $\sec^2\theta$. Note that this is the derivative of $\tan\theta$ and therefore sets up the substitution $u = \tan\theta$. Setting aside $\sec^2\theta$ leaves an even power of $\sec\theta$, which can be expressed in terms of $\tan\theta$ using the Pythagorean identity

$$\sec^2\theta = 1 + \tan^2\theta.$$

Then the substitution $u = \tan\theta$ allows the computation of the integral.

Example

Compute

$$\int \tan^2\theta \sec^6\theta \, d\theta.$$

(See Answer 7)

m is even, n is odd

If neither of the above cases holds, then the integral is a bit more difficult. It typically requires a bit of algebra and several applications of a reduction formula (or integration by parts). A general method is to rewrite the even power of tangent entirely in terms of secant by using the Pythagorean identity

$$\tan^2\theta = \sec^2\theta - 1.$$

This gives an integral which is sums of powers of $\sec\theta$. Each of these can be solved using the reduction formula for secant:

$$\int \sec^n\theta \, d\theta = \frac{1}{n-1}\sec^{n-2}\theta\tan\theta + \frac{n-2}{n-1}\int \sec^{n-2}\theta \, d\theta,$$

along with the fact that

$$\int \sec\theta \, d\theta = \ln|\sec\theta + \tan\theta| + C.$$

Example

Compute

$$\int \tan^2\theta \sec\theta \, d\theta.$$

(See Answer 8)

28.3 Product of sine and cosine with constants

Finally, consider the integral

$$\int \sin(m\theta)\cos(n\theta)\, d\theta.$$

This integral requires some algebra to simplify the integrand. One can verify using the sum and difference formulas for sine that

$$\sin(m\theta)\cos(n\theta) = \frac{1}{2}\left(\sin((m+n)\theta) + \sin((m-n)\theta)\right).$$

This expression can be integrated term by term to find

$$\begin{aligned} \int \sin(m\theta)\cos(n\theta)\, d\theta &= \int \frac{1}{2}\left(\sin((m+n)\theta) + \sin((m-n)\theta)\right)\, d\theta \\ &= \frac{1}{2}\left(-\frac{\cos((m+n)\theta)}{m+n} - \frac{\cos((m-n)\theta)}{m-n}\right) + C. \end{aligned}$$

There are similar formulas for related integrals:

$$\begin{aligned} \int \sin(m\theta)\sin(n\theta)\, d\theta &= -\frac{\sin((m+n)\theta)}{2\,(m+n)} + \frac{\sin((m-n)\theta)}{2\,(m-n)} + C \\ \int \cos(m\theta)\cos(n\theta)\, d\theta &= \frac{\sin((m+n)\theta)}{2\,(m+n)} + \frac{\sin((m-n)\theta)}{2\,(m-n)} + C. \end{aligned}$$

These formulas need not be memorized, but be aware they exist and look them up when necessary.

Example

Compute

$$\int \sin(3\theta)\cos(4\theta)\, d\theta.$$

(See Answer 9)

28.4 Additional examples

Example

Compute

$$\int_{\theta=-\pi/2}^{\pi/2} \cos^n\theta \, d\theta.$$

Use the fact (which is proven using integration by parts) that

$$\int \cos^n \theta \, d\theta = \frac{\cos^{n-1}\theta \sin\theta}{n} + \frac{n-1}{n} \int \cos^{n-2}\theta \, d\theta.$$

(See Answer 10)

Example

Compute

$$\int \tan^3 \theta \, d\theta$$

(See Answer 11)

28.5 EXERCISES

Compute the following indefinite integrals. You may need to use reduction formulae or coordinate changes.

- $\int \sin^2 x \cos^2 x \, dx$
- $\int \sin^3 \frac{x}{2} \cos^3 \frac{x}{2} \, dx$
- $\int \frac{x^3 \, dx}{\sqrt{9 - x^2}}$
- $\int 5 \tan^5 x \sec^3 x \, dx$
- $\int 7 \tan^4 x \sec^4 x \, dx$
- $\int 9 \sin^3 3x \, dx$
- $\int \cos^4 x \, dx$
- $\int \sin x \sec x \tan x \, dx$
- $\int \tan^5 2x \sec^4 2x \, dx$
- $\int \cos x \sqrt{1 - \sin x} \, dx$
- $\int \frac{x^2 \, dx}{\sqrt{1 + x^2}}$
- $\int \tan^4(2x) \sec^4(2x) \, dx$

28.6 Answers to Selected Examples

1. Following the above outline, set aside one factor of $\sin x$, which gives

$$\int \sin^3 x \cos x \, dx = \int (\sin^2 x)(\cos x)(\sin x) \, dx.$$

Now, there is an even power of sine remaining, which can be rewritten using the Pythagorean identity

$$\sin^2 x = 1 - \cos^2 x.$$

This gives

$$\begin{aligned}\int \sin^3 x \cos x \, dx &= \int (\sin^2 x)(\cos x) \sin(x) \, dx \\ &= \int (1 - \cos^2 x)(\cos x) \sin(x) \, dx.\end{aligned}$$

Now, the integral can be handled by letting $u = \cos(x)$ (and $du = -\sin(x) dx$).

$$\begin{aligned}\int (1 - \cos^2 x)(\cos x)(\sin x) \, dx &= \int (1 - u^2) u (-du) \\ &= \int (u^3 - u) \, du \\ &= \frac{u^4}{4} - \frac{u^2}{2} + C \\ &= \frac{\cos^4 x}{4} - \frac{\cos^2 x}{2} + C.\end{aligned}$$

(Return)

2. Following the procedure outlined above, we set aside a factor of cosine and use the Pythagorean identity, which gives

$$\begin{aligned}\int \sin^2(x) \cos^3(x) \, dx &= \int \sin^2(x) \cos^2(x) \cos(x) \, dx \\ &= \int \sin^2(x)(1 - \sin^2 x) \cos(x) \, dx.\end{aligned}$$

Now, we are ready to make the substitution $u = \sin(x)$, $du = \cos(x) \, dx$. This gives

$$\begin{aligned}\int \sin^2(x)(1 - \sin^2 x) \cos(x) \, dx &= \int u^2(1 - u^2) \, du \\ &= \int u^2 - u^4 \, du \\ &= \frac{u^3}{3} - \frac{u^5}{5} + C \\ &= \frac{\sin^3 x}{3} - \frac{\sin^5 x}{5} + C.\end{aligned}$$

(Return)

3. Using the first power reduction formula gives

$$\begin{aligned}\int \sin^2 x\,dx &= \int \frac{1}{2}(1-\cos(2x))\,dx \\ &= \frac{1}{2}\left(x - \frac{\sin(2x)}{2}\right) + C.\end{aligned}$$

(Return)

4. Using both the power reduction formulas and doing some algebra gives

$$\begin{aligned}\int \sin^2\theta\cos^2\theta\,d\theta &= \int \frac{1}{2}\left(1-\cos 2\theta\right)\frac{1}{2}\left(1+\cos 2\theta\right)\,d\theta \\ &= \frac{1}{4}\int (1-\cos 2\theta)(1+\cos 2\theta)\,d\theta \\ &= \frac{1}{4}\int (1-\cos^2 2\theta)\,d\theta \\ &= \frac{1}{4}\int \sin^2 2\theta\,d\theta \\ &= \frac{1}{4}\int \frac{1}{2}\left(1-\cos 4\theta\right)\,d\theta \\ &= \frac{1}{8}\left(\theta - \frac{\sin 4\theta}{4}\right) + C.\end{aligned}$$

(Return)

5. Using the second power reduction formula (and then again in a later step) gives

$$\begin{aligned}\int \cos^4(x)\,dx &= \int (\cos^2(x))^2\,dx \\ &= \int (\frac{1}{2}(1+\cos(2x))^2\,dx \\ &= \frac{1}{4}\int \left(1+2\cos(2x)+\cos^2(2x)\right)\,dx \\ &= \frac{1}{4}\int \left(1+2\cos(2x)+\frac{1}{2}(1+\cos(4x))\right)\,dx \\ &= \frac{1}{4}\left(x+\sin(2x)+\frac{1}{2}x+\frac{1}{8}\sin(4x)\right)+C\end{aligned}$$

(Return)

6. Since the power of tangent is odd, we set aside a factor of $\tan\theta\sec\theta$, and use the Pythagorean identity to find

$$\begin{aligned}\int \tan^3\theta\sec\theta\,d\theta &= \int \tan^2\theta(\tan\theta\sec\theta)\,d\theta \\ &= \int (\sec^2\theta - 1)(\tan\theta\sec\theta)d\theta.\end{aligned}$$

Now, we can make the substitution

$$\begin{aligned}u &= \sec\theta \\ du &= \sec\theta\tan\theta\,d\theta,\end{aligned}$$

which gives

$$\begin{aligned}\int (\sec^2\theta - 1)(\tan\theta\sec\theta)d\theta &= \int (u^2-1)\,du\\ &= \frac{1}{3}u^3 - u + C\\ &= \frac{1}{3}\sec^3\theta - \sec\theta + C.\end{aligned}$$

(Return)

7. Because the power of secant is even, we set aside $\sec^2\theta$ and use the Pythagorean identity to find

$$\begin{aligned}\int \tan^2\theta\sec^6\theta\,d\theta &= \int \tan^2\theta\sec^4\theta(\sec^2\theta)\,d\theta\\ &= \int \tan^2\theta(1+\tan^2\theta)^2(\sec^2\theta)\,d\theta.\end{aligned}$$

Now, we are prepared for a substitution of

$$\begin{aligned}u &= \tan\theta\\ du &= \sec^2\theta\,d\theta.\end{aligned}$$

Making this substitution and simplifying gives

$$\begin{aligned}\int \tan^2\theta(1+\tan^2\theta)^2(\sec^2\theta)\,d\theta &= \int u^2(1+u^2)^2\,du\\ &= \int u^2(1+2u^2+u^4)\,du\\ &= \int (u^2+2u^4+u^6)\,du\\ &= \frac{1}{3}u^3 + \frac{2}{5}u^5 + \frac{1}{7}u^7 + C\\ &= \frac{1}{3}\tan^3\theta + \frac{2}{5}\tan^5\theta + \frac{1}{7}\tan^7\theta + C.\end{aligned}$$

(Return)

8. Using the Pythagorean identity gives

$$\begin{aligned}\int \tan^2\theta\sec\theta\,d\theta &= \int (\sec^2\theta - 1)\sec\theta\,d\theta\\ &= \int (\sec^3\theta - \sec\theta)\,d\theta\\ &= \int \sec^3\theta\,d\theta - \int \sec\theta\,d\theta.\end{aligned}$$

Now, using the reduction formula on the first of these integrals gives

$$\int \sec^3\theta\,d\theta = \frac{1}{2}\sec\theta\tan\theta + \frac{1}{2}\int \sec\theta\,d\theta.$$

Combining this with the above expression and using the integral of secant, we find

$$\begin{aligned}\int \tan^2\theta \sec\theta \, d\theta &= \int \sec^3\theta \, d\theta - \int \sec\theta \, d\theta \\ &= \left(\frac{1}{2}\sec\theta\tan\theta + \frac{1}{2}\int \sec\theta \, d\theta\right) - \int \sec\theta \, d\theta \\ &= \frac{1}{2}\sec\theta\tan\theta - \frac{1}{2}\int \sec\theta \, d\theta \\ &= \frac{1}{2}\sec\theta\tan\theta - \frac{1}{2}\ln|\sec\theta + \tan\theta| + C.\end{aligned}$$

(Return)

9. Using the formula given, we have

$$\begin{aligned}\int \sin(3\theta)\cos(4\theta)\, d\theta &= -\frac{\cos(7\theta)}{14} - \frac{\cos(-\theta)}{-2} + C \\ &= -\frac{\cos(7\theta)}{14} + \frac{\cos\theta}{2} + C,\end{aligned}$$

where we have used the fact that cosine is even to simplify the final expression.

(Return)

10. Applying the limits of integration in the above reduction formula gives

$$\int_{\theta=-\pi/2}^{\pi/2} \cos^n\theta \, d\theta = \left.\frac{\cos^{n-1}\theta\sin\theta}{n}\right|_{-\pi/2}^{\pi/2} + \frac{n-1}{n}\int_{\theta=-\pi/2}^{\pi/2} \cos^{n-2}\theta \, d\theta.$$

Now, notice that

$$\left.\frac{\cos^{n-1}\theta\sin\theta}{n}\right|_{-\pi/2}^{\pi/2} = 0$$

because cosine is 0 at $\pm\pi/2$. Therefore,

$$\int_{\theta=-\pi/2}^{\pi/2} \cos^n\theta \, d\theta = \frac{n-1}{n}\int_{\theta=-\pi/2}^{\pi/2} \cos^{n-2}\theta \, d\theta.$$

This is itself a reduction formula. By computing the base cases $n=0$ and $n=1$, respectively, we find

$$\begin{aligned}\int_{\theta=-\pi/2}^{\pi/2} d\theta &= \theta\Big|_{-\pi/2}^{\pi/2} \\ &= \pi,\end{aligned}$$

and

$$\begin{aligned}\int_{\theta=-\pi/2}^{\pi/2} \cos\theta \, d\theta &= \sin\theta\Big|_{-\pi/2}^{\pi/2} \\ &= 2.\end{aligned}$$

Now the value of the integral for any higher value of n can be found by repeatedly using the above formula until the integral reduces to one of the base cases above. Using induction, one finds

$$\int_{\theta=-\pi/2}^{\pi/2} \cos^n \theta \, d\theta = \begin{cases} \frac{1\cdot 3\cdot 5\cdots(n-1)}{2\cdot 4\cdot 6\cdots n}\pi & \text{if } n \text{ is even} \\ \frac{2\cdot 4\cdot 6\cdots(n-1)}{3\cdot 5\cdot 7\cdots n}\cdot 2 & \text{if } n \text{ is odd.} \end{cases}$$

(Return)

11. Here the power of tangent is odd, so the method calls for setting aside a factor of $\sec\theta\tan\theta$. However, there is no factor of secant in this integral! It turns out that this is not a problem; we can multiply the top and the bottom by secant to introduce a factor of secant, and the algebra works out:

$$\begin{aligned} \int \tan^3\theta \, d\theta &= \int \frac{\tan^2\theta}{\sec\theta}(\sec\theta\tan\theta)\, d\theta \\ &= \int \frac{\sec^2\theta - 1}{\sec\theta}(\sec\theta\tan\theta)\, d\theta \\ &= \int \left(\sec\theta - \frac{1}{\sec\theta}\right)(\sec\theta\tan\theta)\, d\theta. \end{aligned}$$

Now, proceed as usual with the substitution

$$\begin{aligned} u &= \sec\theta \\ du &= \sec\theta\tan\theta\, d\theta. \end{aligned}$$

Thus,

$$\begin{aligned} \int \left(\sec\theta - \frac{1}{\sec\theta}\right)(\sec\theta\tan\theta)\, d\theta &= \int \left(u - \frac{1}{u}\right) du \\ &= \frac{1}{2}u^2 - \ln|u| + C \\ &= \frac{1}{2}\sec^2\theta - \ln|\sec\theta| + C. \end{aligned}$$

(Return)

29 Tables And Computers

This module discusses some of the shortcuts available by using tables of integrals and mathematical software. Many integrals can be easily computed by a computer algebra system, but it is still important to know the underlying concepts so as to be able to use these tools efficiently and accurately.

29.1 Tables of integrals

Most calculus textbooks have an appendix containing one hundred or more integral formulas. All of these formulas can be derived using the techniques of the previous modules (possibly with some additional technical algebra), and it is a good exercise to try to derive some of them.

Using the table is sometimes difficult, because finding the correct form can be tricky. And even with the correct form, an integral may not match the form precisely. It may take some algebra and a u-substitution to match the form.

Example

Use the formula

$$\int \frac{dx}{x^2\sqrt{x^2-a^2}} = \frac{\sqrt{x^2-a^2}}{a^2x} + C$$

to evaluate the integral

$$\int \frac{dx}{(4x^2+4x+1)\sqrt{4x^2+4x-3}}.$$

(See Answer 1)

Other table entries are more inductive in nature, like the reduction formulas mentioned in the integration by parts module.

Example

Use the formulas

$$\int \sec ax\, dx = \frac{1}{a}\ln|\sec ax + \tan ax| + C,$$

and (for $n \geq 2$)

$$\int \sec^n ax\, dx = \frac{\sec^{n-2} ax \tan ax}{a(n-1)} + \frac{n-2}{n-1}\int \sec^{n-2} ax\, dx + C$$

to find

$$\int \sec^3(x)\,dx.$$

(See Answer 2)

29.2 Mathematical software

Expensive computer algebra systems such as *Maple* and *Mathematica* can quickly and accurately dispense with most integrals that can be done by hand. One free alternative, from the makers of Mathematica, is Wolfram Alpha, which for most purposes is as good as its more costly relatives, and in many cases it can explain the intermediate steps of longer computations (though it now only provides three such explanations per day for a user without a paid subscription).

Note that the form of the answer given by computer systems may look different from what one gets by hand or by a table, so care should be taken when comparing answers.

Example

Compute $\int \sec^3(x)\,dx$ using Wolfram Alpha, or other computer algebra system. Note the syntax of the entry (though it is pretty good at parsing other forms of entry). Also note that the answer given is in a different form than that found in the earlier example.

Answer

Example

There are limits to what a computer algebra system can do. Consider the integral

$$\int_{x=0}^{\pi/2} \frac{\sin^n x}{\sin^n x + \cos^n x}\,dx.$$

It turns out that the value of this integral is $\frac{\pi}{4}$ for all n, although Wolfram Alpha is not able to compute it.

But if we enter small particular values of n into Wolfram Alpha, then it does give the answer, although sometimes only in decimal form.

29.3 Answers to Selected Examples

1. Although it is not exactly in the correct form, completing the square should get it closer. Indeed, factoring and completing the square gives

$$\int \frac{dx}{(4x^2+4x+1)\sqrt{4x^2+4x-3}} = \int \frac{dx}{(2x+1)^2\sqrt{(2x+1)^2-4}}.$$

Now, making the u-substitution $u = 2x+1$ (hence $dx = \frac{1}{2}\,du$) gives

$$\begin{aligned}\int \frac{dx}{(2x+1)^2\sqrt{(2x+1)^2-4}} &= \frac{1}{2}\int \frac{du}{u^2\sqrt{u^2-4}} \\ &= \frac{1}{2}\frac{\sqrt{u^2-4}}{4u} + C \\ &= \frac{1}{2}\cdot\frac{\sqrt{(2x+1)^2-4}}{4\,(2x+1)} + C.\end{aligned}$$

(Return)

2. Reducing using the second formula, and then using the first formula gives

$$\begin{aligned}\int sec^3 x &= \frac{\sec x \tan x}{2} + \frac{1}{2}\int \sec x \, dx \\ &= \frac{\sec x \tan x}{2} + \frac{1}{2}\ln|\sec x + \tan x| + C.\end{aligned}$$

(Return)

30 Simple Areas

We know the basic standard formulae for the area of basic shapes, but why are they true? From the point of view of calculus, area A is the integral of dA, the *area element*.

In this chapter, we will use the following procedure to determine a quantity U:
1. Determine the differential element dU.
2. Integrate to compute $U = \int dU$.

30.1 Length of an interval

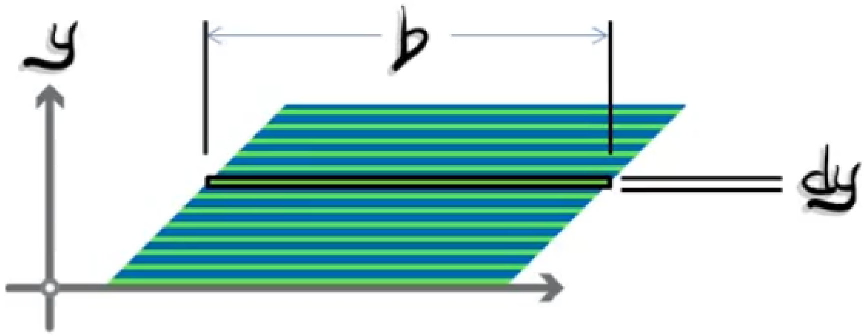
Before getting to areas, first consider how this method works for computing the length L of the interval from a to b. If the length is denoted L, then the length element will be denoted dL, and $L = \int dL$. In this context, the appropriate length element would be dx if we're working along the x-axis.

So, we want to integrate dx as x goes from a to b.
The length,

$$\begin{aligned} L &= \int dL \\ &= \int_{x=a}^{b} dx \\ &= x\Big|_{x=a}^{b} \\ &= b - a \end{aligned}$$

30.2 Parallelogram

The formula for the area of a parallelogram is base $\times$ height (bh). Consider the following rearrangement into differential elements, where we carve the parallelogram into parallel horizontal strips of width b and height dy, where y is the y-axis.



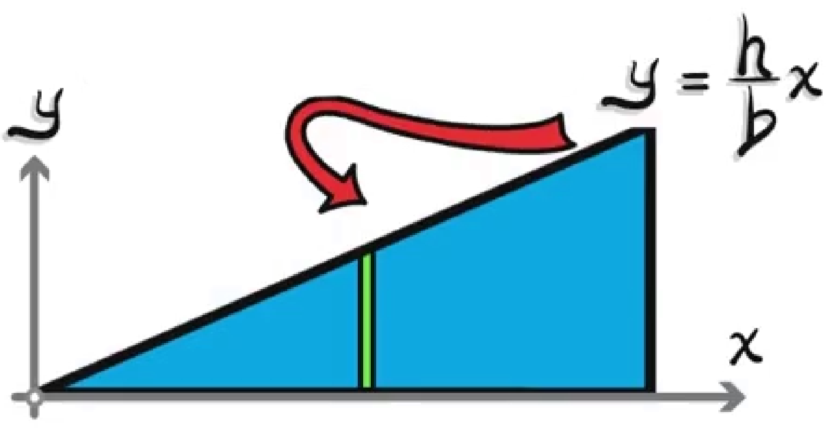
In this case, the area element, $dA = b\,dy$, is the area of this infinitesimal rectangle. The limits on y should go from 0 to the height, h of the parallelogram.
The area,

$$\begin{aligned} A &= \int dA \\ &= \int_{y=0}^{h} b\,dy \\ &= by\Big|_{y=0}^{h} \\ &= bh \end{aligned}$$

We have our familiar answer bh. This means that we've done a rearrangement in terms of infinitesimal strips. Shearing that parallelogram preserves the area element and hence, the area. That is why a parallelogram has the same area as the corresponding rectangle.

30.3 Triangle

The formula for the area of a triangle is $\frac{1}{2}\times$ base $\times$ height ($\frac{1}{2}bh$). Let's think in terms of a differential area element. Given the fact that we can shear and preserve the area element, and thus the area, let's present our triangle as having a hypotenuse modeled by the line $y = \frac{h}{b}x$.



To compute the area element, let's use a vertical strip.

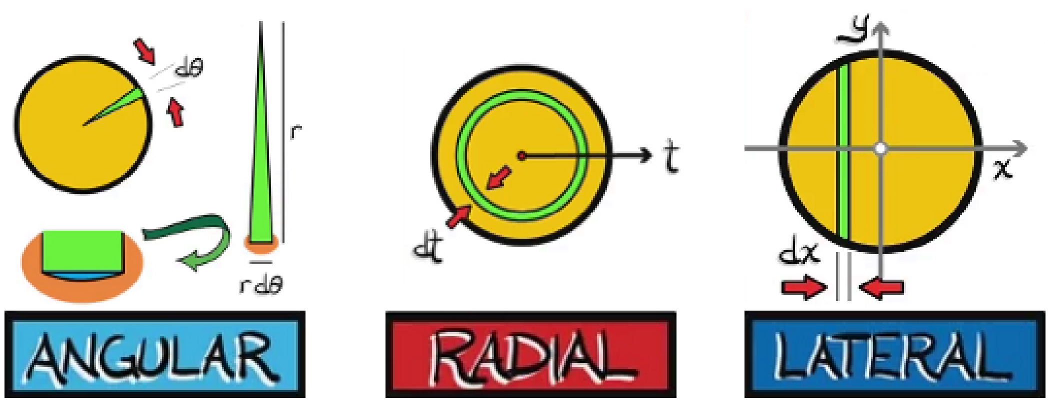
$$dA = \frac{h}{b}x\,dx$$

where the height of that vertical strip is $\frac{h}{b}x$ and the width is the length element dx.
The area,

$$\begin{aligned} A &= \int dA \\ &= \int_{x=0}^{b} \frac{h}{b}x\,dx \\ &= \frac{h}{b}\frac{x^2}{2}\bigg|_{x=0}^{b} \\ &= \frac{hb^2}{2b} \\ &= \frac{1}{2}bh \end{aligned}$$

30.4 Disc

We will use three ways to find the area of a circular disc of radius r:
1. Using an angular area element.
2. Using a radial variable.
3. Using a lateral, or a vertical rectangular strip.



1. **Angular** In this case, we'll use an angular area element. We will take a wedge with angle $d\theta$. If we look at that close up, it's modeled fairly well as a triangle. It's not a perfect triangle, there's a bit of curvature at the end. This is a triangle with two sides of length r whose included angle is $d\theta$. Such a triangle has area $\frac{1}{2}r^2\sin(d\theta) \approx \frac{1}{2}r^2 d\theta$, since $d\theta$ is very small. If we model that as a triangle with height r, and width $r\,d\theta$, we can ignore the higher order terms in the Taylor expansion of that area. We obtain an area element $dA = \frac{1}{2}r(r\,d\theta)$.

   Integrating to get the area, θ has to spin all the way around the circle from 0 to 2π.

The area,

$$
\begin{aligned}
A &= \int dA \\
&= \int_{\theta=0}^{2\pi} \frac{1}{2} r^2 \, d\theta \\
&= \frac{1}{2} r^2 \int_{\theta=0}^{2\pi} d\theta \\
&= \frac{1}{2} r^2 \theta \Big|_{\theta=0}^{2\pi} \\
&= \frac{1}{2} r^2 (2\pi) \\
&= \pi r^2.
\end{aligned}
$$

2. **Radial** Let's consider a radial variable. We can sweep out the area of the circular disk using annuli with a radial coordinate t. Then, we're looking at an annular strip of width dt. The corresponding area element is the circumference $(2\pi t)$ $\times$ thickness (dt).

$$dA = 2\pi t \, dt$$

Integrating this from 0 to the radius r gives us the area.

$$
\begin{aligned}
A &= \int dA \\
&= \int_{t=0}^{r} 2\pi t \, dt \\
&= \pi t^2 \Big|_{t=0}^{r} \\
&= \pi r^2.
\end{aligned}
$$

3. **Lateral** We will use a vertical rectangular strip. Again, it is not a perfect rectangle and there's a little bit of curvature at the end. But, these are higher order terms, and we just care about the differential element. So, using a vertical strip with width dx, and knowing that the formula for the boundary circle is $x^2 + y^2 = r^2$, we solve for y along the upper and lower branches.

$$y = \pm\sqrt{r^2 - x^2}$$

We then obtain an area element that is the area of this rectangular strip.

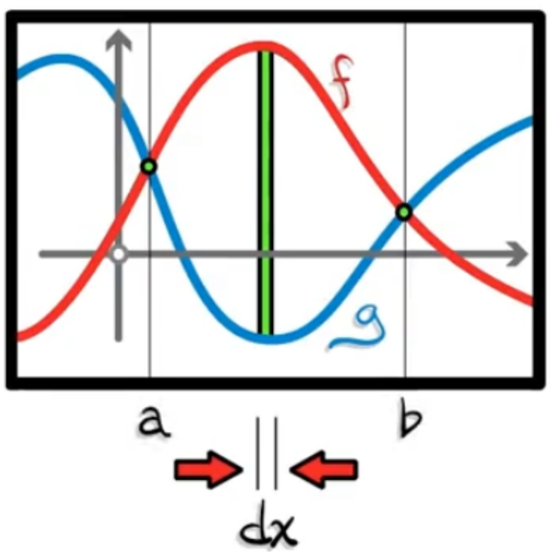
$$dA = 2\sqrt{r^2 - x^2} \, dx$$

In the case of strips, assume the circle is centered at the origin, and let x keep track of where the strip intersects the x-axis. Thus, x ranges from $-r$ to r. Integrating, and using a trigonometric substitution

$x = r \sin u$ gives

$$
\begin{aligned}
A &= \int dA \\
&= \int_{-r}^{r} 2\sqrt{r^2 - x^2}\, dx \\
&= 2\int_{-\pi/2}^{\pi/2} \sqrt{r^2(1 - \sin^2 u)}\; r \cos u \, du \\
&= 2r^2 \int_{-\pi/2}^{\pi/2} \cos^2 u \, du \\
&= 2r^2 \int_{-\pi/2}^{\pi/2} \frac{1}{2}(1 + \cos(2u)) du \\
&= r^2 (u + \frac{1}{2} \sin\, 2u)\Big|_{-\pi/2}^{\pi/2} \\
&= \pi r^2
\end{aligned}
$$

30.5 The area between two curves



Let's say the f is on top and the g is below. Then as we sweep a vertical strip from left to right, we obtain the area. In this case, the area element is a vertical rectangle of width dx and of height $f(x) - g(x)$, the length of the interval between the two.

$$dA = (f(x) - g(x))dx$$

The general formula for the area between two curves $f(x)$ and $g(x)$,

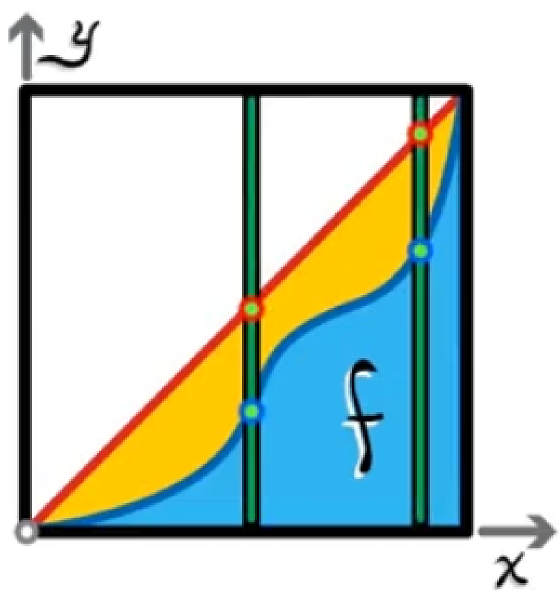
$$\begin{aligned} A &= \int dA \\ &= \int_{x=a}^{b} (f(x) - g(x))dx \end{aligned}$$

Example

Find the area of the region bounded above by $f(x) = 4 + x - x^2$ and below by $g(x) = 1 - x$. (See Answer 1)

30.6 Gini Index (An application of area formula)

In economics, this ratio is used to quantify income inequality in a population.



Let $f(x)$ = Fraction of total income earned by the lowest x fraction of the populace, $0 < x < 1$.

The Gini index quantifies how far f is from a "flat" distribution. This means that f(0) = 0, f(1) = 1.

$$f$$

is probably going to be below the flat distribution where $y = x$, the lowest x fraction earns the lowest x fraction.

The Gini index, $G(f)$ is measuring the difference between these two distributions, in terms of area. It's the ratio of the area between the flat distribution $y = x$ and the given population's income distribution $y = f(x)$. One normalizes that by the area between the flat distribution $y = x$ and $y = 0$, namely the area of that triangle, or

$\frac{1}{2}$.

$$G(f) = \frac{\text{Area between the } y = x \text{ and } y = f(x)}{\text{Area between } y = x \text{ and } y = 0}$$
$$= 2\int_{x=0}^{1} (x - f(x))\, dx.$$

Example

Compute G for a power law distribution $f(x) = x^n$. (See Answer 2)

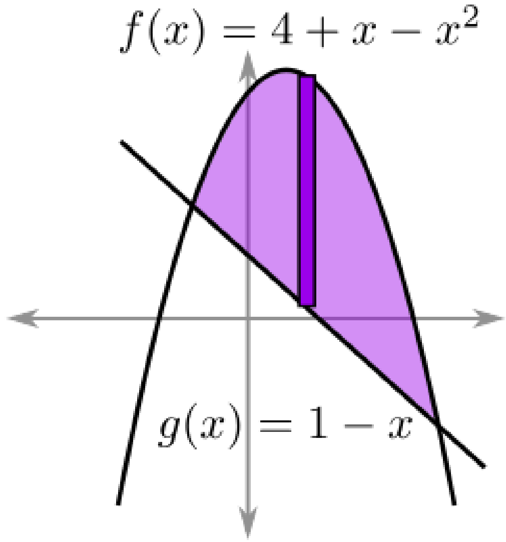
The Gini Index doesn't tell you the income distribution, but we could approximate it in the assumption of a power law. For example, in the year 2010, in the state of New York in USA, the Gini Index was very close to $\frac{1}{2}$. If we assume that it went by a power law distribution, that would imply a cubic distribution of income.

30.7 EXERCISES

- What is the area between the curve $f(x) = \sin^3 x$ and the x-axis from $x = 0$ to $x = \dfrac{\pi}{3}$?
- Find the area of the bounded region enclosed by the curves $y = \sqrt{x}$ and $y = x^2$.
- What is the area between the curve $y = \sin x$ and the x-axis for $0 \leq x \leq \pi$?
- Calculate the Gini index of a country where the fraction of total income earned by the lowest x fraction of the populace is given by
$$f(x) = \frac{2}{5}x^2 + \frac{3}{5}x^3.$$
- Compute the area between the curves $f(x) = e^x \sec^2 x$ and $g(x) = e^x \tan^2 x$ for $0 \leq x \leq \pi$.
- Consider a cone of height h with base a circular disc of radius r. Let's compute the "surface area" – the area of the "outside" of the cone, not including the bottom. Following how we computed the area of a circular disc (which is, indeed, such a cone with $h = 0$), we can decompose its area into infinitesimal triangles with base $r d\theta$ and height the slant length $L = \sqrt{h^2 + r^2}$. The area element dA is then the area of this infinitesimal triangle. Integrating dA from $\theta = 0$ to $\theta = 2\pi$ gives the "surface area" of the cone. What is its value?
- Compute the area between the curves $\sin(x)$ and $\cos(x)$ from $x = 0$ to $x = \pi/2$.
- Compute the area of a triangle with vertices at (0,0), (2,1), (3,6)

30.8 Answers to Selected Examples

1. The logical choice for area element is a vertical strip:



The height of this strip is $f(x) - g(x) = 3 + 2x - x^2$, and the width of the strip is dx. So the area element is $dA = (3 + 2x - x^2)dx$. To find the intersection points, set the curves equal, which gives $1 - x = 4 + x - x^2$. This implies $x^2 - 2x - 3 = 0$, which factors to $(x + 1)(x - 3) = 0$. Thus, the intersections are $x = -1$ and $x = 3$. It follows that

$$\begin{aligned} A &= \int dA \\ &= \int_{-1}^{3} (3 + 2x - x^2) dx \\ &= \left(3x + x^2 - \frac{1}{3}x^3\right)\Bigg|_{-1}^{3} \\ &= (9 + 9 - 9) - \left(-3 + 1 + \frac{1}{3}\right) \\ &= \frac{32}{3}. \end{aligned}$$

(Return)

2.

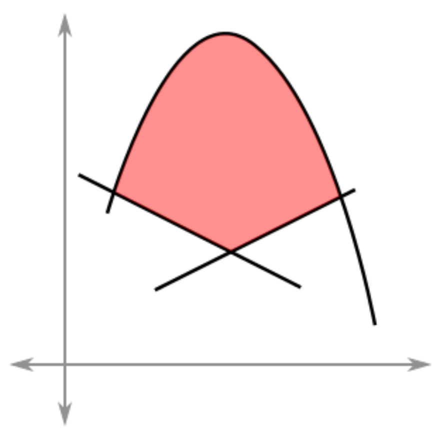
$$\begin{aligned} G(f) &= 2\int_{x=0}^{1} (x - f(x))\, dx \\ &= 2\int_{x=0}^{1} (x - x^n)\, dx \\ &= 2\left(\frac{x^2}{2} - \frac{x^{n+1}}{n+1}\right)\Bigg|_{x=0}^{1} \\ &= 1 - \frac{2}{n+1} \\ &= \frac{n-1}{n+1} \end{aligned}$$

(Return)

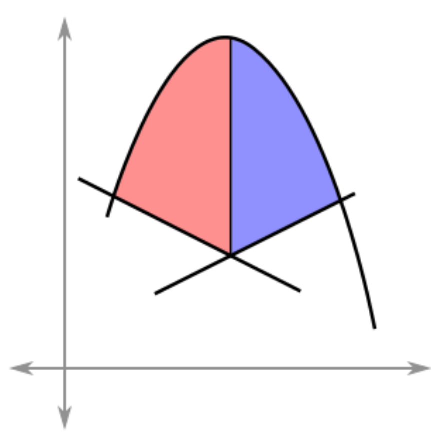
31 Complex Areas

31.1 Complex regions

Some regions in the plane are more complicated and cannot be evaluated with a single integral. This happens when the area element is not bounded by the same curves throughout the region. For instance, consider the region bounded by a parabola and two lines:

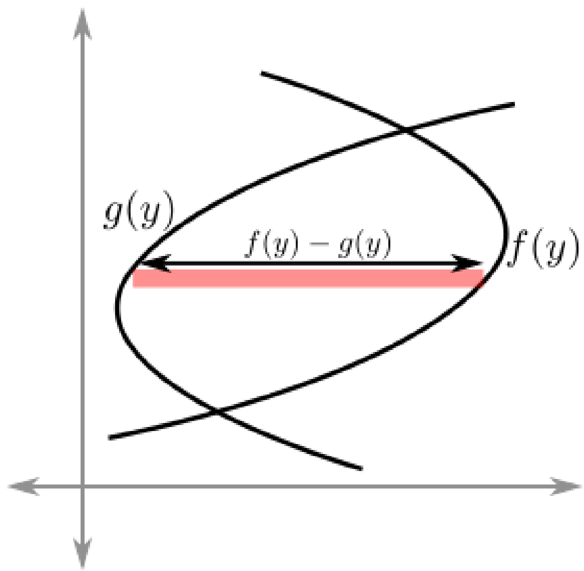


In this case, the only way to find the area of the region is to divide it into regions which can be integrated separately:



31.2 Horizontal strips

Other regions are difficult to integrate using vertical strips as the area element, but work well with horizontal strips as the area element. For example, consider the following region bounded on the left by $x = g(y)$ and on the right by $x = f(y)$:



In this case, the area of a horizontal strip is a function of y, namely $(f(y) - g(y))dy$, where $x = f(y)$ is the curve on the right and $x = g(y)$ is the curve on the left.

Example

Find the area between the curves

$$y - x = 0$$
$$y^2 + x = 2.$$

(See Answer 1)

Example

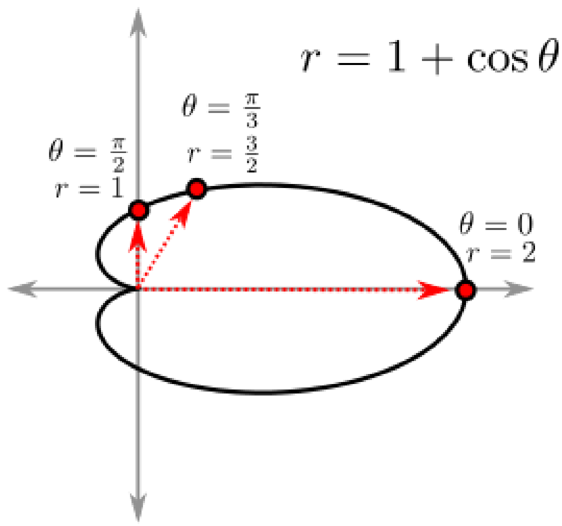
Find the area of the region bounded by $x = 3y$ and $x = y^2 - 4$. (See Answer 2)

Example

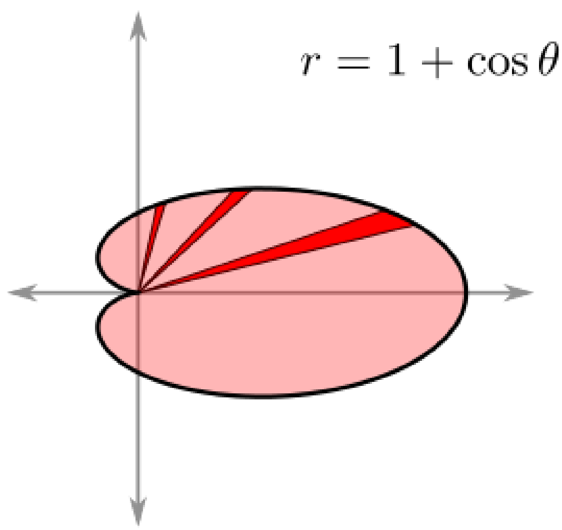
Find the area of the region bounded by $y = \ln(x)$ and the lines $y = 0$, $y = 1$, and $x = 0$. (See Answer 3)

31.3 Polar shapes

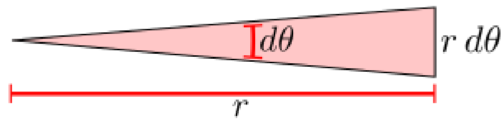
A polar shape is the graph of a polar function $r = f(\theta)$. Here, the input to the function is θ, which is the angle formed with the positive x-axis (known as the *pole*). The output r is the distance from the origin (or *radial* *distance*). For example, the following shows the graph of the polar function $r = 1 + \cos\theta$, which is known as a *cardioid*:



The area of such a region is not usually easy to compute by integrating with respect to x or y (for one thing, the polar equation would need to be expressed in terms of x and y first!). Instead, the way to integrate over such regions is to use a *polar area element*, which is a wedge shaped region. Here are several examples of the polar area element for various values of θ:



To compute what the polar area element is in terms of θ, r, and $d\theta$, note that the region is roughly triangular (the curved portion at the base of the triangle can be ignored since it is a higher order term). The angle at the tip of the triangle is $d\theta$, the height of the triangle is r, and the base of the triangle is $r\,d\theta$:



Thus, the polar area element is

$$dA = \frac{1}{2}(r)(r\,d\theta) = \frac{1}{2}(f(\theta))^2\,d\theta,$$

since $r = f(\theta)$. Thus, the area of a polar region defined by $r = f(\theta)$, where $a \le \theta \le b$, is

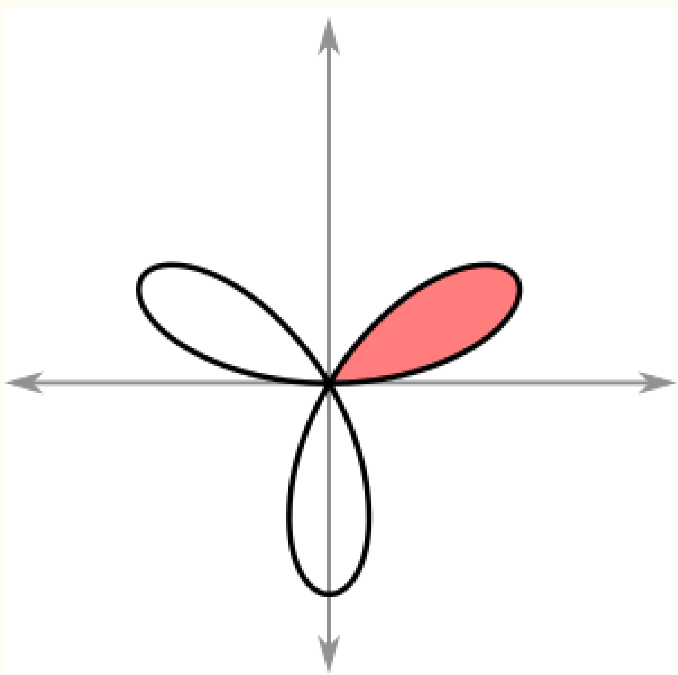
$$A = \int_{\theta=a}^{b} \frac{1}{2}(f(\theta))^2\,d\theta.$$

Example

Compute the area of the cardioid $r = 1 + \cos\theta$. (See Answer 4)

Example

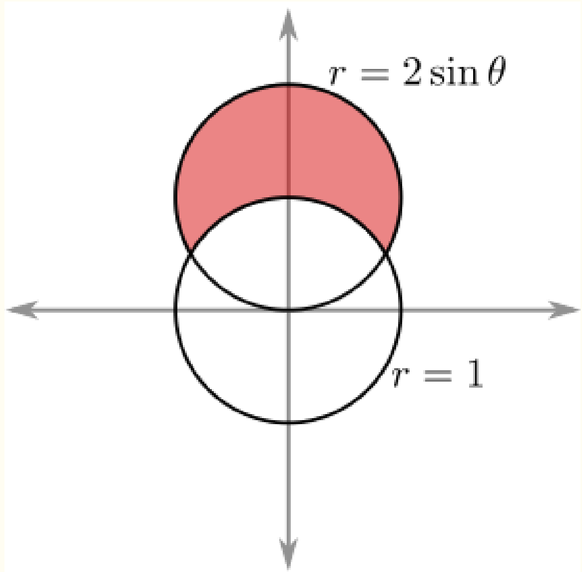
Find the area of a single petal of the polar curve $r = \sin(3\theta)$:



Hint: To find the bounds on θ, compute when $r = 0$. (See Answer 5)

Example

Find the area inside the circle $r = 2\sin\theta$ and outside the circle $r = 1$:



(See Answer 6)

31.4 EXERCISES

- Find the area enclosed by the curves $y = 1$, $x = 1$, and $y = \ln x$.
- Find the area of the bounded region enclosed by the x-axis, the lines $x = 1$ and $x = 2$ and the hyperbola $xy = 1$.
- Compute the area in the bounded (i.e., finite) regions between $y = x(x-1)(x-2)$ and the x-axis.
- Find the area of the sector of a circular disc of radius r (centered at the origin) given by $1 \le \theta \le 3$ (as usual, θ is in radians).
- Use polar coordinates to find an area within $r = 3 - 2cos(\theta)$ and outside $r = 3$.
- Find the area of the overlap between two circles of radius 2 that pass through each others' centers. You can do so with either cartesian or polar coordinates (though one might be easier than the other!).
- Find the area bound by the curves $y = \cos^2 x$ and $y = \dfrac{8x^2}{\pi^2}$.
- Kepler's First Law states that the orbit of every planet is an ellipse with the Sun at one of its two foci. If we think of the Sun as being situated at the origin, we can describe the orbit with the equation:

$$r = \frac{p}{1 + \varepsilon \cos\theta}$$

  The point at which the planet is closest to the Sun (the perihelion) corresponds to $\theta = 0$, while the planet is furthest away from the Sun at $\theta = \pi$ (the aphelion). Knowing the distance between the Sun and the planet at these two points would allow you to fix the values of the constants p and ε. Notice that $\varepsilon = 0$ describes a perfect circle, so that the "eccentricity" ε measures how far the orbit is from being a circle.

Kepler's Second Law states that the line joining a planet and the Sun sweeps out equal areas during equal intervals of time. Another way of expressing this fact is by saying that the "areal velocity"

$$v_A = \frac{dA}{dt}$$

of that line is constant in time.

Express the area element dA in terms of the angle element $d\theta$ and use Kepler's Second Law to deduce the differential equation governing the time evolution of θ.

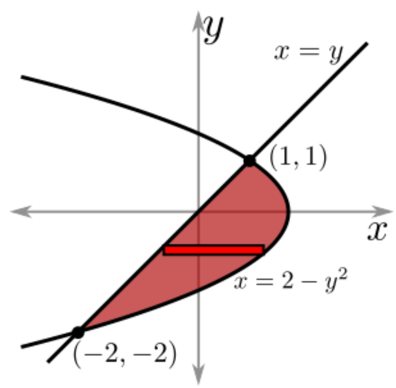
- Let C_1 be the circle given by $r = \sin(\theta)$. Let C_2 be the circle given by $r = \cos(\theta)$. Find the area of region in C_2 that is not in C_1.

31.5 Answers to Selected Exercises

1. Expressing these curves as functions of y, we find

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

Graphing these curves, one finds the bounded region:



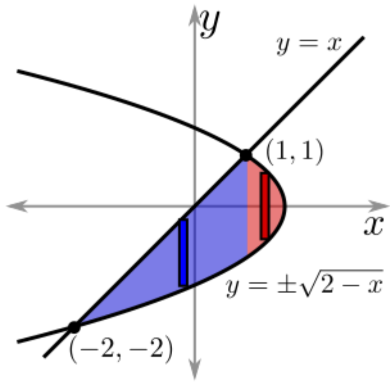
To find the intersections, set the curves equal to one another. This gives

$$y = 2 - y^2.$$

A rearranging and factoring gives

$$y^2 + y - 2 = (y - 1)(y + 2) = 0,$$

and so we find that the intersection points are at $y = 1$ and $y = -2$ (the x-coordinates are the same, since they are on the line $x = y$). Note that using a vertical rectangle as the area element here would not be so easy, because the area element depends on the value of x. Sometimes the strip goes from the parabola below to the line above, as shown in blue, and sometimes the strip goes from parabola to parabola, shown in red:



In particular, the area element for a vertical strip is

$$dA = \begin{cases} (x + \sqrt{2-x})\, dx & \text{if } -2 \le x \le 1 \\ 2\sqrt{2-x}\, dx & \text{if } 1 \le x \le 2 \end{cases}$$

But using a horizontal strip as the area element works much better because throughout the region the strip is always going from the line on the left to the parabola on the right. So using a horizontal strip gives the area element

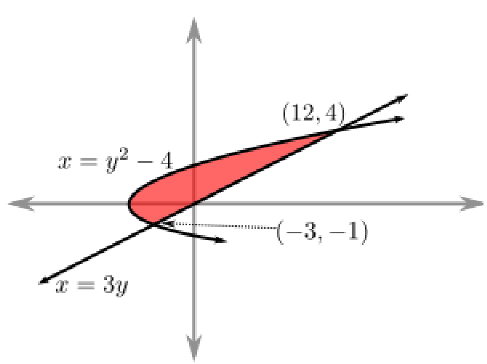
$$dA = ((2 - y^2) - y)\, dy.$$

Integrating this over the range of $-2 \le y \le 1$ gives the area:

$$\begin{aligned} A &= \int dA \\ &= \int_{y=-2}^{1} 2 - y^2 - y\, dy \\ &= 2y - \frac{y^3}{3} - \frac{y^2}{2}\bigg|_{y=-2}^{1} \\ &= \left(2 - \frac{1}{3} - \frac{1}{2}\right) - \left(-4 + \frac{8}{3} - 2\right) \\ &= \frac{9}{2}. \end{aligned}$$

(Return)

2. The region looks roughly as in the following:



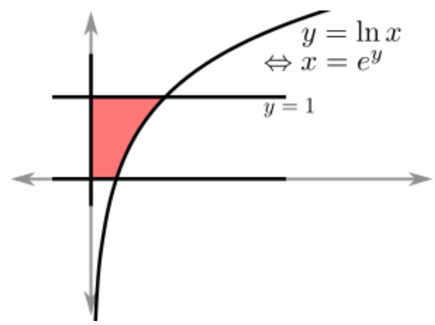
By setting $3y = y^2 - 4$, collecting like terms, and factoring, one finds the intersection points at $y = -1$ and $y = 4$, as indicated in the figure. The area element is a horizontal rectangle, which has area $dA = (3y - (y^2 - 4))dy$.

