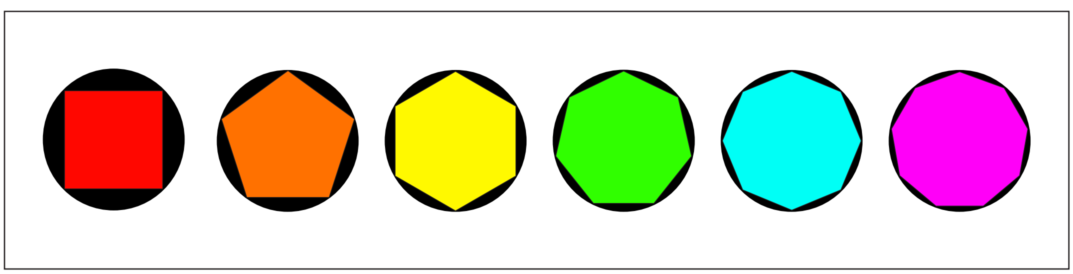
Math 138

Calculus II
for Honours Mathematics

Course Notes

Barbara A. Forrest and Brian E. Forrest



Version 1.53

August 1, 2023

Author Contact Information:

Barbara Forrest (baforres@uwaterloo.ca)

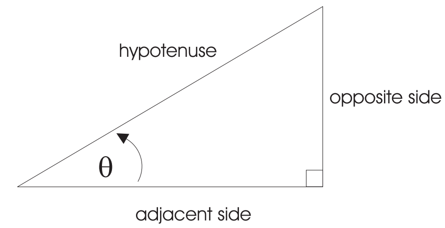
Brian Forrest (beforres@uwaterloo.ca)

QUICK REFERENCE PAGE 1

Right Angle Trigonometry

$\sin\theta = \frac{opposite}{hypotenuse}$ $\cos\theta = \frac{adjacent}{hypotenuse}$ $\tan\theta = \frac{opposite}{adjacent}$

$\csc\theta = \frac{1}{\sin\theta}$ $\sec\theta = \frac{1}{\cos\theta}$ $\cot\theta = \frac{1}{\tan\theta}$

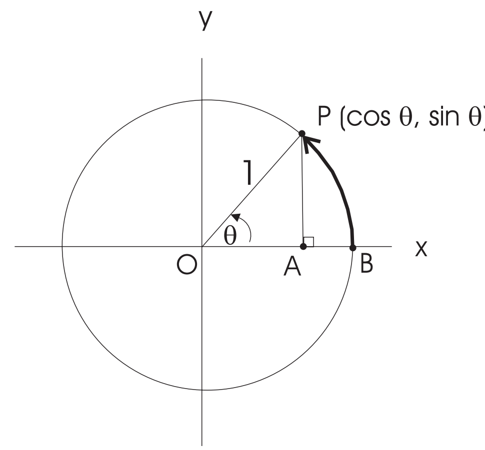


Radians

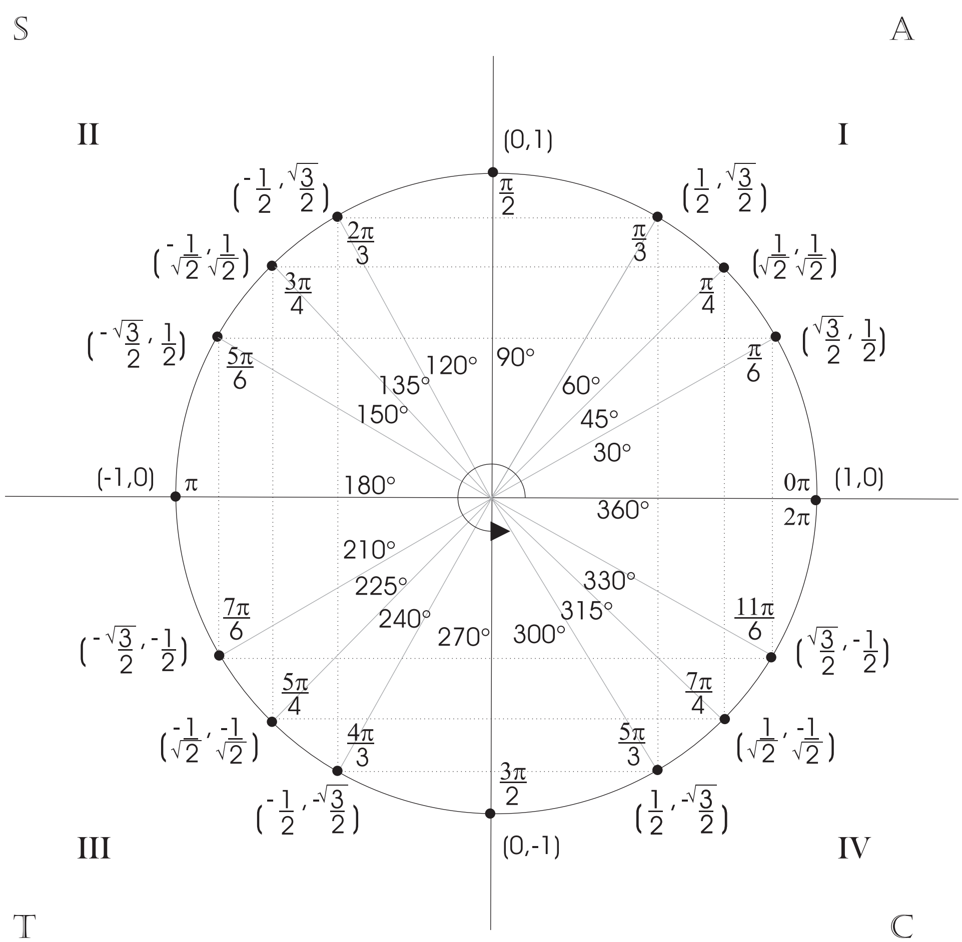
The angle θ in radians equals the length of the directed arc BP, taken positive counter-clockwise and negative clockwise. Thus, π radians = 180° or 1 $rad = \frac{180}{\pi}$.

Definition of Sine and Cosine

For any θ, $\cos\theta$ and $\sin\theta$ are defined to be the $x-$ and $y-$ coordinates of the point P on the unit circle such that the radius OP makes an angle of θ radians with the positive $x-$ axis. Thus $\sin\theta = AP$, and $\cos\theta = OA$.



The Unit Circle



QUICK REFERENCE PAGE 2

Trigonometric Identities

Pythagorean Identity	$\cos^2\theta + \sin^2\theta = 1$
Range	$-1 \le \cos\theta \le 1$ $-1 \le \sin\theta \le 1$
Periodicity	$\cos(\theta \pm 2\pi) = \cos\theta$ $\sin(\theta \pm 2\pi) = \sin\theta$
Symmetry	$\cos(-\theta) = \cos\theta$ $\sin(-\theta) = -\sin\theta$

Sum and Difference Identities
$\cos(A+B) = \cos A\cos B - \sin A\sin B$
$\cos(A-B) = \cos A\cos B + \sin A\sin B$
$\sin(A+B) = \sin A\cos B + \cos A\sin B$
$\sin(A-B) = \sin A\cos B - \cos A\sin B$

Complementary Angle Identities
$\cos(\frac{\pi}{2} - A) = \sin A$
$\sin(\frac{\pi}{2} - A) = \cos A$

Double-Angle Identities	$\cos 2A = \cos^2 A - \sin^2 A$ $\sin 2A = 2\sin A\cos A$

Half-Angle Identities	$\cos^2\theta = \frac{1+\cos 2\theta}{2}$ $\sin^2\theta = \frac{1-\cos 2\theta}{2}$

Other	$1 + \tan^2 A = \sec^2 A$

QUICK REFERENCE PAGE 3

Differentiation Rules

Function	Derivative
$f(x) = cx^a$, $a \neq 0$, $c \in \mathbb{R}$	$f'(x) = cax^{a-1}$
$f(x) = \sin(x)$	$f'(x) = \cos(x)$
$f(x) = \cos(x)$	$f'(x) = -\sin(x)$
$f(x) = \tan(x)$	$f'(x) = \sec^2(x)$
$f(x) = \sec(x)$	$f'(x) = \sec(x)\tan(x)$
$f(x) = \arcsin(x)$	$f'(x) = \dfrac{1}{\sqrt{1-x^2}}$
$f(x) = \arccos(x)$	$f'(x) = -\dfrac{1}{\sqrt{1-x^2}}$
$f(x) = \arctan(x)$	$f'(x) = \dfrac{1}{1+x^2}$
$f(x) = e^x$	$f'(x) = e^x$
$f(x) = a^x$ with $a > 0$	$f'(x) = a^x \ln(a)$
$f(x) = \ln(x)$ for $x > 0$	$f'(x) = \dfrac{1}{x}$

Table of Integrals

$\int x^n\,dx = \dfrac{x^{n+1}}{n+1} + C$
$\int \dfrac{1}{x}\,dx = \ln(\lvert x \rvert) + C$
$\int e^x\,dx = e^x + C$
$\int \sin(x)\,dx = -\cos(x) + C$
$\int \cos(x)\,dx = \sin(x) + C$
$\int \sec^2(x)\,dx = \tan(x) + C$
$\int \dfrac{1}{1+x^2}\,dx = \arctan(x) + C$
$\int \dfrac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x) + C$
$\int \dfrac{-1}{\sqrt{1-x^2}}\,dx = \arccos(x) + C$
$\int \sec(x)\tan(x)\,dx = \sec(x) + C$
$\int a^x\,dx = \dfrac{a^x}{\ln(a)} + C$

Inverse Trigonometric Substitutions

Integral	Trig Substitution	Trig Identity
$\int \sqrt{a^2 - b^2x^2}\,dx$	$bx = a\sin(u)$	$\sin^2(x) + \cos^2(x) = 1$
$\int \sqrt{a^2 + b^2x^2}\,dx$	$bx = a\tan(u)$	$\sec^2(x) - 1 = \tan^2(x)$
$\int \sqrt{b^2x^2 - a^2}\,dx$	$bx = a\sec(u)$	$\sec^2(x) - 1 = \tan^2(x)$

Additional Formulas

Integration by Parts	$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx$
Areas Between Curves	$A = \int_a^b \lvert g(t) - f(t) \rvert\,dt$
Volumes of Revolutions: Disk I	$V = \int_a^b \pi f(x)^2\,dx$
Volumes of Revolutions: Disk II	$V = \int_a^b \pi(g(x)^2 - f(x)^2)\,dx$
Volumes of Revolutions: Shell	$V = \int_a^b 2\pi x(g(x) - f(x))\,dx$
Arc Length	$S = \int_a^b \sqrt{1 + (f'(x))^2}\,dx$

Taylor Series (Maclaurin Series)

$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \cdots$	$R = 1$
$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$	$R = \infty$
$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$	$R = \infty$
$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	$R = \infty$

Differential Equations

Separable	$y' = f(x)g(y)$
Solve	$g(y) \stackrel{?}{=} 0$, $\int \frac{1}{g(y)}dy = \int f(x)dx$
FOLDE	$y' = f(x)y + g(x)$
Solve	$y = \frac{\int g(x)I(x)dx}{I(x)}$, $I(x) = e^{-\int f(x)\,dx}$

QUICK REFERENCE PAGE 4

LIST of THEOREMS:

Chapter 1: Integration

Integrability Theorem
for Continuous Functions
Properties of Integrals Theorem
Integrals over Subintervals Theorem
Average Value Theorem
(Mean Value Theorem for Integrals)
Fundamental Theorem of Calculus (Part 1)
Extended Version of the
Fundamental Theorem of Calculus
Power Rule for Antiderivatives
Fundamental Theorem of Calculus (Part 2)
Change of Variables Theorem

Chapter 2: Techniques of Integration

Integration by Parts Theorem
Integration of Partial Fractions
p-Test for Type I Improper Integrals
Properties of Type I Improper Integrals
The Monotone Convergence Theorem
for Functions
Comparison Test
for Type I Improper Integrals
Absolute Convergence Theorem
for Improper Integrals
p-Test for Type II Improper Integrals

Chapter 3: Applications of Integration

Area Between Curves
Volumes of Revolution: Disk Methods
Volumes of Revolution: Shell Method
Arc Length

Chapter 4: Differential Equations

Theorem for Solving
First-order Linear Differential Equations
Existence and Uniqueness Theorem
for FOLDE

Chapter 5: Numerical Series

Geometric Series Test
Divergence Test
Arithmetic for Series Theorems
The Monotone Convergence Theorem
for Sequences
Comparison Test for Series
Limit Comparison Test
Integral Test for Convergence
p-Series Test
Alternating Series Test (AST) and
the Error in the AST
Absolute Convergence Theorem
Rearrangement Theorem
Ratio Test
Polynomial versus Factorial Growth Theorem
Root Test

Chapter 6: Power Series

Fundamental Convergence Theorem
for Power Series
Test for the Radius of Convergence
Equivalence of Radius of Convergence
Abel's Theorem: Continuity of Power Series
Addition of Power Series
Multiplication of Power Series by $(x - a)^m$
Power Series of Composite Functions
Term-by-Term Differentiation of Power Series
Uniqueness of Power Series Representations
Term-by-Term Integration of Power Series
Taylor's Theorem
Taylor's Approximation Theorem I
Convergence Theorem for Tayor Series
Binomial Theorem
Generalized Binomial Theorem

Table of Contents

Chapter 1

Integration

Many operations in mathematics have an inverse operation: addition and subtraction; multiplication and division; raising a number to the nth power and finding its nth root; taking a derivative and finding its antiderivative. In each case, one operation "undoes" the other. In this chapter, we begin the study of the integral and integration. Soon you will understand that *integration is the inverse operation of differentiation*.

1.1 Areas Under Curves

The two most important ideas in calculus - differentiation and integration - are both motivated from geometry. The problem of finding the tangent line led to the definition of the *derivative*. The problem of finding area will lead us to the definition of the *definite integral*.

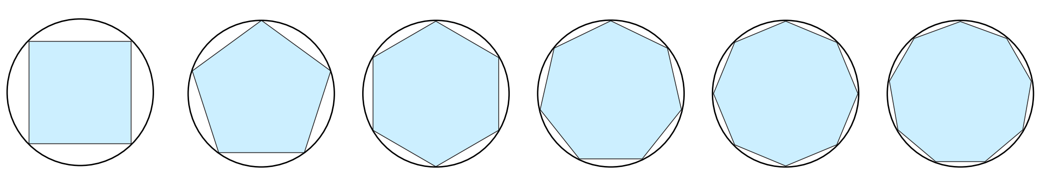
1.1.1 Estimating Areas

Our objective is to find the area under the curve of some function.

What do we mean by the *area under a curve?*

The question about how to calculate areas is actually thousands of years old and it is one with a very rich history. To motivate this topic, let's first consider what we know about finding the area of some familiar shapes. We can easily determine the area of a rectangle or a right-angled triangle, but how could we explain to someone why the area of a circle with radius r is πr^2?

The problem of calculating the area of a circle was studied by the ancient Greeks. In particular, both Archimedes and Eudoxus of Cnidus used the *Method of Exhaustion* to calculate areas. This method used various regular inscribed polygons of known area to approximate the area of an enclosed region.

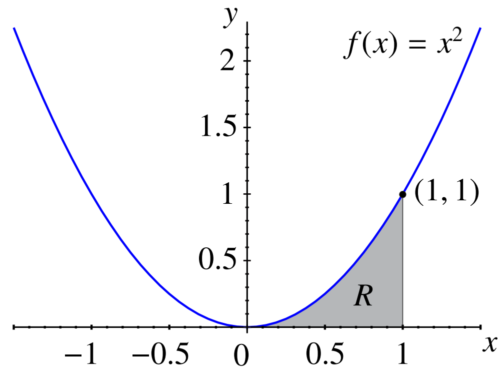


In the case of a circle, as the number of sides of the inscribed polygon increased, the error in using the area of the polygon to approximate the area of the circle decreased. As a result, the Greeks had effectively used the concept of a **limit** as a key technique in their calculation of the area.

1.1.2 Approximating Areas Under Curves

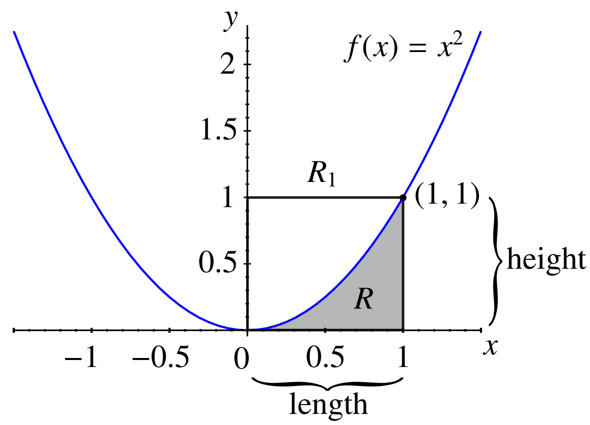
Let's use the ideas from the Method of Exhaustion and try to find the area underneath a parabola by using rectangles as a basis for the approximation.

Suppose that we have the function $f(x) = x^2$. Consider the region R bounded by the graph of f, by the x-axis, and by the lines $x = 0$ and $x = 1$.

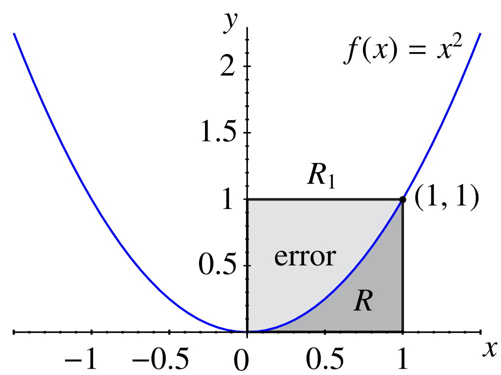


How could we determine the area of this irregular region?

For our first estimate, we can approximate the area of R by constructing a rectangle R_1 of length 1 (from $x = 0$ to $x = 1$) and height 1 ($y = f(1) = 1^2 = 1$). This rectangle (in this case a square) has area $\textit{length} \times \textit{height} = 1 \times 1 = 1$.



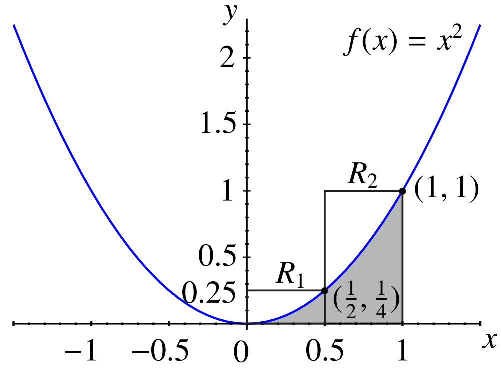
The diagram shows that the area of rectangle R_1 is larger than the area of region R. Moreover, the **error** is actually quite large.



We can find a better estimate if we split the interval $[0, 1]$ into 2 equal subintervals, $[0, \frac{1}{2}]$ and $[\frac{1}{2}, 1]$.

Using these intervals, two rectangles are constructed. The first rectangle R_1 has its length from $x = 0$ to $x = \frac{1}{2}$ with height equal to $f(\frac{1}{2}) = \frac{1}{2^2} = \frac{1}{4}$.

The second rectangle R_2 has its length from $x = \frac{1}{2}$ to $x = 1$ with height $f(1) = 1^2 = 1$.



The area of rectangle R_1 is equal to

$$R_1 = length \times height = \frac{1}{2} \times \frac{1}{2^2} = \frac{1}{8}$$

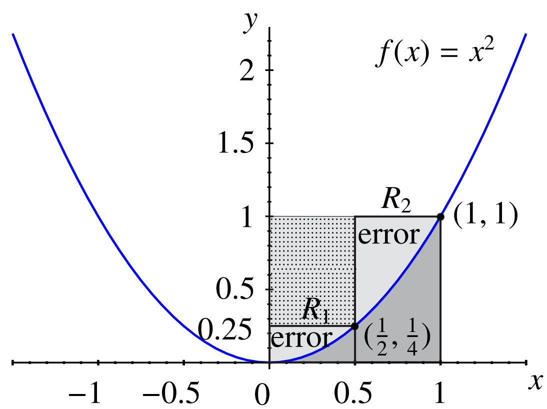
while the area of rectangle R_2 is equal to

$$R_2 = length \times height = \frac{1}{2} \times 1 = \frac{1}{2}$$

Our second estimate for the area of the original region R is obtained by adding the areas of these two rectangles to get

$$R_1 + R_2 = \frac{1}{8} + \frac{1}{2} = \frac{5}{8} = 0.625$$

Observe from the diagram that our new estimate using two rectangles for the area under $f(x) = x^2$ on the interval $[0, 1]$ is much better than our first estimate since the error is smaller. The region containing the dashed lines indicates the improvement in our estimate (this is the amount by which we have reduced the error from our first estimate).



To improve our estimate even further, divide the interval $[0, 1]$ into five equal subintervals of the form

$$[\frac{i-1}{5}, \frac{i}{5}]$$

where i ranges from 1 to 5.

This produces the subintervals

$$[0, \frac{1}{5}], [\frac{1}{5}, \frac{2}{5}], [\frac{2}{5}, \frac{3}{5}], [\frac{3}{5}, \frac{4}{5}], [\frac{4}{5}, \frac{5}{5}]$$

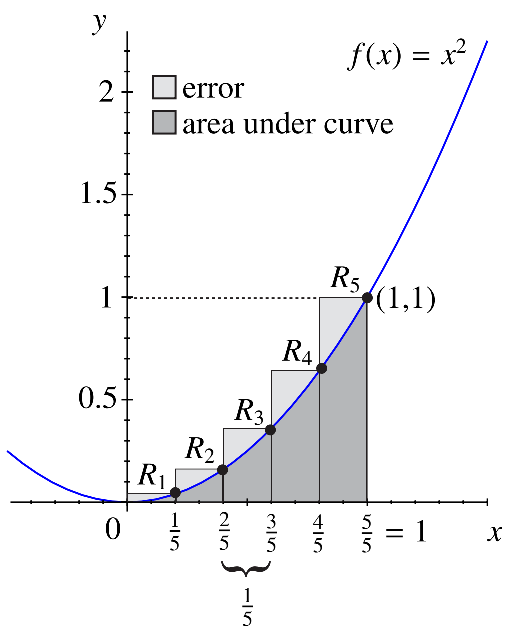
each having equal lengths of $\frac{1}{5}$.