Thus, the area between the curves is

$$\begin{aligned} A &= \int dA \\ &= \int_{-1}^{4} (3y - y^2 + 4)dy \\ &= \frac{3}{2}y^2 - \frac{1}{3}y^3 + 4y \Big|_{-1}^{4} \\ &= \frac{125}{6} \end{aligned}$$

(Return)

3. The region looks roughly like the following:



Note that using vertical rectangles would not be ideal because this would require two integrals (for x from 0 to 1 and from 1 to e). Instead, one can express the curve $y = \ln x$ as $x = e^y$. Now, using horizontal rectangles gives an area element of $dA = e^y dy$. Thus

$$\begin{aligned} A &= \int dA \\ &= \int_{0}^{1} e^y dy \\ &= e^y \Big|_{0}^{1} \\ &= e - 1. \end{aligned}$$

(Return)

4. In this case, $f(\theta) = 1 + \cos\theta$, and so the area element is

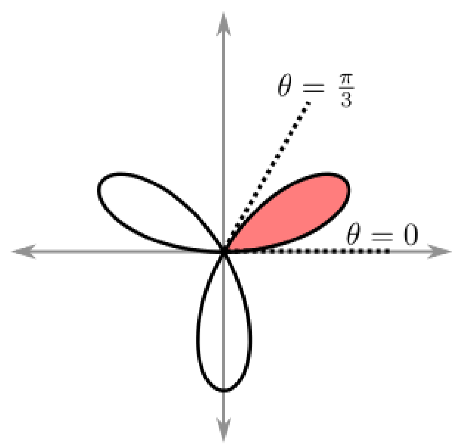
$$\begin{aligned} dA &= \frac{1}{2}(1 + \cos\theta)^2 \, d\theta \\ &= \frac{1}{2}(1 + 2\cos\theta + \cos^2\theta) \, d\theta \end{aligned}$$

Because θ ranges from 0 to 2π to trace out the entire cardioid, it follows that the area is

$$\begin{aligned}
A &= \int dA \\
&= \frac{1}{2}\int_{\theta=0}^{2\pi} (1 + 2\cos\theta + \cos^2\theta)\, d\theta \\
&= \frac{1}{2}\int_{\theta=0}^{2\pi} \left(1 + 2\cos\theta + \frac{1}{2}(1 + \cos(2\theta))\right) d\theta \\
&= \frac{1}{2}\left(\theta + 2\sin\theta + \frac{1}{2}\theta + \frac{1}{4}\sin(2\theta)\right)\Bigg|_{\theta=0}^{2\pi} \\
&= \frac{3\pi}{2}.
\end{aligned}$$

(Return)

5. The area element is $dA = \frac{1}{2}\sin^2(3\theta)\, d\theta$. To find the bounds on θ, set $r = 0$, which gives $\sin(3\theta) = 0$. The smallest values of θ for which this occurs is $\theta = 0$ and $\theta = \frac{\pi}{3}$:



Thus, the area of a single petal is

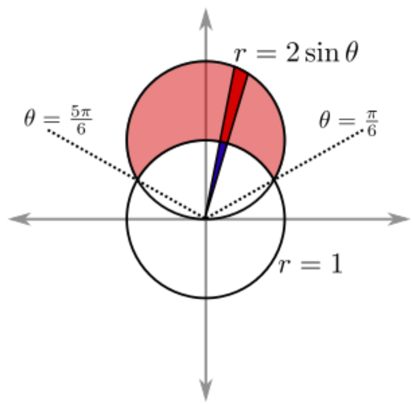
$$\begin{aligned}
A &= \int dA \\
&= \int_{\theta=0}^{\pi/3} \frac{1}{2}\sin^2(3\theta)\, d\theta \\
&= \frac{1}{2}\int_{\theta=0}^{\pi/3} \frac{1}{2}(1 - \cos(6\theta))\, d\theta \\
&= \frac{1}{4}\left(\theta - \frac{1}{6}\sin(6\theta)\right)\Bigg|_{\theta=0}^{\pi/3} \\
&= \frac{\pi}{12}
\end{aligned}$$

(Return)

6. First, we find the intersections by setting the curves equal, which gives

$$2\sin\theta = 1 \quad \Rightarrow \quad \sin\theta = \frac{1}{2},$$

and so the intersections are at $\frac{\pi}{6}$ and $\frac{5\pi}{6}$. The area element of the region is the polar area element of the circle $r = 2\sin\theta$ minus the polar area element of the circle $r = 1$:



So we have that

$$dA = \left(\frac{1}{2}(2\sin\theta)^2 - \frac{1}{2}(1)^2\right)\, d\theta.$$

Thus, the area is

$$\begin{aligned} A &= \int dA \\ &= \frac{1}{2}\int_{\theta=\pi/6}^{5\pi/6} (4\sin^2\theta - 1)\, d\theta \\ &= \frac{1}{2}\int_{\theta=\pi/6}^{5\pi/6} 2(1-\cos(2\theta)) - 1\, d\theta \\ &= \frac{1}{2}\left(\theta - \sin(2\theta)\right)\Big|_{\pi/6}^{5\pi/6} \\ &= \frac{\pi}{3} + \frac{\sqrt{3}}{2}. \end{aligned}$$

From the second to the third line above, we used the power reduction formula for sine:

$$\sin^2\theta = \frac{1}{2}(1-\cos(2\theta)).$$

(Return)

32 Volumes

32.1 Finding the volume element

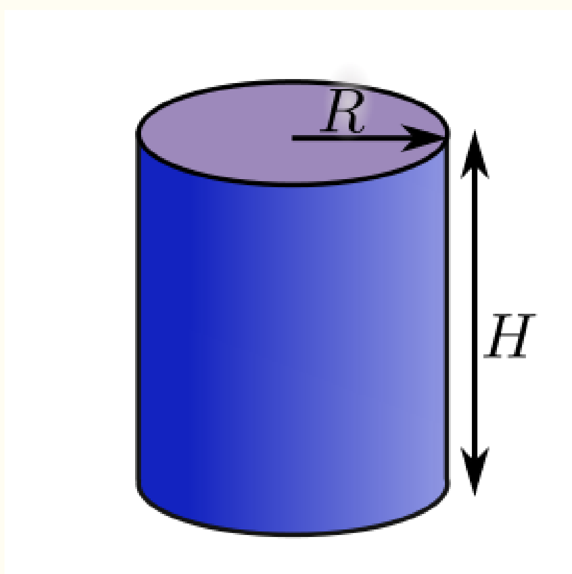
Just as area was computed by finding the area element and integrating, volume is computed by determining the volume element (i.e. the volume of a slice) and then integrating:

$$V = \int dV.$$

The difficulty is in finding a suitable volume element dV. Once that is chosen, the rest is a matter of evaluating the integral.

Example

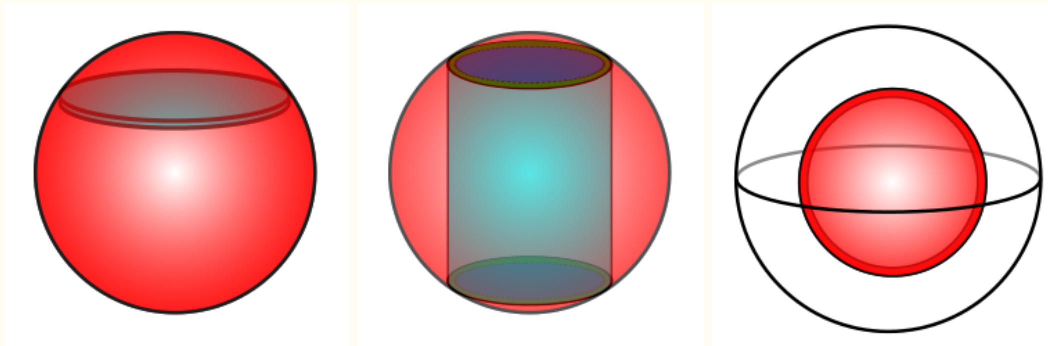
Compute the volume of a cylinder of radius R and height H using several different volume elements dV:



(See Answer 1)

Example

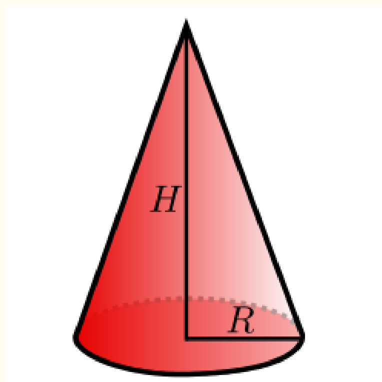
Find the volume of a sphere of radius R. First, by using discs as the volume element (shown on left below). Then use cylindrical shells as the volume element (shown in the middle below). Finally, use a spherical shell for the volume element, as shown in the third diagram.



(See Answer 2)

Example

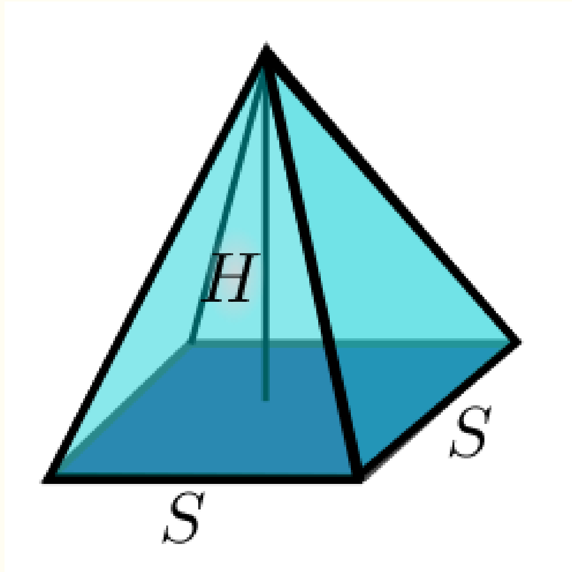
Find the volume of a cone of base radius R and height H.



(See Answer 3)

Example

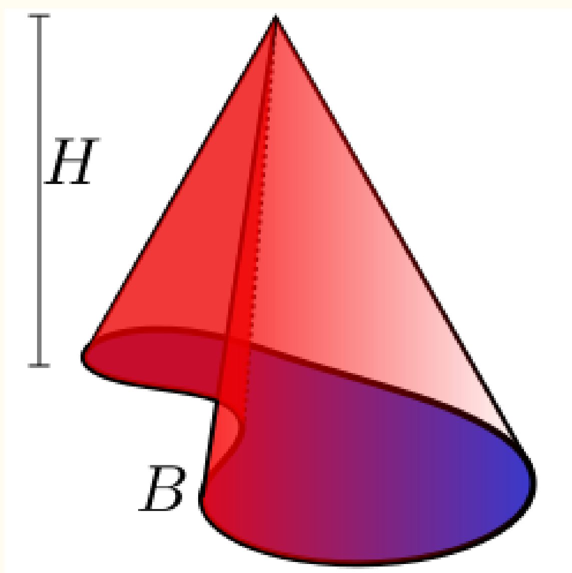
Find the volume of a square pyramid of base edge S and height H.



(See Answer 4)

Example

Show that the volume of a generalized cone of base area B and height H is $\frac{1}{3}BH$. Explain the reason for the factor of $\frac{1}{3}$.



(See Answer 5)

32.2 EXERCISES

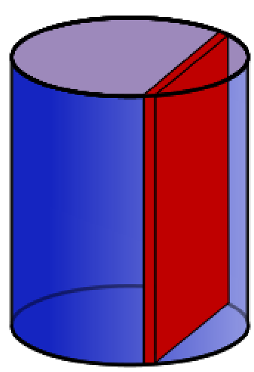
- Find the volume of the following solid: for $1 \le x < \infty$, the intersection of the this solid with the plane perpendicular to the x-axis is a circular disc of radius e^{-x}.
- The base of a solid is given by the region lying between the y-axis, the parabola $y = x^2$, and the line $y = 16$ in the first quadrant. Its cross-sections perpendicular to the y-axis are equilateral triangles. Find the volume of this solid.
- The base of a solid is given by the region lying between the y-axis, the parabola $y = x^2$, and the line $y = 4$. Its cross-sections perpendicular to the y-axis are squares. Find the volume of this solid.
- Find the volume of the solid whose base is the region enclosed by the curve $y = \sin x$ and the x-axis from $x = 0$ to $x = \pi$ and whose cross-sections perpendicular to the x-axis are semicircles.
- Consider a cone of height h over a circular base of radius r. We computed the volume by slicing parallel to the base. What happens if instead we slice orthogonal to the base? What is the volume element obtained by taking a wedge at angle θ of thickness $d\theta$?
- Consider the following solid, defined in terms of polar coordinates: $0 \le r \le R$; $0 \le \theta \le 2\pi$; $0 \le z \le r$. Can you describe this shape? Compute its volume.
- Consider the following solid, defined in terms of polar coordinates: $0 \le r \le R$; $0 \le \theta \le 2\pi$; $0 \le z \le \theta$. Can you describe this shape? Compute its volume.
- Challenge: compute the volume intersection of the (infinite) cylinders of radius R centered along the x and y axes in 3-d. That is, compute the volume of intersection of

$$x^2 + z^2 \le R^2$$
$$x^2 + z^2 \le R^2$$

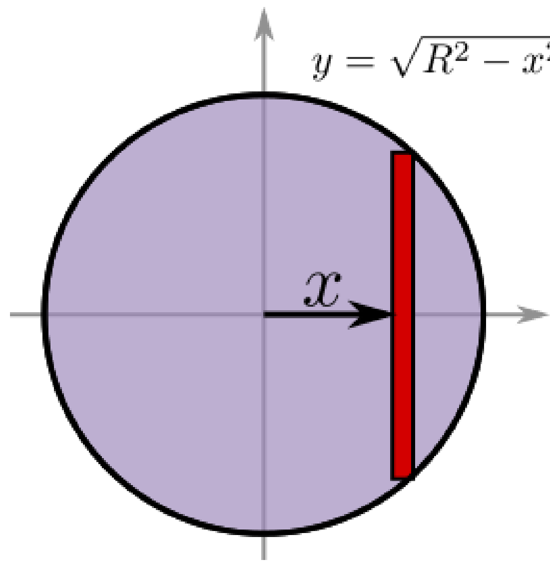
  in the 3-dimensional (x, y, z) space.

32.3 Answers to Selected Examples

1. First, consider making a slice perpendicular to the base of the cylinder:



This gives a rectangular slice whose height is H, the same as the cylinder. The width of the rectangle can be determined by looking at an overhead view of the cylinder. Let x be the distance of the slice from the center of the cylinder (so x ranges from $-R$ to R as the slice sweeps across the cylinder):

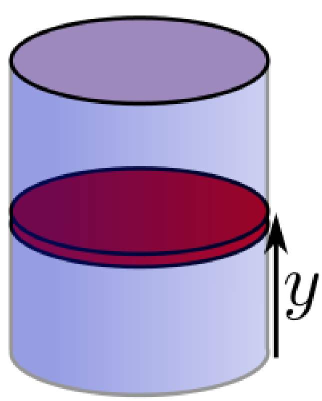


Doing a little algebra, we find that the width of the rectangle is $2\sqrt{R^2 - x^2}$. Finally, the thickness of the slice is dx, and so the volume element in this case is

$$dV = 2H\sqrt{R^2 - x^2}\,dx.$$

Integrating this requires the trigonometric substitution $x = R\sin\theta$. There are easier volume elements we could choose, as we shall see.

Another way to slice is to make cuts parallel to the base of the cylinder. Let y denote the distance of the slice from the base of the cylinder:



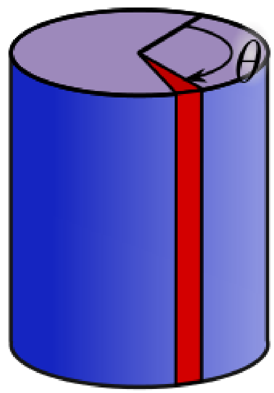
Then each slice is a circle of radius R and thickness dy. Thus

$$dV = \pi R^2\,dy,$$

and y ranges from 0 to H, so the volume is

$$\begin{aligned}
V &= \int dV \\
&= \int_{y=0}^{H} \pi R^2 \, dy \\
&= \pi R^2 y \Big|_{y=0}^{H} \\
&= \pi R^2 H.
\end{aligned}$$

Another possible choice is a wedge shaped volume element. Let θ be the angle that the wedge forms with a fixed axis (so θ ranges from 0 to 2π):



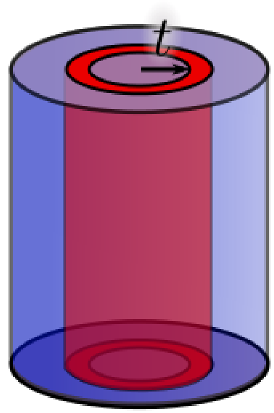
Here, the area of the sector of the circle is $\frac{1}{2}R^2 \, d\theta$. Thus the volume of the wedge is

$$dV = \frac{1}{2} R^2 H \, d\theta.$$

Thus the volume is

$$\begin{aligned}
V &= \int dV \\
&= \int_{\theta=0}^{2\pi} \frac{1}{2} R^2 H \, d\theta \\
&= \frac{1}{2} R^2 H\theta \Big|_{\theta=0}^{2\pi} \\
&= \pi R^2 H.
\end{aligned}$$

One final option is to use cylindrical shells. Let t be the radius of the shell, so that t ranges from 0 to R as the shells sweep through the cylinder.



The height of the cylindrical shell is H and the thickness of the shell is dt. Recalling that the lateral surface area of a cylinder is $2\pi RH$, we have

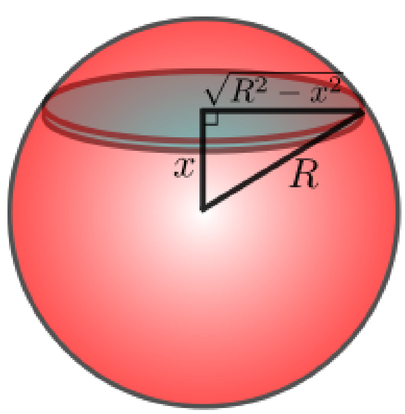
$$dV = 2\pi t H\, dt.$$

Integrating gives that the volumes is

$$\begin{aligned} V &= \int dV \\ &= \int_{t=0}^{R} 2\pi t H\, dt \\ &= 2\pi H \frac{t^2}{2}\bigg|_{t=0}^{R} \\ &= \pi R^2 H. \end{aligned}$$

(Return)

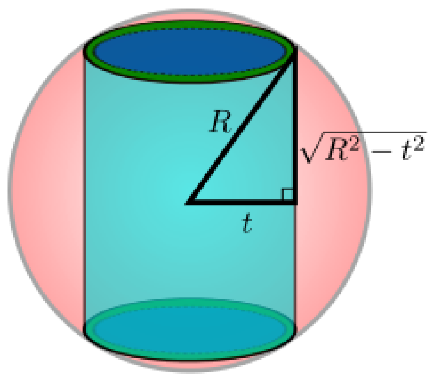
2. Let x be the distance from the center of the disc to the center of the sphere (so x ranges from $-R$ to R as the discs sweep across the sphere). Then drawing a right triangle shows that the radius of the disc is $\sqrt{R^2 - x^2}$ (since the radius of the sphere is R). See the diagram below:



The thickness of the disc is dx, and so the volume of the disc is $\pi(\sqrt{R^2 - x^2})^2 dx$ (the area of the disc times its thickness), and so the volume of the sphere is

$$\begin{aligned} V &= \int dV \\ &= \int_{x=-R}^{R} \pi(R^2 - x^2)dx \\ &= \pi\left(R^2 x - \frac{x^3}{3}\right)\Bigg|_{x=-R}^{R} \\ &= \pi\left(\left(R^3 - \frac{R^3}{3}\right) - \left(-R^3 + \frac{R^3}{3}\right)\right) \\ &= \frac{4}{3}\pi R^3. \end{aligned}$$

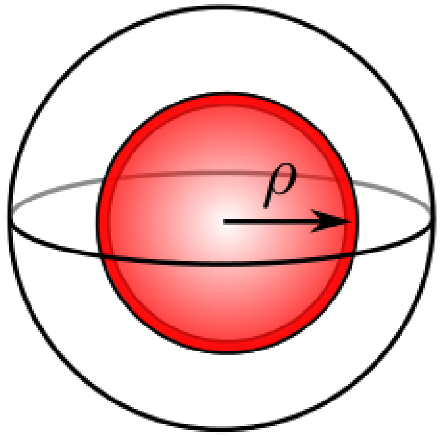
For the cylindrical shell, let t be the radius of the cylinder (so t ranges from 0 to R as the cylinders sweep out the sphere). Then by drawing in a right triangle, one finds that the height of the cylinder is $2\sqrt{R^2 - t^2}$:



Recall that the lateral surface area of a cylinder with radius r and height h is $2\pi rh$. Thus, the lateral surface area of the cylinder is $4\pi t\sqrt{R^2 - t^2}$. The thickness of the shell is dt, and so the volume element is $4\pi t\sqrt{R^2 - t^2}dt$. It follows (after making the u-substitution $u = R^2 - t^2$) that the volume of the sphere is

$$\begin{aligned} V &= \int dV \\ &= \int_{t=0}^{R} 4\pi t\sqrt{R^2 - t^2}dt \\ &= 4\pi \int_{u=R^2}^{0} \frac{-1}{2}\sqrt{u}du \\ &= -2\pi\left(\frac{2}{3}u^{3/2}\Big|_{u=R^2}^{0}\right) \\ &= -2\pi\left(0 - \frac{2}{3}R^3\right) \\ &= \frac{4}{3}\pi R^3. \end{aligned}$$

Finally, for the spherical shell, let ρ denote the radius of the spherical shell:



Recall that the surface area of a sphere of radius ρ is $4\pi\rho^2$. Therefore, the volume of the spherical shell (i.e. our volume element) is

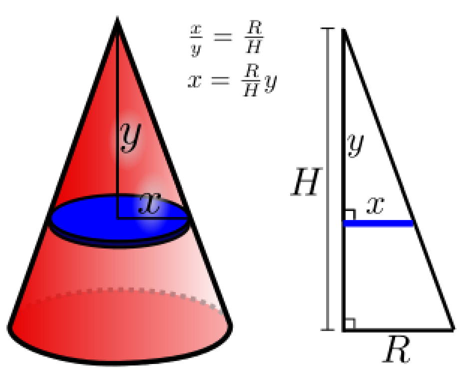
$$dV = 4\pi\rho^2\, d\rho.$$

Note that to sweep over the entire sphere, ρ must range from 0 to R. Therefore,

$$\begin{aligned} V &= \int dV \\ &= \int_{\rho=0}^{R} 4\pi\rho^2\, d\rho \\ &= 4\pi\frac{1}{3}\rho^3\Big|_{\rho=0}^{R} \\ &= \frac{4}{3}\pi R^3. \end{aligned}$$

(Return)

3. The easiest choice for volume element is a slice parallel to the base of the cone, which gives a disc. Let y be the distance from the tip of the cone to the center of the disc (so y ranges from 0 to H as the disc sweeps across the cone), and x be the radius of the disc:



The volume element is the area of the disc, πx^2, times the thickness of the disc, dy. It remains to find x in terms of y. In the cutaway in the figure on the right above, one sees that by similar triangles, $\frac{x}{y} = \frac{R}{H}$, and so $x = \frac{R}{H}y$. Thus, the volume element is

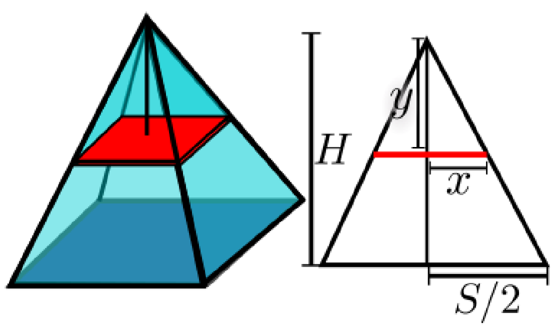
$$\begin{aligned} dV &= \pi x^2 dy \\ &= \pi \left(\frac{R}{H}y\right)^2 dy. \end{aligned}$$

Thus, the volume of the cone is

$$\begin{aligned} V &= \int dV \\ &= \int_{y=0}^{H} \pi \frac{R^2}{H^2} y^2 dy \\ &= \frac{\pi R^2}{H^2} \left(\frac{y^3}{3}\bigg|_{y=0}^{H} \right) \\ &= \pi \frac{R^2}{H^2} \cdot \frac{H^3}{3} \\ &= \frac{1}{3}\pi R^2 H. \end{aligned}$$

(Return)

4. Again, use slices parallel to the base. Let y be the distance from the tip of the cone to the center of the slice (so y ranges from 0 to H), and let x be half of the side length of the slice.



As shown in the above cutaway, one finds by similar triangles that $\frac{x}{y} = \frac{S/2}{H}$, and so $x = \frac{S}{2H}y$. Therefore, the area of a slice is $(2x)^2 = \frac{S^2}{H^2}y^2$, and the thickness of a slice is dy, so the volume element is

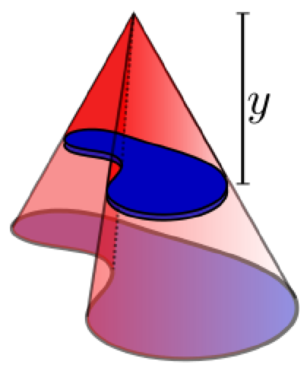
$$dV = \frac{S^2}{H^2} y^2 dy.$$

And so the volume of the pyramid is

$$\begin{aligned} V &= \int dV \\ &= \int_{y=0}^{H} \frac{S^2}{H^2} y^2 dy \\ &= \frac{S^2}{H^2} \left(\frac{y^3}{3} \bigg|_{y=0}^{H} \right) \\ &= \frac{S^2}{H^2} \cdot \frac{H^3}{3} \\ &= \frac{S^2 H}{3}. \end{aligned}$$

(Return)

5. Let y be the distance from the tip of the cone to the slice.



Because the linear dimensions of the slice grow proportionally with y (e.g. the length of the slice is $\frac{y}{H}$ times the length of the base), the area of the slice will grow proportionally with the square of y. This means that

$$\text{Area of the slice } = \left(\frac{y}{H}\right)^2 B$$

Thus, the volume element is $dV = B\frac{y^2}{H^2}dy$, and it follows that the volume of the cone is

$$\begin{aligned} V &= \int dV \\ &= \int_{y=0}^{H} B\frac{y^2}{H^2} dy \\ &= \frac{B}{H^2} \left(\frac{y^3}{3} \bigg|_{y=0}^{H} \right) \\ &= \frac{B}{H^2} \cdot \frac{H^3}{3} \\ &= \frac{1}{3} BH. \end{aligned}$$

The factor of $\frac{1}{3}$ comes from the fact that the volume element is proportional to the square of y, hence the integral has a y^2, which produces a factor of $\frac{1}{3}$ by the power rule.

(Return)

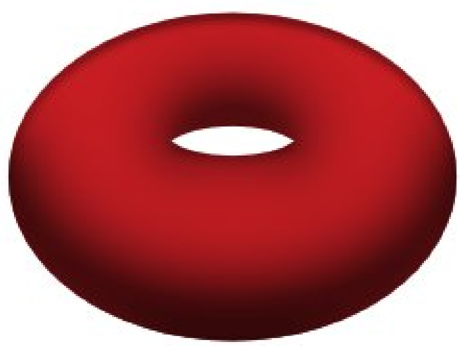
33 Volumes Of Revolution

33.1 Volume element for solid of revolution

Consider a region R in the plane and a line L. The *solid of revolution* of R about the axis L is the solid which results from taking the region R and revolving it around the line L:

(Solid of Revolution Animated GIF)

The result is typically something doughnut shaped. The question of this module is to find the volume of the solid:



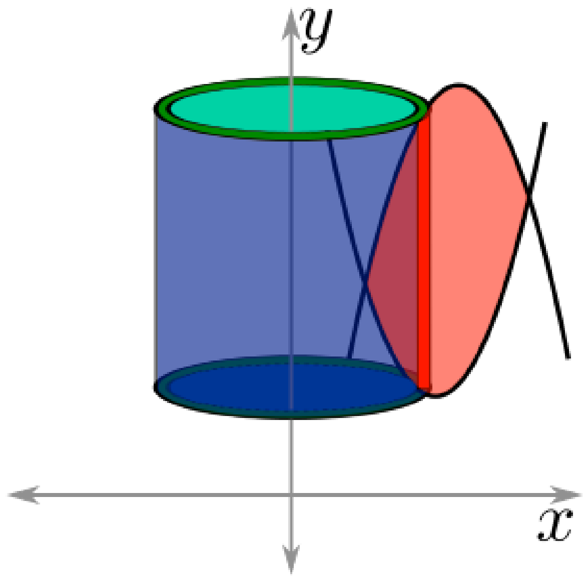
The method is the same as the previous modules: find the volume element (the contribution of a small slice of the region to the total volume) and integrate. Just as area can be computed using vertical or horizontal slices, volume can be computed using corresponding methods: shells or washers, respectively.

The basic outline of finding the volume element for a solid of revolution is

1. Find a convenient area element for the region R in the plane
2. Determine the volume as that area element gets revolved around the axis L.

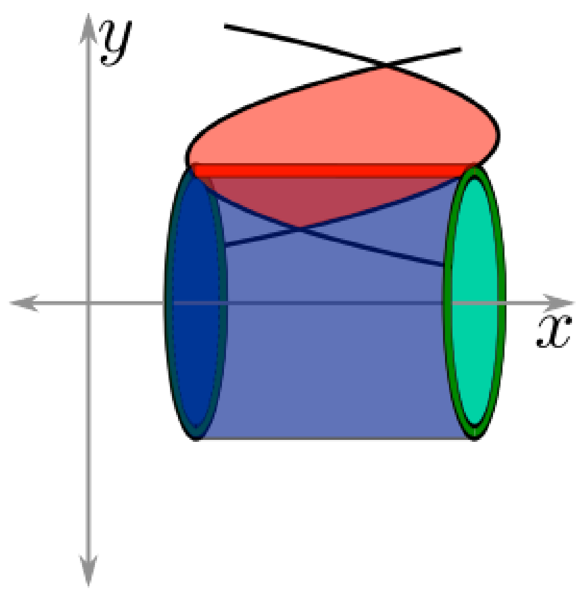
33.2 Cylindrical shells

When the area element is parallel to the axis of rotation, the volume element is a cylindrical shell. Here, the region is bounded by two parabolas. The natural area element for such a region is a vertical rectangle (shown in red). As the region is revolved about the y-axis, the volume element traces out a cylindrical shell, whose volume becomes the volume element of the solid of revolution.



Recall that a cylinder has lateral surface area $2\pi rh$. The thickness of the cylindrical shell is dx (if the axis of rotation is a vertical line) or dy (if the axis of rotation is a horizontal line). Here r and h will generally be functions of x or y (again, depending on whether the axis of rotation is vertical or horizontal).

If a horizontal rectangle is the natural area element (for instance, the region between two horizontal parabolas), and the axis of revolution is the x-axis, then cylindrical shells again arise naturally as the volume element:

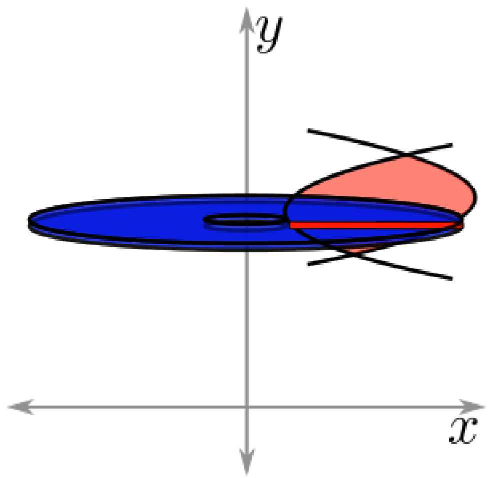


Example

Suppose the region bounded by $y = 3x - x^2$ and $y = x$ is revolved about the y-axis. What is the volume of the resulting solid? (See Answer 1)

33.3 Washers

When the area element is perpendicular to the axis of rotation, the volume element is a washer. So when the area element is a horizontal rectangle (as in a region bounded by horizontal parabolas) and the axis of revolution is vertical, the region traced out by the rectangle is a washer:



A washer is just an annulus (a circle with a circle cut out of it) which has been thickened. The volume of the washer is the area of the annulus times the thickness of the washer. The area of the annulus is $\pi R^2 - \pi r^2$, where R is the radius of the outer circle and r is the radius of the inner circle. The thickness of the washer is dx or dy (depending on the orientation of the washer. Thus, the volume element when using washers is

$$dV = \pi(R^2 - r^2)\, dx \text{ or } dy.$$

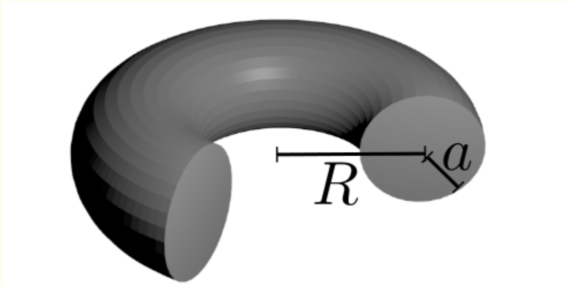
Example

Given the region bounded by $y = 3x - x^2$ and $y = x$, find the volume of the solid resulting from revolving the region about the x-axis. (See Answer 2)

33.4 Additional Examples

Example

Find the volume of a doughnut formed by rotating a disc of radius a about the y-axis. Let the radius of the doughnut be R, as shown in this cutaway:



Use a vertical area element (which leads to a cylindrical shell). (See Answer 3)

Example

Compute the volume of the doughnut again, this time using a horizontal area element (which leads to a washer). (See Answer 4)

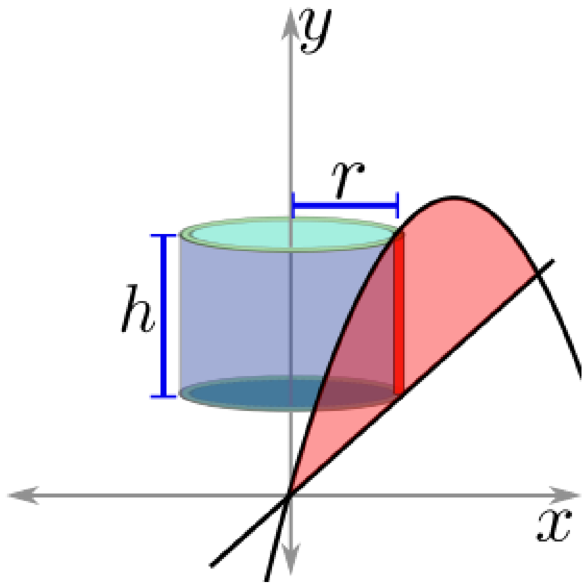
33.5 EXERCISES

- Let D be the region bounded by the curve $y = x^3$, the x-axis, the line $x = 0$ and the line $x = 2$. Find the volume of the region obtained by revolving D about the x-axis.
- Let D be the same region as above. What is the volume of the region formed by rotating this D about the line $x = 3$?
- Let D be the region between the curve $y = -(x-2)^2 + 1$ and the x-axis. Find the volume of the region obtained by revolving D about the y-axis.
- Find the volume obtained by revolving the region between the curves $y = x^3$ and $y = \sqrt[3]{x}$ in the first quadrant about the x-axis.
- Let D be the region under the curve $y = \ln \sqrt{x}$ and above the x-axis from $x = 1$ to $x = e$. Find the volume of the region obtained by revolving D about the x-axis.
- Let D be the region bounded by the graph of $y = 1 - x^4$, the x-axis and the y-axis in the first quadrant. Which of the following integrals can be used to compute the volume of the region obtained by revolving D around the line $x = 5$?
- Challenge: compute the volume of the region obtained by rotating the disc $x^2 + y^2 \leq \epsilon^2$ about the axis given by $y = 1 - x$ for $\epsilon \leq \frac{1}{2}$.
- Let D be the region under the curve $\sqrt{x-1}$ above the x-axis from $x = 1$ to $x = 2$. Compute the volume of solid obtained by rotating D about the x-axis. Compute the volume twice, using both methods.

33.6 Answers to Selected Examples

1. The first step in such a calculation is to draw a decent picture of the region. Then determine whether a vertical or horizontal rectangle would make the best area element. In this case, a vertical rectangle is the best choice.

   Since a vertical rectangle is being revolved about a vertical axis, the result is a cylindrical shell:



The radius of the shell is x (the distance from the y-axis), and the height of the shell is the distance from the top curve to the bottom curve: $h = (3x - x^2) - x = 2x - x^2$. The thickness of the shell is dx. Recalling that the surface area of a cylinder is $2\pi rh$, it follows that the volume element is just the surface area multiplied by the thickness dx:

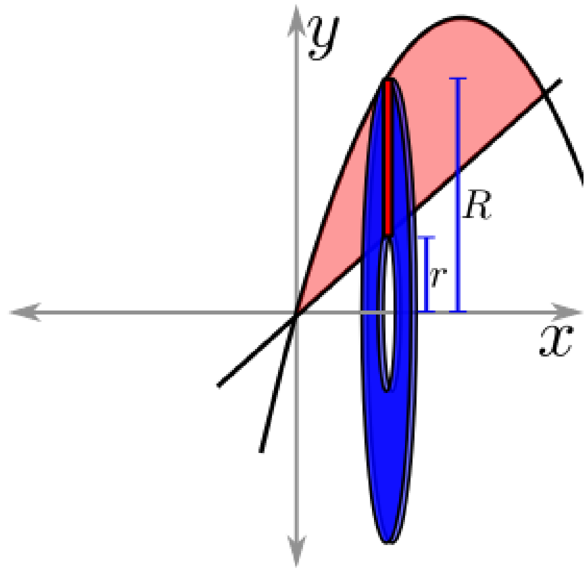
$$\begin{aligned} dV &= 2\pi r h\, dx \\ &= 2\pi x(2x - x^2)\, dx. \end{aligned}$$

A little algebra shows that the intersections occur at $x = 0$ and $x = 2$, so the volume is

$$\begin{aligned} V &= \int dV \\ &= \int_{x=0}^{2} 2\pi x(2x - x^2)\, dx \\ &= 2\pi \int_{x=0}^{2} (2x^2 - x^3)\, dx \\ &= 2\pi \left(\frac{2}{3}x^3 - \frac{1}{4}x^4\right)\Bigg|_{x=0}^{2} \\ &= \frac{8}{3}\pi. \end{aligned}$$

(Return)

2. As in the previous example, the optimal area element is a vertical rectangle. A vertical rectangle revolved about a horizontal axis results in a washer:



The outer radius R is the upper curve: $R = 3x - x^2$, and the inner radius r is the inner curve: $r = x$. The thickness of the washer is dx, and so

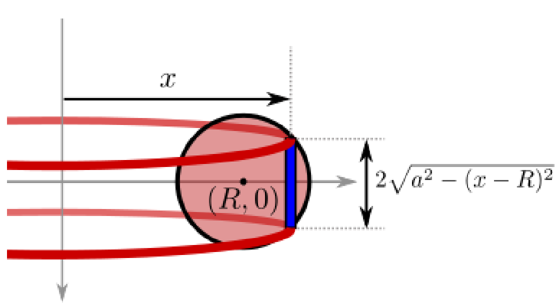
$$\begin{aligned} dV &= \pi(R^2 - r^2)\,dx \\ &= \pi((3x - x^2)^2 - x^2)\,dx. \end{aligned}$$

As in the last example, the intersections are at $x = 0$ and $x = 2$, so

$$\begin{aligned} V &= \int dV \\ &= \int_{x=0}^{2} \pi(9x^2 - 6x^3 + x^4 - x^2)\,dx \\ &= \pi\left(\frac{8}{3}x^3 - \frac{3}{2}x^4 + \frac{1}{5}x^5\right)\Big|_{x=0}^{2} \\ &= \frac{56}{15}\pi. \end{aligned}$$

(Return)

3. First, suppose we use a vertical area element. Since we are rotating about a vertical axis, the area element and axis of rotation are parallel, and so the resulting volume element is a cylindrical shell. Let x be the distance of the area element (the rectangle) from the y-axis. This also happens to be the radius of the cylindrical shell:



The equation of the circle is

$$(x - R)^2 + y^2 = a^2,$$

and solving for y gives

$$y = \pm\sqrt{a^2 - (x - R)^2}.$$

Therefore, the height of the area element (and hence the height of the cylindrical shell) is

$$2\sqrt{a^2 - (x - R)^2}.$$

Now, the volume of the cylindrical shell (our volume element) is

$$\begin{aligned} dV &= 2\pi r h \, dx \\ &= 2\pi x (2\sqrt{a^2 - (x - R)^2}) \, dx \\ &= 4\pi x \sqrt{a^2 - (x - R)^2} \, dx. \end{aligned}$$

Note that x ranges from $R - a$ to $R + a$ as it sweeps across the circle. Therefore the volume is

$$\begin{aligned} V &= \int dV \\ &= \int_{x=R-a}^{R+a} 4\pi x \sqrt{a^2 - (x - R)^2} \, dx. \end{aligned}$$

This is a bit messy, but with a substitution of

$$\begin{aligned} u &= x - R \\ du &= dx, \end{aligned}$$

we find

$$\begin{aligned} \int_{x=R-a}^{R+a} 4\pi x \sqrt{a^2 - (x - R)^2} \, dx &= \int_{u=-a}^{a} 4\pi (u + R) \sqrt{a^2 - u^2} \, du \\ &= \int_{u=-a}^{a} 4\pi u \sqrt{a^2 - u^2} \, du + \int_{u=-a}^{a} 4\pi R \sqrt{a^2 - u^2} \, du. \end{aligned}$$

Here, we have used linearity to split the integral into two integrals. Notice that the first integrand is an odd function of u, and since it is integrated over a symmetric interval, the first integral is 0:

$$\int_{u=-a}^{a} 4\pi u \sqrt{a^2 - u^2} \, du = 0.$$

The second integral can be found by noting that

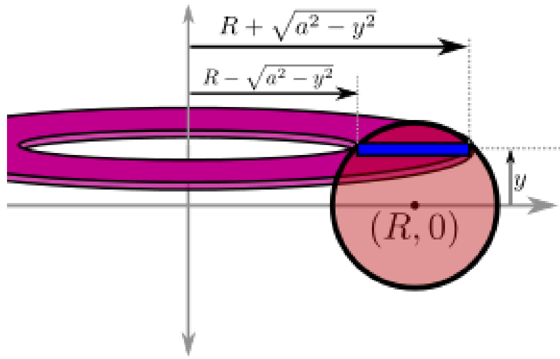
$$\int_{u=-a}^{a} 2\sqrt{a^2 - u^2} \, du$$

gives the area of a disc of radius a (this was an integral done in the simple areas module). Therefore, the volume is

$$\begin{aligned} \int_{u=-a}^{a} 4\pi R \sqrt{a^2 - u^2} \, du &= 2\pi R \int_{u=-a}^{a} 2\sqrt{a^2 - u^2} \, du \\ &= 2\pi R (\pi a^2) \\ &= 2\pi^2 R a^2. \end{aligned}$$

(Return)

4. Carefully drawing and labeling the outer and inner radii of the washer gives the following diagram:



The outer and inner radii can be found by solving the equation of the circle for x:

$$(x-R)^2+y^2=a^2 \quad \Rightarrow \quad x=R\pm\sqrt{a^2-y^2}.$$

Thus, the volume element is the area of the washer times its thickness, dy. Computing this and doing a little algebra gives

$$\begin{aligned} dV &= \left[\pi(R+\sqrt{a^2-y^2})^2-\pi(R-\sqrt{a^2-y^2})^2\right]\,dy \\ &= 4\pi R\sqrt{a^2-y^2}\,dy. \end{aligned}$$

Note that y ranges from $-a$ to a, and so the volume integral is the same one arrived at above:

$$\begin{aligned} V &= \int dV \\ &= \int_{y=-a}^{a} 4\pi R\sqrt{a^2-y^2}\,dy \\ &= 2\pi R(\pi a^2) \\ &= 2\pi^2 Ra^2. \end{aligned}$$

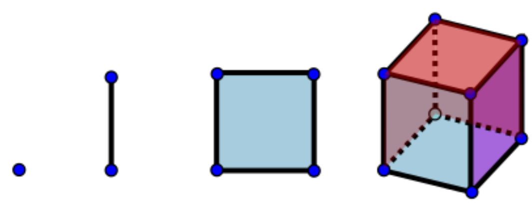
(Return)

34 Volumes In Arbitrary Dimension

The motivation for this module is to find the volume (often referred to as *hypervolume*) of an object in dimension n. This has physical meaning for $n \leq 3$, but what happens for $n \geq 4$?

34.1 The cube in dimension n

Consider the unit cube (i.e. the cube of side length 1) in n dimensions, sometimes called the n-hypercube or just the n-cube. Formally, this is defined to be the set of n-tuples (i.e. lists of length n) $(x_1, x_2, \ldots, x_n)$ such that $0 \leq x_i \leq 1$ for all $1 \leq i \leq n$. For $n = 0, 1, 2, 3$, these are familiar figures: the point, line segment, square, and cube, respectively.



Now, consider some of the various measurements for each of these cubes.

Volume of the cube

For $n = 0$, the cube is just a point, and volume is defined to just be the number of points. So a single point has volume 1.

For $n = 1$, the cube is a line segment. The volume in one dimension is just length, so the one dimension cube has volume 1.

For $n = 2$, the cube is a square of side length 1. In two dimensions, volume is area, so the cube in two dimensions has volume $w \times h = 1 \times 1 = 1$.

For $n = 3$, the cube is a (traditional) cube of side length 1, which has (traditional) volume $l \times w \times h = 1 \times 1 \times 1 = 1$.

For higher values of n, this pattern continues. The intuition is that each additional dimension adds an extra factor of 1, thus the volume of each unit n-cube is 1.

Surface area of cubes

Consider the *surface area* of the cube in dimension n. As with volume, this has physical meaning for $n = 2$ and $n = 3$.

For $n = 2$, the surface area of a square is really its perimeter, which is 4.

For $n = 3$, the surface area is the total area of the faces which bound the cube. There are 6 faces each with area 1, so the surface area is 6.

In general, the n dimension cube will have $2n$ boundary faces, and each face is a cube of dimension $n - 1$, so the surface area (really the hypervolume of the boundary) is $2n$.

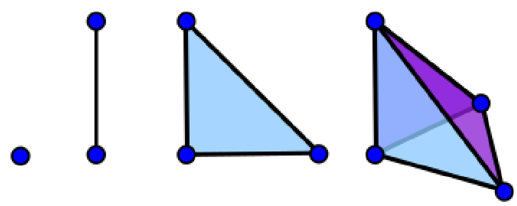
Other features

The *diagonal* of the n-cube can be defined to be the distance from $(0, 0, \ldots, 0)$ to $(1, 1, \ldots, 1)$. Using the distance formula, one finds that the diagonal of the n-cube is $\sqrt{n}$.

The number of corners is fairly easy to count. For $n = 0, 1, 2, 3$, the number of corners is 1, 2, 4, and 8 respectively. Since the n-cube can be thought of as two copies of the $(n - 1)$-cube, one can show by induction that there are 2^n corners in the n-cube.

34.2 Simplex

A *simplex* is a generalization of a triangle or a pyramid. In dimension n, the simplex is defined to be the set of n-tuples $(x_1, x_2, \ldots, x_n)$ such that $0 \leq x_i \leq 1$ and $\sum x_i \leq 1$. This can be thought of as the corner of the n dimension cube where the sum of the coordinates is less than 1. Here are the simplices of dimension $n = 0, 1, 2, 3$:



34.3 Volume of spheres in arbitrary dimension

Now, consider a sphere of radius r in n dimensions. This is the set of points $(x_1, x_2, \ldots, x_n)$ such that $x_1^2 + x_2^2 + \ldots + x_n^2 \leq r^2$. Let $V_n(r)$ be the volume of the sphere of radius r in n dimensions (as above, volume means length, area, volume, hypervolume for $n = 1, 2, 3, \ldots$, respectively). With some careful integration and induction, one finds that

$$V_n(r) = \begin{cases} \frac{\pi^k}{k!} r^n & \text{if } n = 2k \\ \frac{2^n \pi^k k!}{n!} r^n & \text{if } n = 2k+1 \end{cases}.$$

Now, note that as $n \to \infty$ (and r stays fixed), the volume goes to 0 (since factorial grows faster than exponentials).

34.4 EXERCISES

- Consider a four-dimensional box (or "rectangular prism") with side-lengths 1, 1/2, 1/3, and 1/4. What is the 4-dimensional volume of this box?
- What is the "diameter" – i.e., the farthest distance between two points – in this 4-d box? Hint: think in terms of diagonals.
- High-dimensional objects are everywhere and all about. Let's consider a very simple model of the space of digital images. Assume a planar digital image (such as that captured by a digital camera), where each pixel is given values that encode color and intensity of light. Let's assume that this is done via an RGB (red/green/blue) model. Though there are many RGB model specifications, let us use one well-suited for mathematics: to each pixel on associates three numbers (R, G, B), each taking a value in $[0, 1]$.