Next we construct five new rectangles where the *ith* rectangle forms its length from $\frac{i-1}{5}$ to $\frac{i}{5}$ and has height equal to the value of the function at the *right-hand endpoint* of the interval. That is, the height of a rectangle is $f(x) = x^2$ where $x = \frac{i}{5}$ or

$$\begin{aligned} f(\frac{i}{5}) &= (\frac{i}{5})^2 \\ &= \frac{i^2}{5^2} \end{aligned}$$

The area of the *ith* rectangle is given by

$$length \times height = \frac{1}{5} \times \frac{i^2}{5^2} = \frac{i^2}{5^3}$$



Our new estimate is the sum of the areas of these rectangles which is

$$\begin{aligned} R_1 + R_2 + R_3 + R_4 + R_5 &= [(\frac{1}{5})^2(\frac{1}{5})] + [(\frac{2}{5})^2(\frac{1}{5})] + [(\frac{3}{5})^2(\frac{1}{5})] + [(\frac{4}{5})^2(\frac{1}{5})] + [(\frac{5}{5})^2(\frac{1}{5})] \\ &= \frac{1^2}{5^3} + \frac{2^2}{5^3} + \frac{3^2}{5^3} + \frac{4^2}{5^3} + \frac{5^2}{5^3} \\ &= \frac{1}{5^3}(1^2 + 2^2 + 3^2 + 4^2 + 5^2) \\ &= \frac{1}{5^3}\sum_{i=1}^{5} i^2 \end{aligned}$$

Note: It can be shown that for any n

$$\sum_{i=1}^{n} i^2 = \frac{(n)(n+1)(2n+1)}{6}$$

This means that the sum of the areas of the rectangles is

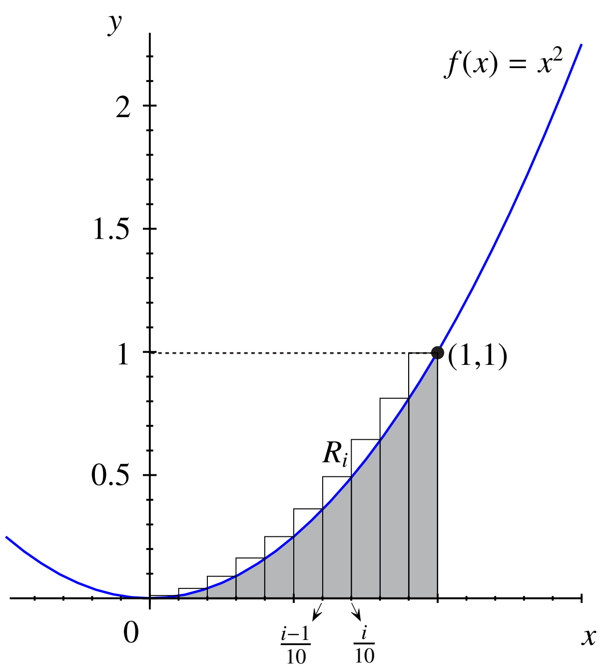
$$\begin{aligned}
\frac{1}{5^3}\sum_{i=1}^{5} i^2 &= \frac{1}{5^3} \times \frac{(5)(5+1)(2(5)+1)}{6} \\
&= \frac{1}{5^3} \times \frac{(5)(6)(11)}{6} \\
&= \frac{11}{25} \\
&= 0.44
\end{aligned}$$

So far the estimates for the area under the curve of $f(x) = x^2$ on the interval $[0, 1]$ are:

Number of Subintervals (Rectangles)	Length of Subinterval (Width of Rectangle)	Estimate for Area under Curve
1	1	1
2	$\frac{1}{2}$	0.625
5	$\frac{1}{5}$	0.44

Observe from the diagram that the estimate for the area is getting better while the error in the estimate is getting smaller.

Let's repeat this process again by using 10 equal subintervals.



In this case, the sum of the areas of the 10 rectangles will be

$$\begin{aligned}
\sum_{i=1}^{10} R_i &= R_1 + R_2 + R_3 + R_4 + R_5 + R_6 + R_7 + R_8 + R_9 + R_{10} \\
&= \frac{1}{10^3} \sum_{i=1}^{10} i^2 \\
&= \frac{1}{10^3} \frac{(10)(10+1)(2(10)+1)}{6} \\
&= \frac{77}{200} \\
&= 0.385
\end{aligned}$$

If we were to use 1000 subintervals, the estimate for the area would be

$$\begin{aligned}
\sum_{i=1}^{1000} R_i &= R_1 + R_2 + R_3 + \ldots + R_{1000} \\
&= \frac{1}{1000^3} \sum_{i=1}^{1000} i^2 \\
&= \frac{1}{1000^3} \frac{(1000)(1000+1)(2(1000)+1)}{6} \\
&= 0.3338335
\end{aligned}$$

You should begin to notice that as we increase the number of rectangles (number of subintervals), the total area of these rectangles seems to be getting closer and closer to the actual area of the original region R. In particular, if we were to produce an accurate diagram that represents 1000 rectangles, we would see no noticeable difference between the estimated area and the true area. For this reason we would expect that our latest estimate of 0.3338335 is actually very close to the true value of the area of region R.

We could continue to divide the interval $[0, 1]$ into even more subintervals. In fact, we can repeat this process with n subintervals for any $n \in \mathbb{N}$. In this generic case, the estimated area R_n would be

$$\begin{aligned}
R_n &= \frac{1}{n^3} \sum_{i=1}^{n} i^2 \\
&= \frac{1}{n^3} \frac{(n)(n+1)(2(n)+1)}{6} \\
&= \frac{\frac{1}{n^3}(2n^3 + 3n^2 + n)}{6} \\
&= \frac{2 + \frac{3}{n} + \frac{1}{n^2}}{6}
\end{aligned}$$

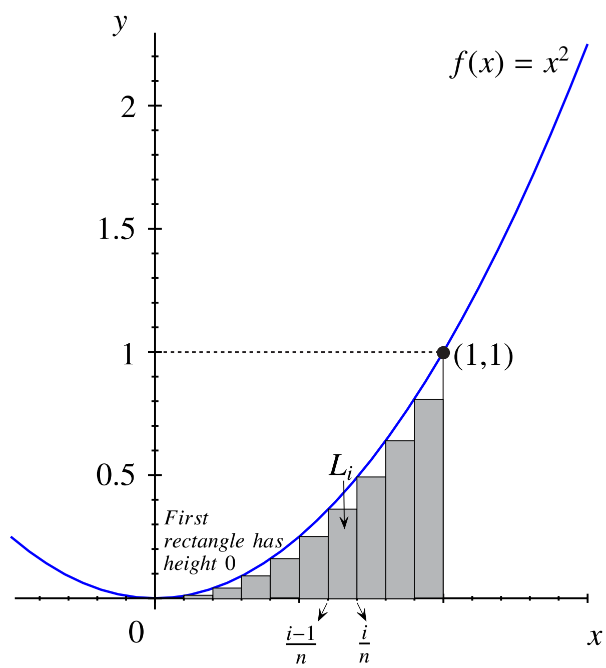
Note that if we let the number of subintervals n approach ∞, then

$$\begin{aligned}\lim_{n\to\infty} R_n &= \lim_{n\to\infty} \frac{2+\frac{3}{n}+\frac{1}{n^2}}{6} \\ &= \frac{2}{6} \\ &= \frac{1}{3}\end{aligned}$$

By calculating the area under the graph of $f(x) = x^2$ using an increasing number of rectangles, we have constructed a sequence of estimates where each estimate is larger than the actual area. Though it appears that the limiting value $\frac{1}{3}$ is a plausible guess for the actual value of the area, at this point the best that we can say is that the area should be less than or equal to $\frac{1}{3}$.

Number of Subintervals (Rectangles)	Length of Subinterval (Width of Rectangle)	Estimate for Area under Curve
1	1	1
2	$\frac{1}{2}$	0.625
5	$\frac{1}{5}$	0.44
10	$\frac{1}{10}$	0.385
1000	$\frac{1}{1000}$	0.3338335
approaches ∞	approaches 0	approaches $\frac{1}{3}$

Alternately, we can use a similar process that would produce an estimate for the area that will be less than the actual value. To do so we again divide the interval $[0, 1]$ into n subintervals of length $\frac{1}{n}$ with the i-th interval $[\frac{i-1}{n}, \frac{i}{n}]$. This interval again forms the length of a rectangle L_i, but this time we will use the *left-hand endpoint* of the interval so that the value $f(\frac{i-1}{n})$ is the height of the rectangle.



In this case, notice that since $f(0) = 0$ the first rectangle is really just a horizontal line with area 0. Then the estimated area L_n for this generic case would be

$$\begin{aligned} L_n &= \frac{1}{n^3}\sum_{i=1}^{n}(i-1)^2 \\ &= \frac{1}{n^3}\frac{(n-1)((n-1)+1)(2(n-1)+1)}{6} \\ &= \frac{1}{n^3}\frac{(n-1)(n)(2n-1)}{6} \\ &= \frac{1}{n^3}\frac{2n^3-3n^2+n}{6} \\ &= \frac{2-\frac{3}{n}+\frac{1}{n^2}}{6} \end{aligned}$$

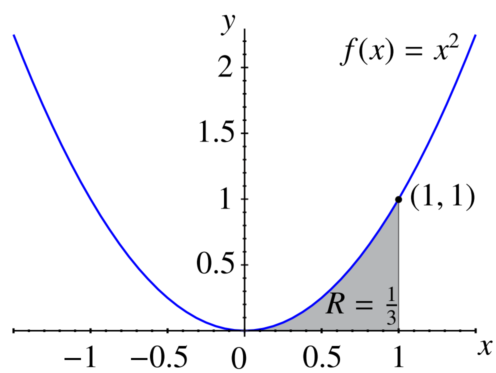
Finally, observe that

$$\lim_{n\to\infty} L_n = \lim_{n\to\infty} \frac{2-\frac{3}{n}+\frac{1}{n^2}}{6} = \frac{1}{3}.$$

In summary, we have now shown that if R is the area of the region under the graph of $f(x) = x^2$, above the x-axis, and between the lines $x = 0$ and $x = 1$, then for each $n \in \mathbb{N}$,

$$L_n \leq R \leq R_n.$$

It would be reasonable to conclude that the area under the graph of $f(x) = x^2$ bounded by the x-axis and the lines $x = 0$ and $x = 1$ is precisely the limit $\frac{1}{3}$. That is, the process of using more and more rectangles to *estimate* the area under the curve gives us a sequence of values that *converge* to the actual area under the curve.



1.1.3 The Relationship Between Displacement and Velocity

In the previous section we looked at a method to determine the area under a curve. In this section we will look at a different problem–finding a geometric relationship between displacement and velocity. Perhaps surprisingly, the problem of finding the area under a curve and the geometric relationship between displacement and velocity are related to one another.

From the study of differentiation, we know that if $s(t)$ represents the *displacement* (or position) of an object at time t and $v(t)$ represents its *velocity*, then

$$\frac{ds}{dt} = s'(t) = v(t)$$

In other words, *the derivative of the displacement (position) function is the velocity function*. By implementing the method we used to calculate area in the last section, we will now see that another relationship exists between displacement and velocity.

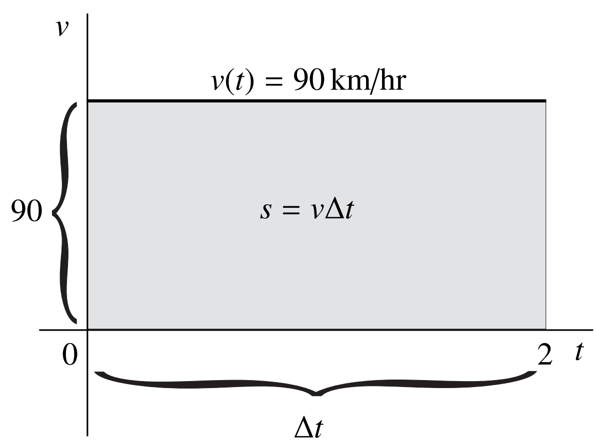
Suppose that we are going to take a trip in a car along a highway. Our task is to determine how far we have travelled after two hours. Unfortunately, the odometer in the car is broken. However, the speedometer is in working condition.

With proper planning, the data from the speedometer can be used to help estimate how far we travelled. To see that this is plausible, suppose that we always travel forward on the highway at a *constant* velocity, say 90 km/hr. We know from basic physics that provided our velocity is constant, if $s = \textit{displacement}$, $v = \textit{velocity}$, and $\Delta t = \textit{time elapsed}$, then

$$s = v\Delta t$$

In our case, the velocity is $v = 90$ km/hr and the elapsed time is $\Delta t = 2$ hrs. Hence, the displacement (or distance travelled since we are always moving forward), is

$$s = v\Delta t = 90 \text{ km/hr} \times 2 \text{ hrs} = 180 \text{ km}$$



Notice that the velocity function $v(t) = 90$ is a constant function and so its graph is a horizontal line. The area below this constant function is just a rectangle with length from $t = 0$ to $t = 2$.

Hence, the area below the curve $v(t) = 90$ on the interval $[0, 2]$ is

$$v \times \Delta t = 90 \text{ km/hr} \times 2 \text{ hrs} = 180 \text{ km} = s$$

Notice that the

area below the constant velocity function is equal to the displacement

during time Δt!

Unfortunately, we will drive through some small towns with heavy traffic so traveling at a constant velocity over the entire trip is impossible. How can we calculate the distance travelled if our velocity varies over the 2 hour duration of the trip?

Consider that it is still reasonable to assume that over any very short interval of time our velocity will be relatively constant.

Let us proceed in the following familiar manner. First we will assume our velocity is always positive and so in this case our displacement will be equal to the distance travelled.

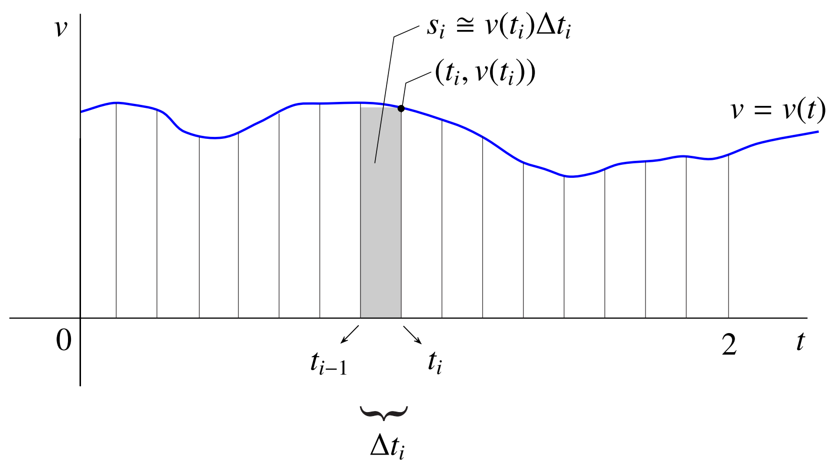
Next let's separate the 2 hour duration of the trip into 120 one minute intervals

$$0 = t_0 < t_1 < t_2 < t_3 < \cdots < t_{i-1} < t_i < \cdots < t_{120} = 2 \text{ hours}$$

so that $t_i = i$ minutes $= \frac{i}{60}$ hours. Let s_i be the distance travelled during time t_{i-1} until t_i. In other words, each s_i is the distance travelled in the i^{th} minute of our trip. Then if s is the total displacement (distance travelled), we have

$$\begin{aligned} s &= (\textit{distance travelled in } 1^{st} \textit{ minute}) + (\textit{distance travelled in } 2^{nd} \textit{ minute}) + \cdots \\ &\quad \ldots + (\textit{distance travelled in } i^{th} \textit{ minute}) + \cdots + (\textit{distance travelled in } 120^{th} \textit{ minute}) \\ &= s_1 + s_2 + s_3 + \cdots + s_i + \cdots + s_{120} \\ &= \sum_{i=1}^{120} s_i \end{aligned}$$

Let $v(t)$ be the function that represents the velocity at time t along the trip, again assuming that $v(t) > 0$. At the end of each minute t_i, the velocity on the speedometer is recorded. That is, $v(t_i)$ is determined.



Next let Δt_i denote the elapsed time between t_{i-1} and t_i so that $\Delta t_i = t_i - t_{i-1}$. However, each interval has the same elapsed time, namely $\frac{1}{60}$ of an hour (or 1 minute). Since it makes sense to assume that the velocity does not vary much over any one minute period, we can assume that the velocity during the interval $[t_{i-1}, t_i]$ was the same as it was at t_i. From this assumption, the previous formula ($s = v\Delta t$) is used to estimate s_i so that

$$s_i \cong v(t_i)\Delta t_i = v(t_i)\frac{1}{60}$$

Finally, we have the estimate for s:

$$s = \sum_{i=1}^{120} s_i \cong \sum_{i=1}^{120} v(t_i)\Delta t_i = \sum_{i=1}^{120} v(t_i)\frac{1}{60}$$

To find an even better estimate, we could measure the velocity every second. This means we would divide the two hour period into equal subintervals $[t_{i-1}, t_i]$ each of length $\frac{1}{3600}$ hours (i.e., 1 hr $\times$ 60 min/hr $\times$ 60 sec/min = 3600 sec/hr and 2 hrs $\times$ 3600 seconds/hr = 7200 seconds). We again let s_i denote the distance travelled over the i^{th} interval. This time we have

$$s_i \cong v(t_i)\Delta t_i = v(t_i)\frac{1}{3600}$$

and

$$s = \sum_{i=1}^{7200} s_i \cong \sum_{i=1}^{7200} v(t_i)\Delta t_i = \sum_{i=1}^{7200} v(t_i)\frac{1}{3600}$$

In fact, for any Natural number $n > 0$, we can divide the interval $[0, 2]$ into n equal parts of length $\frac{2}{n}$ by choosing

$$0 = t_0 < t_1 < t_2 < t_3 < \cdots < t_{i-1} < t_i < \cdots < t_n = 2$$

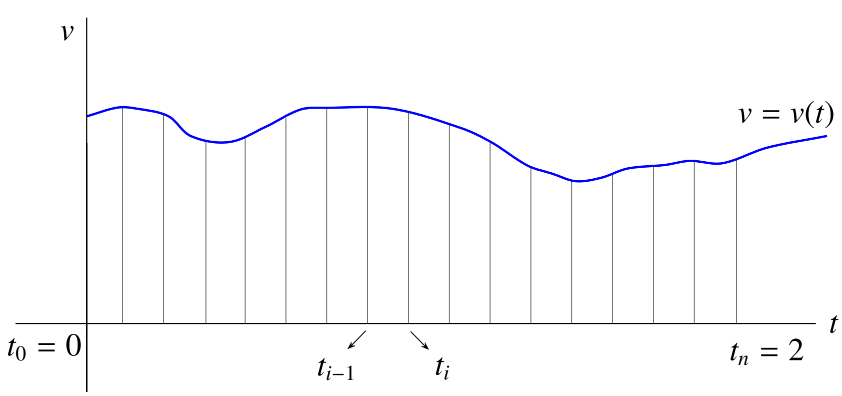
where $t_i = \frac{2i}{n}$ for each $i = 1, 2, 3, \ldots, n$. If we let

$$S_n = \sum_{i=1}^{n} v(t_i)\Delta t_i = \sum_{i=1}^{n} v(t_i)\frac{2}{n}$$

then it can be shown that the sequence $\{S_n\}$ converges and that

$$\lim_{n\to\infty}\{S_n\} = s$$

Let's consider what this last statement means geometrically. The diagram shows the graph of velocity as a function of time over the interval $[0, 2]$ partitioned into n equal subintervals.



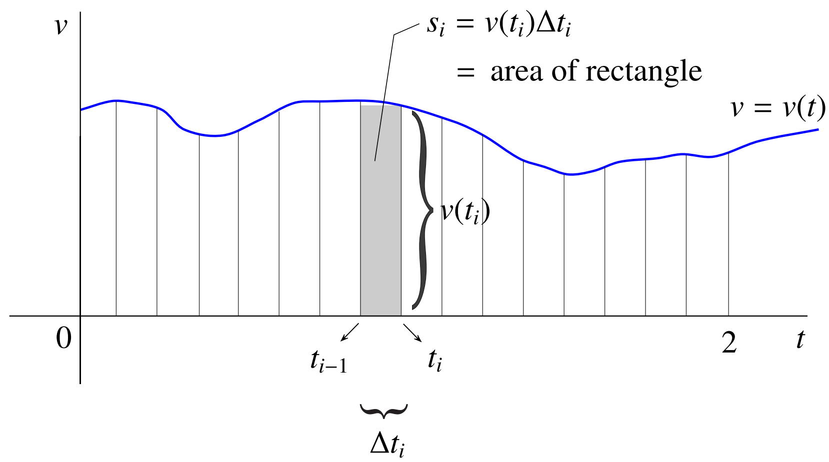
We have that

$$s_i \cong v(t_i)\Delta t_i$$

but

$$v(t_i)\Delta t_i$$

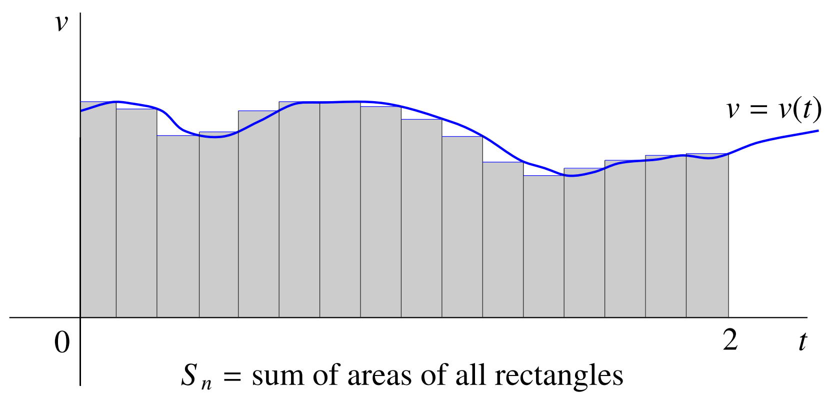
is just the area of the shaded rectangle with height $v(t_i)$ and length Δt_i.



Moreover, if

$$S_n = \sum_{i=1}^{n} v(t_i)\Delta t_i,$$

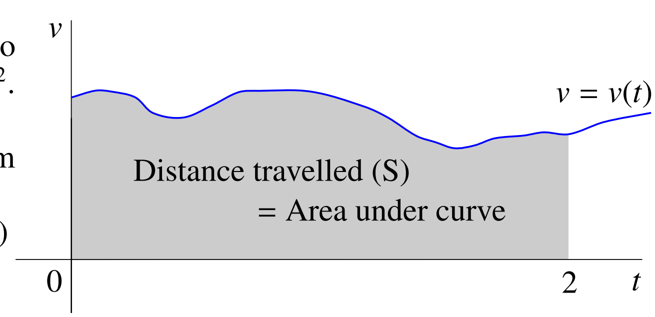
then S_n is the sum of the areas of all of the rectangles in the diagram.