  Since the red/green/blue values are independent, each pixel has associated to it a 3-d cube of possible color values. Consider a (fairly standard) 10-megapixel camera. If I were to consider the "space of all images" that my camera can capture, what does the space look like? How many dimensions does it have? Note: there's no calculus in this problem...just counting!
- Consider an n-dimensional "hypercube" C of all side-lengths equal to 1. Its n-dimensional volume is, clearly, 1. Now consider what happens when you shrink the hypercube's side-lengths by 1 percent (concentrically, so that the shrunken cube has the same center as the original) and remove it from the original cube. By subtracting the n-dimensional volume of this slightly smaller hypercube, conclude how much volume remains in the 1-percent outer "shell."
- In the previous question, what happens to the volume of the 1-percent shell as $n \to \infty$?
- We have seen that the n-dimensional volume of a unit radius ball in dimension n converges to zero as $n \to \infty$. But what about a really large ball? For a ball of radius $R = 10^{10}$ meters in dimension n, what is the limit as $n \to \infty$ of its volume? (in unit of meters-to-the-n^{th})
- For the brave: so, as $n \to \infty$, the volume of the n-ball all concentrates near the surface shell. OK, you've got that. Now answer this: what proportion of the volume is concentrated along the "equatorial plane"? Let's make that specific. Recall, we computed the volume V_n as $I_n \cdot V_{n-1}$, where

  $$I_n = \int_{\theta=-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n \theta \, d\theta.$$

  We can compute the volume of the equatorial slice of thickness 2ϵ (for some small but fixed $\epsilon > 0$) as

  $$V_{n,2\epsilon} = V_{n-1} \int_{-\epsilon}^{\epsilon} \cos^n \theta \, d\theta.$$

  So, here is the (hard!) problem. Compute the limit as $n \to \infty$ of the ratio of $V_{n,2\epsilon}$ to V_n:

  $$\lim_{n\to\infty} \frac{V_{n,2\epsilon}}{V_n} = \lim_{n\to\infty} \frac{1}{I_n} \int_{-\epsilon}^{\epsilon} \cos^n \theta \, d\theta.$$

  If you can do this (a very big if...) you will get a surprise...

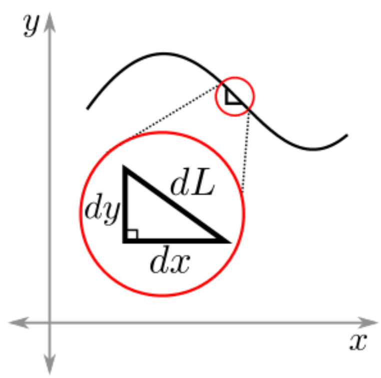
35 Arclength

Consider the graph of a function $y = f(x)$ for $a \leq x \leq b$. The purpose of this module is to find the length of this piece of the curve, known as the *arclength* of the function f from a to b.

As in previous modules, the basic method is to find the arclength element dL and then integrate it:

$$L = \int dL.$$

By zooming in on a portion of the curve, it begins to look like a straight line. Then one can express dL in terms of the infinitesimal horizontal change dx and vertical change dy:



Now, by the Pythagorean theorem one finds that $dL = \sqrt{dx^2 + dy^2}$. A little algebra and the chain rule gives that

$$\begin{aligned} dL &= \sqrt{dx^2 + dy^2} \\ &= \sqrt{dx^2 + \left(\frac{dy}{dx}dx\right)^2} \\ &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\,dx \\ &= \sqrt{1 + (f'(x))^2}\,dx. \end{aligned}$$

So the arclength of the function f from a to b is given by

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

Example

Find the arclength of the curve

$$y = \ln \sin x; \quad \frac{\pi}{4} \le x \le \frac{\pi}{2}.$$

Hint: recall the facts that

$$1 + \cot^2 x = \csc^2 x$$
$$\int \csc x \, dx = -\ln|\csc x + \cot x| + C.$$

(See Answer 1)

Example

Find the arclength of the curve

$$y = x^2 - \frac{1}{8}\ln(x); \quad 1 \le x \le 4.$$

(See Answer 2)

35.1 Parametric curves

If a curve is defined parametrically, i.e. $x = x(t)$ and $y = y(t)$ for $a \le t \le b$, then the arclength element can be written as

$$\begin{aligned} dL &= \sqrt{dx^2 + dy^2} \\ &= \sqrt{\left(\frac{dx}{dt}dt\right)^2 + \left(\frac{dy}{dt}dt\right)^2} \\ &= \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt \\ &= \sqrt{x'(t)^2 + y'(t)^2}\,dt. \end{aligned}$$

So the arclength of a parametric curve $(x(t), y(t))$ for $a \le t \le b$ is given by

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

Example

Find the arclength for a circle of radius r. (See Answer 3)

Example

Find the arclength for the spiral $x(t) = t\cos(t)$, $y(t) = t\sin(t)$ for $0 \le t \le 2\pi$. (See Answer 4)

35.2 Additional Examples

Example

Compute the arclength of the curve

$$y = \frac{2}{3}x^{3/2}; \quad 0 \leq x \leq 3$$

(See Answer 5)

Example

A *catenary* is the curve that is formed by hanging a cable between two towers. It is a fact that the rate of change of the slope of a hanging cable is proportional to the rate of change of arclength with respect to x. Mathematically,

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) = \kappa \cdot \frac{dL}{dx},$$

for some constant κ. Use this fact to find the equation of the catenary. Then find the length of the catenary for $-l \leq x \leq l$. (See Answer 6)

Example

Show that the spiral

$$x = \frac{1}{t}\cos t$$
$$y = \frac{1}{t}\sin t$$

for $2\pi \leq t$ has infinite arclength. (See Answer 7)

35.3 EXERCISES

- Compute the arclength of $y = \frac{x^3}{3} + \frac{1}{4x}$, from $x = 1$ to $x = 2$.

35.4 Answers to Selected Exercises

1. Computing the arclength element from the above formula gives

$$\begin{aligned} dL &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= \sqrt{1 + \left(\frac{1}{\sin x}\cos x\right)^2}\, dx \\ &= \sqrt{1 + \cot^2 x}\, dx \\ &= \sqrt{\csc^2 x}\, dx \\ &= \csc x\, dx. \end{aligned}$$

Therefore, we find that the arclength is

$$
\begin{aligned}
L &= \int dL \\
&= \int_{x=\pi/4}^{\pi/2} \csc x \, dx \\
&= -\ln|\csc x + \cot x|\Big|_{x=\pi/4}^{\pi/2} \\
&= -\ln(1+0) + \ln(\sqrt{2}+1) \\
&= \ln(1+\sqrt{2}).
\end{aligned}
$$

(Return)

2. First, one finds $\frac{dy}{dx} = 2x - \frac{1}{8x}$. So, with some careful algebra one sees that

$$
\begin{aligned}
\sqrt{1+\left(\frac{dy}{dx}\right)^2} &= \sqrt{1+\left(2x-\frac{1}{8x}\right)^2} \\
&= \sqrt{1+(2x)^2 - 2\frac{2x}{8x} + \frac{1}{(8x)^2}} \\
&= \sqrt{1+(2x)^2 - \frac{1}{2} + \frac{1}{(8x)^2}} \\
&= \sqrt{(2x)^2 + \frac{1}{2} + \frac{1}{(8x)^2}}
\end{aligned}
$$

Now note that by reversing the cancellation done in an earlier step when simplifying $-2\frac{2x}{8x} = -\frac{1}{2}$, one finds that $\frac{1}{2} = 2\frac{2x}{8x}$. And so, continuing the computation, one finds

$$
\begin{aligned}
\sqrt{1+\left(\frac{dy}{dx}\right)^2} &= \sqrt{(2x)^2 + 2\frac{2x}{8x} + \frac{1}{(8x)^2}} \\
&= \sqrt{\left(2x+\frac{1}{8x}\right)^2} \\
&= 2x + \frac{1}{8x}.
\end{aligned}
$$

Thus, $dL = \left(2x + \frac{1}{8x}\right) dx$, and it follows that

$$
\begin{aligned}
L &= \int dL \\
&= \int_{x=1}^{4} \left(2x + \frac{1}{8x}\right) dx \\
&= \left(x^2 + \frac{1}{8}\ln(x)\right)\Big|_{x=1}^{4} \\
&= (16 + \frac{1}{8}\ln(4)) - (1 + \frac{1}{8}\ln(1)) \\
&= 15 + \frac{\ln(4)}{8}.
\end{aligned}
$$

(Return)

3. A simple parametrization for the circle of radius r is

$$
\begin{aligned}
x &= r\cos t\\
y &= r\sin t.
\end{aligned}
$$

Note that t ranges from 0 to 2π. Using the above formula, we find that the arclength element is

$$
\begin{aligned}
dL &= \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\,dt\\
&= \sqrt{(-r\sin t)^2 + (r\cos t)^2}\,dt\\
&= \sqrt{r^2(\sin^2 t + \cos^2 t)}\,dt\\
&= \sqrt{r^2}\,dt\\
&= r\,dt.
\end{aligned}
$$

(we used the Pythagorean identity $\sin^2 t + \cos^2 t = 1$ from line three to line four). Therefore,

$$
\begin{aligned}
L &= \int dL\\
&= \int_{t=0}^{2\pi} r\,dt\\
&= r\cdot t\Big|_{t=0}^{2\pi}\\
&= 2\pi r,
\end{aligned}
$$

as desired.

(Return)

4. First, compute $x'(t) = \cos(t) - t\sin(t)$ and $y'(t) = \sin(t) + t\cos(t)$. Then

$$
\begin{aligned}
\sqrt{x'(t)^2 + y'(t)^2} &= \sqrt{(\cos(t) - t\sin(t))^2 + (\sin(t) + t\cos(t))^2}\\
&= \sqrt{\cos^2(t) - 2t\cos(t)\sin(t) + t^2\sin^2(t) + \sin^2(t) + 2t\cos(t)\sin(t) + t^2\cos^2(t)}\\
&= \sqrt{1+t^2}.
\end{aligned}
$$

Thus, $dL = \sqrt{1+t^2}dt$. So one finds that

$$
L = \int_0^{2\pi} \sqrt{1+t^2}dt.
$$

This integral was computed in the Trigonometric Substitution module. The answer becomes

$$
\begin{aligned}
L &= \left(\frac{1}{2}\sinh^{-1}(t) + \frac{1}{2}t\sqrt{1+t^2}\right)\Big|_0^{2\pi}\\
&= \frac{1}{2}\sinh^{-1}(2\pi) + \pi\sqrt{1+4\pi^2}.
\end{aligned}
$$

(Return)

5. Computing the arclength element, we find

$$\begin{aligned} dL &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= \sqrt{1 + \sqrt{x}^2}\, dx \\ &= \sqrt{1+x}\, dx. \end{aligned}$$

Therefore, the arclength is

$$\begin{aligned} L &= \int dL \\ &= \int_{x=0}^{3} \sqrt{1+x}\, dx \\ &= \frac{2}{3}(1+x)^{3/2}\Big|_{x=0}^{3} \\ &= \frac{16}{3} - \frac{2}{3} \\ &= \frac{14}{3}. \end{aligned}$$

(Return)

6. Using the formula for the arclength element, the fact tells us that

$$\frac{d}{dx}\left(\frac{dy}{dx}\right) = \kappa\sqrt{1 + \left(\frac{dy}{dx}\right)^2}$$

Now, making a substitution of

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

simplifies the equation to become

$$\frac{du}{dx} = \kappa\sqrt{1+u^2}.$$

This is a separable differential equation. Separating and integrating gives

$$\int \frac{du}{\sqrt{1+u^2}} = \int \kappa\, dx.$$

The left side can be handled with either a trigonometric or hyperbolic trigonometric substitution. We take the latter approach, and let

$$\begin{aligned} u &= \sinh t \\ du &= \cosh t\, dt. \end{aligned}$$

So we have (remembering the Pythagorean identity for hyperbolic trigonometric functions from the trigonometric substitution module)

$$\begin{aligned}
\int \frac{du}{\sqrt{1+u^2}} &= \int \frac{\cosh t}{\sqrt{1+\sinh^2 t}}\, dt \\
&= \int \frac{\cosh t}{\sqrt{\cosh^2 t}}\, dt \\
&= \int \frac{\cosh t}{\cosh t}\, dt \\
&= \int dt \\
&= t \\
&= \operatorname{arcsinh} u
\end{aligned}$$

(we leave the constant of integration off for now since we will be integrating on the right side as well). On the right side, we have

$$\int \kappa\, dx = \kappa x + C.$$

Putting it together, we have

$$u = \sinh(\kappa x + C).$$

If we pick our coordinates so that $x = 0$ occurs at the low point of the catenary, then note that at this point, we have

$$u = \frac{dy}{dx} = 0,$$

since the slope of the catenary is 0 at the low point. Using this fact and plugging in $x = 0$ into the earlier equation gives

$$u = \sinh(C) = 0$$

and so $C = 0$. This gives

$$u = \frac{dy}{dx} = \sinh(\kappa x).$$

Now integrating both sides gives

$$y = \frac{1}{\kappa}\cosh(\kappa x) + C,$$

where $C = y_0$ is the y value of the low point of the catenary.

To find the length of the catenary, we have

$$
\begin{aligned}
L &= \int dL \\
&= \int_{x=-l}^{l} \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx \\
&= \int_{x=-l}^{l} \sqrt{1+\sinh^2 \kappa x}\, dx \\
&= \int_{x=-l}^{l} \cosh \kappa x\, dx \\
&= \frac{1}{\kappa} \sinh \kappa x \Big|_{x=-l}^{l} \\
&= \frac{1}{\kappa} \left(\sinh \kappa l - \sinh \kappa(-l)\right) \\
&= \frac{2}{\kappa} \sinh \kappa l.
\end{aligned}
$$

since hyperbolic sine is an odd function. This grows very quickly as l increases, because

$$\frac{2}{\kappa} \sinh \kappa l \approx \frac{1}{\kappa} e^{\kappa l}.$$

(Return)

7. Computing

$$
\begin{aligned}
\frac{dx}{dt} &= \frac{-t \sin t - \cos t}{t^2} \\
\frac{dy}{dt} &= \frac{t \cos t - \sin t}{t^2}.
\end{aligned}
$$

Plugging these into the formula for the arclength of a parametric curve and noting the cancellation of cross terms, we have

$$
\begin{aligned}
dL &= \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt \\
&= \sqrt{\left(\frac{-t \sin t - \cos t}{t^2}\right)^2 + \left(\frac{t \cos t - \sin t}{t^2}\right)^2}\, dt \\
&= \sqrt{t^2 \cdot \frac{\sin^2 t + \cos^2 t}{t^4} + \frac{\cos^2 t + \sin^2 t}{t^4}}\, dt \\
&= \sqrt{\frac{1}{t^2} + \frac{1}{t^4}}\, dt \\
&= \frac{\sqrt{t^2+1}}{t^2}\, dt.
\end{aligned}
$$

Therefore, the arclength is

$$\int_{t=2\pi}^{\infty} \frac{\sqrt{t^2+1}}{t^2}\, dt.$$

This integral is difficult to compute exactly, but we only want to show it diverges, which is not as difficult. Note that

$$\frac{\sqrt{t^2+1}}{t^2} \geq \frac{\sqrt{t^2}}{t^2} = \frac{t}{t^2} = \frac{1}{t}.$$

And so by the dominance of definite integrals,

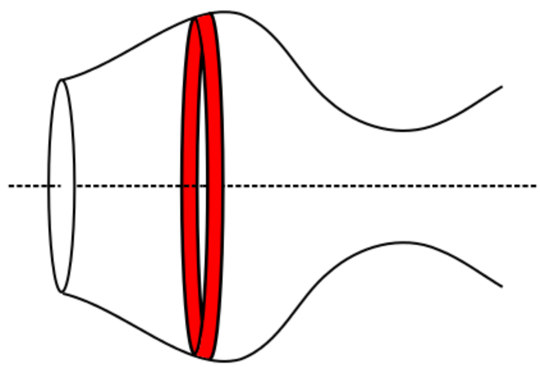
$$\int_{t=2\pi}^{\infty} \frac{\sqrt{t^2+1}}{t^2}\,dt \geq \int_{t=2\pi}^{\infty} \frac{1}{t}\,dt$$

but the integral on the right diverges to infinity by our earlier discussions of p-integrals. Thus, our integral on the left, being larger, also diverges to infinity.

(Return)

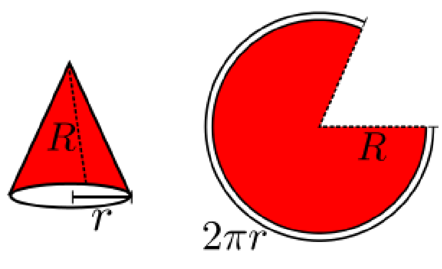
36 Surface Area

This module deals with the surface area of solids of revolution. Consider the portion of a curve $y = f(x)$ for $a \leq x \leq b$ revolved about a horizontal axis to create a solid. In earlier modules the goal was to find the volume of such a solid, but now the focus is on finding the surface area. As always, the method will be to find the surface area element and integrate it. The surface area element which works well is the thin band shown here:



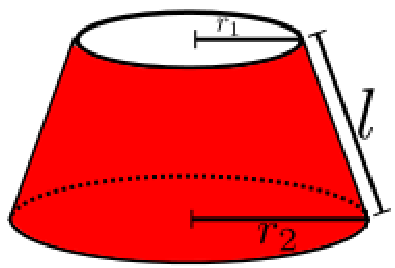
36.1 Surface area of a cone

The first step towards finding the surface area element is to find the lateral surface area of a more simple solid: the cone. Consider a cone whose base has radius r and lateral height R (the lateral height is the distance from the tip of the cone to a point on the circumference of the base; see the left diagram below).



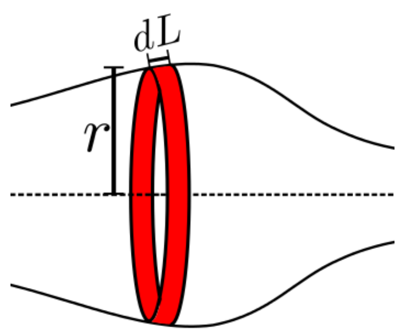
To find the area, consider cutting the cone along the straight dotted line from base circumference to tip and unrolling the cone. The result is a portion of a circle whose radius is R, as shown on the right in the diagram above. Note that the circumference of the base of the cone, $2\pi r$, becomes the length of arc of the unrolled cone. This means that the unrolled cone is a fraction of the full circle of radius R, and that fraction is $\frac{2\pi r}{2\pi R}$ (the ratio of the circumference of the partial circle to the circumference of the whole circle). Thus the surface area of the cone is $\frac{2\pi r}{2\pi R}\pi R^2 = \pi r R$.

The surface area of a cone can be used to find the area of a frustum of a cone whose top radius is r_1, bottom radius is r_2, and lateral height l (as in the below diagram). The area of this frustum is $\pi(r_1 + r_2)l$. Expressed another way, the area is $2\pi r l$, where $r = \frac{r_1+r_2}{2}$ is the average of the two radii of the frustum.



36.2 Surface area element

Now, the surface area element can be found. When the curve is partitioned into sufficiently small pieces, the surface area element is just the area of the frustum formed by rotating the arclength element about the axis (see the diagram):



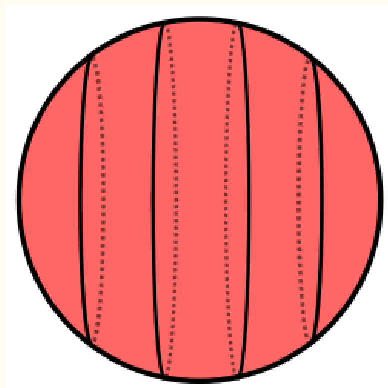
Thus, the surface area element is $dS = 2\pi r dL$, where r is the distance from the curve to the axis of rotation, and dL is the arclength element (i.e. $dL = \sqrt{1 + (f'(x))^2}dx$). In the (common) case where the axis of rotation is the x-axis, one finds that $r = f(x)$.

Thus, the surface area resulting from revolving the curve $y = f(x)$ for $a \leq x \leq b$ about the x-axis is given by

$$\begin{aligned} S &= \int_a^b 2\pi r \, dL \\ &= 2\pi \int_a^b f(x)\sqrt{1 + (f'(x))^2}\, dx. \end{aligned}$$

Example

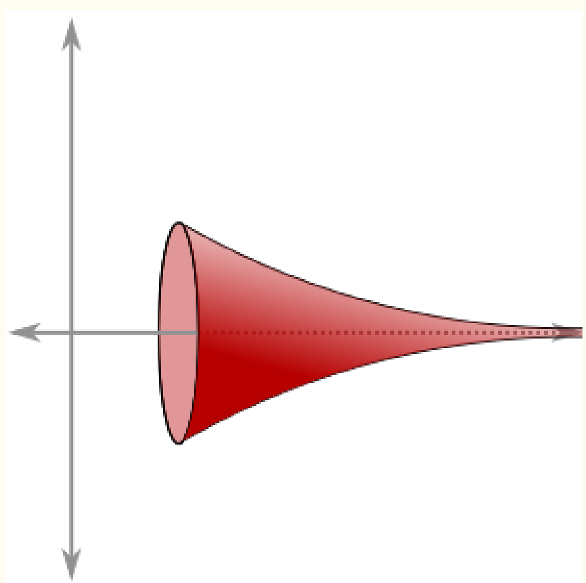
Consider the sphere of radius r. If the sphere is cut into slices of equal width, which slice has the most surface area?



(See Answer 1)

Example

Consider the surface generated by revolving the curve $y = \frac{1}{x^p}$ for $x \geq 1$ about the x-axis.



Find the values of p for which the surface has finite surface area. Then find the values of p for which the solid of revolution has finite volume. (See Answer 2)

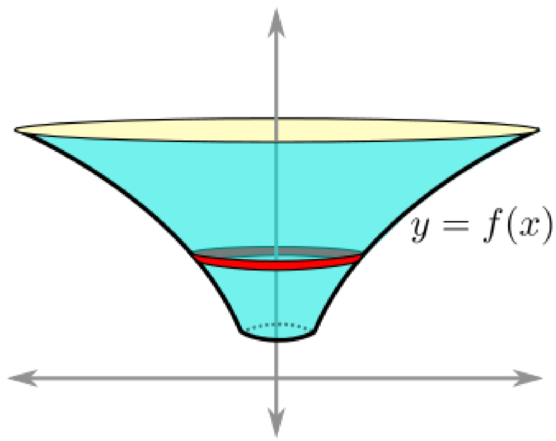
36.3 Rotations about the y-axis

Suppose we want to know the surface area which results from revolving the curve

$$y = f(x); a \leq x \leq b$$

about the y-axis. There are two main ways one can go about finding this surface area:

1. Express everything as a function of y (including range of inputs), and then use the above formula but with the roles of x and y switched.
2. Leave things in terms of x, but adjust the formula slightly.



The first method expresses the curve as

$$x = f^{-1}(y); c \leq y \leq d,$$

where $c = f^{-1}(a)$ and $d = f^{-1}(b)$. Then express the surface area element as

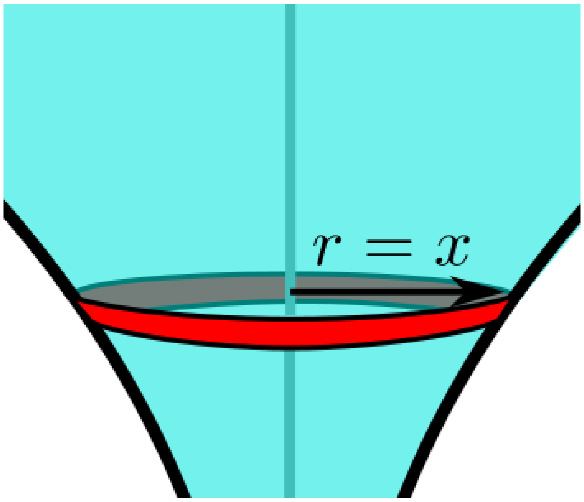
$$\begin{aligned} dS &= 2\pi r\, dL \\ &= 2\pi f^{-1}(y)\sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy. \end{aligned}$$

Putting it together, the surface area can be expressed as

$$\begin{aligned} S &= \int dS \\ &= 2\pi \int_{y=c}^{d} f^{-1}(y)\sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy. \end{aligned}$$

Again, this is really just a reuse of the original formula, with the roles of x and y flipped.

The second method is sometimes simpler to apply because it involves less algebra. The main observation to make is that the radius in the surface area element is simply x when the curve is revolved around the y-axis:



So the surface area element can be written

$$\begin{aligned} dS &= 2\pi r\, dL \\ &= 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx. \end{aligned}$$

This integral is with respect to x, and so it should be integrated over the original range of x:

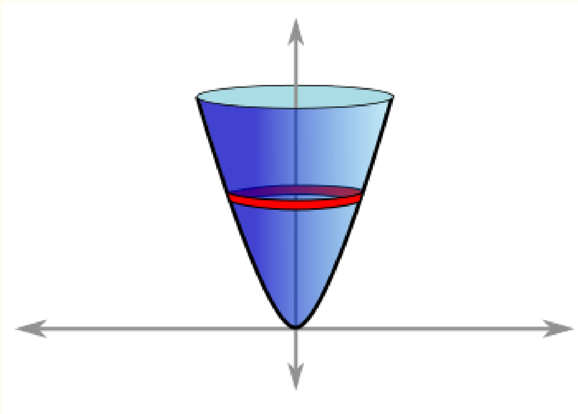
$$S = 2\pi \int_{x=a}^{b} x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx.$$

Example

Compute the surface area of the surface resulting from revolving the curve

$$y = \frac{1}{2}x^2; \quad 0 \leq x \leq 4.$$

about the y-axis:



(See Answer 3)

36.4 EXERCISES

- Compute the surface area resulting from revolving the curve $f(x) = \cosh(x)$, $0 \leq x \leq 2$ about the x-axis.

36.5 Answers to Selected Examples

1. If we center the sphere at the origin, we can think of the sphere as the surface of revolution obtained by revolving the curve
$$y = \sqrt{r^2 - x^2}; \quad -r \leq x \leq r$$
about the x-axis. First, we compute the arclength element:
$$\begin{aligned} dL &= \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx \\ &= \sqrt{1 + \left(\frac{-x}{\sqrt{r^2 - x^2}}\right)^2}\, dx \\ &= \sqrt{1 + \frac{x^2}{r^2 - x^2}}\, dx \\ &= \sqrt{\frac{r^2}{r^2 - x^2}}\, dx \\ &= \frac{r}{\sqrt{r^2 - x^2}}\, dx. \end{aligned}$$
Plugging this into the surface area element, we find
$$\begin{aligned} dS &= 2\pi y\, dL \\ &= 2\pi\sqrt{r^2 - x^2} \cdot \frac{r}{\sqrt{r^2 - x^2}}\, dx \\ &= 2\pi r\, dx. \end{aligned}$$
Note that this is independent of x! This means that every slice of the sphere has equal surface area.
For example, if we were to slice the sphere into four slices of equal thickness, then a middle slice goes from $x = 0$ to $x = \frac{r}{2}$, and its surface area
$$\begin{aligned} \int_{x=0}^{r/2} 2\pi r\, dx &= 2\pi r x\Big|_{x=0}^{r/2} \\ &= 2\pi r \cdot \frac{r}{2} \\ &= \pi r^2. \end{aligned}$$
The end-cap slice, on the other hand, goes from $x = \frac{r}{2}$ to $x = r$, so its surface area is
$$\begin{aligned} \int_{x=r/2}^{r} 2\pi r\, dx &= 2\pi r x\Big|_{x=r/2}^{r} \\ &= 2\pi r\left(r - \frac{r}{2}\right) \\ &= 2\pi r \cdot \frac{r}{2} \\ &= \pi r^2. \end{aligned}$$

So we see that the pieces have equal surface area.

(Return)

2. The surface area, in terms of p, is

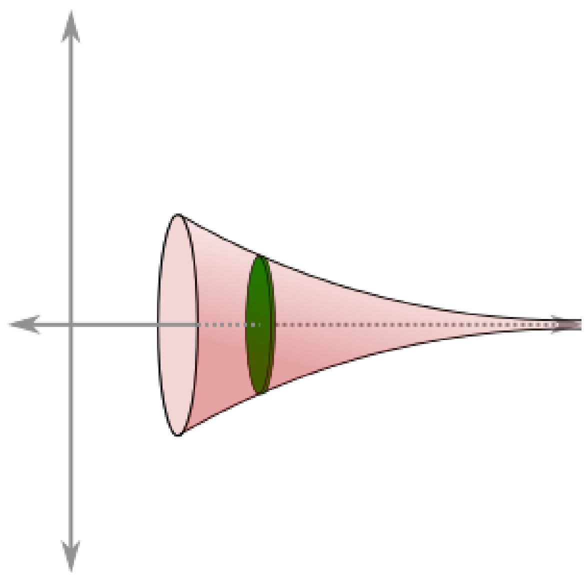
$$S = 2\pi \int_1^\infty \frac{1}{x^p}\sqrt{1+(-px^{-p-1})^2}dx$$
$$= 2\pi \int_1^\infty \frac{1}{x^p}\sqrt{1+\frac{p^2}{x^{2p+2}}}dx$$

Unfortunately, this integral is not computable using standard methods, but we can use a binomial expansion to determine the leading order term of the integrand, which will tell us whether the integral converges or not. We see that

$$\frac{1}{x^p}\sqrt{1+\frac{p^2}{x^{2p+2}}} = \frac{1}{x^p}\left(1+\frac{p^2}{x^{2p+2}}\right)^{1/2}$$
$$= \frac{1}{x^p}\left(1+\frac{1}{2}\cdot\frac{p^2}{x^{2p+2}} + O\left(\frac{1}{x^{4p+4}}\right)\right)$$
$$= \frac{1}{x^p} + O\left(\frac{1}{x^{3p+2}}\right).$$

Therefore, the leading order term in this integral is $\frac{1}{x^p}$, which we know converges for $p > 1$ and diverges for $p \leq 1$ (from our study of p-integrals). So this surface of revolution has finite area if and only if $p > 1$.

Turning to the volume of this solid, it is best to use slices perpendicular to the x-axis, which leads to discs whose radius is y:



The volume element is therefore

$$dV = \pi y^2\, dx$$
$$= \pi \left(\frac{1}{x^p}\right)^2\, dx$$
$$= \pi \cdot \frac{1}{x^{2p}}\, dx.$$

Thus, the volume is

$$\begin{aligned} V &= \int dV \\ &= \pi \int_{x=1}^{\infty} \frac{1}{x^{2p}}\, dx. \end{aligned}$$

We know this is convergent if $2p > 1$, i.e. $p > \frac{1}{2}$. So the volume of the solid is finite if $p > \frac{1}{2}$. This leads to the surprising fact that for

$$\frac{1}{2} < p \leq 1,$$

the volume of the solid is finite, but the surface area is infinite.

(Return)

3. Using the first method requires some algebra. The curve becomes

$$x = \sqrt{2y}; \quad 0 \leq y \leq 8.$$

So the area element is

$$\begin{aligned} dS &= 2\pi r\, dL \\ &= 2\pi r \sqrt{1 + \left(\frac{dx}{dy}\right)^2}\, dy \\ &= 2\pi \sqrt{2y} \sqrt{1 + \left(\frac{1}{\sqrt{2y}}\right)^2}\, dy \\ &= 2\pi \sqrt{2y+1}\, dy. \end{aligned}$$

So the surface area is

$$\begin{aligned} S &= \int dS \\ &= 2\pi \int_{y=0}^{8} \sqrt{2y+1}\, dy \\ &= 2\pi (2y+1)^{3/2} \cdot \frac{1}{3} \Big|_{y=0}^{8} \\ &= \frac{2}{3}\pi \left(17^{3/2} - 1\right). \end{aligned}$$

Using the second method, we have

$$\begin{aligned} dS &= 2\pi r\, dL \\ &= 2\pi x \sqrt{1 + \left(\frac{dy}{dx}\right)}\, dx \\ &= 2\pi x \sqrt{1 + x^2}\, dx. \end{aligned}$$

So the surface area is

$$\begin{aligned} S &= \int dS \\ &= \pi \int_{x=0}^{4} 2x\sqrt{1+x^2}\,dx \\ &= \pi(1+x^2)^{3/2} \cdot \frac{2}{3}\Big|_{x=0}^{4} \\ &= \frac{2}{3}\pi\left(17^{3/2}-1\right). \end{aligned}$$

So we get the answer with (perhaps) slightly less algebra involved.

(Return)

37 Work

Recall that *work* is the amount of energy required to perform some action. When the amount of force is constant, work is simply

$$\text{work} = \text{force} \times \text{distance}.$$

For example, if a book weighing 22 Newtons (about 5 pounds) is lifted 2 meters, the total work done is $22N \times 2m = 44J$ (J is the Joule, which equals one Newton-meter).

Consider a situation where the force is not constant. For instance, if one were to lift a weight using a non-negligible rope, there is less rope being pulled up (and hence less force) as the weight goes further up. It is in situations like these that we need a better formula to compute work.

37.1 Work element

Computing work when the force is not constant requires integration. As in previous sections, the first step is to determine the *work element* dW, and then integrate:

$$W = \int dW.$$

Because work arises in a variety of situations, there is not one simple formula for the work element. For different applications the work element will look different. In some situations, it is best to consider a small movement dx, where the force can be thought of as constant for that small movement, which allows the work element to be expressed as $dW = F \cdot dx$.

Springs

The force required to displace a spring varies with the displacement. The further the spring is stretched, the more resistant it becomes to being stretched further. Consider three types of springs:

- **Linear**. A spring is linear if the force of resistance grows linearly with the displacement. That is,

$$F(x) = \kappa x.$$

for some constant κ, which represents the stiffness of the spring.

- **Hard**. A spring is hard if the force of resistance grows faster than linearly with the displacement:

$$F(x) = \kappa x + O(x^2).$$

- **Soft**. A spring is soft if the force of resistance grows slower than linearly with the displacement:

$$F(x) = \kappa x - O(x^2).$$

Consider for any of these springs what the work element dW is. When the spring is stretched to x, the force of resistance is $F(x)$. For the next infinitesimal amount of stretching dx, the force can be presumed to be constant:

(Stretching Spring Animated GIF)

Therefore, the work element (i.e. the amount of work to stretch the spring the additional amount dx) is

$$dW = F(x)\, dx.$$

Example

Compute the amount of work it takes to stretch a linear spring from rest (when $x = 0$) to $x = a$. (See Answer 1)

Example

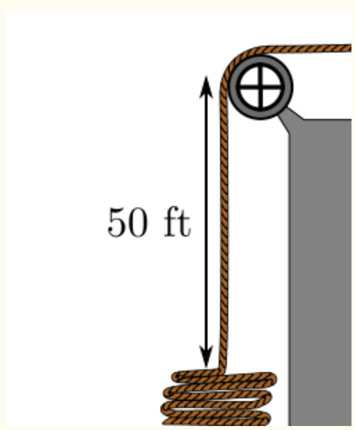
Consider a nonlinear, soft spring which exerts a force of $F(x) = 3x - x^2$ Newtons when the spring is stretched to x meters. Determine how much work is required to stretch the spring from 1 meter to 3 meters. (See Answer 2)

Pulling up a rope

In some situations, one must do a little work to determine what $F(x)$ is, and then one can integrate, as in the above examples.

Example

Consider a rope which is 100 feet long and density 1 pound/foot. It hangs from a wall which is 50 feet high (so 50 feet of rope runs down the length of the wall and the remaining 50 feet is coiled at the bottom of the wall). How much work (in foot-pounds) is required to pull the rope to the top of the wall?



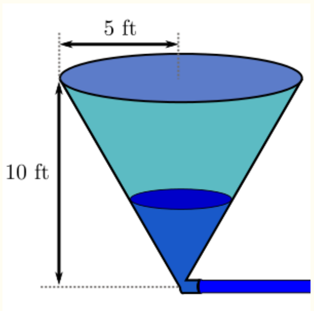
(See Answer 3)

37.2 Work element by slices

In other situations, such as pumping liquid, digging a hole, or piling gravel, a fruitful method for determining the work involved is to consider a slice of the material which is being moved. Determining the weight of the slice, and multiplying by the distance the slice has to be lifted gives the amount of work required for that slice. That is precisely the work element. Integrating over all the slices in the object gives the total amount of work to move that object.

Example

Pumping Liquid Consider an inverted conical tank (so the tip of the cone points downward) with base radius 5 feet and height 10 feet. Water is pumped into the tank through a valve at the tip of the cone:



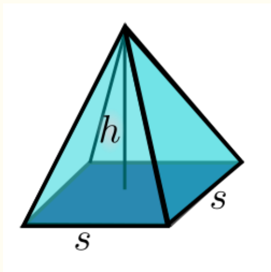
How much work is required to fill the tank with water? Leave the weight density of water as the constant ρ. (See Answer 4)

Example

Digging a Hole Consider two workers digging a hole. How deep should the first worker dig so that each does the same amount of work? Let the weight density of the dirt be the constant ρ, the depth of the hole is D, and the cross-sectional area of the hole is the constant A (so we assume that the hole does not get any wider or narrower as the workers dig). (See Answer 5)

Example

Gravel Pyramid Compute the amount of work required to build a pyramid of gravel. Assume the gravel is infinitesimal with weight density ρ, and that the pyramid has a square base of side length s, and height h:



(See Answer 6)

Example

Rope Revisited Consider the rope example from above, but this time suppose l total feet of rope are hanging from a h foot building, where $l \geq h$, and let ρ be the weight density of the rope. Compute the work required to lift the rope to the top of the building.
For a different perspective, this time, use a work element which equals the amount of work required to lift an infinitesimal length of rope to the top of the building (this will depend on whether the infinitesimal length of rope is hanging at the beginning or is part of the coil at the bottom of the building). Then integrate along the entire length of rope. (See Answer 7)

37.3 EXERCISES

- Consider a conical tank of height 10m. The vertex of the cone is at the bottom, and the base of cone (which is at height 10m) has radius 2m. Let ρ denote the weight density of water. The water inside the tank has height 4m. How much work would it take to pull all the water to the top of the tank?

37.4 Answers to Selected Examples

1. As shown above, the work element (i.e. the amount of work to stretch the spring a short distance dx) is

$$dW = F(x)\,dx = \kappa x\,dx.$$

It follows that

$$\begin{aligned} W &= \int dW \\ &= \int_{x=0}^{a} \kappa x\,dx \\ &= \kappa \frac{x^2}{2}\bigg|_{x=0}^{a} \\ &= \kappa \frac{a^2}{2}. \end{aligned}$$

(Return)

2. The work element is

$$\begin{aligned} dW &= F(x)\,dx \\ &= (3x - x^2)\,dx. \end{aligned}$$

Thus, the total work to stretch the spring from 1 meter to 3 meters is

$$\begin{aligned}
W &= \int dW \\
&= \int_{x=1}^{3} (3x - x^2)\, dx \\
&= \frac{3}{2}x^2 - \frac{1}{3}x^3 \Big|_{x=1}^{3} \\
&= \left(\frac{27}{2} - 9\right) - \left(\frac{3}{2} - \frac{1}{3}\right) \\
&= \frac{10}{3} \text{ Joules.}
\end{aligned}$$

(Return)

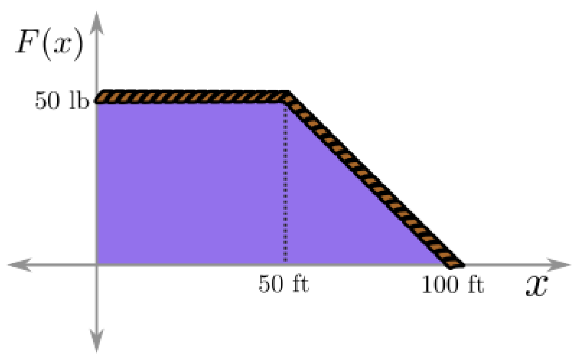
3. As the first 50 feet of rope are brought up, there is always precisely 50 feet of rope hanging from the building (because every foot of rope brought onto the top of the building is replaced by a rope which is coiled below). These 50 feet of rope weigh 50 lbs, so that is the force required to support them. If x denotes the amount of rope which has been taken onto the roof, then

$$F(x) = 50; \quad 0 \leq x \leq 50.$$

After the first 50 feet of rope have been brought to the roof, there is now 50 feet of rope dangling with nothing left coiled below. Therefore, as we bring up these last 50 feet, there is less and less rope hanging, and so the weight of the rope (and hence the force we exert) is decreasing. It decreases linearly, since the rope has constant density. Each foot of rope we bring up decreases the weight by 1 lb, and so

$$F(x) = 100 - x; \quad 50 \leq x \leq 100$$

(to see that this is right, note that it is linear and matches at the endpoints). We can graph the force as a function of the amount of rope we have brought up:



Now, we can find the work by integrating the work element

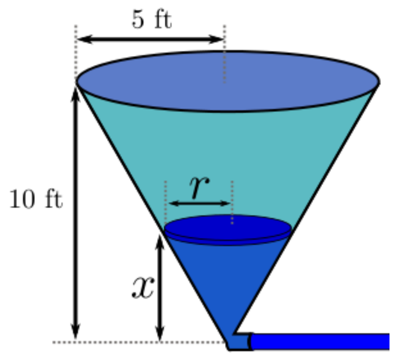
$$\begin{aligned}
W &= \int dW \\
&= \int_{x=0}^{100} F(x)\, dx.
\end{aligned}$$

Note that this is the area under the graph of the force (highlighted above), which is easier to compute than to do it algebraically. Splitting it into a square and a triangle, the area (and hence the work) is

$$50 \text{ lb} \times 50 \text{ ft} + \frac{1}{2} 50 \text{ lb} \times 50 \text{ ft} = 3750 \text{ ft-lb}.$$

(Return)

4. Consider a slice of the water in the tank. Let x be the distance of the slice from the tip of the tank. That is, x is the distance that the slice of water has to be lifted. Let r be the radius of the slice:



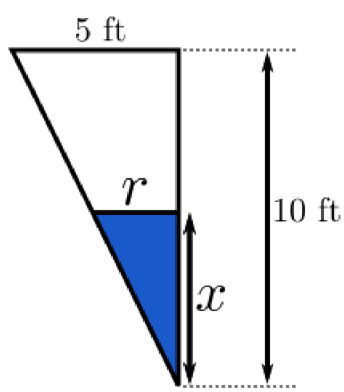
Above, we said that it is the weight of the slice multiplied by the distance the slice had to be moved. But the weight of a slice is just the volume of the slice times the density of the slice. Letting ρ denote the weight density of the substance (in this case water), we have

$$\begin{aligned} dW &= \text{ weight of slice } \times distanceslicetravels \\ &= \rho \cdot dV \cdot \text{ distance slice travels }. \end{aligned}$$

In the problem at hand, we have

$$dV = \pi r^2 \, dx,$$

and the distance the slice is lifted is x, by the way we labeled our diagram. To finish, we must get r in terms of x, which requires a little bit of geometry. If we flatten our cone and look at it from the side, we get similar triangles:



Therefore,

$$\frac{r}{x} = \frac{5}{10} = \frac{1}{2},$$

and so $r = \frac{x}{2}$. Putting this together, we have

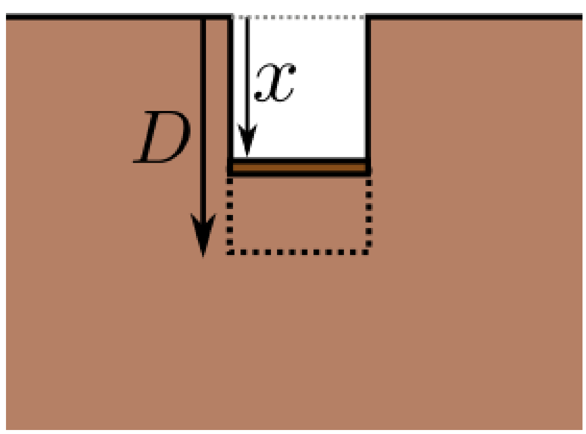
$$\begin{aligned} dW &= \rho \cdot dV \cdot \text{ distance slice travels} \\ &= \rho(\pi r^2\, dx)x \\ &= \pi\rho\frac{1}{4}x^3\, dx. \end{aligned}$$

Note that x ranges from 0 to 10, so the work required to fill the tank is

$$\begin{aligned} W &= \int dW \\ &= \int_{x=0}^{10} \pi\rho\frac{1}{4}x^3\, dx \\ &= \frac{\pi\rho}{4}\frac{x^4}{4}\bigg|_{x=0}^{10} \\ &= 625\pi\rho. \end{aligned}$$

(Return)

5. Let x be the distance down to to the layer of dirt currently being dug:



This is convenient because this is the distance that the current slice of dirt has to be lifted to get out of the hole. The area of the slice of dirt is A, its thickness is dx, and the density is ρ, so we have

$$\begin{aligned} dW &= \text{ weight of slice } \times \text{ distance slice moves} \\ &= (\rho \cdot dV) \cdot x \\ &= (\rho A\, dx) \cdot x \\ &= \rho A x\, dx. \end{aligned}$$

Note that x varies from 0 to D as the hole gets dug. Thus, the total work required to dig the hole is

$$\begin{aligned} W &= \int dW \\ &= \int_{x=0}^{D} \rho A x \, dx \\ &= \frac{1}{2}\rho A D^2. \end{aligned}$$

To find the depth $\tilde{D}$ where the work done is half, we set

$$\int_{x=0}^{\tilde{D}} \rho A x \, dx = \frac{1}{2}W = \frac{1}{4}\rho A D^2.$$

Computing the integral on the left, we find

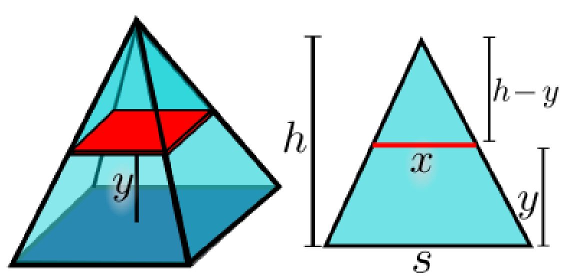
$$\frac{1}{2}\rho A \tilde{D}^2 = \frac{1}{4}\rho A D^2.$$

Solving for $\tilde{D}$ gives

$$\tilde{D} = \frac{1}{\sqrt{2}} D.$$

(Return)

6. If we think of building the pyramid slice by slice, let y be the distance from the base of the pyramid to the slice. This is convenient because this is the distance that the slice must be lifted to be put in place. Also, let x be the side length of the slice:



Then using similar triangles, as shown on the right above, we find that

$$\frac{x}{s} = \frac{h-y}{h}.$$

So we find that

$$x = \frac{h-y}{h} s.$$

Thus, the volume of a slice is just the area x^2 multiplied by the thickness dy, and so we have

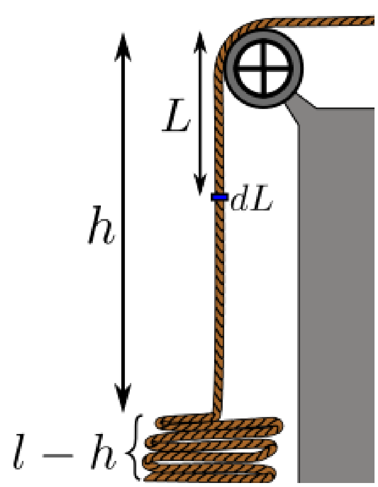
$$\begin{aligned} dW &= \rho dV\, y \\ &= \rho(x^2\, dy)\, y \\ &= \rho\left(\frac{h-y}{h}s\right)^2 y\, dy \\ &= \frac{\rho s^2}{h^2}(h-y)^2 y\, dy. \end{aligned}$$

Because y ranges from 0 to h, we have

$$\begin{aligned} W &= \int dW \\ &= \frac{\rho s^2}{h^2}\int_{y=0}^{h}(h-y)^2 y\, dy \\ &= \frac{\rho s^2}{h^2}\int_{y=0}^{h}(h^2y-2hy^2+y^3)\, dy \\ &= \frac{\rho s^2}{h^2}\left(\frac{1}{2}h^2y^2-\frac{2}{3}hy^3+\frac{1}{4}y^4\right)\Bigg|_{y=0}^{h} \\ &= \frac{\rho s^2}{h^2}\left(\frac{1}{2}h^4-\frac{2}{3}h^4+\frac{1}{4}h^4\right) \\ &= \frac{\rho s^2}{h^2}\cdot\frac{1}{12}h^4 \\ &= \frac{\rho s^2 h^2}{12}. \end{aligned}$$

(Return)

7. Let L be the distance along the rope of the infinitesimal piece being considered:



So L is the distance that the infinitesimal piece must be lifted to get to the top of the building. The weight of the infinitesimal piece is density multiplied by length, and so the work element for a piece of rope which is hanging is

$$dW = \rho L\, dL; \quad 0 \le L \le h.$$

For a piece of rope which is part of the coil at the bottom, the distance it must be lifted is always h, so the work element there is

$$dW = \rho h \, dL; \quad h \le L \le l.$$

So the work can be computed by integrating these work elements over their respective ranges and then adding:

$$\begin{aligned} W &= \int dW \\ &= \int_{L=0}^{h} \rho L \, dL + \int_{L=h}^{l} \rho h \, dL \\ &= \rho \frac{1}{2} L^2 \Big|_{L=0}^{h} + \rho h L \Big|_{L=h}^{l} \\ &= \frac{\rho h^2}{2} + \rho h(l-h). \end{aligned}$$

Another way to think about this is to treat the coiled rope at the bottom of the wall as a single solid object. The rope in the coil has length $l - h$, and so its weight is $\rho(l-h)$. The distance the coil (as a unit) must be lifted is h. It follows that the work to lift the coiled portion of the rope is $\rho h(l-h)$, the result of the second integral above.

(Return)

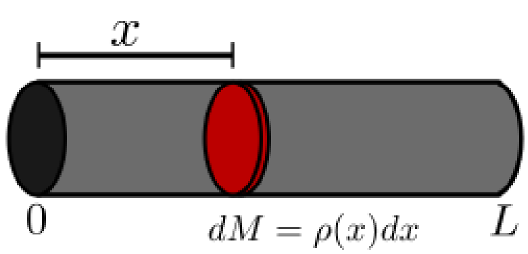
38 Elements

This module deals with various problems that can be modeled using integral calculus. As in the previous sections, the problem will be to find the total accumulation of some quantity U, and the method will be to determine a slice of the quantity, the U element dU, and integrate.