Notice that S_n closely approximates the area bounded by the graph of $v = v(t)$, the t-axis, the line $t = 0$ and the line $t = 2$. If n approaches ∞, we are once again led to conclude that the

displacement (distance travelled) equals the area under the graph of the velocity function.

Note: This is the same process we used to find the area under the graph of $f(x) = x^2$. This example shows geometrically that the displacement (distance travelled) from t_{i-1} through t_i is equal to the area under the graph of the velocity function $v = v(t)$ bounded by the t-axis, $t = t_{i-1}$ and $t = t_i$.



1.2 Riemann Sums and the Definite Integral

In this section, the notion of a *Riemann sum* is introduced and it is used to define the *definite integral*.[1]

Suppose that we have a function f that is bounded on a closed interval $[a, b]$. We begin the construction of a Riemann sum by first choosing a *partition* P for the interval $[a, b]$. By a partition we mean a finite increasing sequence of numbers of the form

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b.$$

This partition subdivides the interval $[a, b]$ into n subintervals

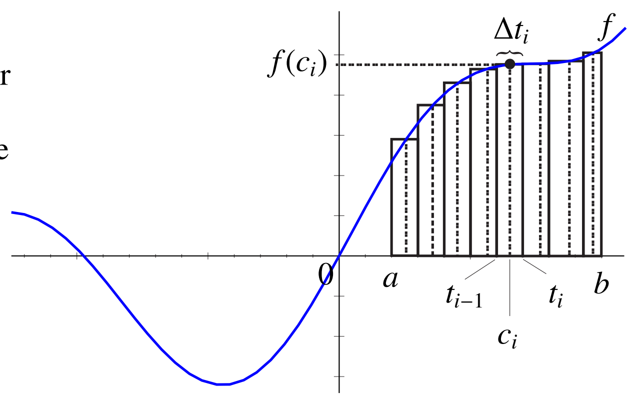
$$[t_0, t_1], [t_1, t_2], \cdots, [t_{i-1}, t_i], \cdots, [t_{n-2}, t_{n-1}], [t_{n-1}, t_n].$$

Note that these subintervals need not be equal in length.

For each $i = 1, 2, \ldots, n$, the length of the *ith* subinterval $[t_{i-1}, t_i]$ is denoted by Δt_i. In other words, $\Delta t_i = t_i - t_{i-1}$.
The *norm of the partition P* is the length of the widest subinterval which we denote by

$$||P|| = \max\{\Delta t_1, \Delta t_2, \ldots, \Delta t_n\}.$$

Now for each $i = 1, 2, \ldots, n$, a point $c_i \in [t_{i-1}, t_i]$ is chosen.[2]



Given these conditions, we can now define a Riemann sum for the partition P.

DEFINITION

Riemann Sum

Given a bounded function f on $[a, b]$, a partition P

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

of $[a, b]$, and a set $\{c_1, c_2, \ldots, c_n\}$ where $c_i \in [t_{i-1}, t_i]$, then a *Riemann sum* for f with respect to P is a sum of the form

$$S = \sum_{i=1}^{n} f(c_i)\Delta t_i.$$

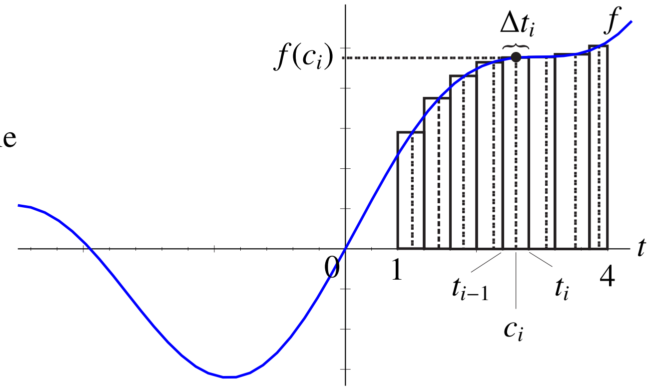
---

[1] Riemann sums are named after the German mathematician Georg Friedrich Bernhard Riemann (1826-1866) who worked on the theory of integration among many other accomplishments in analysis, number theory and geometry.

[2] For a generic Riemann sum, the widths of the subintervals do *not* have to be equal to one another and each c_i need not be the midpoint of each subinterval.

The next diagram represents a Riemann sum for a function f defined on the interval $[1, 4]$.

Since the function is positive on this interval, the terms $f(c_i)\Delta t_i$ represent the area of the rectangle with length equal to the subinterval $[t_{i-1}, t_i]$ and height given by $f(c_i)$. In the diagram, the dashed lines represent the location of the points c_i.



Notice the similarity between these sums and the sums we used in the previous section to determine the area under the graph of $f(x) = x^2$. This similarity occurs because the latter sums were actually special types of Riemann sums.

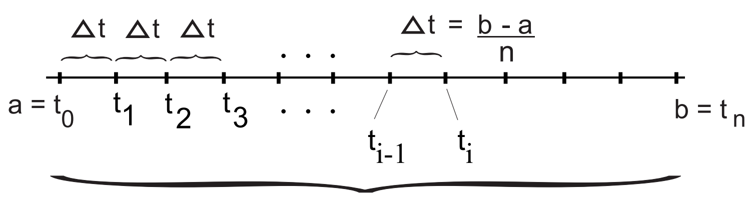
DEFINITION

Regular n-Partition

Given an interval $[a, b]$ and an $n \in \mathbb{N}$, the *regular n−partition* of $[a, b]$ is the partition $P^{(n)}$ with

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

of $[a, b]$ where each subinterval has the *same* length $\Delta t_i = \frac{b-a}{n}$.



"n" equal subintervals in [a,b]

In this case,

$$\begin{aligned}
a = t_0 &= a + 0 \cdot (\frac{b-a}{n}), \\
t_1 &= a + 1 \cdot (\frac{b-a}{n}), \\
t_2 &= a + 2 \cdot (\frac{b-a}{n}), \\
&\vdots \\
t_i &= a + i \cdot (\frac{b-a}{n}), \\
&\vdots \\
t_n &= a + n \cdot (\frac{b-a}{n}) = b.
\end{aligned}$$

DEFINITION

Right-hand Riemann Sum

The *right-hand Riemann sum* for f with respect to the partition P is the Riemann sum R obtained from P by choosing c_i to be t_i, the right-hand endpoint of $[t_{i-1}, t_i]$. That is

$$R = \sum_{i=1}^{n} f(t_i)\Delta t_i.$$

If $P^{(n)}$ is the regular n-partition, we denote the right-hand Riemann sum by

$$\begin{aligned} R_n = \sum_{i=1}^{n} f(t_i)\Delta t_i &= \sum_{i=1}^{n} f(t_i)\frac{b-a}{n} \\ &= \sum_{i=1}^{n} f\left(a + i\left(\frac{b-a}{n}\right)\right)\left(\frac{b-a}{n}\right) \end{aligned}$$

DEFINITION

Left-hand Riemann Sum

The *left-hand Riemann sum* for f with respect to the partition P is the Riemann sum L obtained from P by choosing c_i to be t_{i-1}, the left-hand endpoint of $[t_{i-1}, t_i]$. That is

$$L = \sum_{i=1}^{n} f(t_{i-1})\Delta t_i.$$

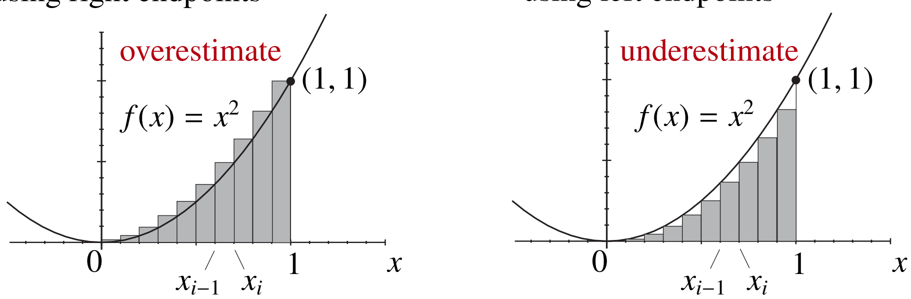
If $P^{(n)}$ is the regular n-partition, we denote the left-hand Riemann sum by

$$\begin{aligned} L_n = \sum_{i=1}^{n} f(t_{i-1})\Delta t_i &= \sum_{i=1}^{n} f(t_{i-1})\frac{b-a}{n} \\ &= \sum_{i=1}^{n} f\left(a + (i-1)\left(\frac{b-a}{n}\right)\right)\left(\frac{b-a}{n}\right) \end{aligned}$$

EXAMPLE 1

(a) Right-hand Riemann sum using right endpoints

(b) Left-hand Riemann sum using left endpoints



A closer look at the examples in the previous section reveals that the sums used to estimate the area under the graph of $f(x) = x^2$ were right-hand Riemann sums. Similarly, the sums used to find the distance travelled were right-hand Riemann sums of the velocity function $v(t)$. Moreover, we saw that if we let n approach ∞, then these sequences of Riemann sums converged to the area under the graph of $f(x) = x^2$ and the total distance travelled, respectively. These examples motivate the following definition:

DEFINITION

Definite Integral

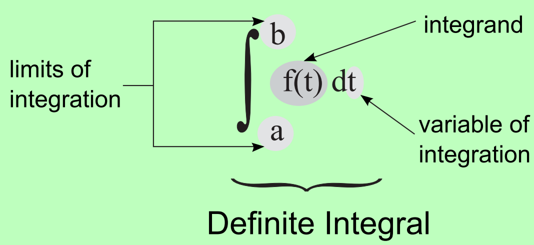
We say that a bounded function f is *integrable* on $[a, b]$ if there exists a unique number $I \in \mathbb{R}$ such that if whenever $\{P_n\}$ is a sequence of partitions with $\lim_{n\to\infty} \|P_n\| = 0$ and $\{S_n\}$ is any sequence of Riemann sums associated with the P_n's, we have

$$\lim_{n\to\infty} S_n = I.$$

In this case, we call I the integral of f over $[a, b]$ and denote it by[3]

$$\int_a^b f(t)\, dt$$

The points a and b are called the *limits of integration* and the function $f(t)$ is called the *integrand*. The variable t is called the *variable of integration*.



NOTE

The variable of integration is sometimes called a *dummy variable* in the sense that if we were to replace t's by x's everywhere, we would not change the value of the integral. ◀

It might seem difficult to find such a number I or even to know if it exists. The next result tells us that if f is continuous on $[a, b]$, then it is integrable. It also shows that the integral can be obtained as a limit of Riemann sums associated with the regular n-partitions.

[3]In the 17th century, Gottfried Wilhelm Leibniz introduced the notation $\int$ for the integral sign which represents an elongated S from the Latin word *summa*.