38.1 Mass

Mass of a rod

Consider the problem of determining the mass of a rod. Suppose the rod's density varies along the length of the rod (but the rod is uniform in cross section). Let $\rho(x)$ denote the linear density (i.e. the mass per unit of length) of the rod at position x:

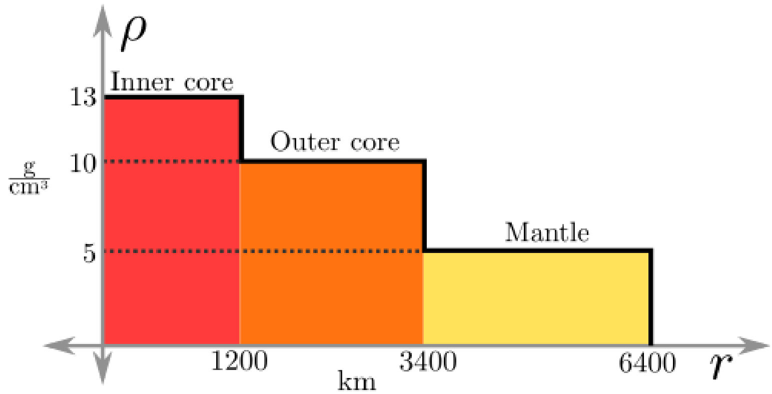


Then the mass element dM is the density $\rho(x)$ times the thickness of the slice dx, as shown above, and it follows that the mass of the rod is

$$\begin{aligned} M &= \int dM \\ &= \int_{x=0}^{L} \rho(x) dx. \end{aligned}$$

Mass of the earth

Consider the problem of finding the mass of the earth. Suppose the density of the earth $\rho(r)$ is given as a function of the distance from the center of the earth. Assume that there are just three layers (inner core, outer core, and mantle) and that the density is constant within each layer.



What is the mass element in this case? It is important to note that in this example we are measuring the contribution of a spherical shell to the mass of the earth. This contribution is the volume of the spherical shell multiplied by the density of the shell. Mathematically,

$$dM = \rho(r) \cdot dV.$$

Recalling that the surface area of a sphere of radius r is $4\pi r^2$, we have that the volume element is

$$dV = 4\pi r^2 \, dr,$$

and so the mass element is

$$dM = 4\pi\rho(r)r^2 \, dr.$$

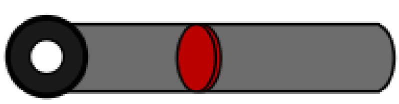
Example

Using the approximate graph of density above, estimate the mass of the earth. (See Answer 1)

38.2 Torque

Imagine a rod of variable density which is attached to a hinge. The torque at the hinge depends not just on the weight of the rod but on the distribution of the weight.

If there were just a mass-less rod with a single point mass, the torque would be Force $\times$ Distance . This can be used to determine the torque element dT by thinking of each slice of the rod as a point mass. What is the torque on such a slice?



$$dT = x\,dF$$
$$dF = g\,dM$$
$$dM = \rho(x)\,dx$$

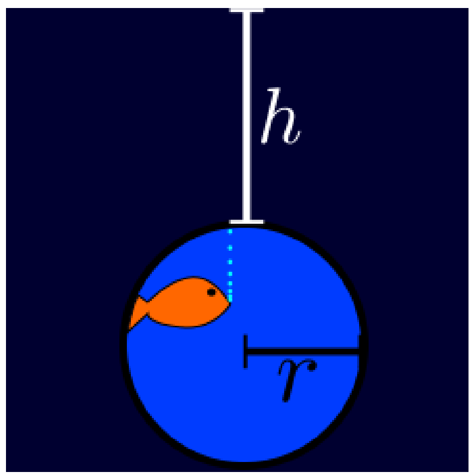
First, the torque element is the distance from the hinge, x, times the force element dF (the force on the slice). The force element is the mass of the slice dM times the gravitational constant g. Finally, as in the previous example, the mass element $dM = \rho(x)\,dx$. Putting it all together, one finds

$$dT = x \cdot g \cdot \rho(x)\,dx.$$

Integrating this over the length of the rod gives the torque.

38.3 Hydrostatic force

The next application is to compute the total force exerted by a tank of fluid on a surface submerged in the tank, often called the *hydrostatic force*. For a tangible example, consider a large aquarium with a circular glass viewing window (see the diagram below). If the viewing window has radius r, and the top of the viewing window is at depth h, then the problem is to find the total force of the water on the viewing window.



As always, the method will be to find the force element dF (the force on a small strip of the window), and then use integration to find the total force.

Recall that if pressure is constant across a surface, the force on the surface is area $\times$ pressure. Hydrostatic pressure is given by

$$P = \text{weight density of fluid} \times \text{depth}.$$

Note the units: $\frac{N}{m^3} \cdot m = \frac{N}{m^2}$, which is the correct unit for pressure (force per unit of area). Since the density of the fluid is assumed to be constant, the pressure only depends on the depth.Therefore, the most logical choice for the force element is a horizontal strip, since the depth, and hence the pressure, will be constant across the strip. Letting dA denote the area of the strip, we find that the force element is given by

$$dF = P\,dA = \rho x\,dA,$$

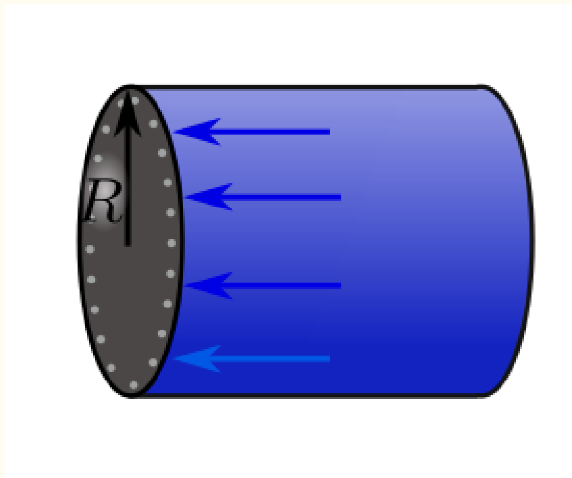
where ρ is the weight density of the fluid and x is the depth of the strip.

Example

Compute the total force exerted on the circular viewing window in the aquarium shown above. (See Answer 2)

Example

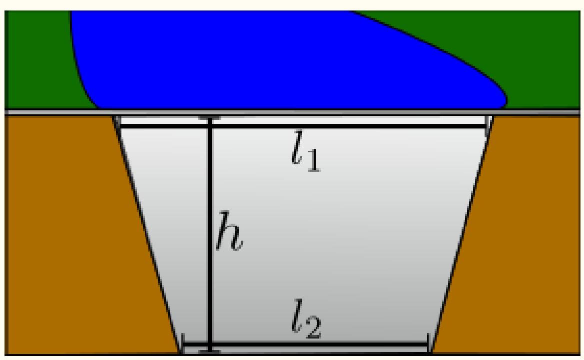
Compute the force on the endcap of a full cylindrical tank of radius R on its side.



(See Answer 3)

Example

Consider a dam in the shape of a trapezoid with height h, top edge l_1 and bottom edge l_2. Find the total force exerted on the dam by the water:



(See Answer 4)

38.4 Present value

Consider the problem of determining the present value of some amount of money at a future time. Turning the problem around, first consider the value of an initial amount of money P_0 at a future time t. Assuming a constant annual nominal interest rate r and continuous compounding, this problem was an example of exponential growth, and had solution

$$P(t) = P_0 e^{rt},$$

where t is the time in years. Given some amount of money, P, at time t, finding its present value is a matter of solving $P = P_0 e^{rt}$ for P_0. In other words, solving for present value in this simple case is the same as finding the initial investment P_0 which yields P after t years of continuous compounding interest. Solving this equation gives that the present value of a future amount P at time t is given by

$$P_0 = Pe^{-rt}.$$

Example

Find the present value of \$1000000 in 30 years, assuming an interest rate of $r = .08$. (See Answer 5)

Now consider an income stream, say from a job. If $I(t)$ is the rate of income at time t, what is the present value of that income stream? Let PV be the present value. Then consider the income earned over a small amount of time t years in the future:

$$dI = I(t)\,dt$$

(the income element). This small bit of income at time t contributes $e^{-rt}I(t)dt$ to the present value of the income stream. Thus the present value element is given by

$$dPV = e^{-rt}I(t)dt.$$

Integrating this over the range of values of t (the time period of the income stream) gives the present value of that income stream.

Example

The Bigbucks lottery has an option of either a single lump sum payment today or an annuity which pays a constant amount each year for 20 years. Suppose the annuity pays \$3 million a year (for 20 years), and that the interest rate will remain steady at $r = .05$. What is the fair lump sum payout today? (See Answer 6)

38.5 EXERCISES

- Consider a dam with the shape of an isosceles triangle. The base of the triangle, which is parallel to the ground, is 5m long, and the height of the triangle is 10m. The weight density of water is given by ρ. Compute the force exerted on the dam by water.

38.6 Answers to Selected Examples

1. Note that our volume is being measured in cubic kilometers, but the density $\rho(r)$ is in grams per cubic centimeter. We need a conversion factor C to make sure the units come out correctly. A little unit conversion gives us that

$$C = 1\frac{\text{g}}{\text{cm}^3} = 10^{12}\frac{\text{kg}}{\text{km}^3}.$$

So we need to multiply by this so that the units are correct (and the final answer will be in kilograms).

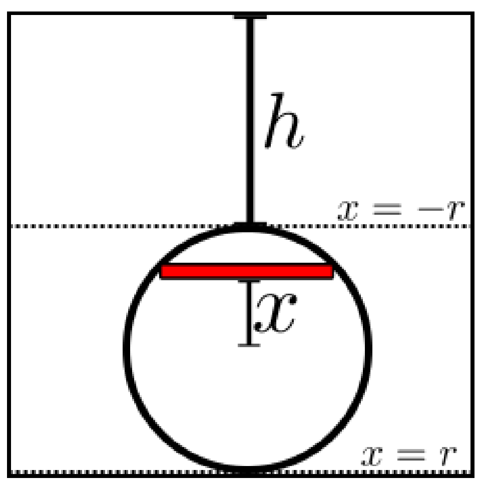
Splitting the integral based on the values of r for which $\rho(r)$ is constant, we find

$$\begin{aligned} M &= \int dM \\ &= C \int_{r=0}^{6400} 4\pi r^2 \rho(r)\, dr \\ &= 4\pi C \left(\int_{r=0}^{1200} r^2 \rho(r)\, dr + \int_{r=1200}^{3400} r^2 \rho(r)\, dr + \int_{r=3400}^{6400} r^2 \rho(r)\, dr \right) \\ &= 4\pi C \left(13 \cdot \frac{r^3}{3}\bigg|_0^{1200} + 10 \cdot \frac{r^3}{3}\bigg|_{1200}^{3400} + 5 \cdot \frac{r^3}{3}\bigg|_{3400}^{6400} \right) \\ &\approx 6.3 \cdot 10^{24} \text{ kilograms} \end{aligned}$$

According to Wolfram Alpha, the mass of the earth is approximately $5.97 \cdot 10^{24}$ kilograms, so our rough estimate is not too far off.

(Return)

2. As mentioned above, we will use horizontal strips of the window as the area element. The force element is the amount of force on that strip of the window. Let x be the distance from the center of the window to the horizontal strip. Let up be negative, down be positive (so the top of the window is $x = -r$ and the bottom of the window is $x = r$:



Then the depth of the strip is $h + r + x$, and the area of the strip is $2\sqrt{r^2 - x^2}\, dx$. Thus the force element in this example is

$$dF = (h + r + x)\rho \cdot 2\sqrt{r^2 - x^2}\, dx$$

where ρ is the weight density of water. So

$$\begin{aligned}
F &= \int dF \\
&= 2\rho \int_{-r}^{r} (h+r+x)\sqrt{r^2-x^2}\,dx \\
&= 2\rho \int_{-r}^{r} (h+r)\sqrt{r^2-x^2}\,dx + 2\rho \int_{-r}^{r} x\sqrt{r^2-x^2}\,dx.
\end{aligned}$$

Now, notice that $x\sqrt{r^2-x^2}\,dx$ is an odd function, so its integral from $-r$ to r is 0. Thus

$$\begin{aligned}
F &= 2\rho(h+r)\int_{-r}^{r} \sqrt{r^2-x^2}\,dx \\
&= 2\rho(h+r)\frac{\pi r^2}{2} \\
&= \rho(h+r)\pi r^2,
\end{aligned}$$

since $\int \sqrt{r^2-x^2}\,dx$ gives half the area of a circle of radius r.

It is worth observing that with a very symmetric window such as the circle in this example, one can take the area of the window, πr^2, and multiply by the pressure at the center of the window $\rho(h+r)$, to find the hydrostatic force:

$$F = \rho\pi r^2(h+r).$$

The reason this works is that the pressure on a horizontal strip above the center of the window averages with the pressure on the strip's mirror image below the center to give the pressure at the center of the window.

(Return)

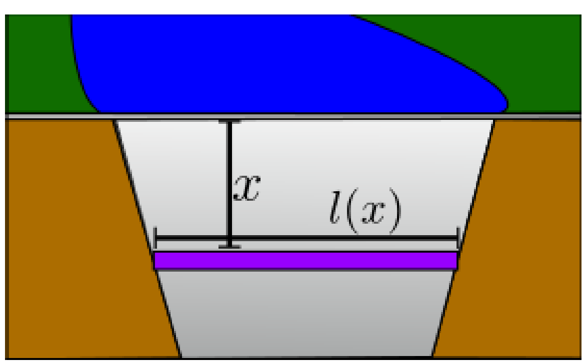
3. Using the knowledge gleaned from the previous example, we can take the area of the endcap, πR^2, and multiply by the hydrostatic pressure at the center of the endcap, which is ρR, to find that the force is

$$F = \pi R^2 \cdot \rho R = \rho\pi R^3.$$

We could also note that this is really a special case of the aquarium window example above, by setting $h = 0$ in that example.

(Return)

4. As above, the force element dF is the force exerted on a horizontal strip. Let x be the distance of the horizontal strip from the top of the dam, and $l(x)$ be the length of the strip



Since the shape is a trapezoid, $l(x)$ is a linear function of x, and from the top and the bottom of the dam, one finds that $l(0) = l_1$ and $l(h) = l_2$. It follows from the slope intercept form of a line that $l(x) = l_1 + \frac{l_2 - l_1}{h} x$.

So the force acting on the strip is $dF = \rho x \, dA$, where ρ is the weight density of the water, x is the depth of the strip, and $dA = (l_1 + \frac{l_2 - l_1}{h} x) \, dx$ is the area of the strip. Putting it all together, one finds

$$\begin{aligned} F &= \int dF \\ &= \int_0^h \rho x \left(l_1 + \frac{l_2 - l_1}{h} x \right) \, dx \\ &= \rho \left(\frac{l_1 x^2}{2} + \frac{l_2 - l_1}{3h} x^3 \right) \Big|_0^h \\ &= \rho \left(\frac{l_1 h^2}{2} + \frac{l_2 - l_1}{3h} h^3 \right) \\ &= \frac{\rho h^2}{6} (l_1 + 2 l_2). \end{aligned}$$

(Return)

5. From the above equation one finds that

$$\begin{aligned} P_0 &= P e^{-rt} \\ &= 1000000 e^{(-.08) \cdot 30} \\ &\approx 90717. \end{aligned}$$

(Return)

6. The income stream $I(t)$ is constant at $3 \cdot 10^6$. Thus,

$$\begin{aligned} PV &= \int dPV \\ &= \int_{t=0}^{20} e^{-rt} I(t) dt \\ &= 3 \cdot 10^6 \int_{t=0}^{20} e^{-.05t} dt \\ &= 3 \cdot 10^6 \left(\frac{1}{-.05} e^{-.05t} \Big|_{t=0}^{20} \right) \\ &= 3 \cdot 10^6 \cdot (-20)(e^{-1} - 1), \end{aligned}$$

which is approximately \$38 million.

(Return)

39 Averages

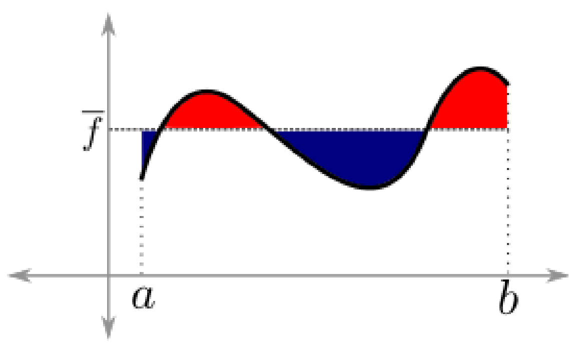
Consider the problem of finding the average test score in a class of 100 students. The answer is to add up all the scores and divide by 100. But what would happen if there were infinitely many students? This module deals with the problem of finding the average value of a function.

39.1 Average value of a function

The definition of the *average value* of a function $f(x)$ over the interval $[a, b]$, denoted $\overline{f}$, is

$$\overline{f} = \frac{\int_a^b f(x)\,dx}{b-a}.$$

One way to interpret the average value is to find the rectangle of length $b-a$ whose area equals the area under the curve f over the interval $[a, b]$. The height of this rectangle is $\overline{f}$. Put another way, $\overline{f}$ is the height of the horizontal line such that the area above the line and below $f(x)$ equals the area which is below the line and above $f(x)$. These areas are shown in red and blue, respectively, in the following diagram:



A better formulation of the average value, which will be useful in other situations and higher dimensions, is

$$\overline{f} = \frac{\int_{x=a}^b f\,dx}{\int_{x=a}^b dx}.$$

This emphasizes that the average value over a region is the integral of the function over the region divided by the volume of that region (in this case, the 1-dimensional volume is just the length of the interval). This generalizes nicely to higher dimensions.

If we go down a dimension to the discrete average, if f_i denotes the ith data point out of n, then the average value of the data is

$$\overline{f} = \frac{\sum_{i=1}^{n} f_i}{n} = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} 1}.$$

This shares a common feature with the earlier formula for average value. Namely, it is the sum (integral) of the function values over a range of inputs divided by the sum (integral) of 1 over that range of inputs.

Example

Compute the average value of $\sin^2 x$ over the interval $[0, 2\pi]$. (See Answer 1)

Example

Compute the average of x^n and e^x over $0 \le x \le T$. Compute the average of $\ln x$ over $1 \le x \le T$. (See Answer 2)

Example

Suppose we are given the density function $\rho(r)$ for the density of the earth at a distance r from the center. Find a formula for the average density of the earth, but do not try to evaluate the integral. (See Answer 3)

39.2 Root mean square

There is another type of average of a function called the *root mean square*. The root mean square of f, denoted f_{RMS} is defined by

$$f_{RMS} = \sqrt{\overline{f^2}}.$$

So the root mean square is the square root of the average value of the square of the function. This is a useful metric when the average value of f is uninteresting.

Example

Compute and compare the average value and the root mean square of $f(x) = \sin x$ on the interval $[0, 2\pi]$. (See Answer 4)

39.3 EXERCISES

- Consider the polar function $f(\theta) = \cos^2(\theta)$. Compute the average value of f from $\theta = 0$ to $\theta = 2\pi$.

39.4 Answers to Selected Exercises

1. From the definition,

$$\begin{aligned}\overline{f} &= \frac{\int_0^{2\pi} \sin^2 x dx}{2\pi} \\ &= \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{2}\left(1-\cos(2x)\right) dx \\ &= \frac{1}{2\pi}\left(\frac{x}{2}-\frac{1}{4}\sin(2x)\right)\Big|_0^{2\pi} \\ &= \frac{1}{2\pi}\cdot\frac{2\pi}{2} \\ &= \frac{1}{2}.\end{aligned}$$

(Return)

2. For $f(x) = x^n$, one finds

$$\begin{aligned}\overline{f} &= \frac{1}{T}\int_0^T x^n \, dx \\ &= \frac{1}{T}\frac{x^{n+1}}{n+1}\Big|_0^T \\ &= \frac{T^n}{n+1}.\end{aligned}$$

For $f(x) = e^x$, the average value is

$$\begin{aligned}\overline{f} &= \frac{1}{T}\int_0^T e^x \, dx \\ &= \frac{e^T - 1}{T}\end{aligned}$$

For $f(x) = \ln x$, recalling the integral using integration by parts, one finds

$$\begin{aligned}\overline{f} &= \frac{1}{T-1}\int_1^T \ln x \, dx \\ &= \frac{1}{T-1}(x\ln x - x)\Big|_1^T \\ &= \frac{1}{T-1}(T\ln T - T + 1).\end{aligned}$$

(Return)

3. Note that we cannot simply integrate the density function and divide by the radius of the earth, for the same reason that we could not integrate the density function to find the mass of the earth in the previous module.

   One way to logically think about it is to note that the average density times the volume of the earth should give the mass of the earth. That is,

$$\overline{\rho}\cdot V = M.$$

Remember that when we found the mass of the earth, we had $dM = \rho\, dV$, where the volume element dV is a spherical shell. So we can write

$$\overline{\rho} = \frac{M}{V} = \frac{\int \rho\, dV}{\int dV}.$$

Then, remembering that the volume of the spherical shell (i.e. the volume element) is $4\pi r^2\, dr$, we have

$$\overline{\rho} = \frac{\displaystyle\int_{r=0}^{R} 4\pi r^2 \rho(r)\, dr}{\displaystyle\int_{r=0}^{R} 4\pi r^2\, dr}.$$

(Return)

4. The average value of $\sin x$ is

$$\begin{aligned}
\overline{f} &= \frac{\displaystyle\int_{x=0}^{2\pi} \sin x\, dx}{2\pi} \\
&= \frac{1}{2\pi}(-\cos x)\Big|_{x=0}^{2\pi} \\
&= \frac{1}{2\pi}(-1-(-1)) \\
&= 0.
\end{aligned}$$

Using the result of an example from above, the root mean square of $\sin x$ is

$$\begin{aligned}
f_{RMS} &= \sqrt{\frac{\displaystyle\int_{x=0}^{2\pi} \sin^2 x\, dx}{2\pi}} \\
&= \sqrt{\frac{1}{2}} \\
&= \frac{1}{\sqrt{2}}.
\end{aligned}$$

(Return)

40 Centroids And Centers Of Mass

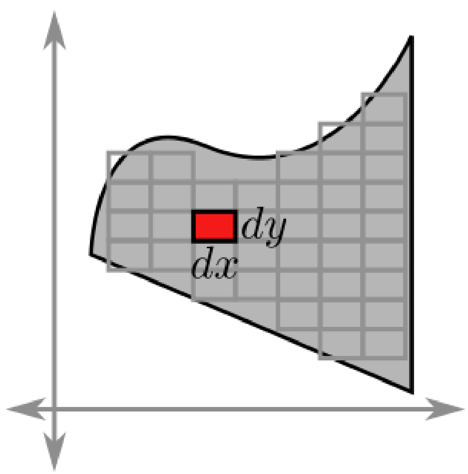
The motivation for this module are the questions:

- what is the average of several locations (e.g. cities on a map)?
- what is the average of an entire region?

The centroid and center of mass give answers to these questions. The formulas for the centroid and the center of mass of a region in the plane seem somewhat mysterious for their apparent lack of symmetry. So before giving the formulas, a brief aside is helpful.

40.1 The area element revisited

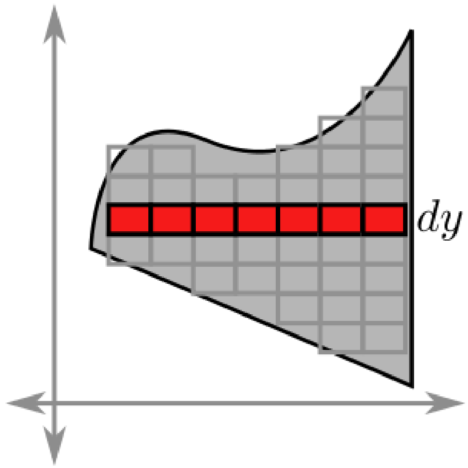
In future courses, the area element of a region will not be a strip of area but a small rectangle with width dx and height dy:



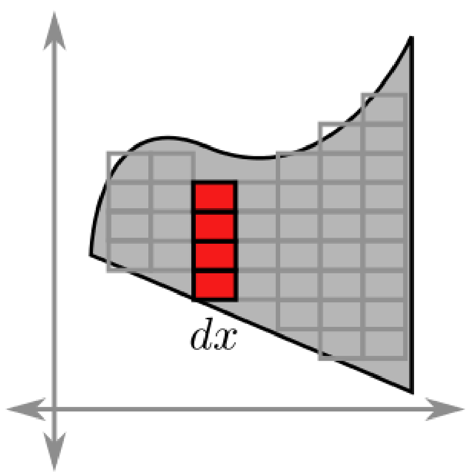
The area of the region, then, is the limit of the sum of the areas of all these small rectangles as the rectangles get infinitely small. The notation used to express this is called a double integral, written

$$\text{Area } = \iint_R dx\, dy.$$

Think of the double integral as a nested integral: $\iint dx\, dy = \int(\int dx)\, dy$. The inner integral is performed first, with respect to x (since the dx is left of the dy). Then the result is integrated with respect to y. Conceptually, the inner integral is adding up the contribution of a row of boxes, and then the outer integral is adding up the rows:



Double integrals can be computed in the other order too: $\iint dy\, dx$. First the inner integral is performed with respect to y, which adds up the contribution of a column of boxes. Then the outer integral adds up the contribution of the columns:



Example

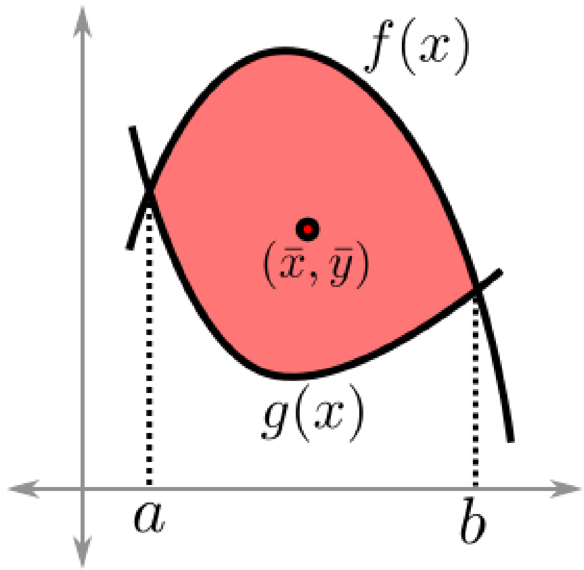
Express the area of the region bounded by the curves $y = x^2 - 4x + 5$ and $y = x + 1$ as a double integral and evaluate the integral. (See Answer 1)

40.2 Centroid

The *centroid* of a region R in the plane is defined to be the point $(\overline{x}, \overline{y})$, where $\overline{x}$ is the average x-coordinate of R and $\overline{y}$ is the average y-coordinate of R. One interpretation is that if the region were cut out of a sheet of uniform density metal and a pin were placed at its centroid, the region would balance on the pin.

The centroid is best expressed mathematically using double integrals:

$$\overline{x} = \frac{\iint_R x\,dx\,dy}{\iint_R\,dx\,dy}$$
$$\overline{y} = \frac{\iint_R y\,dx\,dy}{\iint_R\,dx\,dy}.$$



Suppose the region R is bounded above by the curve $y = f(x)$ and below by the curve $y = g(x)$, and the intersection points are at $x = a$ and $x = b$. Then integration is easier in the $dy\,dx$ order, and the centroid can be written more explicitly as

Centroid of a region

The centroid of the region bounded above by $y = f(x)$ and below by $y = g(x)$ is given by

$$\begin{aligned}\overline{x} &= \frac{\int_a^b \int_{g(x)}^{f(x)} x dy dx}{\int_a^b \int_{g(x)}^{f(x)} dy dx} \\ &= \frac{\int_a^b x(f(x) - g(x))dx}{\int_a^b (f(x) - g(x))dx}.\end{aligned}$$

Similarly,

$$\begin{aligned}\overline{y} &= \frac{\int_a^b \int_{g(x)}^{f(x)} y\,dy\,dx}{\int_a^b \int_{g(x)}^{f(x)}\,dy\,dx} \\ &= \frac{\int_a^b \frac{1}{2}(f(x)^2 - g(x)^2)\,dx}{\int_a^b (f(x) - g(x))\,dx}.\end{aligned}$$

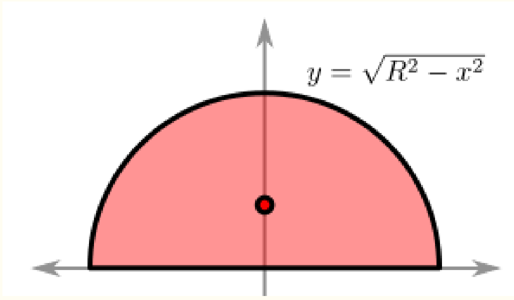
Note that the denominator in each case is the area of the region.

Example

Find the centroid of a triangle with vertices at $(a, 0), (b, 0)$, and $(0, c)$. (See Answer 2)

Example

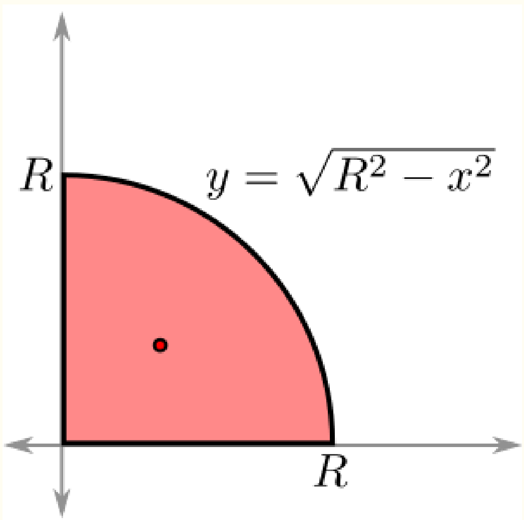
Compute the centroid of the upper half circle of radius R.



(See Answer 3)

Example

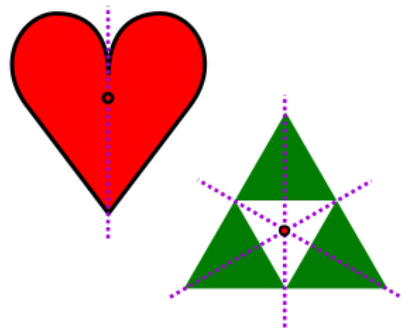
Compute the centroid of the quarter circle of radius R:



(See Answer 4)

Symmetry

It is important to note that centroids respect symmetry. What that means is that if there is an axis of symmetry (i.e. a line where if we reflect the region about the line we get the same region back), then the centroid must lie on the axis of symmetry. If there is more than one axis of symmetry, then the centroid will lie at the intersection of these axes:



40.3 Center of mass

Now consider a region R of the plane cut from a sheet of metal of variable density $\rho(x, y)$. Again, the problem is to find the balancing point $(\overline{x}, \overline{y})$, but in this context it is called the *center of mass*. Again, it is expressed as a double integral:

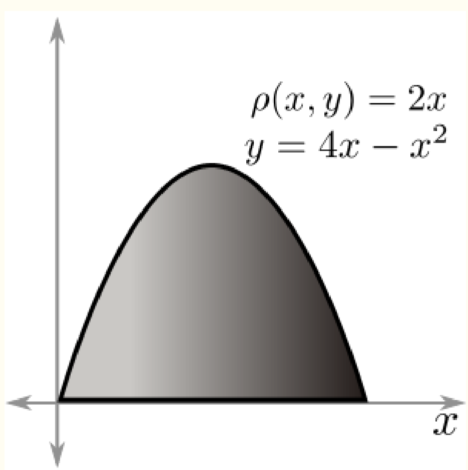
$$\overline{x} = \frac{\iint_R \rho(x, y) x \, dx \, dy}{\iint_R \rho(x, y) \, dx \, dy}$$
$$\overline{y} = \frac{\iint_R \rho(x, y) y \, dx \, dy}{\iint_R \rho(x, y) \, dx \, dy}.$$

The only difference between these and the centroid formulas is that instead of the area element $dA = dx\, dy$, the mass element $dM = \rho(x, y)\, dx\, dy$ is used (multiplying the area element by the density of that point gives the mass contributed by that small rectangle). Indeed, if the density is constant, then $\rho(x, y) = \rho$ factors out of both the numerator and denominator and cancel, leaving the formula for centroid.

Note that the denominator for both $\overline{x}$ and $\overline{y}$ is the mass of the region.

Example

Compute the center of mass of the region bounded above by $y = 4x - x^2$ and below by the x-axis, where the density function is given by $\rho(x, y) = 2x$:



(See Answer 5)

40.4 Centroids using point masses

Given a complex region which consists of the union of simpler regions, there is a method for finding the centroid:

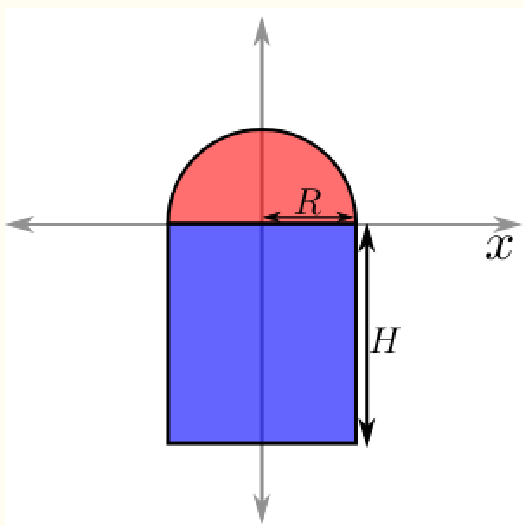
1. Find the centroid of each simple region.
2. Replace each region with a point mass at its centroid, where the mass is the area of the region.
3. Find the centroid of these point masses (this is done by taking a weighted average of their x and y coordinates).

(Centroids and Point Masses Animated GIF)

This is easiest to see with an example:

Example

Find the centroid of a region consisting of a rectangle of width $2R$ and height H which has a semicircle of radius R on one end:



(See Answer 6)

40.5 Application: Pappus' theorem

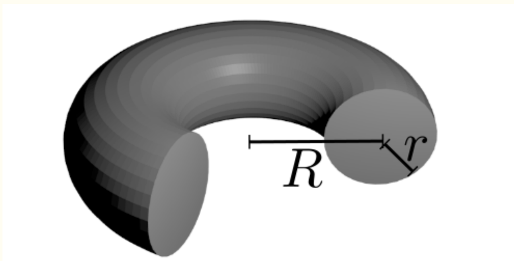
One application of the centroid is known as *Pappus' theorem*, after the Greek mathematician Pappus of Alexandria. It uses the centroid to find the volume and surface area of a solid of revolution.

Pappus' theorem

Consider the solid which results from rotating the plane region R about the axis L.
The volume of this solid is equal to the area of R times the distance the centroid travels (as it gets revolved around the axis).
The surface area of the solid is equal to the perimeter of R times the distance the centroid travels.

Example

Find the volume and surface area of a torus (i.e. a doughnut) with cross sectional radius r and main radius R:

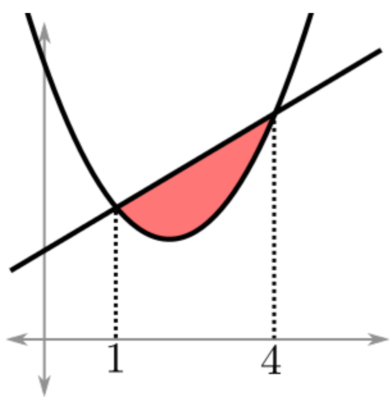


(See Answer 7)

40.6 EXERCISES

- Compute the area of region bounded by curves $x = (y-2)^2 + 2$ and $y = x - 2$ using double integrals.
- Consider the region under the graph $y = x^2$, above the x-axis, from $x = 0$ to $x = 1$. Let S be the solid obtained by revolving this region about the y-axis. Compute the average height (average y-coordinate) of S.
- Let $R1$ denote the region inside the triangle with vertices at (0,1), (-2,0), (0,-1). Given a unit circle centered at the origin, let $R2$ denote the region inside the semicircle for $x \geq 0$. Let R denote the union $R1$ and $R2$. compute the centroid of R.

40.7 Answers to Selected Exercises

1. 

The easier order of integration is $dy\,dx$ because every vertical strip is bounded on top by $y = x+1$ and bounded below by $y = x^2 - 4x + 5$; whereas a horizontal strip would sometimes be bounded on the left by $y = x+1$, and other times be bounded by $y = x^2 - 4x + 5$.

Setting the curves equal gives the intersections at $x = 1$ and $x = 4$. So the area can be found by computing

$$\begin{aligned}
\int_{x=1}^{x=4} \int_{y=x^2-4x+5}^{y=x+1} dy\,dx &= \int_{x=1}^{x=4} \left(y \Big|_{x^2-4x+5}^{x+1} \right) dx \\
&= \int_{x=1}^{x=4} (x+1-(x^2-4x+5))\,dx \\
&= \int_{x=1}^{x=4} (-x^2+5x-4)\,dx \\
&= -\frac{x^3}{3} + \frac{5}{2}x^2 - 4x \Big|_1^4 \\
&= \frac{9}{2}.
\end{aligned}$$

(Return)

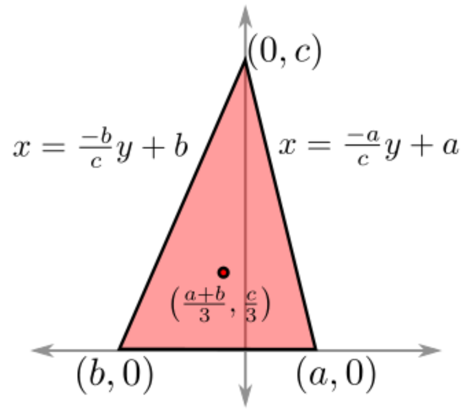
2. The easier order of integration is $dx\,dy$ because a horizontal strip is always bounded on the left by $x = \frac{-b}{c}y + b$ and on the right by $x = \frac{-a}{c}y + a$ (see the diagram below). So one finds that

$$\begin{aligned}
\overline{x} &= \frac{\int_{y=0}^{y=c} \int_{x=\frac{-b}{c}y+b}^{x=\frac{-a}{c}y+a} x\,dx\,dy}{\int_{y=0}^{y=c} \int_{x=\frac{-b}{c}y+b}^{x=\frac{-a}{c}y+a} dx\,dy} \\
&= \frac{\int_{y=0}^{y=c} \int_{x=\frac{-b}{c}y+b}^{x=\frac{-a}{c}y+a} x\,dx\,dy}{\text{Area}}
\end{aligned}$$

Noting that the area of the triangle is $\frac{1}{2}(a-b)c$, one finds

$$\begin{aligned}
\overline{x} &= \frac{2}{(a-b)c} \int_{y=0}^{y=c} \int_{x=\frac{-b}{c}y+b}^{x=\frac{-a}{c}y+a} x\,dx\,dy \\
&= \frac{2}{(a-b)c} \int_{y=0}^{y=c} \frac{1}{2} \left((\frac{-a}{c}y+a)^2 - (\frac{-b}{c}y+b)^2 \right) dy \\
&= \frac{1}{(a-b)c} \cdot \frac{1}{3} \left((\frac{-a}{c}y+a)^3 \frac{-c}{a} - (\frac{-b}{c}y+b)^3 \frac{-c}{b} \right) \Big|_0^c \\
&= \frac{1}{3(a-b)c}(a^2c - b^2c) \\
&= \frac{1}{3(a-b)c}c(a+b)(a-b) \\
&= \frac{1}{3}(a+b).
\end{aligned}$$

A similar computation gives that $\overline{y} = \frac{c}{3}$.



More generally, the centroid of a triangle with coordinates (x_0, y_0), (x_1, y_1), and (x_2, y_2) is

$$(\overline{x}, \overline{y}) = \left(\frac{x_0 + x_1 + x_2}{3}, \frac{y_0 + y_1 + y_2}{3}\right).$$

In other words, the centroid of a triangle is the average of the x coordinates and the average of the y coordinates.

(Return)

3. By the symmetry about the y-axis, the x-coordinate of the centroid is 0.

   To find the y-coordinate, note that the equation of the curve is $y = \sqrt{R^2 - x^2}$. Also, note that the area of the region is $\frac{1}{2}\pi R^2$. Thus,

$$\begin{aligned}\overline{y} &= \frac{2}{\pi R^2} \int_{x=-R}^{R} \frac{1}{2}(\sqrt{R^2 - x^2})^2 \, dx \\ &= \frac{1}{\pi R^2} \left(R^2 x - \frac{1}{3}x^3\right)\Big|_{x=-R}^{R} \\ &= \frac{1}{\pi R^2} \cdot \frac{4R^3}{3} \\ &= \frac{4R}{3\pi}.\end{aligned}$$

   (Return)

4. We know that the area of the region is $\frac{1}{4}\pi R^2$. So we have that

$$\begin{aligned}\overline{x} &= \frac{1}{A} \int x(f(x) - g(x)) \, dx \\ &= \frac{4}{\pi R^2} \int_{x=0}^{R} x(\sqrt{R^2 - x^2} - 0) \, dx.\end{aligned}$$

   Making a substitution of

$$\begin{aligned}u &= R^2 - x^2 \\ du &= -2x \, dx\end{aligned}$$

gives

$$\frac{4}{\pi R^2}\int_{x=0}^{R} x\sqrt{R^2-x^2}\,dx = \frac{4}{\pi R^2}\int_{u=R^2}^{0} -\frac{1}{2}\sqrt{u}\,du$$
$$= \frac{-2}{\pi R^2}\frac{2}{3}u^{3/2}\Big|_{u=R^2}^{0}$$
$$= \frac{2}{\pi R^2}\cdot\frac{2}{3}R^3$$
$$= \frac{4R}{3\pi}.$$

Because the region is symmetric about the line $y = x$, we predict that $\overline{y} = \frac{4R}{3\pi}$ as well. We can verify this by integrating:

$$\overline{y} = \frac{1}{A}\int \frac{1}{2}(f(x)^2 - g(x)^2)\,dx$$
$$= \frac{2}{\pi R^2}\int_{x=0}^{R}(R^2-x^2)\,dx$$
$$= \frac{2}{\pi R^2}\left(R^2x - \frac{1}{3}x^3\right)\Big|_{x=0}^{R}$$
$$= \frac{2}{\pi R^2}\left(R^3 - \frac{1}{3}R^3\right)$$
$$= \frac{2}{\pi R^2}\cdot\frac{2}{3}R^3$$
$$= \frac{4R}{3\pi},$$

as claimed.

(Return)

5. Setting $y = 0$, we find that the curve intersects the x-axis at $x = 0$ and $x = 4$. First, we compute the mass of the region, which is the denominator for both $\overline{x}$ and $\overline{y}$. It is easier to integrate in the $dy\,dx$ order, so we will do that here, and in the integrals that follow.

$$M = \iint_R \rho(x,y)\,dy\,dx$$
$$= \int_{x=0}^{4}\int_{y=0}^{4x-x^2} 2x\,dy\,dx$$
$$= \int_{x=0}^{4}\left(2xy\Big|_{y=0}^{4x-x^2}\right)dx$$
$$= \int_{x=0}^{4} 2x(4x-x^2)\,dx$$
$$= \int_{x=0}^{4}(8x^2-2x^3)\,dx$$
$$= \frac{8}{3}x^3 - \frac{1}{2}x^4\Big|_{x=0}^{4}$$
$$= \frac{512}{3} - 128 = \frac{128}{3}.$$

So to compute $\overline{x}$, we find

$$
\begin{aligned}
\overline{x} &= \frac{1}{M}\iint_R \rho(x,y)x\,dy\,dx \\
&= \frac{3}{128}\int_{x=0}^{4}\int_{y=0}^{4x-x^2}(2x)x\,dy\,dx \\
&= \frac{3}{128}\int_{x=0}^{4}\left(2x^2y\Big|_{y=0}^{4x-x^2}\right)dx \\
&= \frac{3}{128}\int_{x=0}^{4}2x^2(4x-x^2)\,dx \\
&= \frac{3}{128}\left(2x^4-\frac{2}{5}x^5\right)\Big|_{x=0}^{4} \\
&= \frac{3}{128}512-\frac{2048}{5} \\
&= \frac{3}{128}\cdot\frac{512}{5}=\frac{12}{5}.
\end{aligned}
$$

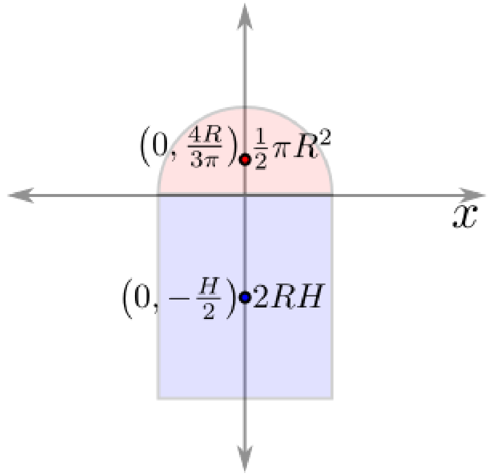
Similarly,

$$
\begin{aligned}
\overline{y} &= \frac{1}{M}\iint_R \rho(x,y)y\,dy\,dx \\
&= \frac{3}{128}\int_{x=0}^{4}\int_{y=0}^{4x-x^2}(2x)y\,dy\,dx \\
&= \frac{3}{128}\int_{x=0}^{4}xy^2\Big|_{y=0}^{4x-x^2}dx \\
&= \frac{3}{128}\int_{x=0}^{4}x(4x-x^2)^2\,dx \\
&= \frac{3}{128}\int_{x=0}^{4}(16x^3-8x^4+x^5)\,dx \\
&= \frac{3}{128}\left(4x^4-\frac{8}{5}x^5+\frac{1}{6}x^6\right)\Big|_{x=0}^{4} \\
&= \frac{3}{128}\cdot 1024\left(1-\frac{8}{5}+\frac{2}{3}\right)=\frac{8}{5}.
\end{aligned}
$$

(Return)

6. From an earlier example, the centroid of the semicircle is $(0,\frac{4R}{3\pi})$, and the weight (the area of the semicircle) is $\frac{1}{2}\pi R^2$.

   The rectangle is symmetric, so its centroid (as it is drawn in the coordinate plane) is $(0,-\frac{H}{2})$, and its weight is $2RH$:



By symmetry,

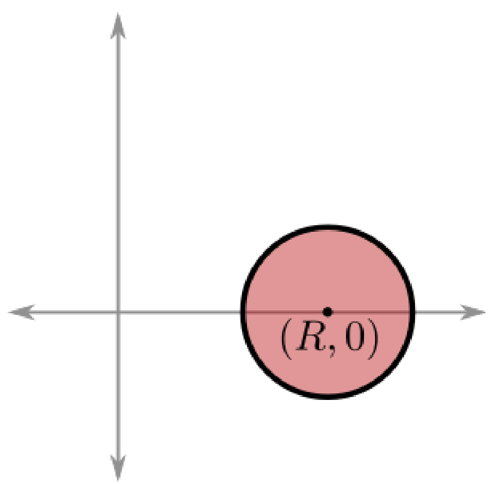
$$\overline{x} = 0.$$

Taking the weighted average of the y-coordinates of the points gives

$$\begin{aligned}\overline{y} &= \frac{\frac{1}{2}\pi R^2 \cdot \frac{4R}{3\pi} + 2RH \cdot \left(-\frac{H}{2}\right)}{\frac{1}{2}\pi R^2 + 2RH} \\ &= \frac{\frac{2}{3}R^3 - H^2 R}{\frac{1}{2}\pi R^2 + 2RH} \\ &= \frac{4R^2 - 6H^2}{3\pi R + 12H}.\end{aligned}$$

We can check that this is reasonable by noting that if $H = 0$ we get the y-coordinate of the centroid of the semicircle, and when $R = 0$ we get the y-coordinate of the centroid of the line segment from $(0, 0)$ to $(0, -H)$.

(Return)

7. 

Here, the region being rotated is a circle, which is easy to work with because a circle's centroid is just its center. For convenience, center the circle at $(R, 0)$ and revolve around the y-axis. Then the distance which the centroid travels is $2\pi R$ (the path of the centroid is just a circle of radius R).

Therefore the surface area of the torus is

$$\begin{aligned} \text{Surface area } &= \text{Perimeter } \cdot \text{ Centroid travel distance} \\ &= (2\pi r) \cdot (2\pi R) \\ &= 4\pi^2 r \cdot R. \end{aligned}$$

And the volume of the torus is

$$\begin{aligned} \text{Volume } &= \text{Area } \cdot \text{ Centroid travel distance} \\ &= (\pi r^2) \cdot (2\pi R) \\ &= 2\pi^2 r^2 R. \end{aligned}$$

(Return)

41 Moments And Gyrations

This module deals with the moment of inertia and the radius of gyration, which are two properties of an object with physical interpretations.

41.1 Moment of inertia

The *moment of inertia* of an object, usually denoted I, measures the object's resistance to rotation about an axis. To get an intuitive understanding of moment of inertia consider swinging a hammer by its handle (higher moment of inertia, harder to swing) versus swinging a hammer by its head (lower moment of inertia, easier to swing). So moment of inertia depends on both the object being rotated and the axis about which it is being rotated.

(Hammer Animated GIF)

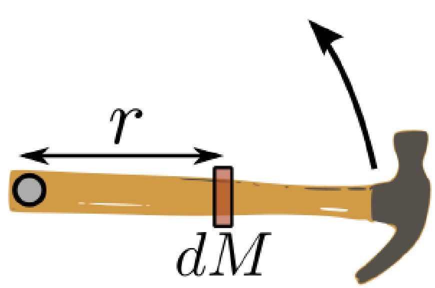
Consider first a particle of mass. The bigger the mass, the more resistant it will be to rotation about an axis. Similarly, the further the particle is from the axis, the more resistant it will be to rotation. For a point mass, the moment of inertia is given by

$$I = r^2 M,$$

where r is the distance of the particle from the axis of rotation, and M is the mass of the particle:

(Particle Animated GIF)

The next question is how to calculate the moment of inertia when all the mass is not at a single point. As in previous modules, the method will be to break the object into slices of mass, and consider the contribution of each slice to the moment of inertia:



Each slice can be thought of as an individual particle of mass which contributes to the moment of inertia. The contribution of the slice becomes the moment of inertia element dI:

$$dI = r^2\, dM.$$

Example

Consider a solid disc of radius R and constant density ρ rotated about its central vertical axis:

(Disk Rotating Around Diameter Animated GIF)

Compute its moment of inertia. (See Answer 1)

Example

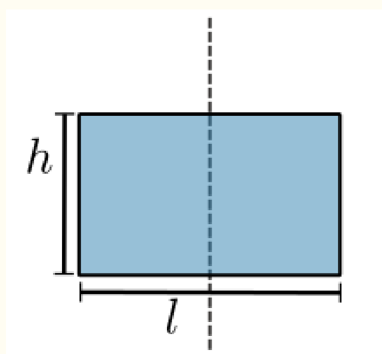
Consider a solid disc of radius R and constant density ρ rotated about its center:

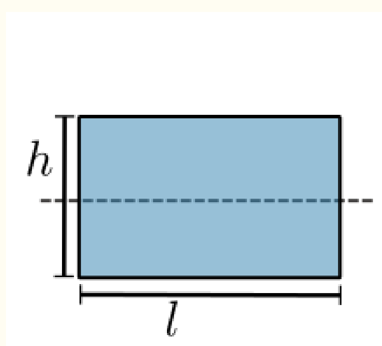
(Disk Rotating Around Center Animated GIF)

Compute its moment of inertia. (See Answer 2)

Example

Consider a rectangle of length l and height h. Compute the moment of inertia about the vertical axis through its center. Then compute the moment of inertia about the horizontal axis through its center. Hint: use symmetry to find the second answer from the first.





(See Answer 3)

41.2 Radius of gyration

Another property of an object, *radius of gyration*, denoted R_g, can be expressed in terms of the moment of inertia. Imagine replacing the object being rotated about an axis by a single point mass being rotated about that same axis. The radius of gyration is the radius at which the point mass has the same moment of inertia as the object. More specifically, $I = MR_g^2$, and solving for R_g gives

$$R_g = \sqrt{\frac{I}{M}}.$$

Note that because

$$I = \int r^2 \, dM$$

we can write

$$\begin{aligned} R_g &= \sqrt{\frac{\int r^2 \, dM}{\int dM}} \\ &= \sqrt{\overline{r^2}} \\ &= r_{RMS} \end{aligned}$$

So the radius of gyration is really the root mean square of the radius.

41.3 Higher mass moments

In the centroid module, we computed $\int x \, dM$ as part of computing the x-coordinate of the center of mass, $\overline{x}$.The moment of inertia I from this module is given by $\int x^2 \, dM$. These are respectively known as the first mass moment and the second mass moment (first and second referring to the powers of x).

There are higher mass moments: $\int x^n \, dM$, for $n \geq 3$, as well as the lower mass moment $\int x^0 dM$, which is just mass. These moments each give more information about how the mass of the object is distributed.

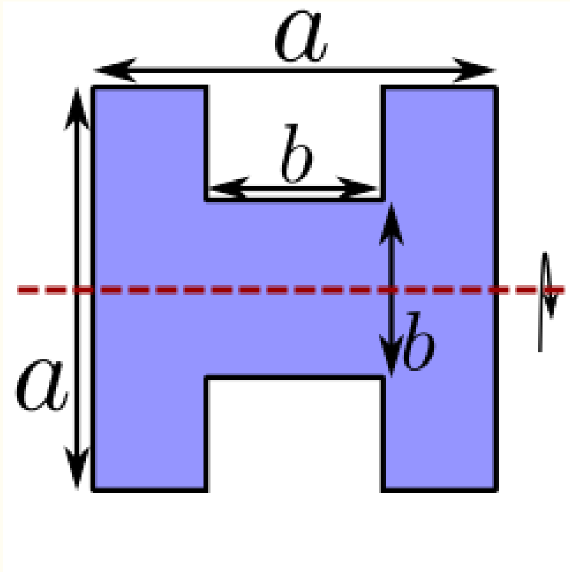
This is similar, in a sense, to how knowledge of the derivative of a function at a point leads to an approximation of the function using the Taylor series. The more derivatives one knows, the better the approximation. A logical question, then, is if one knows all the mass moments of an object, can one perfectly describe the distribution of mass?

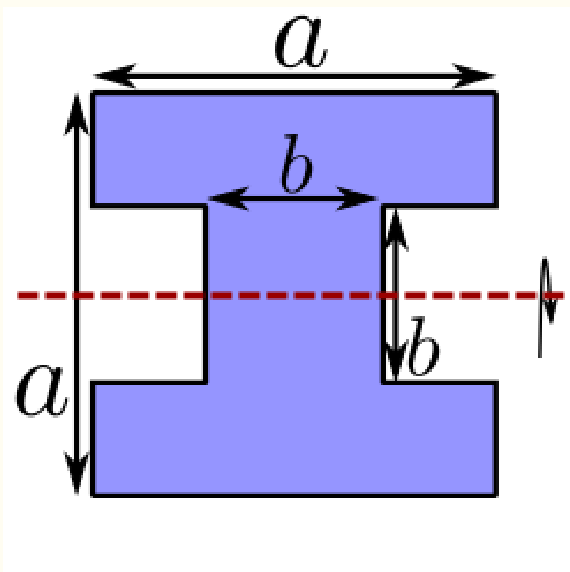
41.4 Additivity of moments

One nice feature of moments is that, being integrals, they are additive. This means that a complex region can be split into simpler regions for which we already know the moment of inertia, and these moments can be added to find the moment of inertia for the entire region.

Example

Compute the moment of inertia for each of the following figures about a horizontal axis through their centers.





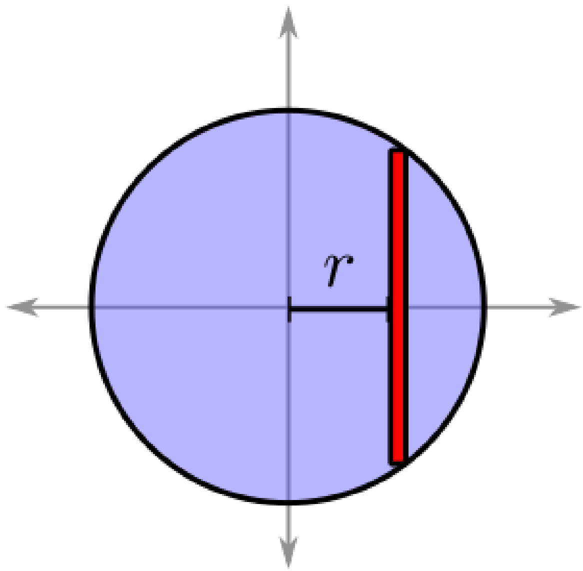
Which has the greater moment of inertia, the I-shaped figure or the H-shaped figure? (See Answer 4)

41.5 EXERCISES

- Consider a right triangle with vertices at (0,0), (5,0), (0,10). Consider rotating the triangle about the y-axis. The density is given by $\rho(x, y) = x$. Compute the moment of inertia. Compute the radius of gyration.

41.6 Answers to Selected Examples

1. Because the distance to the axis is part of the inertia element, a good area element to use is a vertical rectangle, where every point has the same distance to the center axis. Let x be the distance from the central axis to the rectangle (thus, $r = x$):



The area of this rectangle, as has been computed several times previously, is $dA = 2\sqrt{R^2 - x^2}\,dx$. Then the mass element $dM = \rho\,dA$, and it follows that

$$\begin{aligned} dI &= r^2\,dM \\ &= 2x^2\rho\sqrt{R^2 - x^2}\,dx. \end{aligned}$$

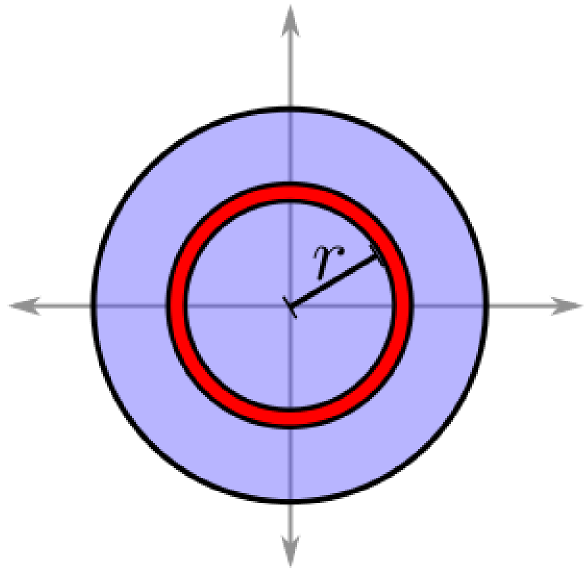
so integrating the inertia element gives

$$\begin{aligned} I &= \int dI \\ &= \int_{x=-R}^{R} 2\rho x^2\sqrt{R^2 - x^2}\,dx \\ &= 4\rho\int_{x=0}^{R} x^2\sqrt{R^2 - x^2}\,dx \end{aligned}$$

(using the fact that the integrand is an even function allows the final step). Now the substitution $x = R\sin\theta$, and some of the trig integral methods gives the answer $\frac{\rho\pi}{4}R^4$, which can also be written $\frac{1}{4}MR^2$, where $M = \pi R^2\rho$ is the mass of the disc.

(Return)

2. In this example, a good area element to use is a ring (also called an annulus), because every point in a ring has the same distance to the origin, which is the axis of rotation (one can imagine the axis sticking out of the page perpendicular to the center of the disc):



As before, $dM = \rho\, dA$. In this case, dA is the area of the ring, which is $2\pi r\, dr$ (the circumference of the ring times the width of the ring). It follows that

$$\begin{aligned} I &= \int dI \\ &= \int_{r=0}^{R} r^2 \rho 2\pi r\, dr \\ &= 2\pi\rho \int_{r=0}^{R} r^3\, dr \\ &= 2\pi\rho \frac{r^4}{4}\bigg|_{r=0}^{R} \\ &= \frac{\pi\rho R^4}{2}. \end{aligned}$$

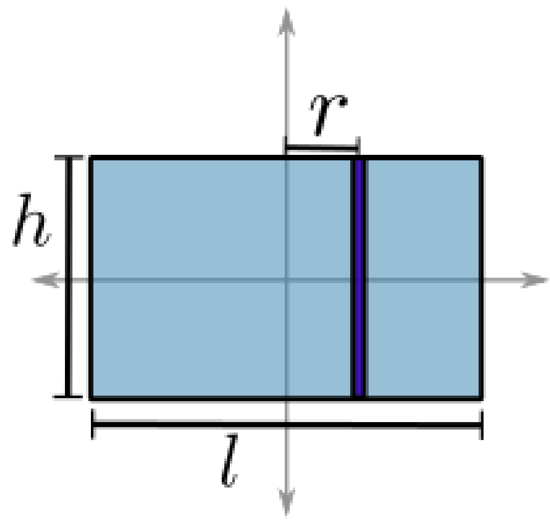
This can be expressed as $\frac{1}{2}MR^2$, where M is again the mass of the disc. Note that the answer in this example is twice that of the previous example. This can be explained (using the answer from the previous example) by noting that $r^2 = x^2 + y^2$ in this example. Therefore,

$$\begin{aligned} I &= \int r^2\, dM \\ &= \int (x^2 + y^2)\, dM \\ &= \int x^2\, dM + \int y^2\, dM, \end{aligned}$$

and these two integrals are, respectively, the moment of inertia about a vertical axis (from the previous example) and the moment of inertia about a horizontal axis. By symmetry, these are equal, which explains why this answer is twice the answer of the previous example.

(Return)

3. Center the rectangle at the origin. About the vertical axis, it is again best to use vertical rectangles. Let r denote the distance of this rectangle from the y-axis:



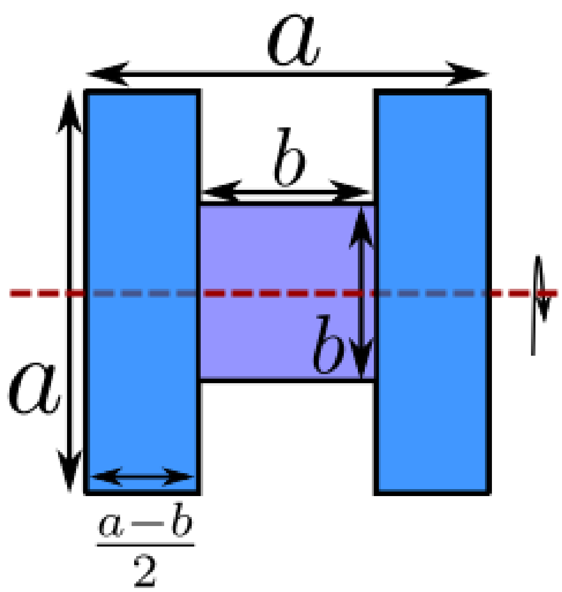
Then $r = x$, and $dM = \rho h\, dx$. Thus

$$\begin{aligned} I &= \int dI \\ &= \rho h \int_{x=-l/2}^{l/2} x^2\, dx \\ &= \rho h \frac{x^3}{3}\bigg|_{x=-l/2}^{l/2} \\ &= \frac{1}{12}\rho h l^3 \\ &= \frac{1}{12} M l^2, \end{aligned}$$

where $M = \rho l h$ is the mass of the rectangle. By symmetry, the moment of inertia about a horizontal axis through the center is $\frac{1}{12} M h^2$.

(Return)

4. For the first figure, we can divide it into two rectangles (in light blue) and a square which are all being rotated about their horizontal center axis:



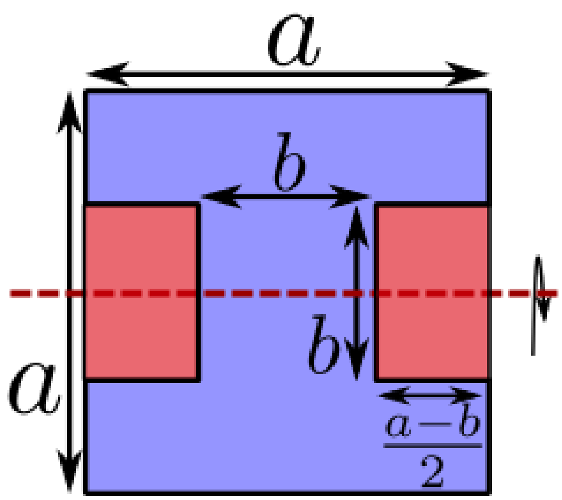
From the above example, we know that the moment of inertia for a rectangle about its horizontal axis is

$$\frac{1}{12}Mh^2 = \frac{1}{12}lh^3,$$

where l is the length and h is the height of the rectangle. So for each of the tall rectangles we have $I = \frac{1}{12}a^3\frac{a-b}{2}$ and for the square in the middle we have $I = \frac{1}{12}b^4$. Putting it together, we have the moment of inertia for the entire region is

$$\begin{aligned} I &= 2 \cdot \frac{1}{12}a^3\frac{a-b}{2} + \frac{1}{12}b^3 b \\ &= \frac{1}{12}(a^4 - a^3 b) + \frac{1}{12}b^4 \\ &= \frac{1}{12}\left(a^4 + b^4 - a^3 b\right). \end{aligned}$$

For the other region, we cannot divide it up into rectangles in the same exact way, because we do not know the moment of inertia for a rectangle rotated about an axis other than one through its center. Instead, we can take the entire square of side length a, and compute its moment of inertia. Then we can subtract off the inertia for the small rectangles we do not want to include, shown in red:



Again, using the fact from the previous example, the moment for the whole square is $\frac{1}{12}a^4$, and the moment for each of the smaller rectangles (which we will subtract) is $\frac{1}{12} \cdot \frac{a-b}{2} \cdot b^3$, so the moment of inertia for the whole region is

$$\begin{aligned} I &= \frac{1}{12}a^4 - 2 \cdot \frac{1}{12} \cdot \frac{a-b}{2} \cdot b^3 \\ &= \frac{1}{12}a^4 - \frac{1}{12}(ab^3 - b^4) \\ &= \frac{1}{12}(a^4 + b^4 - ab^3). \end{aligned}$$

So the I-shaped figure has the greater moment of inertia.

This is important when considering whether to use an H-beam or and I-beam in construction. According to a fact mentioned in higher derivatives, the deflection $u(x)$ (the amount the beam sags at location x) satisfies the equation

$$EI\frac{d^4u}{dx^4} = q(x),$$

where E is the elasticity of the material (a constant), and $q(x)$ is a static load at location x along the beam. Because the I-beam has the greater moment of inertia, it follows that their deflection will be less, and so I-beams are more common in building construction.

(Return)

42 Fair Probability

Probability is the study of the likelihood of certain events occurring in a random experiment. A simple example is a coin flip. There are two outcomes: heads (H) or tails (T). If the coin is fair, then the probability of each outcome is $\frac{1}{2}$, written $P(H) = P(T) = \frac{1}{2}$. Another example is a roll of a standard die. There are six outcomes: 1 through 6. If the die is fair then the probability of each outcome is $\frac{1}{6}$.

In these types of problems, one can find the probability of an event occurring by counting the number of desired outcomes and dividing by the total number of outcomes.

Example

What is the probability that a pair of dice sums to seven or eleven? (See Answer 1)

Example

Alice and Bob play a game where they take turns flipping a fair coin, with Alice going first. The first player to get heads wins. What is the probability that Alice wins?
Hint: find the probability that Alice wins on her first flip, and the probability that she wins on her second flip, and the probability that she wins on her third flip, etc. Add up all these (infinitely many) probabilities to find the probability that she wins.
Second hint: For Alice to win on her second flip, it means that both Alice and Bob got tails on their respective first flips (otherwise the game would have ended in the first round). So the probability of Alice winning on her second flip is

$$P(\text{A got tails}) \cdot P(\text{B got tails}) \cdot P(\text{A got heads}) = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \left(\frac{1}{2}\right)^3.$$

(See Answer 2)

42.1 Uniform distribution

The above examples refer to a fair coin and a fair die. A discrete experiment (i.e. the possible outcomes can be listed) is said to have the *uniform distribution* if the experiment is fair in the sense that every outcome is equally likely.

What about experiments which are not discrete? For instance, a spinner gives a point along the circumference of a circle, and the individual points of the circle cannot be enumerated. Throwing a dart at a circular dartboard likewise has as many outcomes as there are points in the interior of the disk. What does it mean for such an experiment to be *fair*, i.e., what does the uniform distribution mean in an experiment with continuous outcomes?

To answer this question, consider the probability of a range of outcomes of the experiment. So, for instance, what is the probability of the spinner landing in the first quarter of the circle? If the experiment is fair, then this probability should be the same as landing in any other quarter of the circle: $\frac{1}{4}$:

(Spinner Animated GIF)

Thus an experiment is *fair* (i.e. has the *uniform distribution*) if for any set of outcomes D,

$$P(D) = \frac{\text{volume of } D}{\text{total volume of all outcomes}}.$$

Here "volume" depends on the dimension of the experiment. For instance, the spinner has dimension 1 (where volume is really just the length) since any point on the circumference can be specified by a single value (say, the angle of the arrow relative to the positive x-axis). So a spinner is considered fair if the probability of the arrow landing in a certain range along the circumference equals the length of that range divided by the total circumference of the circle.

Length

Example

Find the probability that a randomly chosen angle θ has $\sin\theta > \frac{1}{2}$? (See Answer 3)

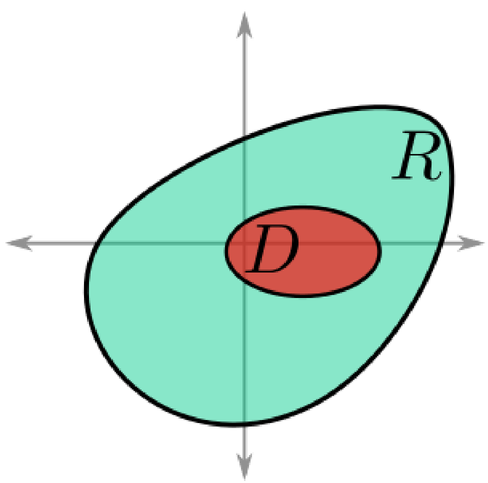
Example

Find the probability that a randomly chosen angle θ has $\tan\theta > 0$. (See Answer 4)

Area

In two dimensions, volume is really area, and so when computing the probability that a randomly chosen point in a region R in the plane lies within the region D, we have

$$P = \frac{\text{Area of } D}{\text{Area of } R}.$$



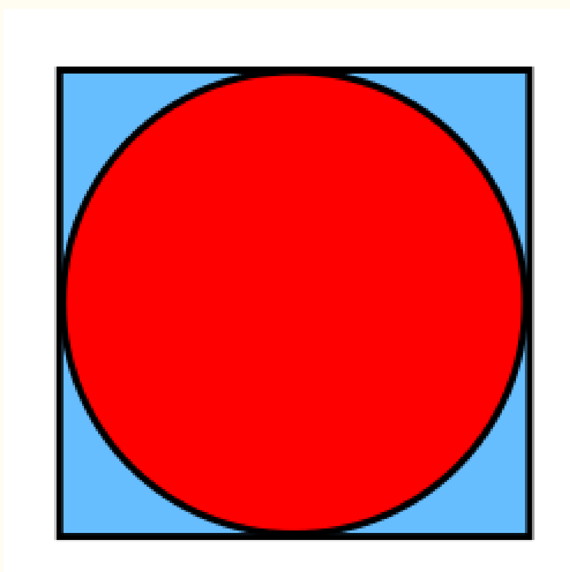
Example

A dartboard is circular with radius 9 inches:

The bullseye is a small circle at the center of the board. Find the radius of the bullseye so that the probability of hitting it is $\frac{1}{100}$ (assuming a throw hits the board uniformly at random). (See Answer 5)

Example

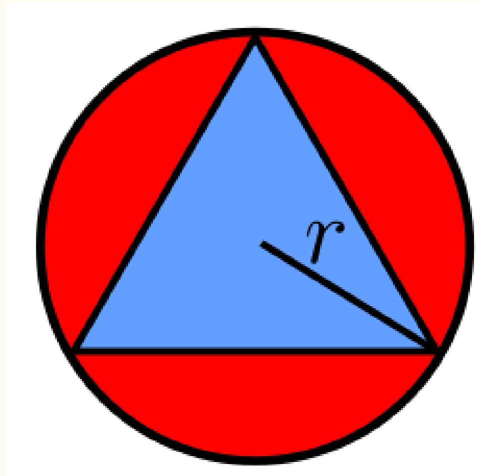
Find the probability that a randomly chosen point in a square lies within the circle inscribed in the square:



(See Answer 6)

Example

Find the probability that a randomly chosen point in a circle lies in the equilateral triangle inscribed in the circle:



Hint: the area of an equilateral triangle of side length s is

$$A = \frac{s^2\sqrt{3}}{4}.$$

(See Answer 7)

There are some probability problems that do not seem geometric in nature but can be solved by graphing the possible outcomes and taking the ratio of the areas.

Example

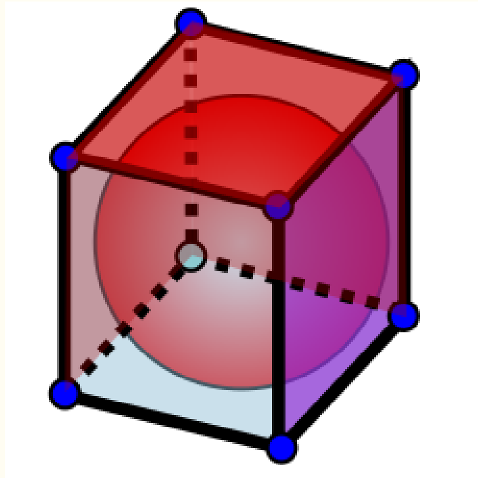
Xander and Yolanda want to meet up to study calculus. Each friend will arrive at the library at some random time between 5 pm and 6pm, wait 20 minutes for the other person, and then leave if the other person does not arrive in that time. Find the probability that the friends successfully meet up.
Hint: Let x be the number of minutes after 5 pm that Xander arrives and y be the number of minutes after 5 pm that Yolanda arrives. Now plot the possible arrival times as a region in the plane and determine the region which corresponds to them successfully meeting up. (See Answer 8)

Volume

Finally, in dimension 3, volume is volume as we traditionally know it. In this case, we imagine picking a point from within a 3D region and know the probability that the point lies within some subset of that region.

Example

Find the probability that a randomly chosen point from within a cube lies within the inscribed sphere:



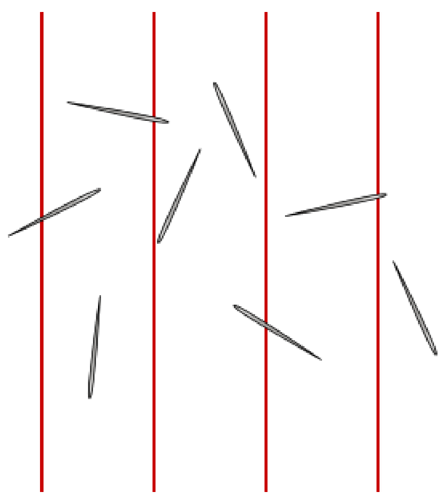
(See Answer 9)

Example

What is the probability that a randomly chosen point in a ball lies within 10% of the boundary (as measured by radius)? (See Answer 10)

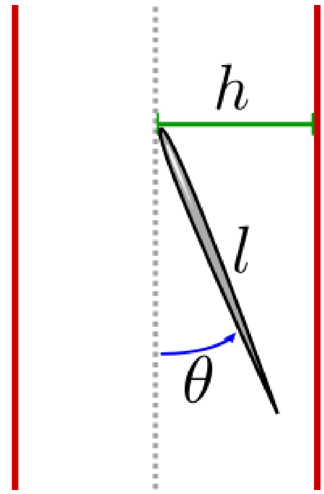
42.2 Buffon needle problem

The Buffon needle problem, named after the Count of Buffon, asks for the probability that a needle of length l, dropped uniformly at random onto a sheet with parallel lines spaced l units apart, will cross a line.

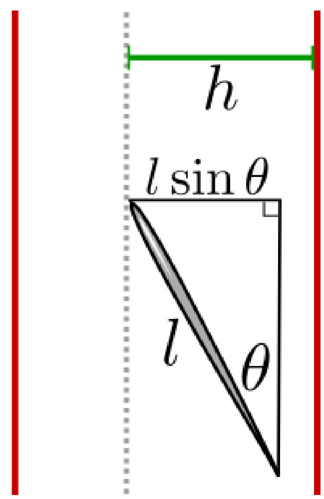


To simplify the problem, consider two parameters which determine whether the needle crosses:

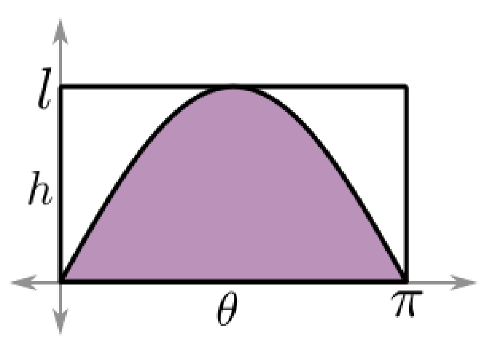
1. h, the distance from the left tip of the needle to the next line to its right
2. θ, the angle that the needle makes with a vertical line:



Note that $0 \le h \le l$ and $0 \le \theta \le \pi$. Now, for what values of h and θ is there a crossing? Note that by right triangle trigonometry, the horizontal distance from the left end of the needle to the right end of the needle is $l \sin\theta$:



Thus, there is a crossing if $h \le l \sin\theta$, and there is no crossing if $h > l \sin\theta$. Graphing this inequality shows that the region below the curve (shown in purple) is where a crossing occurs. The region above the curve is where a crossing does not occur.



Dropping a needle at random is like randomly picking a point in this rectangle. Thus, the probability of a random needle creating a crossing equals the probability of randomly picking a point below the curve in the above rectangle. That probability is given by dividing the area under the curve by the area of the entire rectangle.

$$\begin{aligned} P(\text{crossing}) &= \frac{\int_0^\pi l \sin\theta \, d\theta}{l\pi} \\ &= \frac{1}{\pi}\left(-\cos\theta \Big|_0^\pi\right) \\ &= \frac{2}{\pi}. \end{aligned}$$

42.3 Answers to Selected Examples

1. By listing the desired outcomes, one finds that $(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)$ are the possible pairings which give 7, and $(6,5)$ and $(5,6)$ are the possible pairings which give 11. So there are 8 desired outcomes. The total number of outcomes is 6×6 (six outcomes for the first die paired with each of the six outcomes for the other die). So the probability is

$$\frac{\#\text{ desired}}{\#\text{ total}} = \frac{8}{36} = \frac{2}{9}.$$

(Return)

2. Proceeding as the hint suggests, we look for a pattern.

$$P(\text{A wins on 1st flip}) = P(\text{A gets heads}) = \frac{1}{2}.$$

And then

$$\begin{aligned} &P(\text{A wins on 2nd flip}) \\ &= P(\text{A gets tails}) \cdot P(\text{B gets tails}) \cdot P(\text{A gets heads}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \\ &= \left(\frac{1}{2}\right)^3. \end{aligned}$$

Next,

$$\begin{aligned} &P(\text{A wins on 3rd flip}) \\ &= P(\text{A gets tails}) \cdot P(\text{B gets tails}) \cdot P(\text{A gets tails}) \cdot P(\text{B gets tails}) \cdot P(\text{A gets heads}) \\ &= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \\ &= \left(\frac{1}{2}\right)^5. \end{aligned}$$

In general, for Alice to win on the nth flip, she must get a head on that flip, *and* both Alice and Bob must have gotten tails on each of their previous $n-1$ flips. Thus, there are a total of $2(n-1)+1 = 2n-1$ coin flips that must come out in a precise way, and the probability of each of these is $\frac{1}{2}$, so we have

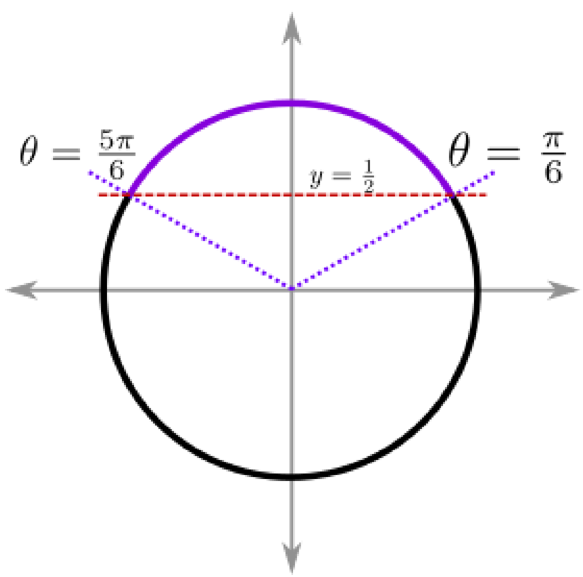
$$P(\text{A wins on } n\text{th flip}) = \left(\frac{1}{2}\right)^{2n-1}.$$

Adding these up for all n, and using the geometric series, gives

$$\begin{aligned} P(\text{A wins}) &= \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^{2n-1} \\ &= \frac{1}{2} + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^5 + \cdots \\ &= \frac{1}{2} \cdot \left(1 + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^4 + \cdots\right) \\ &= \frac{1}{2} \cdot \left(1 + \frac{1}{4} + \left(\frac{1}{4}\right)^2 + \cdots\right) \\ &= \frac{1}{2} \cdot \frac{1}{1 - 1/4} \\ &= \frac{2}{3}. \end{aligned}$$

(Return)

3. We can visualize the sine of the angle by considering a unit circle, and noting that sine is the y-coordinate of a point on the circle:



Then one finds that the angles for which $\sin\theta > \frac{1}{2}$ are

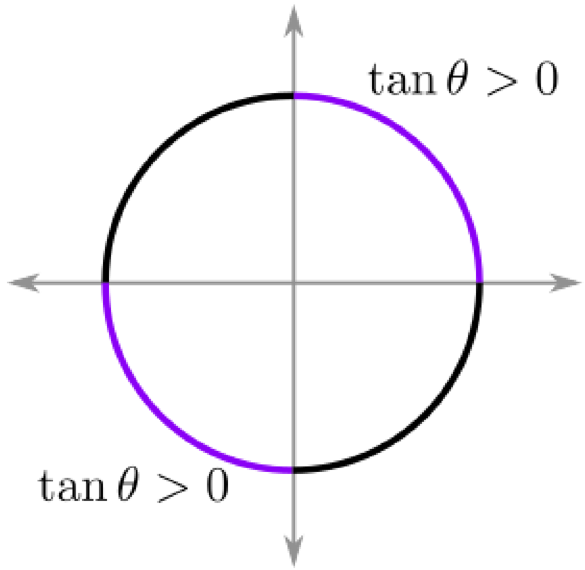
$$\frac{\pi}{6} < \theta < \frac{5\pi}{6}.$$

The length of this portion of the circumference of the circle is $\frac{4\pi}{6}$, and so the probability of a random angle θ satisfying $\sin\theta > \frac{1}{2}$ is

$$\begin{aligned} P &= \frac{4\pi/6}{2\pi} \\ &= \frac{1}{3}. \end{aligned}$$

(Return)

4. Note that tangent is positive when sine and cosine have the same sign, i.e. if sine and cosine are both positive or if sine and cosine are both negative. This corresponds to the first and third quadrants of the unit circle:



The length of each these arcs is $\frac{\pi}{2}$, and so the probability that $\tan\theta > 0$ is

$$P = \frac{2 \cdot \pi/2}{2\pi} = \frac{1}{2}$$

(Return)

5. Ignoring the unnecessary detail of the dartboard, let the radius of the bull's eye be r. Then

$$P\left(\text{(bullseye)}\right) = \frac{A(\text{ bullseye })}{A(\text{ board })} = \frac{\pi r^2}{\pi \cdot 9^2} = \frac{r^2}{81}.$$

Setting equal to $\frac{1}{100}$ and solving gives $r = 0.9$ inches. (In reality, the bullseye is much smaller, but the numbers worked out nicer in this example).

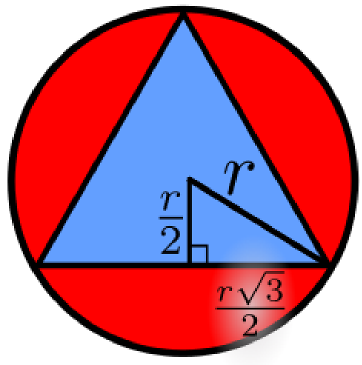
(Return)

6. If the radius of the circle is r, then the side length of the square is $2r$. Thus, the area of the circle is πr^2 and the area of the square is $(2r)^2 = 4r^2$. And so the probability that a point chosen at random within the square also lies within the circle is

$$P = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}.$$

(Return)

7. By doing a little bit of right triangle trigonometry:



we find that the side length of the triangle is

$$s = r\sqrt{3}.$$

Therefore, the area of the triangle is

$$\begin{aligned} A &= \frac{s^2\sqrt{3}}{4} \\ &= \frac{3r^2\sqrt{3}}{4}. \end{aligned}$$

And so the probability of a point within the circle being within the triangle is the ratio of the areas:

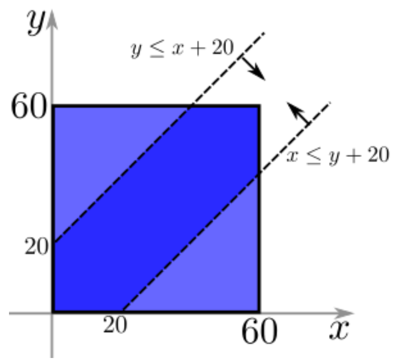
$$\begin{aligned} \frac{\text{Area of triangle}}{\text{Area of circle}} &= \frac{1}{\pi r^2}\frac{3r^2\sqrt{3}}{4} \\ &= \frac{3\sqrt{3}}{4\pi} \end{aligned}$$

(Return)

8. The possible outcomes form a square for $0 \leq x \leq 60$ and $0 \leq y \leq 60$. For the friends to meet, we must have that Yolanda arrives no later than 20 minutes after Xander *and* that Xander arrives no later than 20 minutes after Yolanda arrives. Mathematically,

$$\begin{aligned} y &\leq x + 20 \\ x &\leq y + 20 \end{aligned}$$

The two will successfully meet if and only if these two conditions are met. Graphing these inequalities, the points they have in common are shown below in dark blue:



So the probability of them meeting is the area of the dark blue region divided by the total area of the square. It is easier to determine the area of the region we do not want and subtract. There are two isosceles right triangles of side length 40, so the area of the region we do not want is

$$\text{bad area } = 2 \cdot \frac{1}{2} \cdot 40 \cdot 40 = 1600$$

Therefore, the probability of the friends meeting is

$$\begin{aligned} P &= \frac{\text{area of dark blue region}}{\text{total area}} \\ &= \frac{\text{total area} - \text{light blue area}}{\text{total area}} \\ &= \frac{3600 - 1600}{3600} \\ &= \frac{2000}{3600} \\ &= \frac{5}{9}. \end{aligned}$$

(Return)

9. If r is the radius of the inscribed sphere, then the side length of the cube is $2r$. Therefore, the volume of the sphere is $\frac{4}{3}\pi r^3$ and the volume of the cube is $(2r)^3 = 8r^3$. So the probability that a random point within in the cube lies within the sphere is

$$\begin{aligned} P &= \frac{(4/3)\pi r^3}{8r^3} \\ &= \frac{\pi}{6}. \end{aligned}$$

(Return)

10. Let r be the radius of the ball. Then the volume of the ball (the volume of all the possible outcomes) is $\frac{4}{3}\pi r^3$.

To find the volume of the desired outcomes, consider the volume of the undesired outcomes: those points which lie within 90% of the center. These points form a ball of radius $\frac{9}{10}r$, hence their volume is

$\frac{4}{3}\pi\left(\frac{9}{10}r\right)^3$. So the desirable outcomes have the complementary volume

$$\begin{aligned}\text{volume of desired outcomes} &= \text{ total volume} - \text{ volume of undesired outcomes}\\ &= \frac{4}{3}\pi r^3 - \frac{4}{3}\pi(\frac{9}{10}r)^3\\ &= \frac{4}{3}\pi r^3\left(1-(9/10)^3\right).\end{aligned}$$

Thus, the probability of a point being within 10% of the boundary is

$$\begin{aligned}\frac{\text{volume of desired outcomes}}{\text{total volume}} &= \frac{\frac{4}{3}\pi r^3(1-(9/10)^3)}{\frac{4}{3}\pi r^3}\\ &= 1-(9/10)^3\\ &= 0.271\end{aligned}$$

(Return)

43 Probability Densities

The last module dealt with the uniform distribution, where any one outcome is as likely as another. This module deals with experiments whose outcomes have different probabilities. For example, consider an unfair coin which has a $\frac{2}{3}$ probability of landing heads and a $\frac{1}{3}$ probability of landing tails. Another example is time spent on hold with customer service, where it is more likely that the call is answered in the first hour than in the second hour.

43.1 Random variable and probability density function (PDF)

A *random variable* X is a function whose output should be thought of as the outcome of an experiment. Associated with a random variable is a *probability density function* (PDF) $\rho(x)$, which is defined by $P(a \leq X \leq b) = \int_a^b \rho(x)dx$. That is, the probability that the random variable falls in a certain range of values is given by integrating the PDF over that range of values.

Phrased another way, we can think of probability P as the quantity we want to compute over a certain range of values, and the probability element is given by

$$dP = \rho(x)\,dx.$$

Example

Consider the spinner from the last module. The outcome of a spin is some angle (relative to the positive x-axis) between 0 and 2π. If X is the random variable which gives the output of a spin, then

$$P(a \leq X \leq b) = \frac{b-a}{2\pi},$$

since the spinner was assumed to be fair. This holds for all $0 \leq a \leq b \leq 2\pi$. Then the associated PDF is

$$\rho(x) = \begin{cases} \frac{1}{2\pi} & \text{if } 0 \leq x \leq 2\pi \\ 0. & \text{otherwise} \end{cases}$$

Note

Sometimes a PDF $\rho(x)$ is only defined on a certain domain D. D can be thought of as the set of all possible outcomes of the experiment X. In this case, it is assumed that $\rho(x) = 0$ for x not in that domain. So another way of defining the PDF for the spinner is $\rho(x) = \frac{1}{2\pi}$ for $0 \leq x \leq 2\pi$.

43.2 Properties of a probability density function

The following are defining properties of a PDF. In other words, a function $\rho(x)$ is a PDF on the domain D if and only if it satisfies these properties.

1. $\rho(x) \geq 0$ for all $x \in D$.
2. $\int_D \rho(x)\,dx = 1$.

The first property is necessary since probabilities must be non-negative. The second property reflects the fact that the random variable X associated with $\rho(x)$ must have *some* outcome in the domain D (since D is the set of all possible outcomes), and so integrating over all of these outcomes should give 1.

Note

If $\rho(x)$ is defined on some specific domain D, then the integral over that specific domain should equal 1. This is because $\rho(x) = 0$ outside of that domain, as mentioned in the above note.

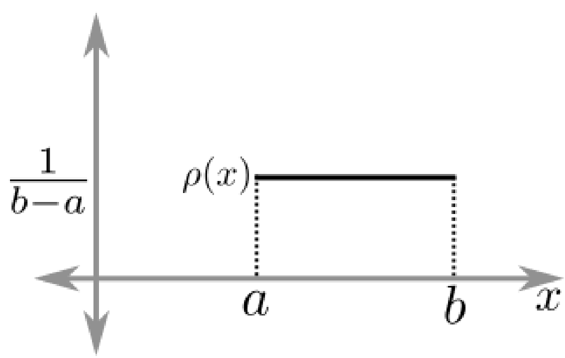
Example

Find the value of the constant c so that $\rho(x) = \frac{c}{1+x^2}$ for all x is a PDF. (See Answer 1)

43.3 Several specific density functions

Uniform density

Hinted at above and in the previous module, the *uniform density function* (or uniform distribution) on $[a, b]$ is given by $\rho(x) = \frac{1}{b-a}$ (and $\rho(x) = 0$ if x is not in $[a, b]$):



More generally, the uniform distribution on the domain D (whatever the dimension) is given by

$$\rho(x) = \frac{1}{\text{Volume of } D}.$$

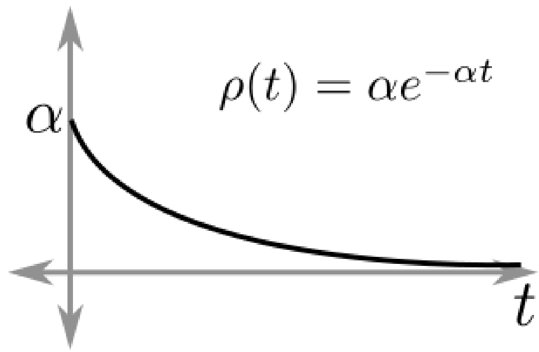
In dimension 0, where outcomes are discrete (as in the rolling of a die or the flipping of a coin), remember that volume is just counting. So in this case the probability of a particular outcome is

$$\rho(x) = \frac{1}{n},$$

where n is the number of outcomes in the domain D (e.g. $n = 6$ for the roll of a die; $n = 2$ for a coin flip).

Exponential density

Another density function used to model many common experiments is the *exponential density function*. This is actually a whole family of density functions given by $\rho(t) = \alpha e^{-\alpha t}$ for $t \geq 0$ and $\alpha > 0$ some constant. The reason a parameter t is used is that the exponential density is often used to model experiments with a time outcome.



Example

Show that the exponential density $\rho(t) = \alpha e^{-\alpha t}$ (for $t \geq 0$) satisfies the properties of a density function. (See Answer 2)

Example

Consider a call made to customer service at Acme company. The number of minutes spent on hold before the call is answered is often modeled with an exponential density function

$$\rho(t) = \alpha e^{-\alpha t}.$$

Find, in terms of α, the probability that the waiting time for a call is less than 30 minutes. (See Answer 3)

Example

Again consider customer service call waiting time at Acme company, and again assume an exponential density function

$$\rho(t) = \alpha e^{-\alpha t}.$$

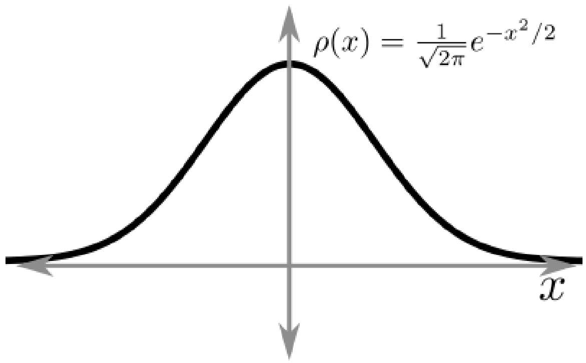
Suppose half of all customers are answered within 5 minutes. Find α and then find the probability that a call takes more than 10 minutes to be answered. (See Answer 4)

Gaussian density

The last probability density function is the '*Gaussian*, or *normal*, density function. This is an important density function and is expanded on in the next module. Like the exponential, the Gaussian density function usually has parameters (see the next module), but in its simplest form, the Gaussian is given by

$$\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

The Gaussian has all real x as its domain, but because it tails off so quickly in both directions, the probability of getting values far from the center (in this case $x = 0$) is very small.



43.4 EXERCISES

- Which of the following are probability density functions?

$$\begin{aligned}
&\text{a. } f(x) = 1/2 \text{ on } D = [0, 2]\\
&\text{b. } f(x) = \frac{\sin(x)}{2} \text{ on } D = [0, 3\pi]\\
&\text{c. } f(x) = 5e^{-2x} \text{ on } D = [0, \infty)\\
&\text{d. } f(x) = \frac{1}{\pi(1 + x^2)} \text{ on } D = (-\infty, \infty)\\
&\text{e. } f(x) = \frac{x}{2} \text{ for } 0 \le x \le 1\\
&\qquad\qquad \frac{1}{2} \text{ for } 1 \le x \le 2\\
&\qquad\qquad \frac{3}{2} - \frac{x}{2} \text{ for } 2 \le x \le 3
\end{aligned}$$

43.5 Answers to Selected Examples

1. As long as $c \ge 0$, the first property for a PDF will be met, since $1 + x^2 > 0$ for all x. To satisfy the second property, compute

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{c}{1 + x^2}\, dx &= c\left(\arctan(x)\Big|_{-\infty}^{\infty}\right)\\
&= c\left(\frac{\pi}{2} - (-\frac{\pi}{2})\right)\\
&= c\pi.
\end{aligned}$$

Since this integral is supposed to be 1, we find that $c = \frac{1}{\pi}$.

(Return)

2. The exponential function is never negative, so one need only check the integral. One finds

$$\begin{aligned}\int_{t=0}^{\infty} \alpha e^{-\alpha t}\, dt &= \alpha \frac{1}{-\alpha} e^{-\alpha t}\Big|_{t=0}^{\infty} \\ &= -(0-1) \\ &= 1,\end{aligned}$$

as desired. So the exponential density is in fact a density.

(Return)

3. To find the probability that $0 \le X \le 30$, use the relationship between probability and the PDF, which is

$$\begin{aligned}P(0 \le X \le 30) &= \int_0^{30} \rho(x)\, dx \\ &= \int_0^{30} \lambda e^{-\lambda x}\, dx \\ &= -e^{-\lambda x}\Big|_0^{30} \\ &= -e^{-30\lambda} - (-1) \\ &= 1 - e^{-30\lambda}.\end{aligned}$$

(Return)

4. Since half of all customers are answered within 5 minutes, we have that

$$P(0 \le X \le 5) = \frac{1}{2}.$$

On the other hand, we know that this can be expressed as the integral of the density function, so we have

$$\begin{aligned}\frac{1}{2} &= \int_{t=0}^{5} \rho(t)\, dt \\ &= \int_{t=0}^{5} \alpha e^{-\alpha t}\, dt \\ &= -e^{-\alpha t}\Big|_{t=0}^{5} \\ &= -e^{-5\alpha} - (-1) \\ &= 1 - e^{-5\alpha}.\end{aligned}$$

So we have that

$$e^{-5\alpha} = \frac{1}{2}.$$

Taking the log of both sides, dividing by -5 and simplifying, we have

$$\begin{aligned}\alpha &= \frac{1}{-5} \ln\left(\frac{1}{2}\right) \\ &= \frac{1}{-5}(-\ln 2) \\ &= \frac{1}{5} \ln 2.\end{aligned}$$

For the second part, we want to know the probability of waiting more than 10 minutes. This is (leaving α as a constant for now)

$$\begin{aligned}
P(X \geq 10) &= \int_{t=10}^{\infty} \rho(t)\,dt \\
&= \int_{t=10}^{\infty} \alpha e^{-\alpha t}\,dt \\
&= -e^{-\alpha t}\Big|_{t=10}^{\infty} \\
&= 0 - \left(-e^{-\alpha \cdot 10}\right).
\end{aligned}$$

Now plugging in the value of α, we have

$$\begin{aligned}
P(X \geq 10) &= e^{-(\ln 2/5)\cdot 10} \\
&= e^{-2\ln 2} \\
&= 2^{-2} \\
&= \frac{1}{4}.
\end{aligned}$$

(Return)

44 Expectation And Variance

When performing an experiment, it is useful to know what the expected outcome will be as well as how much variation one can expect among the outcomes. The notions of expected outcome and variation are made formal in this module by the terms *expectation*,*variance*, and *standard deviation*.

This module will also show some of the connections of these statistical metrics with the applications of the previous modules.

44.1 Expectation

Consider a random variable X with probability density function (PDF) $\rho(x)$ defined on some domain D. The *expectation of* X, denoted by $\mathbb{E}$, is defined by

$$\begin{aligned}\mathbb{E} &= \int_D x\rho(x)\,dx \\ &= \int_D x\,dP,\end{aligned}$$

where dP is the probability element. The expectation of X is sometimes called the *mean of* X, the *expected value*, or the *first moment*. In some books it is denoted μ_X. It is best to think of the expectation as the number one gets by repeating the experiment many times and taking the average of the outputs.