THEOREM 1

Integrability Theorem for Continuous Functions

Let f be continuous on $[a, b]$. Then f is integrable on $[a, b]$. Moreover,

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} S_n$$

where

$$S_n = \sum_{i=1}^{n} f(c_i)\Delta t_i$$

is any Riemann sum associated with the regular n-partitions. In particular,

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} R_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i)\frac{b-a}{n}$$

and

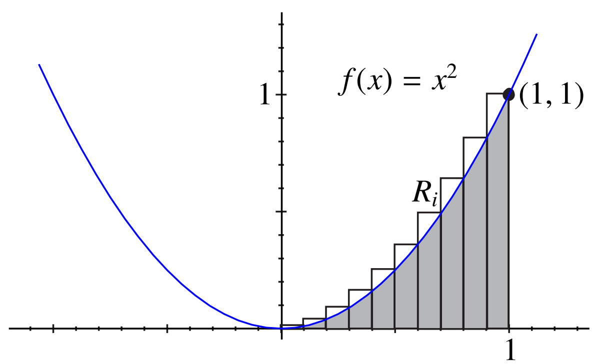
$$\int_a^b f(t)\,dt = \lim_{n\to\infty} L_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_{i-1})\frac{b-a}{n}$$

REMARK

This theorem also holds if f is bounded and has finitely many discontinuities on $[a, b]$. The proof of this theorem is beyond the scope of this course. ◀

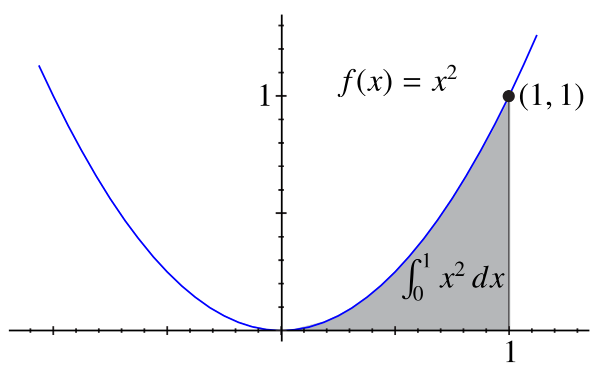
EXAMPLE 2 We have already seen that if $f(x) = x^2$ on $[0, 1]$, then

$$\begin{aligned} R_n &= \frac{1}{n^3}\sum_{i=1}^{n} i^2 \\ &= \frac{1}{n^3}\frac{(n)(n+1)(2(n)+1)}{6} \\ &= \frac{2+\frac{3}{n}+\frac{1}{n^2}}{6} \end{aligned}$$



It follows that

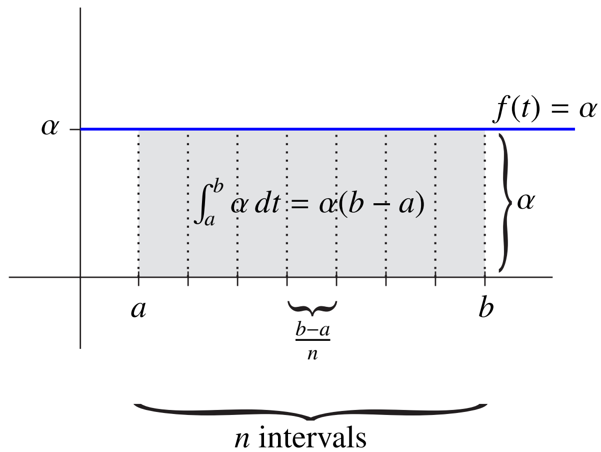
$$\begin{aligned} \int_0^1 x^2\,dx &= \lim_{n\to\infty} \frac{2+\frac{3}{n}+\frac{1}{n^2}}{6} \\ &= \frac{2}{6} \\ &= \frac{1}{3} \end{aligned}$$



Soon we will see how to calculate integrals by means other than using limits of Riemann sums. However, before ending this section, consider the following important example.

EXAMPLE 3 Let $f(t) = \alpha$ for each $t \in [a, b]$.
Then

$$\begin{aligned} R_n &= \sum_{i=1}^{n} f(t_i)\Delta t_i \\ &= \sum_{i=1}^{n} \alpha(\frac{b-a}{n}) \\ &= \alpha \sum_{i=1}^{n} \frac{b-a}{n} \\ &= \alpha(b-a) \end{aligned}$$



Since $R_n = \alpha(b - a)$ for each n, it follows that $\displaystyle\int_a^b \alpha\, dt = \alpha(b - a)$.

In other words, if f is any constant function (for example, α), then the integral of f over the limits of integration from a to b is just α times the length of the interval $[a, b]$ or $\alpha(b - a)$.

1.3 Properties of the Definite Integral

Since the integral is a limit of a sequence, we would expect many of the limit laws to hold. The next theorem shows that this is indeed the case.

THEOREM 2 **Properties of Integrals**

Assume that f and g are integrable on the interval $[a, b]$. Then:

i) For any $c \in \mathbb{R}$, $\int_a^b c\, f(t)\, dt = c \int_a^b f(t)\, dt$.

ii) $\int_a^b (f + g)(t)\, dt = \int_a^b f(t)\, dt + \int_a^b g(t)\, dt$.

iii) If $m \le f(t) \le M$ for all $t \in [a, b]$, then $m(b - a) \le \int_a^b f(t)\, dt \le M(b - a)$.

iv) If $0 \le f(t)$ for all $t \in [a, b]$, then $0 \le \int_a^b f(t)\, dt$.

v) If $g(t) \le f(t)$ for all $t \in [a, b]$, then $\int_a^b g(t)\, dt \le \int_a^b f(t)\, dt$.

vi) The function $|f|$ is integrable on $[a, b]$ and $|\int_a^b f(t)\, dt| \le \int_a^b |f(t)|\, dt$.

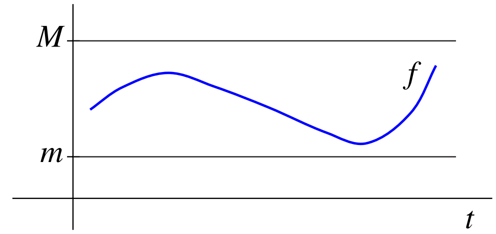
Properties (i) and (ii) in the previous theorem follow immediately from the rules of arithmetic for convergent sequences. Property (iv) can be deduced from Property (iii) and Property (v) can be obtained from Properties (i), (ii) and (iv).

Let's consider why Property (iii) is true.

Assume that

$$m \leq f(t) \leq M$$

for all $t \in [a, b]$.

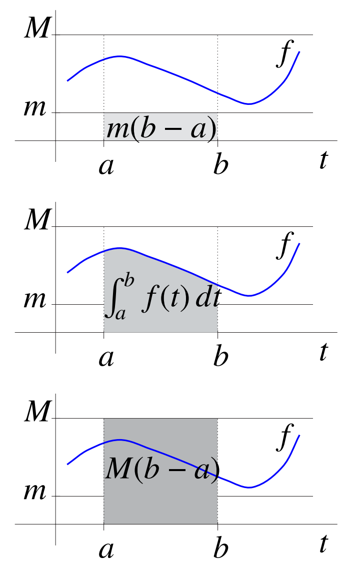


Let

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

be any partition of $[a, b]$. We first observe that $\sum\limits_{i=1}^{n} \Delta t_i = b - a$. Then since $m \leq f(t_i) \leq M$,

$$m(b-a) = \sum_{i=1}^{n} m\Delta t_i \leq \sum_{i=1}^{n} f(t_i)\Delta t_i \leq \sum_{i=1}^{n} M\Delta t_i = M(b-a).$$



It then follows

$$m(b-a) \leq \int_a^b f(t)\,dt \leq M(b-a)$$

as expected.

Property (vi) can be derived by applying the triangle inequality to the Riemann sums associated with $\int_a^b f(t)\,dt$.

1.3.1 Additional Properties of the Integral

Up until now, in defining the definite integral we have always considered integrals of the form

$$\int_a^b f(t)\,dt$$

where $a < b$. However, it is necessary to give meaning to

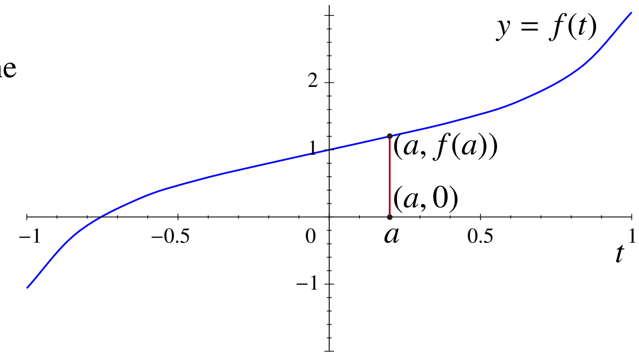
$$\int_a^a f(t)\,dt$$

and to

$$\int_b^a f(t)\,dt.$$

How do we define $\int_a^a f(t)\,dt$?

If we assume that $f(a) > 0$, we can again view this integral as the "area" of the region below the graph of $y = f(t)$, but this time the interval is "from $x = a$ to $x = a$." This is a *degenerate* rectangle that is just the line segment joining $(a, 0)$ and $(a, f(a))$.



We can see that the line segment has height $f(a)$ but length 0. As such it makes sense to define its "area" to be 0. In keeping with our theme that the integral of a positive function represents area, we are led to the following definition.

DEFINITION $\int_a^a f(t)\,dt$ **[Identical Limits of Integration]**

Let $f(t)$ be defined at $t = a$. Then we define

$$\int_a^a f(t)\,dt = 0.$$

Recall the convention that moving to the right represents a positive amount and moving to the left represents a negative amount. In the definition of

$$\int_a^b f(t)\,dt$$

where $a < b$, we began at the left-hand endpoint a of an interval $[a, b]$ and moved to the right towards b. In the case of the integral

$$\int_b^a f(t)\,dt$$

where $a < b$, we are suggesting that using the interval $[a, b]$ we move from b to the left towards a. This is the opposite or *negative* of the original orientation. For this reason, we define:

DEFINITION $\int_b^a f(t)\,dt$ **[Switching the Limits of Integration]**

Let f be integrable on the interval $[a, b]$ where $a < b$. Then we define

$$\int_b^a f(t)\,dt = -\int_a^b f(t)\,dt.$$

EXAMPLE 4 Recall that we have already seen

$$\int_0^1 x^2\,dx = \frac{1}{3}.$$

It follows that

$$\int_1^0 x^2\,dx = -\frac{1}{3}.$$

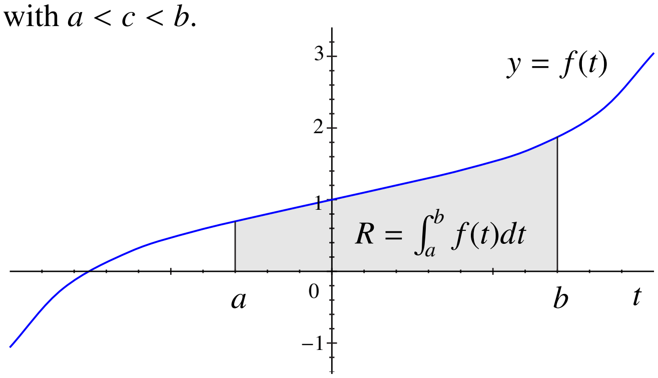
◀

The next property shows us how to separate an integral over one region into the sum of integrals over two or more regions. To motivate this rule, assume that f is continuous and positive on $[a,b]$ with $a<c<b$.

We know that the integral

$$\int_a^b f(t)\,dt$$

represents the area of the region R bounded by the graph of $y=f(t)$, the t-axis, and the lines $t=a$ and $t=b$.



However, the line $t=c$ separates the region R into two subregions, which we denote by R_1 and R_2.

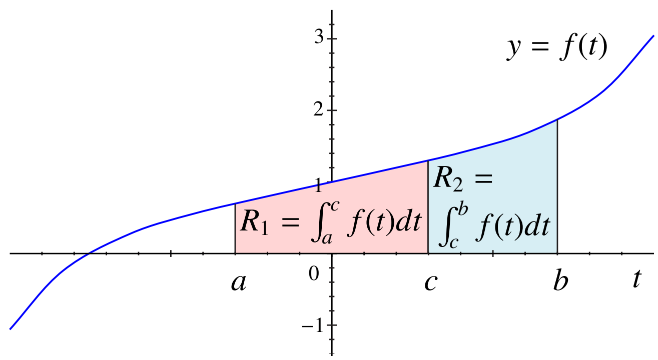
We also have that

$$\int_a^c f(t)\,dt$$

and

$$\int_c^b f(t)\,dt$$

represent the areas of regions R_1 and R_2, respectively.



The diagrams show that the area of R is the sum of the areas of R_1 and R_2. In other words, $R = R_1 + R_2$. But this suggests that

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt.$$

The definition of the integral can be used to prove this fact.

THEOREM 3 **Integrals over Subintervals**

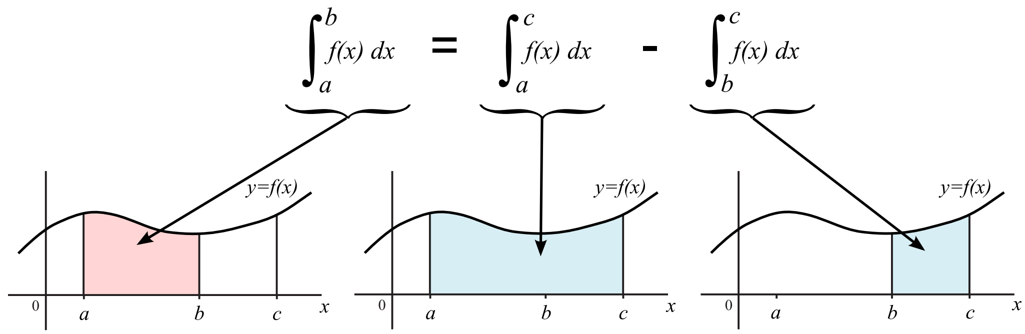
Assume that f is integrable on an interval I containing a, b and c. Then

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt.$$

Note: The proof of this theorem is not part of this course.

This theorem also holds when c lies outside the interval $[a, b]$. Consider the following example.

EXAMPLE 5 Assume that f is integrable on the interval $[a, c]$ where $a < b < c$. Then we have that



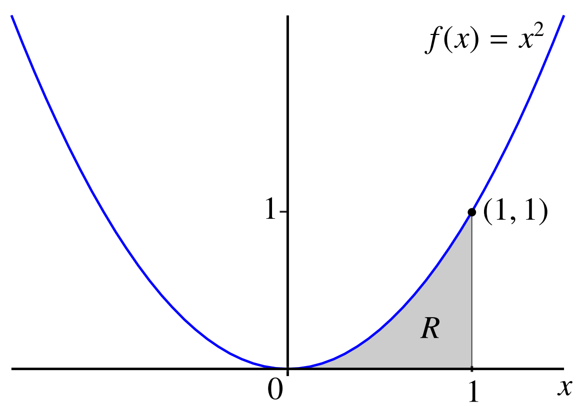
Since $\int_b^c f(x)\,dx = -\int_c^b f(x)\,dx$, then the previous theorem holds. That is,

$$\begin{aligned}\int_a^b f(x)\,dx &= \int_a^c f(x)\,dx - \int_b^c f(x)\,dx \\ &= \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.\end{aligned}$$

◀

1.3.2 Geometric Interpretation of the Integral

We have already seen from our study of Riemann sums that the area of the region R bounded by the graph of $f(x) = x^2$, by the x-axis, and by the lines $x = 0$ and $x = 1$ is $\int_0^1 x^2\,dx$.

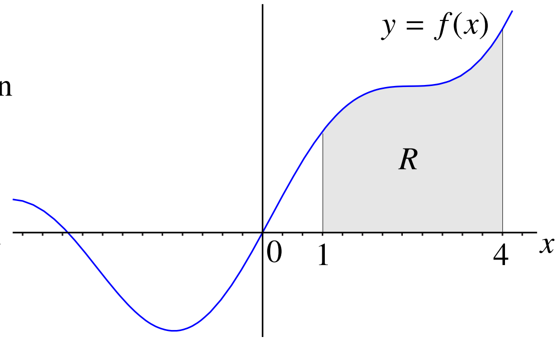


In fact, whenever $\mathbf{f(x) \geq 0}$ on all of $[a, b]$, the area under $f(x)$ and above the x-axis bounded by the lines $x = a$ and $x = b$ will be $\int_a^b f(x)\,dx$. However, what happens if

$$f(x) \leq 0$$

on some part of $[a, b]$?

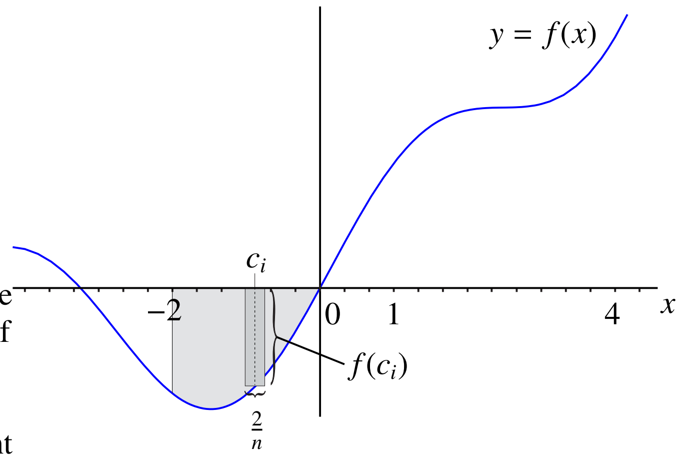
For example, assume that f is as shown in the diagram. Then $\int_1^4 f(x)\,dx$ is simply the area of the region R bounded by the graph of f, the x-axis, and the lines $x = 1$ and $x = 4$.



Suppose instead that we wanted to calculate $\int_{-2}^0 f(x)\,dx$.

Notice that the function f is *negative* on the interval $[-2, 0]$. Consider a term in a generic Riemann sum from the regular $n-$partition:

$$f(c_i)\frac{2}{n}$$



Then $\frac{2}{n}$ is the length of the rectangle (i.e., the length of the interval $[-2, 0]$ is 2 and we divide 2 into n intervals, so the length of each subinterval is $\frac{2}{n}$). However, in this case $f(c_i) < 0$ (negative) and it is the negative of the height of the rectangle since the graph of f lies below the x-axis in this interval.
It follows that the Riemann sum

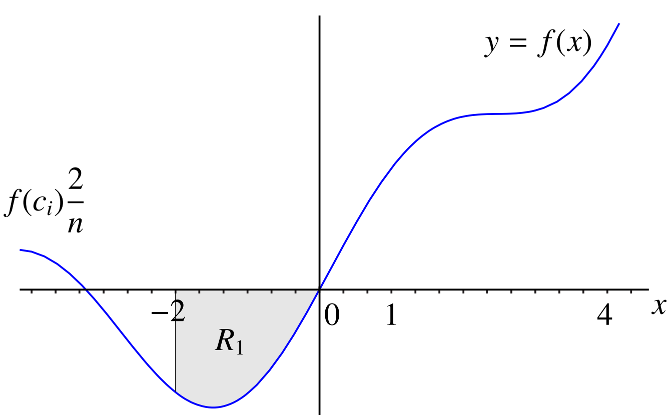
$$S_n = \sum_{i=1}^{n} f(c_i)\frac{2}{n}$$

approximates the *negative* of the area bounded by the graph of f, the x-axis, and the lines $x = -2$ and $x = 0$.

If we let $n \to \infty$, then

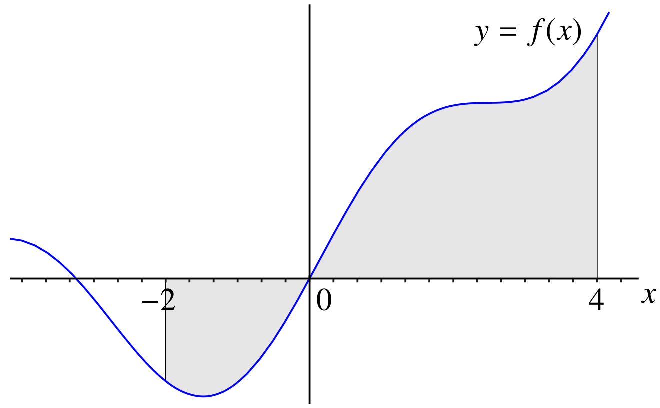
$$\int_{-2}^0 f(x)\,dx = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} f(c_i)\frac{2}{n}$$

is the *negative* of the area of the region R_1.



Suppose now that we wanted to evaluate

$$\int_{-2}^{4} f(x)\, dx.$$



Then, by the properties of the integral, we can write

$$\int_{-2}^{4} f(x)\, dx = \int_{-2}^{0} f(x)dx + \int_{0}^{4} f(x)dx.$$

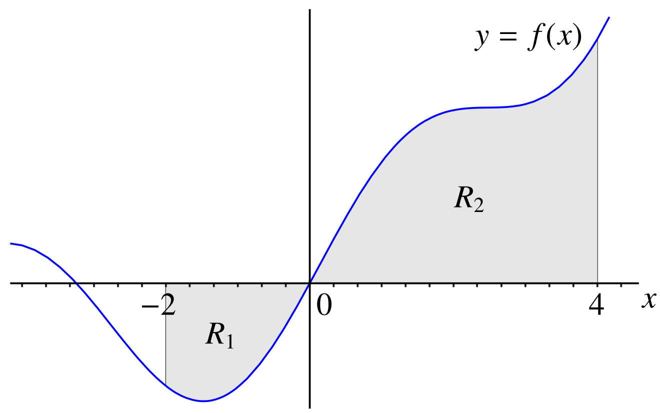
But we have just seen that

$$\int_{-2}^{0} f(x)\, dx$$

represents the *negative* of the area of region R_1 and we also know that

$$\int_{0}^{4} f(x)\, dx$$

represents the area of the region R_2.



It follows that $\int_{-2}^{4} f(x)\, dx$ represents the area of region R_2 ***minus*** the area of region R_1.

$$\int_{-2}^{4} f(x)\, dx = R_2 - R_1$$

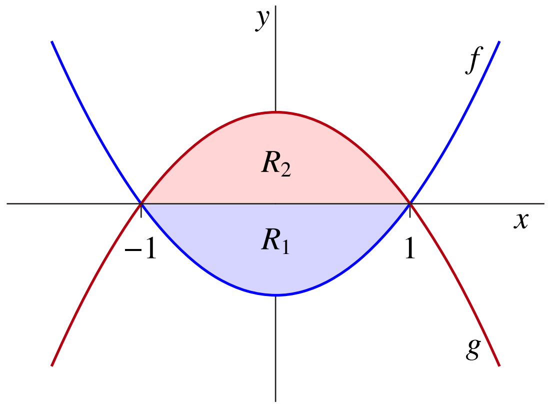
In general, if f is a continuous function on the interval $[a, b]$, then

$$\int_{a}^{b} f(x)\, dx$$

represents the area of the region under the graph of f that lies *above* the x-axis between $x = a$ and $x = b$ ***minus*** the area of the region above the graph of f that lies *below* the x-axis between $x = a$ and $x = b$.

If you are not yet convinced that for $f(x) \le 0$ on $[a, b]$, $\int_a^b f(x)\ dx$ is simply the negative of the area of the region above the graph of f, below the x-axis, and between $x = a$ and $x = b$, then the following example may convince you.