The notion of expectation is more general than the mean because one can also take the expectation of a function of X. The *expectation of* $f(X)$ is defined by

$$\mathbb{E}[f(X)] = \int_D f(x)\rho(x)\,dx.$$

Example

Find the expectation of X, where X is uniformly distributed on the interval $[a, b]$. (See Answer 1)

Example

Recall that the random variable X is said to have the *exponential distribution* if the PDF associated with X is $\rho(t) = \alpha e^{-\alpha t}$ for $t \geq 0$, where $\alpha > 0$ is some constant. Find the expectation of the exponential distribution (in terms of α). (See Answer 2)

44.2 Variance

Consider a random variable X with PDF $\rho(x)$. The *variance of* X, denoted $\mathbb{V}$, is defined by

$$\begin{aligned}\mathbb{V} &= \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2.\end{aligned}$$

In the notation of the lecture,

$$\begin{aligned}\mathbb{V} &= \int_D (x - \mathbb{E})^2 \, dP \\ &= \int_D x^2 \, dP - \mathbb{E}^2.\end{aligned}$$

Note: it requires some calculation to show the second equality above holds. Either of the above expressions may be taken as the definition of variance, and the second one might be slightly simpler for the sake of computation. (See Justification 3)

Example

Compute the variance of the exponential density function $\rho(x) = \alpha e^{-\alpha x}$. (See Answer 4)

44.3 Standard deviation

Consider a random variable X with PDF $\rho(x)$. Then the *standard deviation of* X, denoted σ_X, is defined by

$$\begin{aligned}\sigma_X &= \sqrt{V[X]} \\ &= \sqrt{E[X^2] - E[X]^2} \\ &= \sqrt{\int_D x^2 \rho(x)\, dx - \left(\int_D x\rho(x)\, dx\right)^2}.\end{aligned}$$

Example

Find the standard deviation of X, where X is uniformly distributed over $[a, b]$. (See Answer 5)

44.4 Interpretations

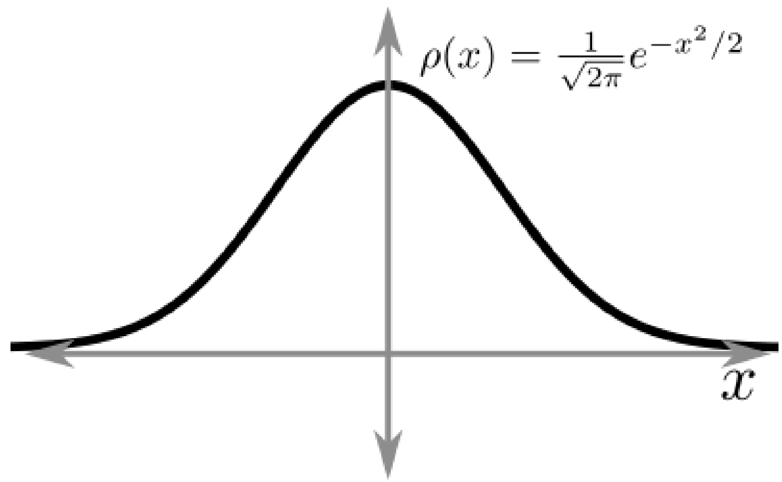
If one interprets the PDF $\rho(x)$ as the density of a rod at location x, then:

1. The mean, $\mu = \int x\rho(x)\, dx$, gives the center of mass of the rod.
2. The variance, $V = \int (x - \mu)^2 \rho(x)\, dx$, gives the moment of inertia about the line $x = \mu$.
3. The standard deviation, $\sigma = \sqrt{V}$, gives the radius of gyration about the line $x = \mu$.

44.5 The normal distribution

A random variable X is said to have the *normal distribution*, or to be *normally distributed*, with mean μ and standard deviation σ if its PDF is of the form

$$\rho(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$



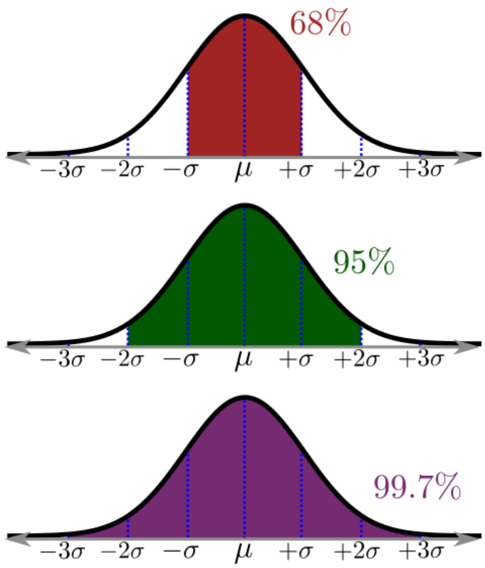
Due to its ubiquity throughout the sciences, the normal distribution is one of the most well-known probability distributions. However, because its PDF does not have an elementary anti-derivative, it is not easy to calculate exact probabilities associated with the normal distribution. Instead, there are is a rule of thumb which can be used.

The 68-95-99.7 rule

Given a random variable X which is normally distributed with mean μ and standard deviation σ, the following hold:

1. $P(\mu - \sigma \leq X \leq \mu + \sigma) \approx .68$.
2. $P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx .95$.
3. $P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) \approx .997$.

In other words, 68% of samples will fall within 1 standard deviation of the mean. 95% of samples will fall within 2 standard deviations of the mean. And 99.7% of samples will fall within 3 standard deviations. These rules, along with the symmetry of the normal PDF, can be used to approximate many probabilities relating to the normal distribution:



Example

The height of men in a certain population is normally distributed with mean $\mu = 70$ inches and standard deviation $\sigma = 2$ inches. If a man is chosen at random from the population, what is the probability that he is taller than 72 inches? (See Answer 6)

44.6 EXERCISES

- Compute the expected value of normally distributed random variable with probability density function $\rho(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ on $-\infty < x < \infty$.

44.7 Answers to Selected Examples

1. Recall that the PDF associated with X is given by $\rho(x) = \frac{1}{b-a}$ for $a \leq x \leq b$. Thus, the mean is given by

$$\begin{aligned}\mathbb{E} &= \int_a^b x \cdot \frac{1}{b-a}\, dx \\ &= \frac{1}{b-a} \frac{x^2}{2}\bigg|_a^b \\ &= \frac{1}{b-a} \cdot \frac{1}{2}(b^2 - a^2) \\ &= \frac{1}{b-a} \cdot \frac{1}{2}(b+a)(b-a) \\ &= \frac{1}{2}(a+b).\end{aligned}$$

(Return)

2. From the definition of expectation, one finds

$$\begin{aligned}\mathbb{E} &= \int_0^\infty t\alpha e^{-\alpha t}\,dt \\ &= \alpha \int_0^\infty te^{-\alpha t}\,dt.\end{aligned}$$

Using integration by parts, with

$$\begin{aligned} u &= t & du &= dt \\ dv &= e^{-\alpha t} & v &= \frac{1}{-\alpha}e^{-\alpha t},\end{aligned}$$

we find that

$$\begin{aligned}\alpha \int_0^\infty te^{-\alpha t}\,dt &= \alpha\left(\frac{t}{-\alpha}e^{-\alpha t} - \int_0^\infty \frac{1}{-\alpha}e^{-\alpha t}\,dt\right) \\ &= \left(-te^{-\alpha t} - \frac{1}{\alpha}e^{-\alpha t}\right)\Big|_0^\infty \\ &= (0-0) - (0 - \frac{1}{\alpha}) \\ &= \frac{1}{\alpha}.\end{aligned}$$

(Return)

3. Expanding out the expression and using the linearity of the integral, we find

$$\begin{aligned}\int_D (x-\mathbb{E})^2\,dP &= \int_D (x-\mathbb{E})^2 \rho(x)\,dx \\ &= \int_D x^2\,dP - \int_D 2x\mathbb{E}\,dP + \int \mathbb{E}^2\,dP \\ &= \int_D x^2\,dP - 2\mathbb{E}\int_D x\,dP + \mathbb{E}^2 \int \,dP \\ &= \int_D x^2\,dP - 2\mathbb{E}\cdot\mathbb{E} + \mathbb{E}^2 \\ &= \int_D x^2\,dP - \mathbb{E}^2.\end{aligned}$$

because $\int x\,dP = \mathbb{E}$ and $\int\,dP = 1$, by the definition of expectation and the definition of the probability density function, respectively.

(Return)

4. The variance requires us to compute

$$\begin{aligned}\mathbb{E}(X^2) &= \int_D x^2\,dP \\ &= \int_{x=0}^\infty x^2 \alpha e^{-\alpha x}\,dx.\end{aligned}$$

Using integration by parts, with

$$
\begin{aligned}
u &= x^2 & du &= 2x\,dx \\
dv &= \alpha e^{-\alpha x}\,dx & v &= -e^{-\alpha x},
\end{aligned}
$$

we find

$$\int_{x=0}^{\infty} x^2 \alpha e^{-\alpha x}\,dx = -x^2 e^{-\alpha x}\Big|_{x=0}^{\infty} + \int_{x=0}^{\infty} 2x e^{-\alpha x}\,dx.$$

This second integral can be done with integration by parts again, or we can use the fact that this is almost the integral for the expectation. Namely, we know

$$\int_{x=0}^{\infty} x\alpha e^{-\alpha x}\,dx = \frac{1}{\alpha},$$

and so by dividing through by α, we have

$$\int_{x=0}^{\infty} x e^{-\alpha x}\,dx = \frac{1}{\alpha^2}.$$

Putting this together, we have

$$
\begin{aligned}
\int_{x=0}^{\infty} x^2 \alpha e^{-\alpha x}\,dx &= -x^2 e^{-\alpha x}\Big|_{x=0}^{\infty} + \frac{2}{\alpha^2} \\
&= (0-0) + \frac{2}{\alpha^2} \\
&= \frac{2}{\alpha^2}.
\end{aligned}
$$

Finally, then, the variance is

$$
\begin{aligned}
\mathbb{V} &= \int_D x^2\,dP - \mathbb{E}^2 \\
&= \frac{2}{\alpha^2} - \left(\frac{1}{\alpha}\right)^2 \\
&= \frac{1}{\alpha^2}.
\end{aligned}
$$

(Return)

5. Again, recall that the PDF for the uniform distribution is $\rho(x) = \frac{1}{b-a}$ for $a \leq x \leq b$. Thus,

$$
\begin{aligned}
E[X^2] &= \int_a^b x^2 \frac{1}{b-a}\,dx \\
&= \frac{1}{b-a}\frac{x^3}{3}\Big|_a^b \\
&= \frac{1}{b-a}\cdot\frac{1}{3}(b^3 - a^3) \\
&= \frac{1}{b-a}\cdot\frac{1}{3}(b-a)(b^2+ba+a^2) \\
&= \frac{b^2+ab+a^2}{3}
\end{aligned}
$$

From the previous example, $E[X] = \mu_X = \frac{a+b}{2}$. Thus,

$$\begin{aligned}
\sigma_X &= \sqrt{E[X^2] - E[X]^2} \\
&= \sqrt{\frac{b^2 + ba + a^2}{3} - \frac{b^2 + 2ba + a^2}{4}} \\
&= \sqrt{\frac{b^2 - 2ab + a^2}{12}} \\
&= \frac{b-a}{\sqrt{12}}.
\end{aligned}$$

(Return)

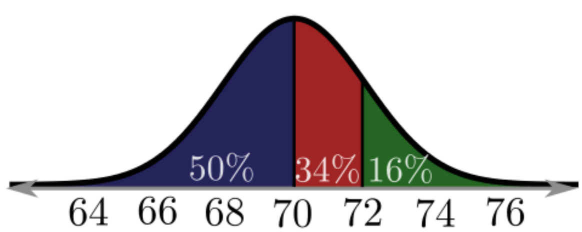
6. Let X be the height of a randomly chosen man. Then $P(68 \leq X \leq 72) = .68$ by the above rule. By symmetry $P(68 \leq X \leq 70) = P(70 \leq X \leq 72) = .34$. Also, by symmetry, $P(X \leq 70) = .5$. Thus,

$$\begin{aligned}
P(X \leq 72) &= P(X \leq 70) + P(70 \leq X \leq 72) \\
&= .5 + .34 \\
&= .84.
\end{aligned}$$

It follows that

$$\begin{aligned}
P(X > 72) &= 1 - P(X \leq 72) \\
&= 1 - .84 \\
&= .16.
\end{aligned}$$

This is best visualized by labeling the various regions under the normal curve with their areas:



So the probability that a randomly chosen man from the population is taller than 72 inches is .16.

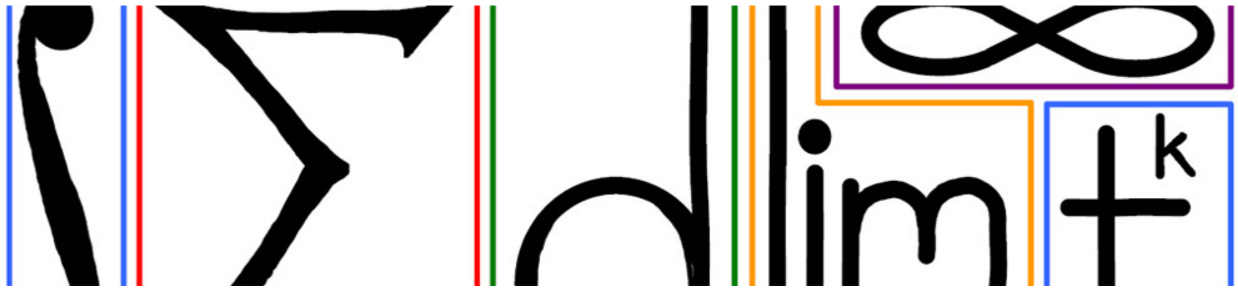
(Return)

The Penn Calc Companion

Part III: Discrete Calculus

About this Document

Contents



45 Sequences

The remainder of the course is a look at discrete calculus, which is a study of all the previous sections (functions, derivatives, integrals) applied to a different kind of function: sequences. A sequence is a function, but instead of taking any real number as input, a sequence takes an integer as input.

45.1 Sequence

A *sequence* a is a function from the non-negative integers $0, 1, 2, 3, \ldots$ to the real numbers $\mathbb{R}$. The usual functional notation $a(n)$ is sometimes replaced with a_n. There are several ways to define a sequence. Here are three of the most common ways, demonstrated on the powers of 2.

1. An *explicit formula* gives a_n as a function of n, i.e. $a_n = f(n)$. This is usually the most convenient, since it typically gives the most information about the sequence. e.g. $a_n = 2^n$ for $n \geq 0$.
2. A *recursion relation* gives a_n as a function of previous terms in the sequence. Note that some initial conditions must be given as well. $a_n = f(a_{n-1}, a_{n-2}, \ldots, a_{n-k})$. e.g. $a_n = 2a_{n-1}$; $a_0 = 1$.
3. Finally, *listing terms* can be used if no explicit or recursive formula is available. This is sometimes used in experimental settings so that one can study the terms and look for a pattern. e.g. $a = (1, 2, 4, 8, 16, 32, \ldots)$.

Example

Write out the first six terms of the sequence defined by $a_n = 2a_{n-1} + 1$; $a_0 = 0$. Look for a pattern to try to find an explicit formula for a_n. (See Answer 1)

45.2 Limits of sequences

Recall that limits of functions came in two flavors. First, there were limits of the form $\lim_{x \to c} f(x)$. The second type of limits were of the form $\lim_{x \to \infty} f(x)$.

Only the second type of limit is sensible for a sequence. (To see why the first type of limit does not make sense for sequences, go back to the definition of a limit.) The definition of $\lim_{n \to \infty} a_n$ for sequences is the same as for continuous functions:

The Limit of a Sequence

We say that

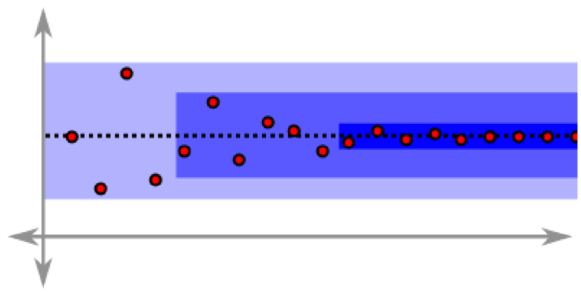
$$\lim_{n\to\infty} a_n = L$$

if for any $\epsilon > 0$ there exists M such that for all $n > M$,

$$|a_n - L| < \epsilon.$$

In other words, the sequence a_n has limit L if a_n gets arbitrarily close to L for sufficiently large n. If the limit L exists, then the sequence a_n is said to *converge* to L.

Intuitively, a sequence a_n has a limit if for any band around L, there is some point where all the terms of a_n are within the band around L:



Recall that Newton's method defined a sequence of numbers defined by the recursion relation

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Here, the sequence *hopefully* converged to a root of the function $f(x)$. This gives an example of an application where it is useful to know about the convergence of a sequence.

Example

Let $a_n = 4 + \frac{(-1)^n}{n}$. Find $\lim_{n\to\infty} a_n$, if the limit exists. If the limit does not exist, explain why. (See Answer 2)

Example

Let $a_n = (-1)^n$. Find $\lim_{n\to\infty} a_n$, if the limit exists. If the limit does not exist, explain why. (See Answer 3)

45.3 Methods for computing limits

Many of the methods for computing limits of continuous functions carry over to computing limits of sequences. In particular, all of the big-O notation still applies.

Example

Compute the limit of the sequence $a_n = 3n - \sqrt{9n^2 + 6n}$. (See Answer 4)

45.4 Monotone, bounded sequences

In general, it can be difficult to find the limit of a sequence, but for certain sequences it is possible to prove that the limit exists.

A sequence is *monotone increasing* if it is non-decreasing, i.e., $a_0 \leq a_1 \leq a_2 \leq \dots$. A sequence is *monotone decreasing* if it is non-increasing, i.e., $a_0 \geq a_1 \geq a_2 \geq \dots$. A sequence that is either monotone increasing or monotone decreasing is *monotone*.

A sequence is *bounded above* if there exists some real number B such that $a_n \leq B$ for all $n \geq 0$. Similarly, a sequence is *bounded below* if there exists a real number C such that $a_n \geq C$ for all $n \geq 0$. A sequence is *bounded* if it is both bounded above and bounded below.

Monotone Convergence Theorem

If a sequence a_n is bounded and monotone, then the sequence converges.

Example

Let a_n be the sequence defined by $a_n = \frac{5+a_{n-1}}{2}$; $a_0 = 1$. Show that the sequence converges by using the Monotone Convergence Theorem. (See Answer 5)

45.5 Recursion relations and limits

When a sequence is defined by a recursion relation and the limit of the sequence exists, one can find the limit by simply taking the limit of both sides of the recursion relation and solving. This is best demonstrated by example.

Example

Find the limit $L = \lim_{n\to\infty} a_n$, where $a_n = \frac{5+a_{n-1}}{2}$; $a_0 = 1$, as in the previous example. (See Answer 6)

Example

Find

$$L = 1 + \cfrac{1}{1 + \frac{1}{1+\frac{1}{1+\dots}}}$$

by expressing L as a limit of a recursively defined sequence a_n which begins $(1, 1 + \frac{1}{1}, 1 + \frac{1}{2}, \dots)$. Assume that the limit exists. (See Answer 7)

Example

Find $\sqrt{1 + \sqrt{1 + \sqrt{1 + \cdots}}}$. Assume the limit exists. (See Answer 8)

The golden ratio, often denoted

$$\varphi = \frac{1+\sqrt{5}}{2},$$

appears in many settings, both man made and natural. The golden ratio is also deeply connected with the Fibonacci numbers, as we will see in this example and in the future.

Example

The *Fibonacci sequence*, F_n is defined by the recursion relation and initial conditions

$$F_{n+2} = F_{n+1} + F_n; \quad F_0 = 0, F_1 = 1.$$

So the first few Fibonacci numbers are

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \ldots$$

Show that

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \varphi.$$

You may assume that the limit of the sequence exists. Hint: Observe that

$$\frac{F_{n+1}}{F_n} = \frac{F_n + F_{n-1}}{F_n} = 1 + \frac{F_{n-1}}{F_n}.$$

(See Answer 9)

45.6 EXERCISES

- Compute the limit of the sequence $a_n = 2n^2 - (8n^6 + 6n^4)^{1/3}$.

45.7 Answers to Selected Examples

1.

$$\begin{aligned}
a_0 &= 0\\
a_1 &= 2a_0 + 1 = 1\\
a_2 &= 2a_1 + 1 = 3\\
a_3 &= 2a_2 + 1 = 7\\
a_4 &= 2a_3 + 1 = 15\\
a_5 &= 2a_4 + 1 = 31.
\end{aligned}$$

One might notice that adding 1 to each term in the sequence gives the sequence $(1, 2, 4, 8, 16, 32, \ldots)$, which look like the powers of 2. So it appears that $a_n = 2^n - 1$ for $n \geq 0$.

(Return)

2. We claim the limit is 4. We know that

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

Therefore, for any $\epsilon > 0$ we can choose M so that $\frac{1}{M} < \epsilon$. Then for any $n > M$, we have

$$\begin{aligned}
|a_n - 4| &= \left|4 + \frac{(-1)^n}{n} - 4\right| \\
&= \left|\frac{(-1)^n}{n}\right| \\
&= \frac{1}{n} \\
&< \frac{1}{M} \\
&< \epsilon,
\end{aligned}$$

as desired. For this course, we will not typically be this formal, but it is useful to see this type of argument at least a few times.

(Return)

3. From the intuitive understanding of a limit, it is clear that the terms of this sequence are not getting closer together, and so the limit does not exist.

   More formally, suppose the limit, say L, existed. Then from the definition of the limit of a sequence, we could find a number M such that

   $$|a_n - L| < \frac{1}{3}$$

   for all $n > M$. Since the terms of the sequence are -1 and 1, this would imply

   $$|-1 - L| < \frac{1}{3}$$

   This implies

   $$\begin{aligned}
   -\frac{1}{3} &< -1 - L < \frac{1}{3} \\
   \frac{2}{3} &< -L < \frac{4}{3} \\
   -\frac{2}{3} &> L > \frac{-4}{3}.
   \end{aligned}$$

   Similarly,

   $$|1 - L| < \frac{1}{3}$$

   which implies by similar algebra that

   $$\frac{4}{3} > L > \frac{2}{3}.$$

   Since L cannot be simultaneously positive and negative, we have reached a contradiction. And so we see that the limit L cannot exist.

   (Return)

4. A little factoring from the radical, and using the binomial series gives

$$\begin{aligned} a_n &= 3n - \sqrt{9n^2 + 6n} \\ &= 3n - 3n\sqrt{1 + \frac{2}{3n}} \\ &= 3n\left(1 - \sqrt{1 + \frac{2}{3n}}\right) \\ &= 3n\left(1 - (1 + \frac{1}{2}\frac{2}{3n} + O(1/n^2))\right) \\ &= -1 + O\left(\frac{1}{n}\right). \end{aligned}$$

So the limit is -1.

(Return)

5. It is not obvious that this sequence is either bounded or monotone. Writing out the first few terms, though, gives $a = (1, 3, 4, 4.5, 4.75, \ldots)$. It appears that the sequence is increasing, and that $a_n \leq 5$ for all $n \geq 0$.

   To show that the sequence is increasing, use induction. The first few terms are increasing, so assume that $a_{n-1} \leq a_n$ for some n. Then adding 5 and dividing by 2 throughout gives $\frac{5+a_{n-1}}{2} \leq \frac{5+a_n}{2}$, which implies $a_n \leq a_{n+1}$ by the recursive definition of a_n. Thus, the sequence is increasing.

   To see that $a_n \leq 5$, again use induction. Assume $a_n \leq 5$ for some n. Then adding 5 and dividing by 2 gives $\frac{5+a_n}{2} \leq \frac{5+5}{2} = 5$. This means $a_{n+1} \leq 5$ by the definition of a_n. Thus, $a_n \leq 5$ for all n. Finally, note that any increasing sequence is bounded below. So the sequence is bounded.

   Thus, by the Monotone Convergence Theorem, the sequence a_n converges.

   (Return)

6. The limit exists by the previous example, so taking limits of both sides of the recursion relation and simplifying gives

$$\begin{aligned} \lim_{n\to\infty} a_n &= \lim_{n\to\infty} \frac{5 + a_{n-1}}{2} \\ L &= \frac{5 + \lim_{n\to\infty} a_{n-1}}{2} \\ L &= \frac{5 + L}{2}. \end{aligned}$$

   (Note that $\lim a_{n-1} = \lim a_n$ because the limit of a sequence does not depend on the indexing of the terms.) Now, solving for L gives $2L = 5 + L$, so $L = 5$.

   (Return)

7. The sequence a_n can be defined recursively by $a_n = 1 + \frac{1}{a_{n-1}}$; $a_0 = 1$. Then taking limits of both sides gives $L = 1 + \frac{1}{L}$. Multiplying through and collecting terms gives $L^2 - L - 1 = 0$, and solving for L gives $L = \frac{1\pm\sqrt{5}}{2}$.

   Note though that $a_n > 0$ for all n, so it must be that $L = \frac{1+\sqrt{5}}{2}$. This is the celebrated *golden ratio*.

   (Return)

8. One can express this as the limit of a recursively defined sequence a_n given by $a_0 = 1$ and $a_n = \sqrt{1 + a_{n-1}}$. Then letting $L = \lim a_n$, and taking limits of the recursion relation gives

$$L = \sqrt{1 + L}.$$

Squaring both sides and collecting terms gives $L^2 - L - 1 = 0$, which is the same equation from the previous example. Thus, the limit is again the golden ratio $L = \frac{1+\sqrt{5}}{2}$.

(Return)

9. Let

$$L = \lim_{n\to\infty} \frac{F_{n+1}}{F_n}.$$

Then, beginning with the hint and taking the limit of both sides, we have

$$\begin{aligned} \lim_{n\to\infty} \frac{F_{n+1}}{F_n} &= \lim_{n\to\infty} \left(1 + \frac{F_{n-1}}{F_n}\right) \\ &= 1 + \lim_{n\to\infty} \frac{F_{n-1}}{F_n} \\ &= 1 + \lim_{n\to\infty} \frac{1}{F_n/F_{n-1}} \\ &= 1 + \frac{1}{\lim\limits_{n\to\infty} F_n/F_{n-1}}. \end{aligned}$$

But this means

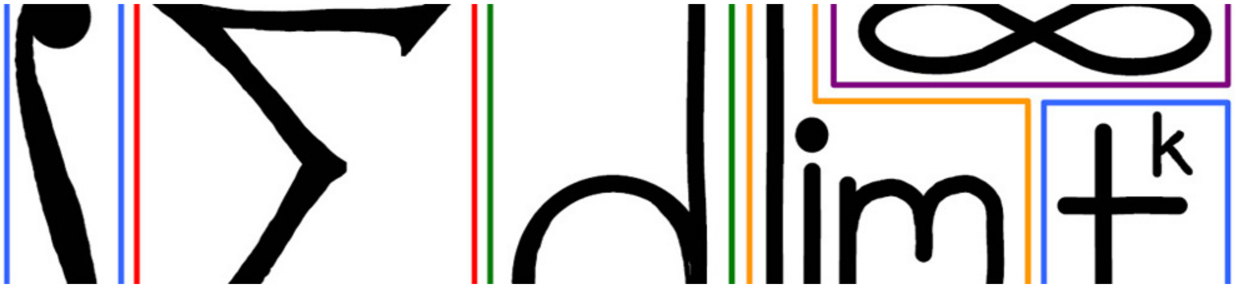
$$L = 1 + \frac{1}{L},$$

which is an equation which we solved in an above example, where we found

$$L = \frac{1+\sqrt{5}}{2},$$

as desired.

(Return)



46 Differences

What is the derivative of a sequence? The original definition will not work because change in input is discrete and so one cannot take the limit as the change in input goes to 0. Instead, using the interpretation of the derivative as a rate of change leads to two different discrete derivatives. They are called *difference operators*, and are defined below.

46.1 Difference operators

The discrete analog of the derivative is the difference operator, defined as follows.

Difference operators

Given a sequence a_n, the *forward difference of a*, denoted $(\Delta a)_n$, is defined by

$$(\Delta a)_n = a_{n+1} - a_n.$$

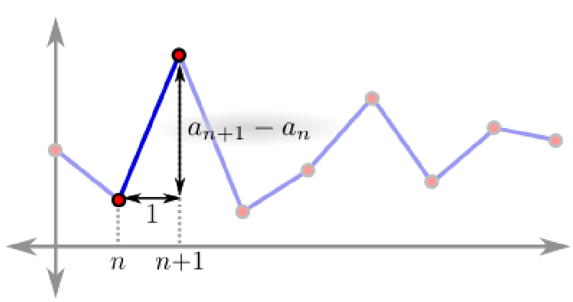
The *backward difference of a*, denoted by $(\nabla a)_n$, is defined by

$$(\nabla a)_n = a_n - a_{n-1}.$$

This can be interpreted as the change in output over the change in input:

$$(\Delta a)_n = \frac{a(n+1) - a(n)}{(n+1) - n} = \frac{a_{n+1} - a_n}{1} = a_{n+1} - a_n,$$

which resembles the definition for the derivative of a continuous function, but without the limit. If we plot the points of the sequence as if we were graphing the function, and then connect the dots with linear segments, then the forward difference can also be interpreted as the slope between adjacent points:



Example

The sequence $(4n) = 0, 4, 8, 12, 16, \ldots$ has forward difference sequence

$$\Delta(4n) = 4, 4, 4, 4, \ldots$$

Example

Find the forward difference sequence of the Fibonacci sequence $F = 0, 1, 1, 2, 3, 5, 8, 13, 21 \ldots$. (See Answer 1)

Example

Find the forward difference of the powers of two: $(2^n) = 1, 2, 4, 8, 16, 32, \ldots$. (See Answer 2)

Product rules

These difference operators have their own versions of the differentiation rules for continuous functions. For example, there is a product rule. For sequences $a = (a_n)$ and $b = (b_n)$, define the new sequence (ab) by $(ab)_n = (a_n b_n)$. Then

$$\Delta(ab)_n = a_n \Delta b_n + b_n \Delta a_n + \Delta a_n \Delta b_n$$
$$\nabla(ab)_n = a_n \nabla b_n + b_n \nabla a_n - \nabla a_n \nabla b_n.$$

This should be reminiscent of the product rule $(fg)' = fg' + f'g$.

46.2 Higher order difference operators

Just as the derivative of a function gives another function, the difference operator of a sequence a gives another sequence Δa. Then one can take the difference operator of Δa which gives yet another sequence $\Delta^2 a$. And so on. Consider, for instance, the sequence $a_n = n^3$ for $n \geq 0$. Then

$$\begin{array}{rcccccccccc}
a & = & 0 & 1 & 8 & 27 & 64 & 125 & 216 & 343 & \ldots \\
\Delta a & = & 1 & 7 & 19 & 37 & 61 & 91 & 127 & \ldots & \\
\Delta^2 a & = & 6 & 12 & 18 & 24 & 30 & 36 & \ldots & & \\
\Delta^3 a & = & 6 & 6 & 6 & 6 & 6 & \ldots & & & \\
\Delta^4 a & = & 0 & 0 & 0 & 0 & \ldots & & & &
\end{array}$$

A little bit of pattern matching shows the following:

$$\begin{aligned}
a_n &= n^3 \\
(\Delta a)_n &= 3n^2 + 3n + 1 \\
(\Delta^2 a)_n &= 6n + 6 \\
(\Delta^3 a)_n &= 6 \\
(\Delta^4 a)_n &= 0
\end{aligned}$$

Just as taking higher derivatives of a polynomial eventually gives 0, taking higher order difference operators of a polynomial eventually gives 0. Moreover, if a is a polynomial of degree p, then $\Delta^{p+1} a = (0)$.

Note that the power rule is not quite the same as for regular derivatives (e.g. $n^3 \mapsto 3n^2 + 3n + 1$). This is an artifact of the binomial expansion. The next section shows a convenient way to avoid these problems.

46.3 Falling powers

The *falling power* $n^{\underline{k}}$ is defined to be

$$n^{\underline{k}} = n(n-1)(n-2)\cdots(n-k+1), \qquad n^{\underline{0}} = 1.$$

The falling power can be thought of as a discrete version of the monomial x^k. One nice feature of the falling power is that

$$\begin{aligned}
\Delta n^{\underline{k}} &= (n+1)^{\underline{k}} - n^{\underline{k}} \\
&= (n+1)n(n-1)\cdots(n-k+2) - n(n-1)\cdots(n-k+1) \\
&= n(n-1)\cdots(n-k+2)\,(n+1-(n-k+1)) \\
&= kn^{\underline{k-1}},
\end{aligned}$$

which is the familiar power rule.

Example

Express n^3 in terms of falling factorials and use the power rule above to find Δn^3. (See Answer 3)

46.4 Discrete e

With the falling power in hand, consider the discretized version of the exponential function, found by replacing the usual monomials x^k with $n^{\underline{k}}$ in the Taylor series for the exponential:

$$\sum_{k=0}^{\infty} \frac{n^{\underline{k}}}{k!} = 1 + n + \frac{n(n-1)}{2!} + \frac{n(n-1)(n-2)}{3!} + \cdots$$

If one evaluates this at $n = 1$ to find the "discrete e", note that all the terms after the first two disappear because of the $n-1$ factor in all the higher terms. Thus, the discrete version of e is 2. This is consistent with the earlier note that 2^n acted like the exponential function (it is its own discrete derivative). Indeed, the above series is 2^n, which can be seen by noting it is simply the binomial series $(1+x)^n$ evaluated at $x = 1$.

46.5 Sequence operators

Just as there were operators on functions (e.g. integration, differentiation, logarithm, exponentiation), there are operators on sequences:

Operator	Notation	What it does	Notes
Identity	I	$(Ia)_n = a_n$	$I^2 = I$
Left shift	E	$(Ea)_n = a_{n+1}$	E^2 shifts twice, etc.
Right shift	E^{-1}	$(E^{-1}a)_n = a_{n-1}$	$E^{-1}E = I$
Forward difference	Δ	$(\Delta a)_n = a_{n+1} - a_n$	$\Delta = E - I$
Backward difference	∇	$(\nabla a)_n = a_n - a_{n-1}$	$\nabla = I - E^{-1}$

Higher derivatives

The expressions for the forward and backward differences in terms of I and E can be used to compute the higher derivatives of sequences more easily than it would be to compute them by hand. For example, the third derivative can be found by expanding the expression as a binomial:

$$\begin{aligned}\Delta^3 a &= (E-I)^3 a \\ &= (E^3 - 3E^2 I + 3EI^2 - I^3)a \\ &= (a_{n+3} - 3a_{n+2} + 3a_{n+1} - a_n).\end{aligned}$$

More generally, the kth derivative can be written similarly:

$$\begin{aligned}\Delta^k &= (E-I)^k \\ &= \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} E^i\end{aligned}$$

Indefinite integral

As shown above, the forward difference can be expressed in terms of the operators E and I:

$$\Delta = E - I.$$

A logical question is to ask whether it is possible to take the anti-difference, just as there is the antiderivative for functions. The claim is that

$$\Delta^{-1} = -(I + E + E^2 + E^3 + E^4 + \cdots)$$

This looks reminiscent of the geometric series. By taking the inverse of Δ (and a few liberties with the algebra), we have

$$\begin{aligned}\Delta^{-1} &= (E-I)^{-1} \\ &= -(I-E)^{-1} \\ &= -(I + E + E^2 + E^3 + \cdots),\end{aligned}$$

by thinking of $(I-E)^{-1}$ as the reciprocal of $1-E$ in a sense (because I acts like 1), and applying the geometric series. This derivation is not rigorous, but the above formula does give the anti-difference (up to an additive constant), so long as the sequence a_n is eventually 0.

This is just a hint of the calculus of sequences, which we explore a little bit more in the modules to come.

46.6 Answers to Selected Examples

1. The difference sequence is

   $$\Delta F = 1, 0, 1, 1, 2, 3, 5, 8, \ldots$$

   Note that this is just the Fibonacci sequence again, but slightly shifted. This makes sense since the difference

   $$(\Delta F)_n = F_{n+1} - F_n = F_{n-1},$$

   by rearranging the Fibonacci recursion relation.

   (Return)

2. The difference sequence is
$$\Delta(2^n) = 1, 2, 4, 8, 16, 32, \ldots$$
This shows that 2^n can be thought of as the discrete analog of the exponential e^x, in the sense that it is its own (discrete) derivative.

(Return)

3. This requires a little bit of algebra. We know that we need a $n^{\underline{3}}$ to get a n^3. Note that
$$n^{\underline{3}} = n(n-1)(n-2) = n^3 - 3n^2 + 2n.$$
So we have our n^3, but we also have some unwanted lower order terms. We must add $3n^{\underline{2}}$ to both sides to cancel the $-3n^2$. Note that
$$n^{\underline{2}} = n(n-1) = n^2 - n,$$
and so
$$\begin{aligned} n^{\underline{3}} + 3n^{\underline{2}} &= (n^3 - 3n^2 + 2n) + 3(n^2 - n) \\ &= n^3 - n. \end{aligned}$$
So to get rid of the final term of $-n$, we note that $n^{\underline{1}} = n$, and so we can add this to both sides to find
$$n^{\underline{3}} + 3n^{\underline{2}} + n^{\underline{1}} = n^3.$$
This tells us that
$$\begin{aligned} \Delta n^3 &= \Delta\left(n^{\underline{3}} + 3n^{\underline{2}} + n^{\underline{1}}\right) \\ &= 3n^{\underline{2}} + 6n^{\underline{1}} + 1 \\ &= 3(n^2 - n) + 6n + 1 \\ &= 3n^2 + 3n + 1, \end{aligned}$$
which matches the observation about $\Delta(n^3)$ in the above section.

(Return)

47 Discrete Calculus

The previous two modules have laid the groundwork for the discretization, or digitization, of calculus:

1. The discrete version of a function is a sequence.
2. The discrete version of the derivative is the difference operator.
3. The discrete version of the integral is the sum.
4. The discrete version of a differential equation is a recurrence relation.

This module mostly deals with #3, the integral of a discrete function. For continuous functions, there was the Fundamental Theorem of Integral Calculus which made computing integrals easy under certain conditions. Essentially, it said that the integral of the derivative is the function itself, evaluated at the endpoints. The discrete version says the same thing:

The Discrete Fundamental Theorem of Integral Calculus (FTIC)

Given a sequence u,

$$\sum_{n=A}^{B} \Delta u = u\Big|_{n=A}^{B+1},$$

and

$$\sum_{n=A}^{B} \nabla u = u\Big|_{n=A-1}^{B}$$

(See Answer 1)

Example

Note that

$$\Delta \frac{1}{n} = \frac{1}{n+1} - \frac{1}{n} = \frac{-1}{n^2+n}.$$

Use this, along with FTIC, to find $\displaystyle\sum_{n=A}^{B} \frac{1}{n^2+n}$. (See Answer 2)

Example

Use the fact that

$$\Delta n! = (n+1)! - n! = (n+1)n! - n! = (n+1-1)n! = n \cdot n!,$$

along with the FTIC to find

$$\sum_{n=A}^{B} n!n$$

(See Answer 3)

Example

Let F denote the Fibonacci sequence defined by

$$F_{n+2} = F_{n+1} + F_n; \quad F_0 = 0, F_1 = 1.$$

Note that

$$\Delta(F_{n+1}) = F_{n+2} - F_{n+1} = F_n,$$

by rearranging the above recursion relation. Use this, along with the FTIC, to find

$$\sum_{n=1}^{k} F_n.$$

(See Answer 4)

Power rule for falling powers

Recall from the previous module that the falling power

$$n^{\underline{k}} = n(n-1)(n-2)\cdots(n-k+1)$$

has a nice power rule for the difference:

$$\Delta(n^{\underline{k}}) = kn^{\underline{k-1}}.$$

By running this in reverse, we can find the anti-difference (or antiderivative) of the falling power, which is very similar to the power rule for integration:

$$\Delta^{-1}(n^{\underline{k}}) = \frac{1}{k+1}n^{\underline{k+1}} + C,$$

where C is a constant. Using this, along with the FTIC, allows one to find closed formulas for the sum of polynomials.

Example

Using the power rule and the FTIC, find

$$\sum_{n=1}^{k} n.$$

(See Answer 5)

Example

Find

$$\sum_{n=1}^{k} n^2.$$

(See Answer 6)

Example

With a little algebra (shown in the previous module), one finds that

$$n^3 = n^{\underline{3}} + 3n^{\underline{2}} + n^{\underline{1}}.$$

Use this fact, along with FTIC, to find

$$\sum_{n=1}^{k} n^3$$

(See Answer 7)

47.1 Integration by parts

There is also a discrete version of integration by parts. First, the following product rule can be established for the forward difference:

$$\Delta(uv) = u\Delta v + Ev\Delta u.$$

(See Details 8) Then, just like for continuous functions, one can integrate both sides (i.e. sum both sides), rearrange, and apply FTIC to find

$$\sum_{n=A}^{B} u\Delta v = uv\Big|_{n=A}^{B+1} - \sum_{n=A}^{B} Ev\Delta u.$$

Example

Find $\displaystyle\sum_{n=0}^{k} n2^n$. (See Answer 9)

47.2 Differential equations

The discrete version of a differential equation is a recurrence relation. This is an equation relating one term of a sequence u with one or more previous terms in the sequence. In this context, the shift operator acts like the derivative.

Example

Consider the linear first-order recurrence relation

$$u_{n+1} = \lambda u_n,$$

which can be written as

$$Eu = \lambda u$$

or, with some rearrangement,

$$(E - \lambda I)u = 0.$$

By inspection, one can see the solution $u_n = C\lambda^n$, where $C = u_0$ is some constant (it can be thought of as an initial condition).

Example

Now, consider using the following linear first-order difference equation

$$\Delta u = \lambda u.$$

This is reminiscent of the differential equation $\frac{dx}{dt} = ax$, considered earlier in the course. Recall that $\Delta = E - I$, and so this difference relation can be written

$$(E - I)u = \lambda u.$$

With a little more rearranging one finds

$$(E - (\lambda + 1)I)u = 0,$$

which is almost identical to the equation from the previous example, but with λ replaced by $\lambda + 1$. Therefore, the solution is $u_n = C(\lambda + 1)^n$.

Compare this to the solution in the continuous case, which was

$$\frac{dx}{dt} = \lambda x \quad \Longrightarrow \quad x = Ce^{\lambda t}.$$

When $\lambda = 1$, we have the solution

$$x = Ce^t.$$

If we let $\lambda = 1$ in the discrete difference equation, we have

$$u = C(\lambda + 1)^n = C \cdot 2^n,$$

which again shows that 2^n is the discrete version of the exponential function e^t.

47.3 Fibonacci numbers revisited

We can take the recurrence relation for the Fibonacci numbers:

$$F_{n+2} = F_{n+1} + F_n$$

and rearrange it to be rewritten in terms of the shift operator:

$$\begin{aligned} F_{n+2} - F_{n+1} - F_n &= 0 \\ (E^2 - E - I)F &= 0. \end{aligned}$$

Note that the solutions to the equation

$$x^2 - x - 1 = 0$$

are

$$\varphi = \frac{1}{2}(1 + \sqrt{5}) \quad \text{and} \quad \psi = \frac{1}{2}(1 - \sqrt{5}).$$

So we can factor the operator $E^2 - E - I$ to see that

$$\begin{aligned}(E^2 - E - I)F &= 0\\ (E - \varphi I)(E - \psi I)F &= 0.\end{aligned}$$

It is a fact that solutions to this equation are combinations (that is, a sum) of solutions to

$$\begin{aligned}(E - \varphi I)F1 &= 0\\ (E - \psi I)F2 &= 0.\end{aligned}$$

The solutions to these equations are

$$\begin{aligned}F1_n &= c_1\varphi^n\\ F2_n &= c_2\psi^n.\end{aligned}$$

So we have that

$$F_n = c_1\varphi^n + c_2\psi^n,$$

for some constants c_1 and c_2, which depend on the initial conditions of the sequence. To find those constants, we can plug in convenient values of n and solve.

Letting $n = 0$, we have

$$F_0 = 0 = c_1 + c_2 \quad \Longrightarrow \quad c_2 = -c_1.$$

Letting $n = 1$, we have

$$\begin{aligned}F_1 = 1 &= c_1\varphi + c_2\psi\\ &= c_1\varphi - c_1\psi\\ &= c_1(\varphi - \psi)\\ &= c_1 \cdot \sqrt{5}.\end{aligned}$$

So we find

$$\begin{aligned}c_1 &= \frac{1}{\sqrt{5}}\\ c_2 &= -\frac{1}{\sqrt{5}}.\end{aligned}$$

This gives an explicit, closed form equation for the Fibonacci numbers:

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n.$$

47.4 EXERCISES

- (a) What is $\Delta(\frac{1}{n^2})$? (b) Using part (a), find $\sum_{n=A}^{B} \frac{-2n-1}{n^2(n+1)^2}$.

47.5 Answers to Selected Examples

1. To see why this holds, evaluate the sum and carefully note the cancellation. This type of sum is called a *telescoping* sum. First, for the forward difference operator, one finds

$$\begin{aligned}\sum_{n=A}^{B} \Delta u &= \sum_{n=A}^{B} u_{n+1} - u_n \\ &= (u_{A+1} - u_A) + (u_{A+2} - u_{A+1}) + (u_{A+3} - u_{A+2}) + \cdots + u_{B+1} - u_B \\ &= -u_A + u_{A+1} - u_{A+1} + u_{A+2} - u_{A+2} + u_{A+3} + \cdots - u_B + u_{B+1} \\ &= -u_A + u_{B+1},\end{aligned}$$

   as desired. Similarly, with the backwards difference operator, one finds

$$\begin{aligned}\sum_{n=A}^{B} \nabla u &= \sum_{n=A}^{B} u_n - u_{n-1} \\ &= u_A - u_{A-1} + u_{A+1} - u_A + u_{A+2} - u_{A+1} + \cdots + u_B - u_{B-1} \\ &= -u_{A-1} + u_B,\end{aligned}$$

   as claimed.

   (Return)

2. By FTIC,

$$\begin{aligned}\sum_{n=A}^{B} \frac{1}{n^2+n} &= -\sum_{n=A}^{B} \Delta \frac{1}{n} \\ &= -\left(\frac{1}{n}\bigg|_{n=A}^{B+1} \right) \\ &= \frac{1}{A} - \frac{1}{B+1} \\ &= \frac{B-A+1}{A(B+1)}\end{aligned}$$

   (Return)

3. By the fact above, we have

$$\begin{aligned}\sum_{n=A}^{B} n!n &= \sum_{n=A}^{B} \Delta n! \\ &= n!\bigg|_{n=A}^{B+1} \\ &= (B+1)! - A!.\end{aligned}$$

   (Return)

4. By the above fact, we have

$$\begin{aligned}
\sum_{n=1}^{k} F_n &= \sum_{n=1}^{k} \Delta(F_{n+1}) \\
&= F_{n+1}\Big|_{n=1}^{k+1} \\
&= F_{k+2} - F_2 \\
&= F_{k+2} - 1.
\end{aligned}$$

(Return)

5. Note that $n = n^{\underline{1}}$, and so

$$\begin{aligned}
\sum_{n=1}^{k} n &= \sum_{n=1}^{k} n^{\underline{1}} \\
&= \frac{1}{2} n^{\underline{2}}\Big|_{n=1}^{k+1} \\
&= \frac{1}{2} n(n-1)\Big|_{n=1}^{k+1} \\
&= \frac{k(k+1)}{2}.
\end{aligned}$$

(Return)

6. Note that

$$n^2 = n^{\underline{2}} + n^{\underline{1}}.$$

Therefore,

$$\begin{aligned}
\sum_{n=1}^{k} n^2 &= \sum_{n=1}^{k} n^{\underline{2}} + n^{\underline{1}} \\
&= \frac{1}{3} n^{\underline{3}} + \frac{1}{2} n^{\underline{2}}\Big|_{n=1}^{k+1} \\
&= \frac{(k+1)\,k(k-1)}{3} + \frac{(k+1)k}{2} \\
&= (k+1)k\left(\frac{k-1}{3} + \frac{1}{2}\right) \\
&= (k+1)k\left(\frac{2k-2+3}{6}\right) \\
&= \frac{k(k+1)(2k+1)}{6}.
\end{aligned}$$

(Return)

7. Using the above fact and FTIC, one finds

$$\begin{aligned}
\sum_{n=1}^{k} n^3 &= \sum_{n=1}^{k} \left(n^{\underline{3}} + 3n^{\underline{2}} + n^{\underline{1}}\right) \\
&= \frac{1}{4} n^{\underline{4}} + n^{\underline{3}} + \frac{1}{2} n^{\underline{2}}\Big|_{n=1}^{k+1}.
\end{aligned}$$

Now, note that $n^{\underline{m}} = 0$ when $n = 1$ for all $m > 1$, because of the factor of $n - 1$. Therefore, the evaluation at the bottom limit is 0. Continuing with the algebra, we find

$$\begin{aligned}
\sum_{n=1}^{k} n^3 &= \frac{1}{4}(k+1)k(k-1)(k-2) + (k+1)k(k-1) + \frac{1}{2}(k+1)k \\
&= k(k+1)\left(\frac{1}{4}(k-1)(k-2) + k - 1 + \frac{1}{2}\right) \\
&= k(k+1)\left(\frac{(k-1)(k-2) + 4(k-1) + 2}{4}\right) \\
&= \frac{k(k+1)}{4}\left(k^2 - 3k + 2 + 4k - 4 + 2\right) \\
&= \frac{k(k+1)}{4}k(k+1) \\
&= \left(\frac{k(k+1)}{2}\right)^2 .
\end{aligned}$$

Recalling that $1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}$, this shows the remarkable fact that

$$1^3 + 2^3 + 3^3 + \cdots + k^3 = (1 + 2 + 3 + \cdots + k)^2 .$$

(Return)

8. By subtracting and adding a common term and rearranging, one finds

$$\begin{aligned}
(\Delta(uv))_n &= u_{n+1}v_{n+1} - u_n v_n \\
&= u_{n+1}v_{n+1} - u_n v_{n+1} + u_n v_{n+1} - u_n v_n \\
&= (u_{n+1} - u_n)v_{n+1} + u_n(v_{n+1} - v_n) \\
&= \left[\Delta u E v + u \Delta v\right]_n .
\end{aligned}$$

(Return)

9. Let $u = n$ and $\Delta v = 2^n$. It follows that $\Delta u = 1$, $v = 2^n$, and $Ev = 2^{n+1}$. Thus, by the integration by parts formula,

$$\begin{aligned}
\sum_{n=0}^{k} n2^n &= n2^n \Big|_{n=0}^{k+1} - \sum_{n=0}^{k} 2^{n+1} \\
&= (k+1)2^{k+1} - (2^{k+2} - 2) \\
&= (k+1)2^{k+1} - 2 \cdot 2^{k+1} + 2 \\
&= (k-1)2^{k+1} + 2.
\end{aligned}$$

(Return)

48 Numerical ODEs

The previous few modules have discretized functions, derivatives, and integrals. This module shows how discrete methods can be used to approximate solutions to problems in the continuous realm. One such application was Newton's method for approximating roots of functions, seen back in Linear Approximations. This module deals with approximating solutions of continuous differential equations.

In certain situations, differential equations can be solved exactly. For example, a separable differential equation

$$\frac{dx}{dt} = f(x)g(t)$$

is solved by rearranging and integrating, as seen in Antidifferentiation. A linear first order differential equation

$$\frac{dx}{dt} = f(t)x + g(t)$$

is solved by an application of the product rule, as seen in More differential equations. However, there are many differential equations of the form

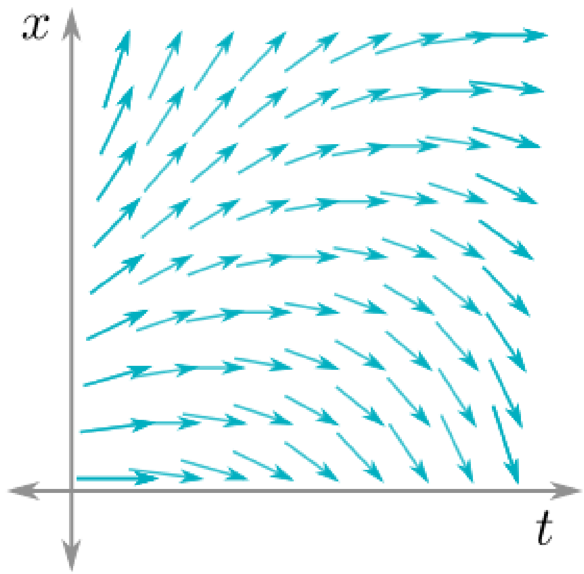
$$\frac{dx}{dt} = f(x, t)$$

which cannot be solved exactly by the above methods. This is the situation where techniques known as *numerical ordinary differential equations* can be used to approximate a solution. There are three methods covered in this module:

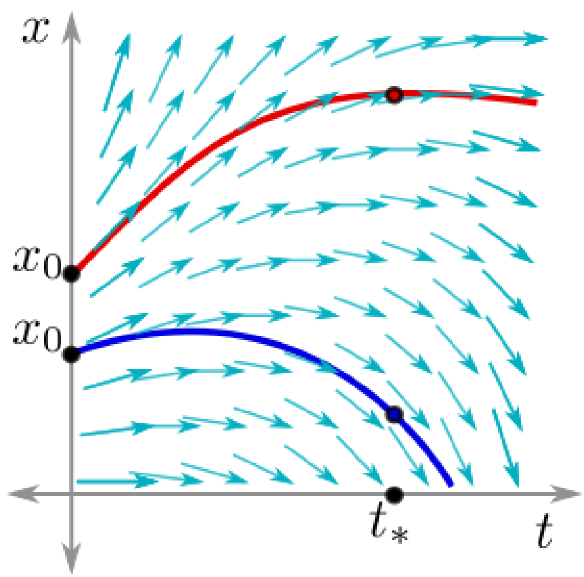
1. Euler's method
2. Midpoint method
3. Runge-Kutta method

48.1 Euler's method

Euler's method uses a difference equation to approximate the solution of an initial value problem. More specifically, given the differential equation $\frac{dx}{dt} = f(x, t)$, and initial value $x_0 = x(t_0)$, Euler's method approximates $x(t_*)$ for some $t_* > t_0$.



One way to visualize the problem is to imagine a river where the water is going in different directions at different locations. If one drops something that floats (perhaps a very small rock, or a duck) at different locations in the river, where would the object end up further down the river?



To compute the approximation, first a positive integer N is chosen, and the t axis is split into N intervals, giving the sequence of time points $t = (t_0, t_1, \ldots, t_N)$, where $t_N = t_*$. The time step is $h = \frac{t_* - t_0}{N}$. Then there is a corresponding sequence $x = (x_0, x_1, \ldots, x_N)$ where $x_0 = x(t_0)$ is the initial condition and each subsequent x_n is an approximation of $x(t_n)$, given by the update rule

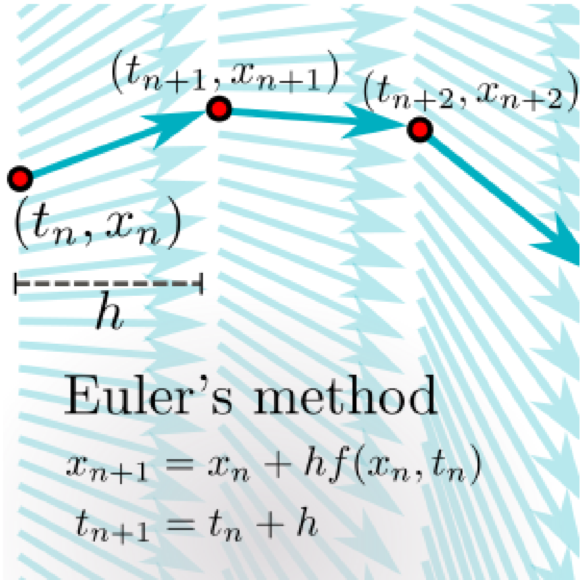
$$
\begin{aligned}
x_{n+1} &= x_n + hf(x_n, t_n) \\
t_{n+1} &= t_n + h.
\end{aligned}
$$

This recurrence is sensible by considering the discretization of the original differential equation:

$$\begin{aligned} \frac{dx}{dt} = f(x,t) \quad &\Rightarrow \quad \frac{\Delta x}{\Delta t} = f(x,t) \\ &\Rightarrow \quad \frac{x_{n+1} - x_n}{h} = f(x_n, t_n), \end{aligned}$$

and then rearranging.

This process is using a linearization at each point (t_n, x_n) to get to the next point (t_{n+1}, x_{n+1}), then this is repeated to get the next point, and so on:



Remember that a linearization is only good when the change in input, h in this case, is small. Therefore, a bigger value of N gives a better approximation, but requires more computation.

Example

Let $\frac{dx}{dt} = x$ with $x_0 = x(0) = 1$. Use Euler's method with N left as a variable to approximate $x(1)$. What happens as $N \to \infty$? (See Answer 1)

Example

Consider the differential equation

$$\frac{dx}{dt} = t + x^2.$$

Use Euler's method to estimate (t_*, x_*), where $t_* = 1$, the initial conditions are $t_0 = 0$ and $x_0 = 1$, and the step size is $h = \frac{1}{2}$. (See Answer 2)

Taylor series perspective

If we think of $x = x(t)$ as a function of t, and expand the Taylor series of x about $t = t_0$, we have

$$\begin{aligned} x(t_0 + h) &= x(t_0) + h \cdot \left.\frac{dx}{dt}\right|_{t=t_0} + O(h^2) \\ &= x_0 + h \cdot f(x_0, t_0) + O(h^2), \end{aligned}$$

so Euler's method can be seen as taking the first order Taylor approximation and using it to form the recursion relation. The above equation shows that the error for a single step is in $O(h^2)$, so the error over all N steps is

$$\begin{aligned} \text{Error} &= N \cdot O(h^2) \\ &= \frac{t_* - t_0}{h} \cdot O(h^2) \\ &= O(h). \end{aligned}$$

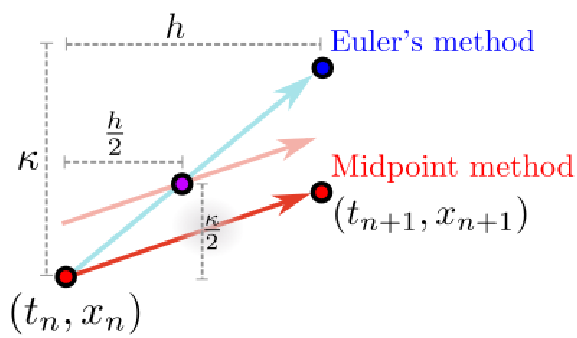
48.2 Midpoint method

Another method for solving the differential equation

$$\frac{dx}{dt} = f(x, y)$$

is known as the *midpoint method*. There is still an update rule which is similar to the one used in Euler's rule, but the function f is evaluated at a different point. The idea is to consider the point where Euler's rule would have taken us, and find the midpoint of that with our starting point. Use that midpoint as the point to evaluate f in the update rule.

As described, it is a little bit complicated, but with some extra notation and a diagram it becomes clearer. Let $\kappa = h \cdot f(x_n, t_n)$. So κ is the quantity which would be added to x_n in the update rule for Euler's rule:



Then the eponymous midpoint which is used in the update rule is

$$\left(t_n + \frac{h}{2}, x_n + \frac{\kappa}{2}\right).$$

So the update rule for the midpoint method is

$$\begin{aligned} x_{n+1} &= x_n + h \cdot f\left(x_n + \frac{\kappa}{2}, t_n + \frac{h}{2}\right) \\ \kappa &= h \cdot f(x_n, t_n). \end{aligned}$$

The midpoint method is a bit more complicated than Euler's method, but it has a benefit. The midpoint method is a second order approximation, and as a result the error turns out to be $O(h^3)$ for an individual step, and hence $O(h^2)$ for the full approximation process. Therefore, it is a more accurate method of approximation.

48.3 Runge-Kutta method

The final method for solving the above differential equation is called the *Runge-Kutta method.* It is a fourth order approximation. Its error for an individual step is $O(h^5)$ and for the whole process is $O(h^4)$. This makes it the most accurate model of the three in this module. It is difficult to describe in an intuitive manner other than to say it is a sort of average of Euler's method, the midpoint method, and other methods. The update rule is as follows:

$$\begin{aligned}
x_{n+1} &= x_n + \frac{1}{6}\left(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4\right) \\
\kappa_1 &= h \cdot f(x_n, t_n) \\
\kappa_2 &= h \cdot f\left(x_n + \frac{\kappa_1}{2}, t_n + \frac{h}{2}\right) \\
\kappa_3 &= h \cdot f\left(x_n + \frac{\kappa_2}{2}, t_n + \frac{h}{2}\right) \\
\kappa_4 &= h \cdot f\left(x_n + \kappa_3, t_n + h\right).
\end{aligned}$$

Note that κ_1 is the increase used in Euler's method, and κ_2 is the increase used in the Midpoint method.

48.4 Comparison of methods

We already know that Euler's method is the most basic, then the Midpoint method, and finally Runge-Kutta. We should expect, then, that Runge-Kutta should give the best approximation, followed by the Midpoint method, followed by Euler's method.

Example

Use each of the three methods to approximate the solution of

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

with initial conditions

$$t_0 = 0 \qquad x_0 = 1$$
$$t_* = 1 \qquad x_* = e$$

(we already know that $x_* = e$ because we have solved this differential equation several times already), using a step size of $h = 1$. (See Answer 3)

48.5 EXERCISES

- Given $\frac{dx}{dt} = x^3 + x^2 t$ with initial condition $t_0 = 0, x_0 = 1$, approximate $x(2)$ using the midpoint method with size step $h = 1$.

48.6 Answers to Selected Examples

1. Note that

$$\begin{aligned} x_{n+1} &= x_n + \frac{x_n}{N} \\ &= x_n(1 + \frac{1}{N}) \end{aligned}$$

(this is independent of t). Iterating this gives $x_N = \left(1 + \frac{1}{N}\right)^N \approx x(1)$. As N increases, the approximation gets better and better, and one finds that

$$\lim_{N\to\infty} x_N = \lim_{N\to\infty} \left(1 + \frac{1}{N}\right)^N.$$

Call this limit L and take the ln of both sides. Then we find that

$$\begin{aligned} \ln L &= \ln\left(\lim_{N\to\infty}\left(1 + \frac{1}{N}\right)^N\right) \\ &= \lim_{N\to\infty} N \ln\left(1 + \frac{1}{N}\right) \\ &= \lim_{N\to\infty} N\left(\frac{1}{N} + O\left(\frac{1}{N^2}\right)\right) \\ &= \lim_{N\to\infty} 1 + O\left(\frac{1}{N}\right) \\ &= 1. \end{aligned}$$

Therefore, $L = e$. So $x(1) = e$, which matches the value from solving the problem by separation of variables.

(Return)

2. Here, we have

$$f(x,t) = t + x^2.$$

So, using the initial conditions, we have

$$f(x_0, t_0) = 0 + 1 = 1.$$

Then we have that

$$\begin{aligned} x_1 &= x_0 + h \cdot f(x_0, t_0) \\ &= 1 + \frac{1}{2}\cdot 1 \\ &= \frac{3}{2}. \end{aligned}$$

And $t_1 = \frac{1}{2}$. So

$$\begin{aligned} f(x_1, t_1) &= t_1 + x_1^2 \\ &= \frac{1}{2} + \left(\frac{3}{2}\right)^2 \\ &= \frac{11}{4}. \end{aligned}$$

Thus,

$$\begin{aligned}
x_2 &= x_1 + h \cdot f(x_1, t_1) \\
&= \frac{3}{2} + \frac{1}{2} \cdot \frac{11}{4} \\
&= \frac{23}{8}
\end{aligned}$$

And thus we have $x_* = \frac{23}{8}$.

(Return)

3. It may seem ill advised to use a step size equal to the distance from the start time to the end time, but this is just for the sake of comparison. Note that in this example,

$$f(x, t) = x.$$

Euler's method gives

$$\begin{aligned}
x_1 &= x_0 + h \cdot f(x_0, t_0) \\
&= 1 + 1 \cdot 1 \\
&= 1 + 1,
\end{aligned}$$

which is a pretty rough estimate of e. From this computation we have that $\kappa = 1$ for the midpoint method.

The **Midpoint method** gives

$$\begin{aligned}
x_1 &= x_0 + h \cdot f\left(x_0 + \frac{\kappa}{2}, t_0 + \frac{h}{2}\right) \\
&= 1 + 1 \cdot f\left(1 + \frac{1}{2}, 0 + \frac{1}{2}\right) \\
&= 1 + 1 \cdot \left(1 + \frac{1}{2}\right) \\
&= 1 + 1 + \frac{1}{2},
\end{aligned}$$

so we see that we have gotten another term closer to e.

For **Runge-Kutta**, we need to compute several different κs. From the computations we did in Euler's method and the Midpoint method, we have that

$$\begin{aligned}
\kappa_1 &= 1 \\
\kappa_2 &= \frac{3}{2}.
\end{aligned}$$

Computing the other values, we have

$$\begin{aligned}
\kappa_3 &= h \cdot f\left(x_0 + \frac{\kappa_2}{2}, t_0 + \frac{h}{2}\right) \\
&= 1 \cdot f\left(1 + \frac{3}{4}, 0 + \frac{1}{2}\right) \\
&= 1 + \frac{3}{4} \\
&= \frac{7}{4}.
\end{aligned}$$

And

$$\begin{aligned}\kappa_4 &= h \cdot f\left(x_0 + \kappa_3, t_0 + h\right) \\ &= 1 \cdot f\left(1 + \frac{7}{4}, 0 + 1\right) \\ &= 1 + \frac{7}{4} \\ &= \frac{11}{4}.\end{aligned}$$

Putting these together (and rearranging the fractions a little bit), we have that

$$\begin{aligned}x_1 &= x_0 + \frac{1}{6}(\kappa_1 + 2\kappa_2 + 2\kappa_3 + \kappa_4) \\ &= 1 + \frac{1}{6}\left(1 + 2 \cdot \frac{3}{2} + 2 \cdot \frac{7}{4} + \frac{11}{4}\right) \\ &= 1 + \frac{1}{6} + \frac{1}{2} + \frac{7}{12} + \frac{11}{24} \\ &= 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24},\end{aligned}$$

which has two more terms of e, making it a very good approximation, despite only having one step.
(Return)

49 Numerical Integration

The topic of this module is the discretization of the definite integral. How does one approximate the definite integral of a function which does not have an easily computable anti-derivative? The answer is with finite sums, which is the discrete analog of the definite integral.

Recall that one interpretation for the definite integral is area under the curve. The goal is to find a finite sum which approximates the area under the curve. There are three common techniques for making this approximation: Riemann sums, trapezoid rule, and Simpson's rule. Each gives an approximation of the integral $\int_a^b f(x)dx$.

Another way to think about this problem (and the practical applications) is to consider some sporadic sample points, thought of as a sequence

$$x = (x_n) = (x_0, x_1, \ldots, x_n).$$

Corresponding to these sample points is the sequence of function values

$$f = (f_n) = (f_0, f_1, \ldots, f_n),$$

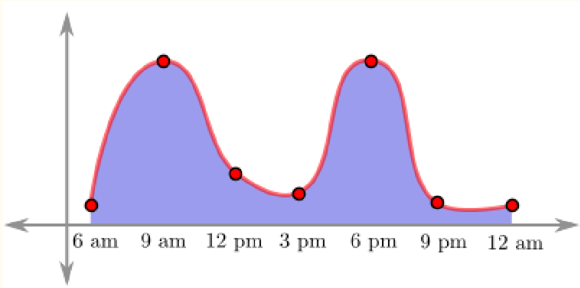
where $f_i = f(x_i)$. The goal is to approximate the definite integral of the underlying function f using these function values. Of course, in real world applications the function may not be continuous, let alone a familiar function with an easily computed anti-derivative. The method of numerical integration gives an approximation of the definite integral in this situation with imperfect information.

Example

Consider the problem of estimating the number of people who pass through a certain busy subway station each day. One could sit in the station for the entire day and count every person, or one could get an estimate by going in periodically and counting the number of people who come in over a short time.
This gives a sampling of the rate of passengers entering the station at different times:

Time	6 am	9 am	12 pm	3 pm	6 pm	9 pm	12 am
Passengers per hour	120	1200	360	240	1200	180	120

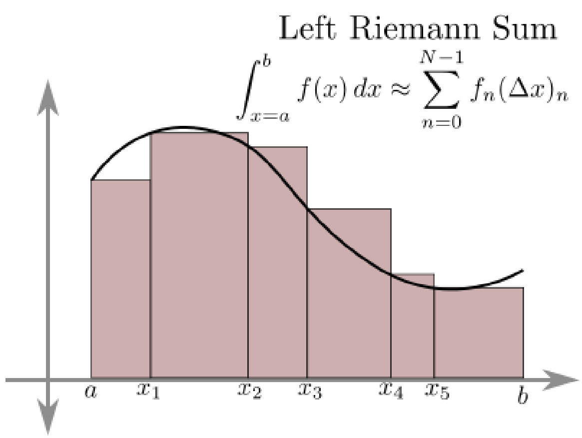
The definite integral of the rate of passenger arrival gives the total number of passengers using the station that day:



49.1 Riemann Sums

The most rudimentary approximation is given by *Riemann sums*, which should be familiar from the definition of the definite integral.

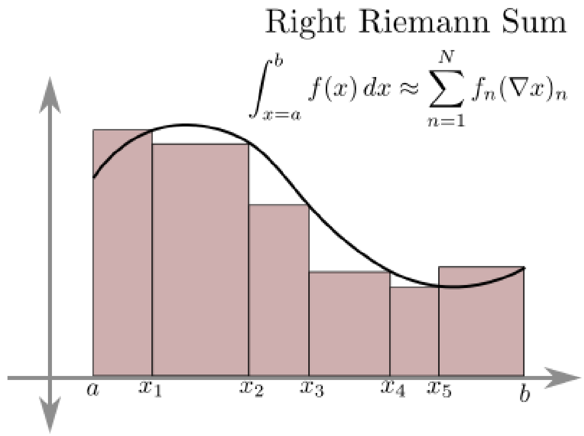
The *left Riemann sum* uses the left endpoint of the ith subinterval as the sample point to compute the height of the the ith rectangle:



Thus, the area of the nth rectangle is. So the *left Riemann sum* is given by

$$\int_a^b f(x)dx \approx \sum_{n=0}^{N-1} f_n \cdot (x_{n+1} - x_n)$$

The *right Riemann sum* uses the right endpoint of the ith subinterval to compute the height of the ith rectangle:



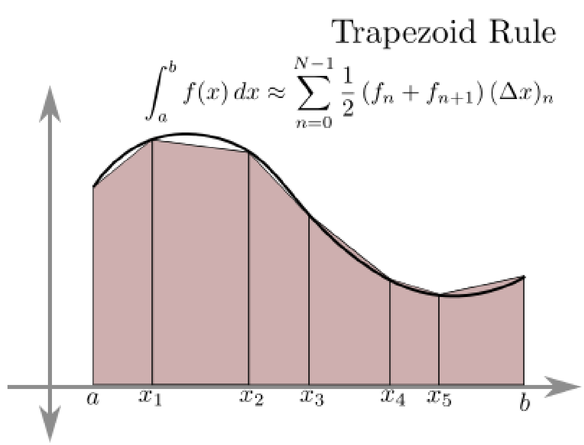
In this case, the area of the nth rectangle is $f_n \cdot (\nabla x)_n$. So the *right Riemann sum* is given by

$$\int_a^b f(x)dx \approx \sum_{n=1}^{N} f_n \cdot (x_n - x_{n-1})$$

An improvement on the left and right Riemann sums, called the trapezoid rule, is given in the next section.

49.2 Trapezoid rule

The *trapezoid rule* uses trapezoids instead of rectangles to approximate the area above each subinterval:



Recall that the area of a trapezoid with bases (parallel segments) of length p and q, with height h has area $\frac{1}{2}h(p+q)$. The area of the nth trapezoid, then, is $\frac{1}{2}(\Delta x)_n(f_n + f_{n+1})$. Thus, the trapezoid rule gives the approximation:

$$\int_a^b f(x)dx \approx \sum_{n=0}^{N-1} \frac{1}{2}(\Delta x)_n(f_n + f_{n+1})$$
$$\approx \sum_{n=0}^{N-1} \frac{1}{2}(f_n + f_{n+1})(x_{n+1} - x_n).$$

If the sample points are evenly spaced uniformly, then the formula for the trapezoid rule simply becomes

$$\int_a^b f(x)dx \approx h\left(\frac{1}{2}(f_0 + f_N) + \sum_{n=1}^{N-1} f_n\right),$$

where $h = \frac{b-a}{N}$ is the width of each trapezoid. Note that the trapezoid rule is the average of the left and right Riemann sums.

49.3 Simpson's rule

Think about how the previous approximations interpolate the function f. The Riemann sum is a piece-wise constant approximation (also called a step function). The trapezoid rule is a piece-wise linear approximation. The logical next step is to use piece-wise quadratic approximations. That is how Simpson's rule works. Another way to think of it is that Simpson's rule will compute the area under a parabola exactly (whereas Riemann sums and trapezoid rule will have errors in general).

One way that Simpson's rule differs from the above rules is that the sample points must be evenly spaced. Let $h = \frac{b-a}{N}$ be the distance between sample points. The *Simpson's rule* approximation is given by

$$\int_a^b f(x)dx \approx \frac{h}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + \ldots + 2f_{N-2} + 4f_{N-1} + f_N\right).$$

Here, N must be even. (See Derivation 1)

Example

Approximate the definite integral $\int_0^2 x^3 dx$ with $N = 4$ using (uniformly spaced) right and left Riemann sums; trapezoid rule; and Simpson's rule. Here is a table of pertinent values to make the computation easier:

n	x_n	f_n
0	0	0
1	.5	.125
2	1	1
3	1.5	3.375
4	2	8

(See Answer 2)

49.4 Errors bounds

With any approximation, it is good to get some idea of how far off the approximation is from the true value. Let E_T be the error using the trapezoidal rule, and E_S be the error from using Simpson's rule.

Using advanced calculus beyond the scope of this course, one can bound these errors as follows.

$$E_T \leq \frac{M(b-a)^3}{12N^2}$$

where M is the maximum value of $|f''(x)|$ on the interval $[a, b]$.

$$E_S \leq \frac{M(b-a)^5}{180N^4}$$

where M is the maximum value of $|f^{(4)}(x)|$ on the interval $[a, b]$.

49.5 EXERCISES

- Using the fact that $\sum_{n=0}^{j} n^2 = \frac{j(j+1)(2j+1)}{6}$, approximate $\int_0^2 x^2 dx$ using the right Riemann sum with N number of intervals.

49.6 Answers to Selected Examples

1. **Note** This derivation is a little bit different from the one in lecture, and perhaps more elementary.

   The idea of Simpson's rule is to fit a parabola to the first three points (x_0, f_0), (x_1, f_1), (x_2, f_2), and then find the area under that parabola. Then, fit a parabola to the next three points (overlapping the endpoints) (x_2, f_2), (x_3, f_3), (x_4, f_4), find the area under that parabola, and so on.

   Consider the parabola $g(x) = ax^2 + bx + c$ determined by the points $(-h, f_{n-1}), (0, f_n), (h, f_{n+1})$. Solving the resulting system of equations

$$\begin{aligned} f_{n-1} &= ah^2 - bh + c \\ f_n &= c \\ f_{n+1} &= ah^2 + bh + c, \end{aligned}$$

   gives the parabola which fits these points. One finds that

$$\begin{aligned} a &= \frac{1}{h^2}\left(\frac{f_{n-1}}{2} - f_n + \frac{f_{n+1}}{2}\right) \\ b &= \frac{f_{n+1} - f_{n-1}}{2h} \\ c &= f_n. \end{aligned}$$

   And so the area

$$\begin{aligned} \int_{x=-h}^{h} g(x)dx &= \int_{x=-h}^{h} ax^2 + bx + c\,dx \\ &= a\frac{x^3}{3} + b\frac{x^2}{2} + cx\Big|_{x=-h}^{h} \\ &= 2a\frac{h^3}{3} + 2ch, \end{aligned}$$

by using symmetry. Plugging in the earlier values for a and c and simplifying gives

$$\int_{x=-h}^{h} g(x)dx = \frac{h}{3}\left(f_{n-1} + 4f_n + f_{n+1}\right).$$

Carrying this out for all each triplet of points gives the approximation

$$\begin{aligned}\int_{x=a}^{b} f(x)dx &\approx \frac{h}{3}\left(f_0 + 4f_1 + f_2 + f_2 + 4f_3 + f_4 + \cdots + f_N\right)\\ &\approx \frac{h}{3}\left(f_0 + 4f_1 + 2f_2 + 4f_3 + \cdots + f_N\right),\end{aligned}$$

as desired.

(Return)

2. Let L, R, T, and S be the respective approximations. Then

$$\begin{aligned}L &= \frac{1}{2}(0 + .125 + 1 + 3.375) = 2.25\\ R &= \frac{1}{2}(.125 + 1 + 3.375 + 8) = 6.25\\ T &= \frac{1}{2}(0 + .125 + 1 + 3.375 + 4) = 4.25\\ S &= \frac{1}{6}(0 + .5 + 2 + 13.5 + 8) = 4.\end{aligned}$$

Note that the true value is

$$\int_{x=0}^{2} x^3\,dx = \frac{1}{4}x^4\Big|_{x=0}^{2} = 4,$$

so Simpson's rule gets the answer exactly and the trapezoid rule is the next closest.

(Return)

50 Series

The previous module discussed finite sums as the discrete analog of definite integrals with finite bounds. Then, logically, the discrete analog of improper integrals with infinite bounds should be infinite sums, referred to as *infinite series* or just *series* when there is no confusion.

50.1 Definition of infinite series

Recall that when computing definite integrals with bounds at infinity, one replaces the infinite bound with a variable and then takes the limit:

$$\int_{x=1}^{\infty} f(x)dx = \lim_{T\to\infty} \int_{x=1}^{T} f(x)dx.$$

For infinite series, the definition is analogous:

$$\sum_{n=1}^{\infty} a_n = \lim_{T\to\infty} \sum_{n=1}^{T} a_n.$$

The expression $\sum_{n=1}^{T} a_n$ is called the Tth *partial sum* of the series $\sum_{n=1}^{\infty} a_n$. Then a series *converges* if the sequence of partial sums converges. If the sequence of partial sums does not converge, we say the series *diverges*.

Example

Give the first few terms of the sequence of partial sums for the infinite series

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

(See Answer 1)

50.2 Taylor series revisited

Recall some of the Taylor series from earlier modules:

$$\begin{aligned}
e^x &= 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots \\
\sin(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \\
\cos(x) &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots \\
\frac{1}{1-x} &= 1 + x + x^2 + x^3 + \dots && (|x| < 1) \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \dots && (|x| < 1)
\end{aligned}$$

These provide many examples of series which not only converge, but can be evaluated exactly.

Example

Compute $\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \frac{1}{81} + \dots$. (See Answer 2)

Example

Compute $1 - \frac{\pi^2}{2!} + \frac{\pi^4}{4!} - \frac{\pi^6}{6!} + \dots$. (See Answer 3)

50.3 Classifying series

There are some series which cannot be evaluated exactly, though it is known that the series converges. For example, the series

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n^3} &= 1 + \frac{1}{8} + \frac{1}{27} + \frac{1}{64} + \dots \\
&\approx 1.2
\end{aligned}$$

converges, but it is not known what the exact value is (though one can calculate as many digits as one likes). This is in contrast with an apparently similar series,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n^2} &= 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots \\
&= \frac{\pi^2}{6},
\end{aligned}$$

for which an exact value is known (though the proof of this value is beyond the scope of this course).

Yet another similar series, called the *harmonic series*,

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \dots,$$

diverges, as will be shown in the next module.

There are two questions then. First, does a series converge or not? Second, if it does converge, to what does it converge? This course deals mostly with the first question, in this module and the next few modules. More advanced analysis classes can help answer the second question.

When intuition fails

Determining the convergence of a series using intuition can be dangerous. As one example of how intuition can fail, consider what happens when we plug $x = 1$ into the series for $\ln(1 + x)$:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots .$$

Although $x = 1$ is not in the radius of convergence for this series, it turns out that this series still converges. Now, consider what happens if we multiply both sides by $\frac{1}{2}$:

$$\frac{1}{2}\ln 2 = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \cdots .$$

Adding this equation together with the above one, some of the terms cancel and some combine to show that

$$\frac{3}{2}\ln 2 = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots$$

Now, notice that this series has all the terms of the series for $\ln 2$, but in slightly different order. And yet, the series evaluates to a different value. That is certainly counter intuitive. Perhaps even more alarming, the terms of the series for $\ln 2$ can be rearranged so that the resulting series evaluates to *any* real number. The takeaway is that we must tread carefully and not trust intuition but instead rely on logic.

50.4 The nth term test for divergence

The first, and usually simplest, test *for divergence* of a series is the nth term test.

The nth term test for divergence

If $\lim_{n\to\infty} a_n \neq 0$, then the series $\sum_{n=0}^{\infty} a_n$ diverges.

(See Proof 4) This test applies to all series, and it is easy to apply (just take the limit of the terms of the series). However, many of the series encountered in this course will have terms which go to 0 in the limit, in which case the test is inconclusive (see the caveat below).

Example

Show that the series

$$\sum_{n=0}^{\infty} \frac{4^n - 3^n}{4^n + 2^n}$$

diverges. (See Answer 5)

Example

Show that the series

$$\sum_{n=0}^{\infty}(-1)^n = 1 - 1 + 1 - 1 + 1 - \dots$$

diverges. (See Answer 6)

Example

Show that the series

$$\sum_{n=1}^{\infty}\cos\left(\frac{1}{n}\right) = \cos(1) + \cos(1/2) + \cos(1/3) + \dots$$

diverges. (See Answer 7)

Caveat

This is not a test for convergence! In particular, if $\lim_{n\to\infty} a_n = 0$, then the test is inconclusive (the series might converge or diverge). If the test is inconclusive, one of the other tests from the upcoming modules must be used.

In logical terms, this says that the *converse* of the nth term test does not hold. On the other hand, the *contrapositive* does hold:

$$\sum_{n=0}^{\infty} a_n \text{ converges} \quad \Rightarrow \quad \lim_{n\to\infty} a_n = 0$$

(The contrapositive of a true statement is always true, but the converse is not always true).

Example

What does the nth term test say about the series $\sum_{n=2}^{\infty}\frac{\ln(n)}{n}$? (See Answer 8)

Example

What does the nth term test say about the series

$$\sum_{n=1}^{\infty}\arctan n?$$

(See Answer 9)

Example

Suppose the series

$$\sum_{n=1}^{\infty} a_n = L,$$

converges, and that $a_n > 0$ for all n. What, if anything, can be said about the convergence of the series

$$\sum_{n=1}^{\infty} \frac{1}{a_n}?$$

(See Answer 10)

50.5 EXERCISES

- Determine whether the following series converges or diverges

$$\sum_{n=2}^{\infty} \frac{n^3}{n \ln(n^{100})}$$

$$\sum_{n=1}^{\infty} \frac{2^n}{n^{100}}$$

50.6 Answers to Selected Examples

1. Adding the first term, then the first two terms, then the first three terms, and so on, gives

$$\frac{1}{2}, \frac{1}{2} + \frac{1}{4}, \frac{1}{2} + \frac{1}{4} + \frac{1}{8}, \ldots$$

which becomes

$$\frac{1}{2}, \frac{3}{4}, \frac{7}{8}, \frac{15}{16}, \ldots .$$

This sequence appears to be converging to 1, which is consistent with what we know this series to be (by the geometric series).

(Return)

2. Note that by the geometric series,

$$\begin{aligned} \frac{1}{1-1/3} &= 1 + \frac{1}{3} + \frac{1}{9} + \ldots \\ &= \frac{3}{2}. \end{aligned}$$

So $\frac{1}{3} + \frac{1}{9} + \ldots = \frac{1}{2}$.

(Return)

3. Note that this is the Taylor series for $\cos(x)$ with $x = \pi$. Thus, the series evaluates to $\cos(\pi) = -1$.

(Return)

4. We give a proof by contrapositive. That is, we prove that if $\sum_{n=0}^{\infty} a_n$ converges, then $\lim_{n\to\infty} a_n = 0$.

Assume $\sum_{n=0}^{\infty} a_n$ converges. Then by the definition of convergence, we know that the sequence of partial sums s_T defined by

$$s_T = \sum_{n=1}^{T} a_n$$

converges also. Therefore,

$$\lim_{T\to\infty} s_T = \lim_{T\to\infty} s_{T-1} = L$$

for some L. Then we have

$$\lim_{T\to\infty} (s_T - s_{T-1}) = 0,$$

by linearity of the limit. But notice that

$$\begin{aligned} s_T - s_{T-1} &= \sum_{n=1}^{T} a_n - \sum_{n=1}^{T-1} a_n \\ &= (a_1 + a_2 + \cdots + a_{T-1} + a_T) - (a_1 + a_2 + \cdots + a_{T-1}) \\ &= a_T. \end{aligned}$$

Putting this together with the above limit, we find

$$\begin{aligned} \lim_{T\to\infty} (s_T - s_{T-1}) &= \lim_{T\to\infty} a_T \\ &= 0, \end{aligned}$$

which is what we were trying to prove.

(Return)

5. Here,

$$a_n = \frac{4^n - 3^n}{4^n + 2^n}.$$

Applying the nth term test, we find

$$\begin{aligned} \lim_{n\to\infty} a_n &= \lim_{n\to\infty} \frac{4^n - 3^n}{4^n + 2^n} \\ &= \lim_{n\to\infty} \frac{4^n(1-(3/4)^n)}{4^n(1+(1/2)^n)} \\ &= \lim_{n\to\infty} \frac{1-(3/4)^n}{1+(1/2)^n} \\ &= 1 \neq 0, \end{aligned}$$

because both $\left(\frac{3}{4}\right)^n \to 0$ and $\left(\frac{1}{2}\right)^n \to 0$. Therefore, by the nth term test, the series diverges.

(Return)

6. Since $\lim_{n\to\infty}(-1)^n$ does not exist (the sequence oscillates), the series diverges by the nth term test.

(Return)

7. Note that $\lim_{n\to\infty} \cos\left(\frac{1}{n}\right) = \cos(0) = 1 \neq 0$. Thus, by the nth term test, the series diverges.

(Return)

8. Note that $\lim_{n\to\infty} \frac{\ln(n)}{n} = 0$, either by using l'Hospital's rule or by recalling that $\ln(n)$ grows much more slowly than n. Thus, the nth term test is **inconclusive**.

   (Return)

9. Note that
$$\lim_{n\to\infty} \arctan n = \frac{\pi}{2} \neq 0,$$
   and so by the nth term test, the series **diverges**.

   (Return)

10. By the contrapositive of the nth term test mentioned above, we know that since the series
$$\sum_{n=1}^{\infty} a_n$$
    converges, we know that
$$\lim_{n\to\infty} a_n = 0.$$
    But this implies that
$$\lim_{n\to\infty} \frac{1}{a_n} = \infty$$
    so by the nth term test, we know that the series
$$\sum_{n=1}^{\infty} \frac{1}{a_n}$$
    diverges.

    (Return)

51 Convergence Tests 1

Unlike the nth term test for divergence from the last module, this module gives several tests which, if successfully applied, give a definitive answer of whether a series converges or not. A common feature of the tests in this module is that they use a comparison.

51.1 Integral test for convergence and divergence

This is a test which can definitively tell whether a series converges or diverges. However, it can be harder to apply.

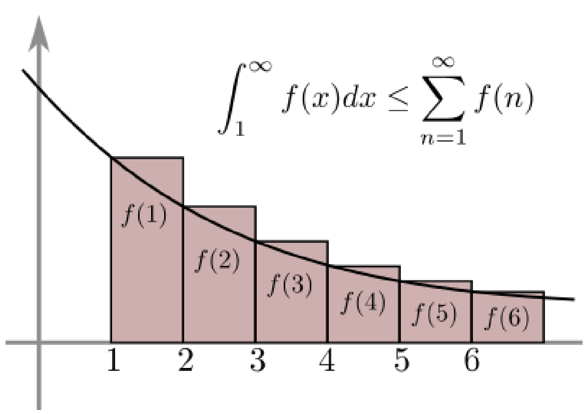
Integral test for convergence and divergence

If $f(x)$ is a positive, decreasing function, then for any integer m,

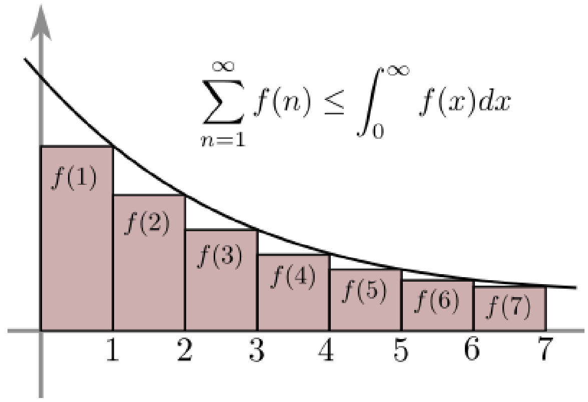
$$\sum_{n=m}^{\infty} f(n) \text{ converges } \iff \int_m^{\infty} f(x)dx \text{ converges.}$$

The double arrow $\iff$ means *if and only if*. So the series and integral either both converge, or they both diverge.

To see why, visualize the series as a sum of rectangles with base 1 and height $f(n)$. If these rectangles are drawn to the right of the curve $f(x)$, then the result is the following figure. Each rectangle is labeled with its area. The combined area of the rectangles completely contains the area under $f(x)$, which establishes the inequality shown:



On the other hand, if the rectangles are drawn to the left of the curve $f(x)$, then all the rectangles lie below the curve. Their combined area is less than the area under the curve $f(x)$, which establishes the inequality shown:



Combining these two inequalities gives

$$\int_1^{\infty} f(x)dx \leq \sum_{n=1}^{\infty} f(n) \leq \int_0^{\infty} f(x)dx.$$

If the integral diverges, then the series diverges (by the first inequality). And if the integral converges, then the series converges (by the second inequality). This establishes the integral test.

Example

Use the integral test to determine if

$$\sum_{n=2}^{\infty} \frac{1}{n \ln(n)}$$

converges or diverges. (See Answer 1)

Example

Use the integral test to determine if

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

converges or diverges. (See Answer 2)

51.2 The p-series test

The next example is important enough that it gets its own name: *the p-series*. This makes use of the p-integrals that we computed earlier.

Example

Find the values of p for which the series

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

converges. (See Answer 3)

The p-series test

The p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$.

Harmonic series

The series $\sum \frac{1}{n}$, known as the *harmonic series*, diverges by to the p-series test. This is a significant fact to keep in mind because the harmonic series diverges even though the terms of the series go to 0. Thus, the harmonic series is a demonstration that the nth term test is a test for *divergence* only and *cannot* be used to show a series converges.

Note that the harmonic series is a sort of boundary between convergence and divergence. The series $\sum \frac{1}{n^{.999}}$ diverges, but the series $\sum \frac{1}{n^{1.001}}$ converges.

Example

The p-series test proves the convergence of two examples from the last module: $\sum \frac{1}{n^2}$ and $\sum \frac{1}{n^3}$.
Example
Determine the values of p for which the series

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^p}$$

converges. (See Answer 4)

51.3 Comparison test

The integral test compared a series to its related integral. This test compares one series to another.

Comparison test

Let a_n and b_n be positive sequences such that $b_n > a_n$ for all n. It follows that

1. if $\sum b_n$ converges, then $\sum a_n$ converges.
2. if $\sum a_n$ diverges, then $\sum b_n$ diverges.

In other words, if a series is smaller than a convergent series, then it converges too. If a series is bigger than a divergent series, then it diverges too.

Caveat

It is critical that the inequality be in the correct direction. A series which is larger than a convergent series might converge or diverge. A series which is smaller than a divergent series might converge or diverge.

Example

Show that $\sum \frac{\ln(n)}{n}$ diverges. (See Answer 5)

Example

Show that

$$\sum_{n=4} \frac{n^3 - 2n^2 - 10}{n^5 + 7}$$

converges. (See Answer 6)

Example

Determine whether

$$\sum_{n=2}^{\infty} \frac{1}{\ln(n!)}$$

converges or diverges. Hint: try for a rough upper bound or lower bound on $n!$ and see which one gives the right comparison. (See Answer 7)

51.4 Limit test

The final test of this module is a slightly different type of comparison. Recall that when comparing two functions f and g to see which is "bigger" asymptotically, one computes the limit

$$\lim_{n\to\infty} \frac{f(n)}{g(n)}.$$

If this limit is infinite, then f is bigger. If the limit is 0, then g is bigger. If the limit is L where $0 < L < \infty$, then the two functions are roughly equal (up to a constant multiple). It is this third case that is used for this test (sometimes called the *Limit comparison test*):

Limit test

Let a_n and b_n be positive series. If $\lim_{n\to\infty} \frac{a_n}{b_n} = L$ and $0 < L < \infty$, then the series $\sum a_n$ and $\sum b_n$ either both converge or both diverge.

The key to the limit test is finding a suitable sequence b_n which is approximately equal to a_n in the limit. Often, a_n will be a ratio, in which case the lower order terms in the numerator and denominator can be dropped, and what is left will be b_n. Ideally, it will be easy to see if $\sum b_n$ converges or diverges. Finally, one must check that $0 < \lim_{n\to\infty} \frac{a_n}{b_n} < \infty$.

Example

Show that

$$\sum \frac{n^2 - n}{n^3 + 7}$$

diverges. (See Answer 8)

Example

Determine whether

$$\sum \sin\left(\frac{1}{n}\right)$$

converges or diverges. (See Answer 9)

51.5 EXERCISES

- Determine whether the following series converges or diverges

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

$$\sum_{n=1}^{\infty} \frac{|\sin(n)^n|}{n^2}$$

51.6 Answers to Selected Examples

1. Computing, one finds (using the u-substitution $u = \ln(x)$) that

$$\int_2^{\infty} \frac{1}{x\ln(x)}\,dx = \int_{\ln(2)}^{\infty} \frac{1}{u}\,du$$
$$= \ln(u)\Big|_{\ln(2)}^{\infty},$$

which diverges since $\ln(u) \to \infty$ as $u \to \infty$. Since the integral diverges, the series also diverges by the integral test.

(Return)

2. The integral test says the series converges if and only if

$$\int_{x=1}^{\infty} \frac{x}{e^x}\,dx = \int_{x=1}^{\infty} xe^{-x}\,dx$$

converges. We know this integral converges (recall that xe^{-x} is the PDF for the exponential distribution). But we can also compute it again.

This integral is a good candidate for integration by parts, with

$$u = x \qquad du = dx$$
$$dv = e^{-x}\,dx \qquad v = -e^{-x}.$$

Thus,

$$\begin{aligned}
\int_{x=1}^{\infty} xe^{-x} &= -xe^{-x}\Big|_{x=1}^{\infty} - \int_{x=1}^{\infty} -e^{-x}\,dx \\
&= -xe^{-x} - e^{-x}\Big|_{x=1}^{\infty} \\
&= 0 - (-e^{-1} - e^{-1}) \\
&= \frac{2}{e}.
\end{aligned}$$

Since the integral converges, the series converges too, by the integral test.

(Return)

3. We know from the integral test that

$$\sum_{n=1}^{\infty} frac1n^p$$

converges if and only if

$$\int_{x=1}^{\infty} \frac{1}{x^p}\,dx$$

converges. But this integral converges if and only if $p > 1$, as we saw in the module on p-integrals. Therefore, the p-series converges if and only if $p > 1$.

(Return)

4. Making the substitution

$$\begin{aligned}
u &= \ln x \\
du &= \frac{1}{x}\,dx,
\end{aligned}$$

we see that

$$\int_{x=2}^{\infty} \frac{1}{x(\ln x)^p}\,dx = \int_{u=\ln 2}^{\infty} \frac{1}{u^p}\,du.$$

This integral (again from our knowledge of the p-integral) converges if and only if $p > 1$. Therefore, by the integral test, the series

$$\sum_{n=2}^{\infty} \frac{1}{n(\ln n)^p}$$

converges if and only if $p > 1$.

(Return)

5. Note that $\frac{\ln(n)}{n} > \frac{1}{n}$. Since $\sum \frac{1}{n}$ diverges, the series $\sum \frac{\ln(n)}{n}$ diverges too, by the comparison test.

(Return)

6. Note that

$$\frac{n^3 - 2n^2 - 10}{n^5 + 7} < \frac{n^3}{n^5} = \frac{1}{n^2},$$

since the numerator on the left is smaller, and the denominator on the left is bigger. So

$$\sum_{n=4}^{\infty} \frac{n^3 - 2n^2 - 10}{n^5 + 7} < \sum_{n=4}^{\infty} \frac{1}{n^2}.$$

$\sum \frac{1}{n^2}$ converges (p-series test from above), and so

$$\sum_{n=4}^{\infty} \frac{n^3 - 2n^2 - 10}{n^5 + 7}$$

converges as well, by the comparison test.

(Return)

7. A lower bound for $n!$ might be 2^n (or any exponential). This gives

$$\begin{aligned}\sum \frac{1}{\ln(n!)} &< \sum \frac{1}{\ln(2^n)} \\ &= \sum \frac{1}{n \ln(2)} \\ &= \frac{1}{\ln 2} \sum \frac{1}{n},\end{aligned}$$

which diverges, since it is a constant multiple of the harmonic series. This comparison does not go in the right direction, since our original series is *smaller* than a divergent series. Thus, we should try going in the other direction to find an upper bound for $n!$.

A rough upper bound for $n!$ is n^n. This gives

$$\begin{aligned}\sum \frac{1}{\ln(n!)} &> \sum \frac{1}{\ln(n^n)} \\ &= \sum \frac{1}{n \ln(n)}.\end{aligned}$$

This series diverges (see the example earlier in this module). Therefore, the original series, which is *bigger* than a divergent series, also diverges.

(Return)

8. By dropping the lower order terms in the numerator and denominator, one finds

$$\frac{n^2 - n}{n^3 + 7} \approx \frac{n^2}{n^3} = \frac{1}{n}.$$

So $b_n = \frac{1}{n}$ is a good choice. Assuming the above approximation is not too rough, the work is done since the harmonic series, $\sum \frac{1}{n}$, diverges.

To make sure the approximation is not too rough, compute the limit

$$\begin{aligned}\lim_{n\to\infty} \frac{a_n}{b_n} &= \lim_{n\to\infty} \frac{\left(n^2 - n\right) / \left(n^3 + 7\right)}{1/n} \\ &= \lim_{n\to\infty} \frac{n(n^2 - n)}{n^3 + 7} \\ &= \lim_{n\to\infty} \frac{n^3 - n^2}{n^3 + 7} \\ &= 1.\end{aligned}$$

Thus, by the limit test, $\sum \frac{n^2-n}{n^3+7}$ and $\sum \frac{1}{n}$ either both converge or both diverge. Since $\sum \frac{1}{n}$ diverges, so too must $\sum \frac{n^2-n}{n^3+7}$.

(Return)

9. With an unusual series like this, the nth term test is a good first thing to try. But it is inconclusive since $\sin(1/n) \to 0$ as $n \to \infty$.

   To get a handle on how this function acts, note that when n is large, $1/n$ is small, and so we can use the Taylor series for $\sin x$ about 0:

$$\sin\left(\frac{1}{n}\right) = \frac{1}{n} + O\left(\frac{1}{n^3}\right).$$

   Thus, $\sin(1/n) \approx \frac{1}{n}$, which means $\frac{1}{n}$ is a good candidate for b_n in the limit test. This will work:

$$\begin{aligned}\lim_{n\to\infty} \frac{a_n}{b_n} &= \lim_{n\to\infty} \frac{\sin(1/n)}{1/n} \\ &= \lim_{n\to\infty} \frac{1/n + O(1/n^3)}{1/n} \\ &= 1.\end{aligned}$$

   Since $\sum b_n = \sum \frac{1}{n}$ diverges (harmonic series), it follows by the limit test that $\sum \sin(1/n)$ diverges also.

   (Return)

52 Convergence Tests 2

This module deals with the root test and the ratio test for convergence. Unlike the tests from the previous module, the tests of this module can be applied without having to find a good series to compare.

Recall that the geometric series

$$1 + x + x^2 + x^3 + \cdots$$

converges to $\frac{1}{1-x}$ provided that $|x| < 1$. This fact is at the heart of both the root test and the ratio test.

52.1 Root test

The Root test

Given a series $\sum a_n$, with all $a_n > 0$, let

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

If $\rho < 1$, then $\sum a_n$ converges. If $\rho > 1$, then $\sum a_n$ diverges. Finally, if $\rho = 1$, then the test is inconclusive.

(See Justification 1) The root test works best when the term a_n involves an nth power, or can be expressed as an nth power, although it can be used in other situations as well.

Example

Determine if

$$\sum \left(\frac{n}{2n+1}\right)^n$$

converges or diverges. (See Answer 2)

Example

Determine if

$$\sum \left(\frac{n}{n+1}\right)^{n^2}$$

converges or diverges. (See Answer 3)

Example

Use the root test on the p-series

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

(See Answer 4)

52.2 Ratio test

The Ratio Test

Given a series $\sum a_n$, with all $a_n > 0$, let

$$\rho = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}.$$

If $\rho < 1$, then the series $\sum a_n$ converges. If $\rho > 1$, then the series $\sum a_n$ diverges. If $\rho = 1$, then the test is inconclusive (the series might converge or diverge).

(See Justification 5) The ratio test works best when a_n involves exponential functions and factorials, since in these situations there is a lot of cancellation. It does not work well with ratios of polynomials, because the test is inconclusive.

Example

Determine if the series

$$\sum \frac{n}{3^n}$$

converges or diverges. (See Answer 6)

Example

Determine if the series

$$\sum \frac{n!}{(2n)!}$$

converges or diverges. (See Answer 7)

Example

Show that the ratio test is inconclusive on the p-series $\sum \frac{1}{n^p}$. (See Answer 8)

Example

Determine if the series

$$\sum_{n=0}^{\infty} \frac{4^n}{(2n)!}$$

converges or diverges. (See Answer 9)

52.3 Summary of methods for a positive series

There is no foolproof method for determining the convergence or divergence of a series. However, here is a rough guide for tests to try. Given a series $\sum a_n$, where $a_n > 0$ for all n:

1. Do the terms go to 0? If $\lim_{n\to\infty} a_n \neq 0$, then by the nth term test, the series diverges. (If $\lim_{n\to\infty} a_n = 0$, then the test is inconclusive).
2. Does a_n involve exponential functions like c^n where c is constant? Does a_n involve factorial? Then the ratio test should be used.
3. Is a_n of the form $(b_n)^n$ for some sequence b_n? Then use the root test.
4. Does ignoring lower order terms make a_n look like a familiar series (e.g. p-series or geometric series)? Then use the comparison test or the limit test.
5. Does the sequence a_n look easy to integrate? Then use the integral test.

52.4 EXERCISES

- Determine whether the following series converges or diverges

$$\sum_{n=2}^{\infty} \frac{1}{\ln^n(n)}$$

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

- What does the ratio test say about the convergence of

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

52.5 Answers to Selected Examples

1. Recall what it means that $\lim \sqrt[n]{a_n} = \rho$. It means that for a given $\epsilon > 0$, there exists an M such that for all $n > M$,

$$|\sqrt[n]{a_n} - \rho| < \epsilon.$$

In other words, for n sufficiently big, $\sqrt[n]{a_n} \approx \rho$, and so $a_n \approx \rho^n$. This says that, roughly speaking, the series is eventually geometric with common ratio ρ. Hence, the series converges for $\rho < 1$ and diverges for $\rho > 1$.

(Return)

2. Computing, one finds that

$$\begin{aligned}\rho &= \lim_{n\to\infty} \sqrt[n]{a_n} \\ &= \lim_{n\to\infty} \sqrt[n]{\left(\frac{n}{2n+1}\right)^n} \\ &= \lim_{n\to\infty} \frac{n}{2n+1} \\ &= \frac{1}{2}.\end{aligned}$$

Since $\rho < 1$, the series converges by the root test.

(Return)

3. Computing, one finds that

$$\begin{aligned}\rho &= \lim_{n\to\infty} \sqrt[n]{\left(\frac{n}{n+1}\right)^{n^2}} \\ &= \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n \\ &= \lim_{n\to\infty} \frac{1}{(1+1/n)^n} \\ &= \frac{1}{e}.\end{aligned}$$

The last step above follows from the fact that $\lim_{n\to\infty}(1+1/n)^n = e$. Thus, $\rho < 1$ and so the series converges by the root test.

(Return)

4. Trying the root test gives

$$\begin{aligned}\rho &= \lim_{n\to\infty} \sqrt[n]{\frac{1}{n^p}} \\ &= \lim_{n\to\infty} \left(\frac{1}{n^{1/n}}\right)^p \\ &= 1,\end{aligned}$$

and so the root test is **inconclusive**.

The last step above follows from the fact that $\lim_{n\to\infty} n^{1/n} = 1$. Let $y = \lim_{n\to\infty} n^{1/n}$. Taking the natural log of both sides gives

$$\begin{aligned}\ln(y) &= \lim_{n\to\infty} \frac{1}{n}\ln(n) \\ &= \lim_{n\to\infty} \frac{\ln(n)}{n} \\ &= 0.\end{aligned}$$

Thus $y = e^0 = 1$, as desired.

(Return)

5. As in the justification for the root test, $\frac{a_{n+1}}{a_n}$ is eventually very close to ρ, and remains close to ρ ever after. Roughly speaking, then $a_{n+1} \approx \rho a_n$ for all sufficiently big n. But that means that after a while, the series becomes roughly geometric with common ratio ρ. Therefore, when $\rho < 1$ the series converges, and when $\rho > 1$ the series diverges.

   (Return)

6. Here, $a_n = \frac{n}{3^n}$. Computing, one finds

$$\begin{aligned} \rho &= \lim_{n\to\infty} \frac{a_{n+1}}{a_n} \\ &= \lim_{n\to\infty} \frac{(n+1) / \left(3^{n+1}\right)}{n/3^n} \\ &= \lim_{n\to\infty} \frac{n+1}{3^{n+1}} \cdot \frac{3^n}{n} \\ &= \lim_{n\to\infty} \frac{1}{3} \frac{n+1}{n} \\ &= \frac{1}{3}. \end{aligned}$$

   $\rho < 1$, and so by the ratio test, the series converges.

   (Return)

7. In this example, $a_n = \frac{n!}{(2n)!}$. Computing gives

$$\begin{aligned} \rho &= \lim_{n\to\infty} \frac{a_{n+1}}{a_n} \\ &= \lim_{n\to\infty} \frac{(n+1)!}{[2(n+1)]!} \cdot \frac{(2n)!}{n!} \\ &= \lim_{n\to\infty} \frac{n+1}{(2n+2)(2n+1)} \\ &= \lim_{n\to\infty} \frac{n+1}{4n^2+6n+2} \\ &= 0. \end{aligned}$$

   So $\rho < 1$, and the series converges by the ratio test.

   (Return)

8. Computing shows that

$$\begin{aligned} \rho &= \lim_{n\to\infty} \frac{1/(n+1)^p}{1/n^p} \\ &= \lim_{n\to\infty} \frac{n^p}{(n+1)^p} \\ &= \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^p \\ &= 1, \end{aligned}$$

   and so the ratio test is inconclusive.

   (Return)

9. There is a quick, tricky way to see that this converges, which is to note that

$$\sum_{n=0}^{\infty} \frac{4^n}{(2n)!} = \sum_{n=0}^{\infty} \frac{2^{2n}}{(2n)!},$$

which is the Taylor series for $\cosh x$ with $x = 2$. We know cosh converges everywhere, so we know this series converges.

Alternatively, we can use the ratio test. We have $a_n = \frac{4^n}{(2n)!}$. Therefore

$$\begin{aligned}
\rho &= \lim_{n\to\infty} \frac{a_{n+1}}{a_n} \\
&= \lim_{n\to\infty} \frac{4^{n+1}}{(2(n+1))!} \cdot \frac{(2n)!}{4^n} \\
&= \lim_{n\to\infty} \frac{4 \cdot 4^n}{(2n+2)\cdot(2n+1)\cdot(2n)!} \cdot \frac{(2n)!}{4^n} \\
&= \lim_{n\to\infty} \frac{4}{(2n+2)\cdot(2n+1)} \\
&= 0.
\end{aligned}$$

Since $\rho < 1$, this series converges by the ratio test, confirming what we already knew.

(Return)

53 Absolute And Conditional

Up until now, the convergence tests covered by this course have only covered series with positive terms. What happens when a series has some positive and some negative terms? This module describes some tools for determining the convergence or divergence of such a series.

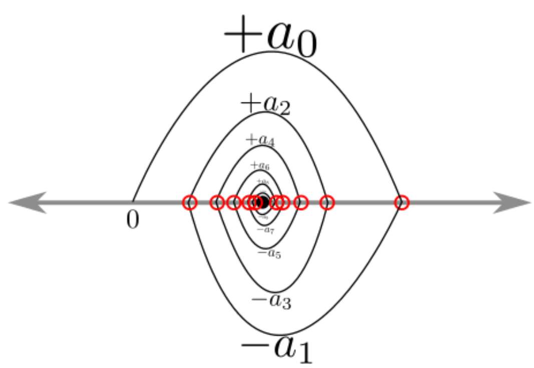
53.1 Alternating series test

One particular type of series is fairly simple to test for convergence. A series $\sum a_n$ is *alternating* if it is of the form $\sum(-1)^n b_n$, where b_n is a positive sequence. In other words, a series is alternating if its terms are alternately positive and negative.

Alternating Series Test

An alternating series $\sum(-1)^n b_n$ with decreasing terms converges if and only if $\lim_{n\to\infty} b_n = 0$.

The intuition behind this test is that if the terms are alternating, decreasing, and go to zero, then the partial sums of the series gradually zero in on the true value. The first partial sum is greater than the true value, the second partial sum is less than the true value, and so on:



Example

The series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

is alternating. Its terms are decreasing, and the terms go to zero, so by the alternating series test, the series converges.

Example

Determine if

$$\sum_{n=0}^{\infty}(-1)^n\frac{1}{2^n}$$

converges. (See Answer 1)

Example

Determine if

$$\sum_{n=2}^{\infty}(-1)^n\frac{\ln(n)}{n}$$

converges. (See Answer 2)

53.2 Conditional and absolute convergence

Consider the two previous examples. Note that both $\sum(-1)^n\frac{1}{2^n}$ and $\sum(-1)^n\frac{\ln(n)}{n}$ converge. However, if these series were not alternating, then $\sum\frac{1}{2^n}$ still converges (it is geometric), whereas $\sum\frac{\ln(n)}{n}$ does not converge (comparison with the harmonic series $\sum\frac{1}{n}$).

These series demonstrate a distinction between two types of convergence.

Absolute and Conditional Convergence

The series $\sum a_n$ *converges absolutely* if $\sum|a_n|$ converges.
A series $\sum a_n$ *converges conditionally* if $\sum a_n$ converges, but $\sum|a_n|$ diverges.
In other words, a series is conditionally convergent if it is convergent but not absolutely convergent.

Example

Using this terminology, the series

$$\sum_{n=0}^{\infty}(-1)^n\frac{1}{2^n}$$

converges absolutely, but

$$\sum_{n=2}^{\infty}(-1)^n\frac{\ln(n)}{n}$$

converges conditionally.

Example

Find the values of p for which the alternating p-series

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

diverges, converges conditionally, and converges absolutely. (See Answer 3)

Example

Determine if

$$\sum_{n=1}^{\infty}(-1)^n \frac{e^n}{n^5}$$

converges, and if so, whether it is conditional or absolute convergence. (See Answer 4)

Example

Determine if

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

converges, and if so, whether it is conditional or absolute convergence. (See Answer 5)

53.3 Absolute convergence test

Some series are not strictly alternating, but have some positive and some negative terms, sporadically. In this situation, it can be difficult to determine whether the series converges directly, but the following test sometimes makes the determination easier.

Absolute Convergence Test

If the series $\sum |a_n|$ converges, then $\sum a_n$ converges. In other words, if a series converges absolutely, then the series converges.

(See Justification 6)

Example

Determine if

$$\sum_{n=1}^{\infty} \frac{\sin(n)}{n^2}$$

converges or diverges. (See Answer 7)

53.4 EXERCISES

- Determine whether the following series converges or not. If it converges, is it conditionally convergent or absolutely convergent?

$$\sum_{n=2}^{\infty}(-1)^n \frac{1}{n\ln(n)}$$

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

- Determine whether the following series is convergent or divergent

$$\sum_{n=1}^{\infty}\cos(n)\frac{\ln(n)}{n!}$$

53.5 Answers to Selected Exercises

1. The series is alternating, the terms are decreasing, and $\lim_{n\to\infty}\frac{1}{2^n} = 0$. Thus, by the alternating series test, the series converges.

   Alternatively, we can observe that this series is geometric:

$$\begin{aligned}\sum_{n=0}^{\infty}(-1)^n\frac{1}{2^n} &= \sum_{n=0}^{\infty}\left(-\frac{1}{2}\right)^n \\ &= \frac{1}{1-(-1/2)} \\ &= \frac{2}{3}.\end{aligned}$$

   (Return)

2. The series is alternating, terms are decreasing, and

$$\lim_{n\to\infty}\frac{\ln(n)}{n} = 0.$$

   And so by the alternating series test, the series converges.

   (Return)

3. When $p \leq 0$, the terms of the series do not go to 0. That is,

$$\lim_{n\to\infty}\frac{1}{n^p} \neq 0.$$

   So by the nth term test for divergence, the series diverges for these values of p.

   When $0 < p \leq 1$, the terms of the series are decreasing, alternating, and going to 0. So by the alternating series test, the series converges for this range of p. However, if absolute values are taken, then the resulting series, $\sum \frac{1}{n^p}$, does not converge by the p-series test. Thus, for $0When1$ (Return)

4. The series is alternating, but notice that the terms do not go to 0, since $e^n > n^5$. Therefore, by the nth term test for divergence, this series diverges.

   (Return)

5. The series is alternating, and note that $n^n > n!$, which leads us to believe that the series converges. In fact, the series converges absolutely, i.e.

$$\sum \frac{n!}{n^n}$$

   converges, as the ratio test shows, along with some careful algebra:

$$\begin{aligned}
\rho &= \lim_{n\to\infty} \frac{(n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{n!} \\
&= \lim_{n\to\infty} \frac{(n+1)n^n}{(n+1)^{n+1}} \\
&= \lim_{n\to\infty} \frac{n^n}{(n+1)^n} \\
&= \lim_{n\to\infty} \left(\frac{n}{n+1}\right)^n \\
&= \lim_{n\to\infty} \frac{1}{\left(\frac{n+1}{n}\right)^n} \\
&= \lim_{n\to\infty} \frac{1}{(1+1/n)^n} \\
&= \frac{1}{e} < 1.
\end{aligned}$$

   So $\rho < 1$, and so the series converges absolutely.

   (Return)

6. Consider the series $\sum(a_n + |a_n|)$. Note that

$$0 \le a_n + |a_n| \le 2|a_n|$$

   So $\sum(a_n + |a_n|) \le \sum 2|a_n|$, and by the comparison test we find $\sum(a_n + |a_n|)$ converges. Then

$$\begin{aligned}
\sum a_n &= \sum((a_n + |a_n|) - |a_n|) \\
&= \sum(a_n + |a_n|) - \sum |a_n|
\end{aligned}$$

   converges too, being the difference of two convergent, positive series.

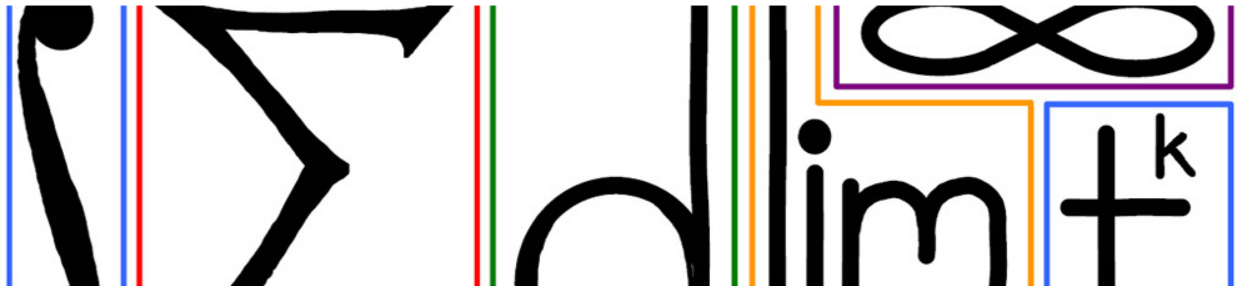
   (Return)

7. $\sin(n)$ is a messy function because it is sometimes positive and sometimes negative, but not in a simple alternating pattern. However, one nice thing about $\sin(n)$ is that, in absolute value, it is bounded by 1. So

$$\sum_{n=1}^{\infty} \left|\frac{\sin(n)}{n^2}\right| \le \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

   Since $\sum \frac{1}{n^2}$ converges by p-series test, it follows by the comparison test that $\sum \left|\frac{\sin(n)}{n^2}\right|$ converges. Hence $\sum \frac{\sin(n)}{n^2}$ converges absolutely, and so it converges by the absolute convergence test.

   (Return)



54 Power Series

Sequences can be thought of as the discretization of a function. This module goes in the opposite direction: turning a sequence into a function called a power series.

54.1 Power series

Given a sequence a_n, the *power series* associated with a_n is the infinite sum $f(x) = \sum_{n=0}^{\infty} a_n x^n$.

Example

The power series associated with the sequence $a_n = 1$ is the function $f(x) = 1 + x + x^2 + x^3 + \ldots$.

Example

All of the Taylor series encountered earlier in the course are power series. For instance, the Taylor series for the exponential is the power series associated with the sequence $a_n = \frac{1}{n!}$.

Example

The *Lucas numbers* L_n are like the Fibonacci numbers but with different initial conditions. $L_0 = 2$ and $L_1 = 1$, and $L_n = L_{n-1} + L_{n-2}$. So the sequence begins $L = (2, 1, 3, 4, 7, 11, 18, \ldots)$.
Find the closed-form function $L(x)$ given by the power series

$$\begin{aligned} L(x) &= L_0 + L_1 x + L_2 x^2 + L_3 x^3 + \cdots \\ &= \sum_{n=0}^{\infty} L_n x^n. \end{aligned}$$

This is known in the field of combinatorics as the *generating function* for the sequence L_n. (See Answer 1)

These power series have lots of useful applications in enumeration and asymptotic analysis, among other things. But the rest of this module will deal with convergence. Given a sequence (a_n), for what values of x does its associated power series $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots$ converge?

54.2 Interval and radius of convergence

Recall that some Taylor series had restrictions on the values of x for which the series equaled the function (e.g. geometric series, $\ln(1+x)$, $\arctan(x)$). In other words, the series converges for some values of x and diverges for other values of x.

In general, given a power series $f(x) = \sum a_n x^n$, the goal is to find the values of x for which the series converges. The following theorem tells us that the set of such values is always an interval:

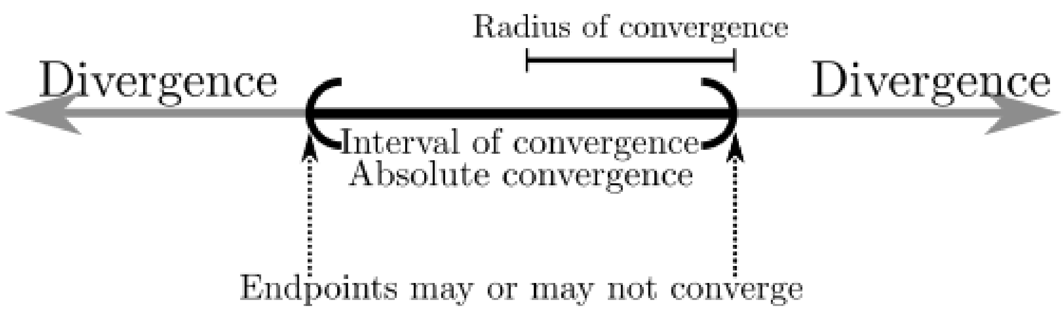
Given

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,$$

there exists some $0 \leq R \leq \infty$ such that

- the series converges absolutely if $|x| < R$;
- the series diverges if $|x| > R$;
- might converge or diverge when $x = R$ or $x = -R$.

The *interval of convergence* is the interval $(-R, R)$, possibly with the endpoints included (they need to be individually checked in general). The *radius of convergence* is half the length of the interval of convergence.



The method for finding the interval of convergence is to use the ratio test to find the interval where the series converges absolutely and then check the endpoints of the interval using the various methods from the previous modules.

Previously, when using the ratio test, one computed ρ and then checked if $\rho < 1$, $\rho > 1$, or $\rho = 1$. Now, the goal is to find the values of x for which the series $\sum a_n x^n$ converges absolutely, i.e. for which $\rho < 1$. So ρ is computed, in terms of x, and is set to be less than 1. This gives an interval of values for x which the series converges absolutely. Once this interval has been determined, the endpoints must be checked for convergence as well.

Radius of convergence

For a general power series $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots$, what is the radius of convergence R in terms of the sequence a_n? Note that

$$\begin{aligned}\rho &= \lim_{n\to\infty} \frac{|a_{n+1}x^{n+1}|}{|a_n x^n|} \\ &= \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}|x|.\end{aligned}$$

And for absolute convergence $\rho < 1$. So in terms of the a_n, we get absolute convergence when

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}|x| < 1,$$

or equivalently,

$$|x| < \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}.$$

Thus, we have shown that the radius of convergence of the series $f(x) = a_0 + a_1 x + a_2 x^2 + \cdots$ is given by

$$R = \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|}.$$

Example

Find the interval and radius of convergence for the power series

$$\sum_{n=1}^{\infty} \frac{n}{2^n} x^n$$

(See Answer 2)

Example

Find the interval of convergence for

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

(See Answer 3)

54.3 Shifted power series

Recall that the Taylor series for a function can be computed at a point c other than 0. In this case the series took the form

$$f(x) = a_0 + a_1(x-c) + a_2(x-c)^2 + a_3(x-c)^3 + \cdots$$

Such a power series is said to be centered at c, since the interval of convergence for the series will be centered at c. To see why, carry out the calculation as above (replacing all the $|x|$'s with $|x-c|$'s) to find that the series converges absolutely when

$$|x-c| < \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} = R.$$

So the radius of convergence is the same (it only depends on the sequence a_n), and it is only the center of the interval that has changed. Thus, the interval of convergence is $(c-R, c+R)$, and again one must individually check the endpoints.

Example

Find the interval of convergence for

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

(See Answer 4)

Example

Find the interval of convergence for

$$\sum_{n=2}^{\infty}\frac{(2x+5)^n}{\ln n}.$$

(See Answer 5)

54.4 Answers to Selected Examples

1. Begin by writing out the series $L(x)$, $xL(x)$, and $x^2L(x)$:

$$\begin{aligned} L(x) &= L_0 + L_1x + L_2x^2 + L_3x^3 + L_4x^4 + \cdots \\ xL(x) &= \qquad L_0x + L_1x^2 + L_2x^3 + L_3x^4 + \cdots \\ x^2L(x) &= \qquad\qquad L_0x^2 + L_1x^3 + L_2x^4 + \cdots \end{aligned}$$

Now note what happens when we take $L(x) - xL(x) - x^2L(x)$ and collect like terms. Because of the recurrence, all of the coefficients of the form $L_n - L_{n-1} - L_{n-2} = 0$, which leaves only two terms in the power series:

$$\begin{aligned} L(x) - xL(x) - x^2L(x) &= L_0 + (L_1 - L_0)x + (L_2 - L_1 - L_0)x^2 + (L_3 - L_2 - L_1)x^3 + \cdots \\ &= L_0 + (L_1 - L_0)x + 0x^2 + 0x^3 + 0x^4 + \cdots \\ &= 2 + (1-2)x \\ &= 2 - x. \end{aligned}$$

Now, solving for $L(x)$ by factoring and dividing gives

$$L(x) = \frac{2-x}{1-x-x^2}.$$

(Return)

2. Using the ratio test for absolute convergence, one computes

$$\begin{aligned}
\rho &= \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| \\
&= \lim_{n\to\infty}\left|\frac{(n+1)x^{n+1}/2^{n+1}}{nx^n/2^n}\right| \\
&= \lim_{n\to\infty}\frac{n+1}{2^{n+1}}\cdot\frac{2^n}{n}|x| \\
&= \lim_{n\to\infty}\frac{1}{2}\frac{n+1}{n}|x| \\
&= \frac{|x|}{2}.
\end{aligned}$$

Setting $\rho < 1$ means $\frac{|x|}{2} < 1$, hence $|x| < 2$. So for all $-2 < x < 2$, the series converges absolutely.

Checking the endpoint $x = -2$ gives the series

$$\begin{aligned}
\sum_{n=1}^{\infty}\frac{n}{2^n}(-2)^n &= \sum_{n=1}^{\infty}\frac{n}{2^n}(-1)^n 2^n \\
&= \sum_{n=1}^{\infty}(-1)^n n,
\end{aligned}$$

which diverges by the nth term test for divergence. Similarly, the endpoint $x = 2$ gives the series

$$\sum_{n=1}^{\infty}\frac{n}{2^n}2^n = \sum_{n=1}^{\infty} n,$$

which diverges, also by the nth term test. Thus, both endpoints diverge and so the interval of convergence is $(-2, 2)$. The radius of convergence is 2.

(Return)

3. Using the ratio test gives

$$\begin{aligned}
\rho &= \lim_{n\to\infty}\left|\frac{x^{n+1}}{n+1}\frac{n}{x^n}\right| \\
&= \lim_{n\to\infty}\frac{|x|n}{n+1} \\
&= |x|.
\end{aligned}$$

Setting this less than 1 gives $|x| < 1$ so the interval is $(-1, 1)$. Checking the endpoint $x = 1$ gives the alternating harmonic series which converges. The other endpoint $x = -1$ gives the harmonic series which diverges. So the interval of convergence is $(-1, 1]$.

(Return)

4. The interval of convergence will be centered at -3. We can take a short cut and just compute the radius of convergence $R = \lim \frac{|a_n|}{|a_{n+1}|}$, where $a_n = \frac{(-1)^n}{n^2}$. This gives $R = 1$.

Then the interval of convergence is $(-3-1, -3+1) = (-4, -2)$, and it remains to check the endpoints. At $x = -4$, we get $\sum \frac{1}{n^2}$, which converges by p-series. At $x = -2$, we get the alternating p-series $\sum(-1)^n \frac{1}{n^2}$, which converges by alternating test. Thus, the interval of convergence is $[-4, -2]$.

(Return)

5. It takes a little rearranging before this takes the form of a shifted series as defined above (the problem is the coefficient of 2 on the x). The 2 can be factored out, and what results is

$$\begin{aligned}\sum_{n=2}^{\infty} \frac{(2x+5)^n}{\ln n} &= \sum_{n=2}^{\infty} \frac{2^n(x+5/2)^n}{\ln n} \\ &= \sum_{n=2}^{\infty} \frac{2^n}{\ln n}(x+5/2)^n,\end{aligned}$$

which is now of the form given above. Note that this series is centered at $-5/2$. The radius of convergence is

$$\begin{aligned}R &= \lim_{n\to\infty} \frac{2^n}{\ln(n)} \cdot \frac{\ln(n+1)}{2^{n+1}} \\ &= \frac{1}{2}\end{aligned}$$

(since the ratio of logs goes to 1). Thus, the interval of convergence is $(-3,-2)$. Checking the endpoints, one finds convergence at $x=-3$ (by alternating series test) and divergence at $x=-2$ (by comparison of $\sum \frac{1}{\ln n}$ with the harmonic series, for example).

So the interval of convergence is $(-3,-2)$.

(Return)

55 Taylor Series Redux

The last module took a sequence (a_n) and turned it into a power series $f(x) = \sum_{n=0}^{\infty} a_n x^n$. This module turns this around and asks can we go from a function $f(x)$ to a sequence (a_n)? The answer is yes, and this is the familiar process of computing the Taylor series for the function. Recalling the Taylor series for f:

$$f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(0)x^n,$$

we see that $a_n = \frac{1}{n!} f^{(n)}(0)$ is the sequence corresponding to $f(x)$. We could also center the Taylor series at a different point, and get a different sequence, but for now let's keep things centered at 0.

Now that we have turned our function into its Taylor series, we come back to the questions deferred from earlier in the course:

1. For what values of x does a function's Taylor series converge?
2. Does the Taylor series converge to the function?

These questions are the topics of this module.

55.1 Taylor series convergence

We now have the tools to see when a power series converges, so the answer to the first question is that the series

$$\sum_{n=0}^{\infty} a_n x^n = \sum_{n=0} \frac{1}{n!} f^{(n)}(0)x^n$$

converges absolutely for $|x| < R$, where

$$\begin{aligned} R &= \lim_{n\to\infty} \frac{|a_n|}{|a_{n+1}|} \\ &= \lim_{n\to\infty} \frac{f^{(n)}(0)}{n!} \cdot \frac{(n+1)!}{f^{(n+1)}(0)} \\ &= \lim_{n\to\infty} (n+1) \frac{f^{(n)}(0)}{f^{(n+1)}(0)}. \end{aligned}$$

Within the interval of convergence, differentiation and integration of a power series are nice, in that they can be done term by term:

1. $\dfrac{df}{dx} = \displaystyle\sum_{n=0}^{\infty} n a_n x^{n-1}$

2. $\displaystyle\int f(x)\,dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} + C$

Why is this useful? Being able to differentiate and integrate term by term allows us to compute the Taylor series for various functions by differentiating or integrating the Taylor series for other functions.

Example

Compute the Taylor series for $\arctan x$ by noting that $\frac{d}{dx} \arctan x = \frac{1}{1+x^2}$. (See Answer 1)

Example

Show that the power series $f(x) = 1 + 2x + 3x^2 + 4x^3 + 5x^4 + \cdots$ can be written as $\frac{1}{(1-x)^2}$ for $|x| < 1$. Hint: try integrating both sides and see what familiar function you get. (See Answer 2)

55.2 Example

The *Fresnel Integral* $C(x)$ is defined by

$$C(x) = \int_{t=0}^{x} \cos(t^2)dt.$$

There is no elementary expression for this integral, but it can be expressed as a series by expanding the series for cos and then integrating term by term:

$$\begin{aligned} C(x) &= \int_{t=0}^{x} \cos(t^2)dt \\ &= \int_{t=0}^{x} \left(\sum_{n=0}^{\infty} (-1)^n \frac{(t^2)^{2n}}{(2n)!} \right) dt \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n+1}}{(2n)!\,(4n+1)} \Bigg|_{t=0}^{x} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+1}}{(2n)!\,(4n+1)}. \end{aligned}$$

55.3 Taylor series convergence to a function

Now, we consider the second question above: when a Taylor series converges, does it always converge to the function? Unfortunately, not always. Even with a smooth f, and x within the interval of convergence, it is possible that the Taylor series does not converge to f. The following definition is used for functions whose Taylor series *do* converge to the functions themselves:

Definition: Real-analytic function

A function f is *real-analytic* at $x = a$ if for x sufficiently close to a,

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{n!} f^{(n)}(a)(x-a)^n.$$

That is, a function f is real-analytic at a if the Taylor series for f converges to f near a.

Almost all the functions we have encountered in this course are real-analytic. However, there are examples of smooth functions which are not real-analytic, as the next example shows.

Example

Consider the function

$$f(x) = \begin{cases} e^{-1/x} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

To show that this function is smooth, we must show that its derivative exists at $x = 0$ (everywhere else it is a composition of nice functions, so we need not worry). We use the definition of the derivative:

$$\begin{aligned} \left.\frac{df}{dx}\right|_{x=0} &= \lim_{h \to 0^+} \frac{f(h) - f(0)}{h} \\ &= \lim_{h \to 0^+} \frac{1}{h} e^{-1/h}. \end{aligned}$$

This can either be thought of as a $0/0$ case for l'Hospital's rule, or we can do a change of variables $t = \frac{1}{h}$, which gives the limit

$$\begin{aligned} \lim_{h \to 0^+} \frac{1}{h} e^{-1/h} &= \lim_{t \to \infty} t e^{-t} \\ &= \lim_{t \to \infty} \frac{t}{e^t} \\ &= 0, \end{aligned}$$

since the exponential beats a polynomial asymptotically. So $f'(0) = 0$. It turns out that all of the higher derivatives of f at 0 are 0 as well. So if we tried to expand this as a Taylor series, we would have

$$\begin{aligned} f(x) &= f(0) + f'(0)x + \frac{1}{2!} f''(0)x^2 + \cdots \\ &= 0 + 0x + 0x^2 + \cdots \\ &= 0. \end{aligned}$$

So the Taylor series for f converges to 0, despite the fact that f is non-zero for $x > 0$.

55.4 Answers to Selected Examples

1. We have that

$$\frac{d}{dx}\arctan x = \frac{1}{1+x^2}$$
$$= 1 - x^2 + x^4 - x^6 + x^8 - \cdots,$$

for $|x| < 1$. Thus, for $|x| < 1$, we can safely integrate both sides of this equation to find

$$\arctan x = \int \left(1 - x^2 + x^4 - x^6 + x^8 - \cdots\right) dx$$
$$= x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \cdots + C.$$

By checking $\arctan(0) = 0$, we find that the integration constant $C = 0$. Thus for $|x| < 1$ we have

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

(Return)

2. Integrating both sides gives

$$\int f(x)dx = \int (1 + 2x + 3x^2 + \cdots)dx$$
$$= x + x^2 + x^3 + \cdots$$
$$= \frac{x}{1-x},$$

by recognizing that this is the geometric series. Now, differentiating both sides gives that

$$f(x) = \frac{1}{(1-x)^2},$$

as desired. Note that this only holds within the interval of convergence for the geometric series, $|x| < 1$.
(Return)

56 Approximation And Error

Given a series that is known to converge but for which an exact answer is not known, how does one find a good approximation to the true value? One way to get an approximation is to add up some number of terms and then stop. But how many terms are enough? How close will the result be to the true answer? That is the motivation for this module.

56.1 Error defined

Given a convergent series

$$s = \sum_{n=0}^{\infty} a_n.$$

Recall that the partial sum s_k is the sum of the terms up to and including a_k, i.e.,

$$\begin{aligned} s_k &= a_0 + a_1 + a_2 + \ldots + a_k \\ &= \sum_{n=0}^{k} a_n. \end{aligned}$$

Then the *error* E_k is the difference between s_k and the true value s, i.e.,

$$\begin{aligned} E_k &= s - s_k \\ &= \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{k} a_n \\ &= a_{k+1} + a_{k+2} + a_{k+3} + \ldots \\ &= \sum_{n=k+1}^{\infty} a_n. \end{aligned}$$

In other words, the error is the sum of all the terms from the infinite series which were not included in the partial sum.

56.2 Alternating series error bound

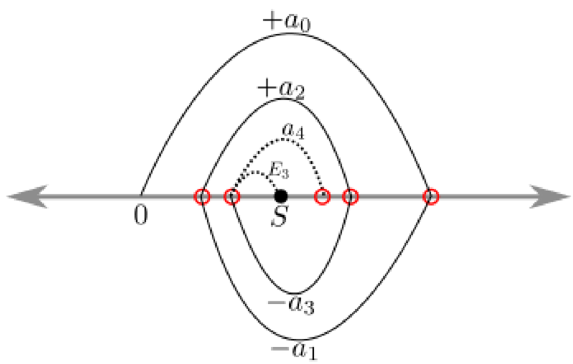
For a decreasing, alternating series, it is easy to get a bound on the error E_k:

$$|E_k| \leq |a_{k+1}|.$$

In other words, the error is bounded by the next term in the series.

Note

If the series is strictly decreasing (as is usually the case), then the above inequality is strict.



To see why the alternating bound holds, note that each successive term in the series overshoots the true value of the series. In other words, if S is the true value of the series,

$$\begin{aligned} a_0 &> S \\ a_0 - a_1 &< S \\ a_0 - a_1 + a_2 &> S. \end{aligned}$$

The above figure shows that if one stops at $a_0 - a_1 + a_2 - a_3$, then the error E_3 must be less than a_4.

Example

What is the minimum number of terms of the series

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

one needs to be sure to be within $\frac{1}{100}$ of the true sum?
The goal is to find k so that $|E_k| \le \frac{1}{100}$. Since $|E_k| \le |a_{k+1}|$, the question becomes for which value of k is $|a_{k+1}| \le \frac{1}{100}$? If $k = 9$, then $|a_{k+1}| = \frac{1}{100}$, and so by the alternating series error bound, $|E_9| \le \frac{1}{100}$. Thus 9 terms are required to be within $\frac{1}{100}$ of the true value of the series.

56.3 Integral test for error bounds

Another useful method to estimate the error of approximating a series is a corollary of the integral test. Recall that if a series $\sum f(n)$ has terms which are positive and decreasing, then

$$\int_{m+1}^{\infty} f(x)dx < \sum_{n=m+1}^{\infty} f(n) < \int_{m}^{\infty} f(x)dx.$$

But notice that the middle quantity is precisely E_m. So

$$\int_{m+1}^{\infty} f(x)dx < E_m < \int_{m}^{\infty} f(x)dx.$$

This bound is nice because it gives an upper bound *and* a lower bound for the error.

Example

How many terms of the series

$$\sum_{n=1}^{\infty} \frac{1}{n^3}$$

must one add up so that the Integral bound guarantees the approximation is within $\frac{1}{100}$ of the true answer? (See Answer 1)

56.4 Taylor approximations

Recall that the *Taylor series* for a function f about 0 is given by

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \\ &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots . \end{aligned}$$

The *Taylor polynomial of degree N* is the approximating polynomial which results from truncating the above infinite series after the degree N term:

$$\begin{aligned} f(x) &\approx \sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!} x^n \\ &= f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(N)}(0)}{N!}x^N. \end{aligned}$$

This is a good approximation for $f(x)$ when x is close to 0. How good an approximation is it? That is the purpose of the last error estimate for this module.

As in previous modules, let $E_N(x)$ be the error between the Taylor polynomial and the true value of the function, i.e.,

$$f(x) = \sum_{n=0}^{N} \frac{f^{(n)}(0)}{n!} x^n + E_N(x).$$

Notice that the error $E_N(x)$ is a function of x. In general, the further away x is from 0, the bigger the error will be.

A first, weak bound for the error is given by

$$E_N(x) \leq Cx^{N+1}$$

for some constant C and x sufficiently close to 0. In other words, $E_N(x)$ is $O(x^{N+1})$. A stronger bound is given in the next section.

Taylor remainder theorem

The following gives the precise error from truncating a Taylor series:

Taylor remainder theorem

The error $E_N(x)$ is given precisely by

$$E_N(x) = \frac{f^{(N+1)}(t)}{(N+1)!}x^{N+1},$$

for some t between 0 and x, inclusive. So if $x < 0$, then $x \leq t \leq 0$, and if $x > 0$, then $0 \leq t \leq x$.

Example

Consider the case when $N = 0$. The Taylor remainder theorem says that

$$f(x) = f(0) + f'(t)x,$$

for some t between 0 and x. Solving for $f'(t)$ gives

$$f'(t) = \frac{f(x) - f(0)}{x - 0},$$

for some $0 < t < x$ if $x > 0$ and $x < t < 0$ if $x < 0$, which is precisely the statement of the Mean value theorem. Therefore, one can think of the Taylor remainder theorem as a generalization of the Mean value theorem.

Taylor error bound

As it is stated above, the Taylor remainder theorem is not particularly useful for actually finding the error, because there is no way to actually find the t for which the equation holds. There is a slightly different form which gives a bound on the error:

Taylor error bound

$$|E_N(x)| \leq \frac{C}{(N+1)!}|x|^{N+1},$$

where C is the maximum value of $|f^{(N+1)}(t)|$ over all t between 0 and x, inclusive.

Example

Estimate $\sqrt{e}$ using

$$e^{1/2} \approx 1 + \frac{1}{2} + \frac{(1/2)^2}{2!} + \frac{(1/2)^3}{3!} \approx 1.64,$$

and bound the error. (See Answer 2)

56.5 Answers to Selected Exercises

1. If one adds up the first m terms, then by the integral bound, the error E_m satisfies

$$\begin{aligned} E_m &< \int_m^\infty \frac{dx}{x^3} \\ &= \left.\frac{x^{-2}}{-2}\right|_m^\infty \\ &= \frac{1}{2m^2}. \end{aligned}$$

Setting $\frac{1}{2m^2} \leq \frac{1}{100}$ gives that $m^2 \geq 50$, so $m \geq 8$. Thus, $m = 8$ is the minimum number of terms required so that the Integral bound guarantees we are within $\frac{1}{100}$ of the true answer.

Note

If you actually compute the partial sums using a calculator, you will find that 7 terms actually suffice. But remember, we want the *guarantee* of the integral test, which only ensures that $\frac{1}{128} < E_7 < \frac{1}{98}$, despite the fact that in reality, $E_7 \approx .009 < .01$.

(Return)

2. The function is $f(x) = e^x$, and the approximating polynomial used here is

$$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!}.$$

Then according to the above bound,

$$|E_3(x)| \leq \frac{C}{4!}|x|^4,$$

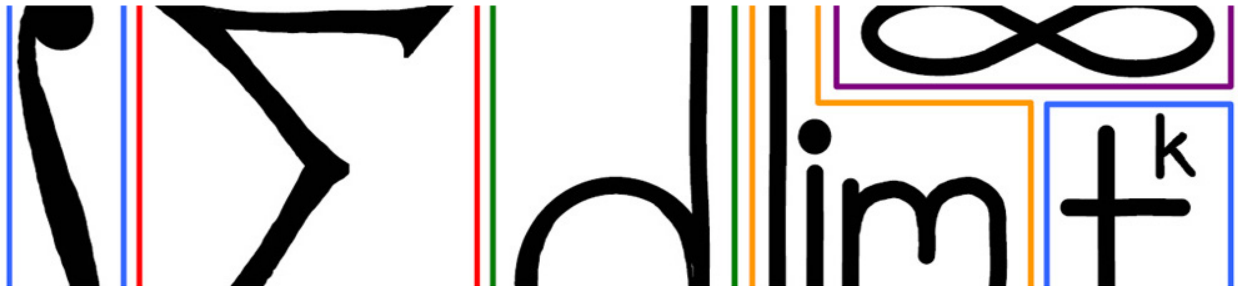
where C is the maximum of $f^{(4)}(t) = e^t$ for $0 \leq t \leq x$. Since e^t is an increasing function, $C = e^x$. Thus,

$$|E_3(x)| \leq \frac{e^x}{4!}x^4.$$

Thus,

$$|E_3(1/2)| \leq \frac{e^{1/2}}{4!}(1/2)^4 < \frac{1}{100}.$$

(Return)



57 Calculus

In this course, we've learned skills in five key areas:

- Limits
- Differentiation
- Integration
- ODEs
- Series

Some of the things we can do are pretty impressive. However, there are many simple-seeming questions in single-variable calculus that show us just how much we have left to learn.

57.1 PROBLEM 1

Why is the standard Gaussian a probability density function? In other words, why is

$$\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\,dx = \sqrt{2\pi} \tag{1}$$

Why is it that with all our methods, we cannot evaluate this integral easily? Let's try...

MORAL

This integral is *easy* with a little bit of multivariable calculus.

57.2 PROBLEM 2

Recall that we began this class with the definition of the exponential function e^x and then, to obtain series for sin and cos, we invoked **Euler's formula**:

$$e^{it} = \cos t + i \sin t$$

Why is this true? We certainly could substitute in our favorite Taylor series and verify that it is true, but wouldn't it be better to have a *principled* reason for why this is so? Let's try...

Let $z = e^{it}$. Then $z' = iz$ by that very familiar differential equation. Now, name the real and imaginary parts of z by x and y respectively. Then $z = x + iy$, and $z' = x' + iy'$. On the other hand, multiplying by i gives

$$\begin{aligned} z &= x + iy \\ iz &= ix + i^2 y = -y + ix. \end{aligned}$$

Therefore, $z' = iz$ becomes $x' + iy' = -y + ix$, and so by equating the real and imaginary parts in this equation we get the system

$$\begin{aligned} x' &= -y \\ y' &= x \end{aligned}$$

This is a system of differential equations which is easy to solve with some multivariable calculus, but for now we are stuck. We can observe that $x = \cos t$ and $y = \sin t$ provides a solution, but we cannot say it is the only solution without more tools.

MORAL

This system of ODEs is *easy* to solve with a little bit of multivariable calculus.

57.3 PROBLEM 3

On several occasions, we have referenced the famous series

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Why is this true? Could we use discrete calculus to derive it? I don't think so...but here is proof using Taylor series and integration.

Let $u = \arcsin(x)$. Consider the integral

$$\int_{u=0}^{\pi/2} u\,du = \left.\frac{u^2}{2}\right|_{u=0}^{\pi/2} = \frac{\pi^2}{8}.$$

On the other hand, the integral in terms of x is

$$\int_{x=0}^{1} \arcsin(x) \frac{1}{\sqrt{1-x^2}}\,dx$$

One can show that the Taylor series for $\arcsin(x)$ is given by

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

Plugging this in gives

$$\frac{\pi^2}{8} = \int_{x=0}^{1} \left(x + \sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \frac{x^{2n+1}}{2n+1} \right) \frac{1}{\sqrt{1-x^2}}\,dx.$$

One can find with careful integration by parts and induction that the inner integral evaluates to

$$\int_{x=0}^{1} \frac{x^{2n+1}}{\sqrt{1-x^2}}dx = \frac{2\cdot 4\cdot 6\cdots(2n)}{3\cdot 5\cdot 7\cdots(2n+1)},$$

(See Details 1)

Plugging this in cancels almost everything, leaving

$$\frac{\pi^2}{8} = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}.$$

This is the sum of the *odd* reciprocals squared. Giving names to these various sums:

$$S = \sum_{n=1}^{\infty} \frac{1}{n^2}$$
$$S_o = \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}$$
$$S_e = \sum_{n=1}^{\infty} \frac{1}{(2n)^2}.$$

Note that $S = S_o + S_e$ (S_o has the odd terms, and S_e has the even terms). Further,

$$\begin{aligned} S_e &= \sum_{n=1}^{\infty} \frac{1}{(2n)^2} \\ &= \frac{1}{4}\sum_{n=1}^{\infty} \frac{1}{n^2} \\ &= \frac{1}{4}S. \end{aligned}$$

Substituting, we find that $S = S_o + S/4$, and so $S = \frac{4}{3}S_o$. Since $S_o = \frac{\pi^2}{8}$ as shown above, it follows that $S = \frac{\pi^2}{6}$, as desired.

MORAL

This series is *easy* to evaluate with a little bit of multivariable calculus. Well, actually, no: it's not easy, but it is a bit simpler. In the end, some math problems are hard.

57.4 Answers to Selected Exercises

1. If we encountered this integral earlier in the course, we would hit it with a trigonometric substitution $x = \sin\theta$ (and $dx = \cos\theta\, d\theta$, which changes the integral to

$$\int_{\theta=0}^{\pi/2} \sin^{2n+1}\theta\, d\theta.$$

This integral can be found inductively using the following reduction formula:

$$\int \sin^n\theta\, d\theta = \frac{-\sin^{n-1}\theta\cos\theta}{n} + \frac{n-1}{n}\int \sin^{n-2}\theta\, d\theta.$$

(See Proof of Reduction Formula 2)

2. Use integration by parts. Let $u = \sin^{n-1}\theta$ and $dv = \sin\theta\, d\theta$. Then $du = (n-1)\sin^{n-2}\theta\cos\theta\, d\theta$ and $v = -\cos\theta$. It follows that

$$\begin{aligned}\int \sin^n\theta\, d\theta &= -\sin^{n-1}\theta\cos\theta + (n-1)\int \sin^{n-2}\theta\cos^2\theta\, d\theta \\ &= -\sin^{n-1}\theta\cos\theta + (n-1)\int \sin^{n-2}\theta(1-\sin^2\theta)\, d\theta \\ &= -\sin^{n-1}\theta\cos\theta + (n-1)\int \sin^{n-2}\theta\, d\theta - (n-1)\int \sin^n\theta\, d\theta.\end{aligned}$$

Now, solving for the integral $\int \sin^n\theta\, d\theta$ (by adding $(n-1)\int \sin^n\theta\, d\theta$ to both sides of the above equation and dividing by n) gives the desired result.

Applying the reduction formula in the situation at hand gives

$$\begin{aligned}\int_0^{\pi/2} \sin^{2n+1}\theta\, d\theta &= \left.\frac{-\sin^{2n}\theta\cos\theta}{2n+1}\right|_0^{\pi/2} + \frac{2n}{2n+1}\int_0^{\pi/2}\sin^{2n-1}\theta\, d\theta \\ &= \frac{2n}{2n+1}\int_0^{\pi/2}\sin^{2n-1}\theta\, d\theta,\end{aligned}$$

since the first quantity evaluates to 0. Now, induction gives the result.

(Return)

(Return)

58 Foreshadowing

In this module we give a hint at what is to come in multivariable calculus.

58.1 Functions

As in single variable calculus, multivariable calculus is primarily a study of functions. But instead of functions with one input and one output, multivariable calculus looks at functions with multiple inputs and outputs. The notation for a function f with n real inputs and m real outputs is

$$f : \mathbb{R}^n \to \mathbb{R}^m.$$

58.2 Matrices

When dealing with multiple inputs and multiple outputs, it becomes necessary to keep track of several pieces of information when dealing with, say, the derivative. The data structure which makes this possible is a *matrix*, which is an array of numbers arranged in rows and columns. For example, a 4×3 matrix has 4 rows and 3 columns, and might look like

$$\begin{bmatrix} 3 & 1 & 4 \\ 1 & 5 & 9 \\ 2 & 6 & 5 \\ 3 & 5 & 8 \end{bmatrix}.$$

A *square matrix* has the same number of rows and columns. A particular square matrix with a special name is the *identity matrix*, which has 1's on the main diagonal and 0's everywhere else. For example, the 3×3 identity matrix is given by

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Matrix algebra

One feature of matrices is that they can be multiplied, by a slightly unintuitive process. Consider the product of a 2×3 matrix with a 3×3 matrix:

$$\begin{bmatrix} 1 & -1 & 4 \\ 2 & 5 & 7 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 0 \\ 4 & 4 & 3 \\ -2 & 7 & 1 \end{bmatrix} = \begin{bmatrix} -10 & 25 & 1 \\ 10 & 71 & 22 \end{bmatrix}$$

To see where these numbers come from, arrange the matrices a little differently:

(Matrix Multiplication Animated GIF)

To get an entry in the resulting matrix (bottom right), take the corresponding row from the matrix to the left and the corresponding column from the matrix above, multiply their corresponding entries together, and add. The above example shows the calculation for two entries in the result.

Note that for this multiplication to be defined, the number of columns in the first matrix must match the number of rows in the second matrix (otherwise there would not be the correct number of entries to multiply together and add).

There are some nice features of matrix multiplication, and some features which are a little bit different than regular multiplication:

- The identity matrix can be thought of as the matrix equivalent of 1, since multiplying by the identity (of the appropriate size) gives back the same matrix with which we began.
- Matrix multiplication is associative (i.e. $(AB)C = A(BC)$ for appropriately sized matrices A, B, C), but it is *not* commutative (i.e. $AB \neq BA$) in general. Indeed, both orders of multiplication is only defined if A and B are square matrices of the same size.
- There is a matrix version of $\sqrt{-1}$ (again thinking of I as 1). Note that

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -I.$$

- It is possible that the product of two non-zero matrices to give the zero matrix:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

58.3 Vectors

Another data structure of importance is a *vector*, which can be thought of as a single row or single column matrix (depending on context). A useful way to visualize a vector is as a difference between two points, or an arrow from one point to another. So a vector can be thought of as a line segment with both a magnitude (the distance between the two points) and a direction (which way the arrow points).

Vectors can be added by visualizing placing the tail of one vector at the head of the other. Vectors can also be multiplied by a matrix to give another vector (the multiplication is just matrix multiplication, again by thinking of the vector as a matrix with just a single column).