EXAMPLE 6 Consider the function f shown in the diagram and let $g(x) = -f(x)$. (In other words, g is a reflection of f in the x-axis.)



Note that since area is preserved by reflection, the area of R_1 and the area of R_2 are equal positive values since area is always a positive number. Suppose that we want to calculate $\int_{-1}^1 f(x)\ dx$. We note that $f(x) \le 0$ on $[-1, 1]$. This means that $g(x) = -f(x) \ge 0$ on $[-1, 1]$. We also have that $\int_{-1}^1 g(x)\ dx$ is equal to the area of region R_2. But

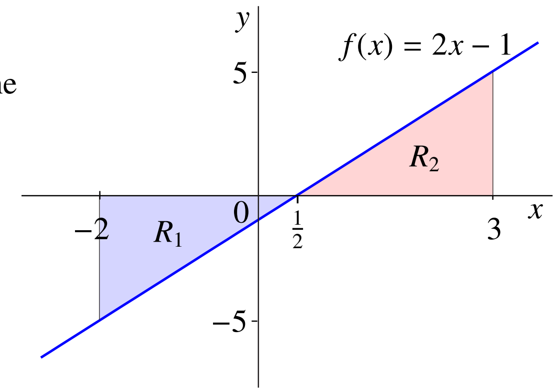
$$\begin{aligned}
\int_{-1}^1 f(x)\ dx &= \int_{-1}^1 (-g(x))\ dx \\
&= -\int_{-1}^1 g(x)\ dx \\
&= -(\text{area of region } R_2) \\
&= -(\text{area of region } R_1)
\end{aligned}$$

since the area of R_1 and the area of R_2 are equal. As such, the example shows that for $f(x) \le 0$ on $[a, b]$, $\int_a^b f(x)\ dx$ is the negative of the area of the region above the graph of f, below the x-axis, and between $x = a$ and $x = b$. ◀

EXAMPLE 7 Find $\int_{-2}^3 (2x - 1)\ dx$.

When calculating definite integrals, it is always advisable to look at the graph of the integrand, in this case $f(x) = 2x - 1$, if possible.

Since $2x - 1 = 0$ when $x = \frac{1}{2}$, the graph of f sits below the x-axis between $x = -2$ and $x = \frac{1}{2}$ and above the x-axis between $x = \frac{1}{2}$ and $x = 3$. This gives us the limits of integration for regions R_1 and R_2.



Then

$$\begin{aligned}\int_{-2}^{3}(2x-1)\,dx &= \int_{-2}^{\frac{1}{2}}(2x-1)\,dx + \int_{\frac{1}{2}}^{3}(2x-1)\,dx \\ &= -\text{area region } R_1 + \text{area region } R_2\end{aligned}$$

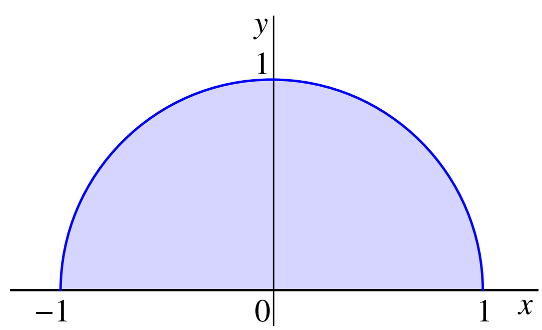
That is, this is the area of the region R_2 *minus* the area of the region R_1. The region R_1 is a right triangle with base extending from $x = -2$ to $x = \frac{1}{2}$. This means that its base is $2.5 = \frac{5}{2}$. Since $f(-2) = -5$, the diagram shows that the height of the triangle is 5. Since the area of a triangle is $\frac{1}{2}$ (*base*) $\times$ (*height*), it follows that the area of region R_1 is $\frac{1}{2}(\frac{5}{2})5 = \frac{25}{4}$. It is again the case that the base of the triangle R_2 is 2.5 and its height is $f(3) = 2(3) - 1 = 5$. It follows that the area of R_2 is also $\frac{25}{4}$ and so

$$\int_{-2}^{3}(2x-1)\,dx = \frac{-25}{4} + \frac{25}{4} = 0.$$

Notice that in this example we avoided using Riemann sums by interpreting the integral geometrically (in this case, the area of two triangles). ◀

EXAMPLE 8 Find $\int_{-1}^{1} \sqrt{1-x^2}\,dx$.

This is the area under the graph of the function $y = \sqrt{1-x^2}$.



The shape of this region is that of a semi-circle with radius 1. To see that this is the case, we note that $y^2 = 1 - x^2$ so $x^2 + y^2 = 1$. The latter equation is the equation of the circle centered at the origin with radius 1. (Since by assumption $\sqrt{1-x^2}$ is the positive square root, we are only interested in the top half of the circle.) A circle of radius 1 has area π, so this half circle has area $\frac{\pi}{2}$. It follows that

$$\int_{-1}^{1} \sqrt{1-x^2}\, dx = \frac{\pi}{2}.$$

◀

Problem!

Unfortunately, this method of evaluating integrals by identifying an easily calculated area has severe limitations. For example, we would not be able to find $\int_{-\frac{1}{2}}^{1} \sqrt{1-x^2}\, dx$ with what we know at present. Instead, there exists a powerful tool that can be used to find the integral of general functions. This tool is called *The Fundamental Theorem of Calculus*. Along with this theorem, you will be required to learn various techniques in order to integrate a variety of functions. The remainder of this chapter will focus on this task.

However, before we can state and prove the Fundamental Theorem of Calculus, we must first investigate what is meant by the *average value of a function* over an interval $[a, b]$.

1.4 The Average Value of a Function

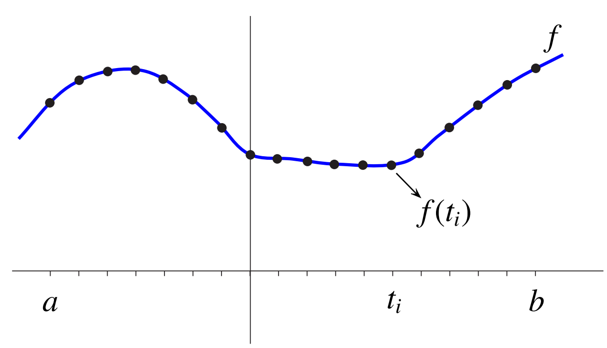
Question: What is meant by "the average value of a continuous function over an interval $[a, b]$?"

We know that the average of n real numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ is

$$\frac{\alpha_1 + \alpha_2 + \ldots + \alpha_n}{n}.$$

But how do you add all of the values of f on $[a, b]$? Is this possible?

One approach would be to take sample values of f and calculate the average of these samples as an estimate of the average value. However, we need a satisfactory method for obtaining such samples. One method to ensure that the choice of sample points is as representative as possible is to use the regular n-partition



$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$$

where $t_i = a + \dfrac{i(b-a)}{n}$ and consider

$$\frac{\sum_{i=1}^{n} f(t_i)}{n}.$$

To acquire an even better sample, more and more points need to be considered. Therefore, it might make sense to define the average of f on $[a, b]$ to be

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(t_i)}{n}$$

if this limit exists.

However, for continuous functions the limit always exists. In fact,

$$\begin{aligned}
\lim_{n\to\infty} \frac{\sum_{i=1}^{n} f(t_i)}{n} &= \lim_{n\to\infty} \frac{1}{b-a} \sum_{i=1}^{n} f(t_i)\frac{(b-a)}{n} \\
&= \frac{1}{b-a} \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i)\frac{(b-a)}{n} \\
&= \frac{1}{b-a} \lim_{n\to\infty} R_n \qquad \text{(where } R_n \text{ is the right-hand Riemann sum)} \\
&= \frac{1}{b-a} \int_a^b f(t)\, dt
\end{aligned}$$

We are led to the following definition:

DEFINITION

Average Value of f

If f is continuous on $[a, b]$, the *average value of f on* $[a, b]$ is defined as

$$\frac{1}{b-a} \int_a^b f(t)\, dt$$

1.4.1 An Alternate Approach to the Average Value of a Function

Recall that the *Extreme Value Theorem* implies that there exists m, M such that

$$m \leq f(x) \leq M$$

for all $x \in [a, b]$. Moreover, there exists $c_1, c_2 \in [a, b]$ such that $f(c_1) = m$, $f(c_2) = M$.

It make sense that the average of f on $[a, b]$ should occur between m and M. Now

$$\int_a^b m\, dx \leq \int_a^b f(x)\, dx \leq \int_a^b M\, dx$$

Therefore

$$m(b-a) \leq \int_a^b f(x)\, dx \leq M(b-a)$$

Equivalently,

$$m \leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq M$$

Let $\alpha = \dfrac{1}{b-a}\displaystyle\int_a^b f(x)\,dx$. Then

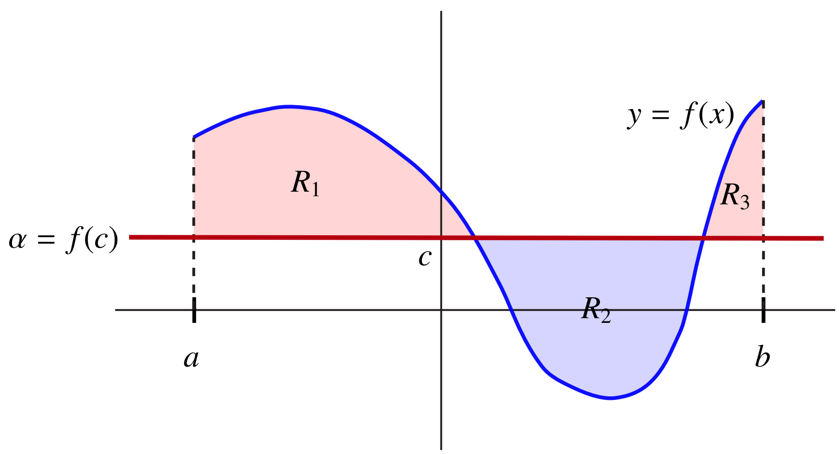
$$f(c_1) \leq \alpha \leq f(c_2).$$

By the *Intermediate Value Theorem*, there exists c between c_1 and c_2 such that

$$f(c) = \alpha = \frac{1}{b-a}\int_a^b f(x)\,dx$$

Geometrically, it follows that

$$Area\ R_1 + Area\ R_3 = Area\ R_2$$



In other words, the area above $\alpha = f(c)$ but below $y = f(x)$ equals the area below $\alpha = f(c)$ but above $y = f(x)$.

Once again it makes sense to say that

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx$$

represents the average value of the function f on $[a, b]$.

The following theorem is established.

THEOREM 4 **Average Value Theorem (Mean Value Theorem for Integrals)**

Assume that f is continuous on $[a, b]$.

Then there exists $a \leq c \leq b$ such that

$$f(c) = \frac{1}{b-a}\int_a^b f(t)\,dt$$

Important Note:

If $b < a$ and if f is continuous on $[b, a]$, then there exists $b < c < a$ with

$$\begin{aligned} f(c) &= \frac{1}{a-b}\int_b^a f(t)dt \\ &= \frac{1}{a-b}\left(-\int_a^b f(t)dt\right) \\ &= \frac{1}{b-a}\int_a^b f(t)dt \end{aligned}$$

so the Average Value Theorem holds even if $b < a$. ◀

You have now been presented with all of the background information required to understand the *The Fundamental Theorem of Calculus*. It is so named because it is one of the most important results in mathematics!

1.5 The Fundamental Theorem of Calculus (Part 1)

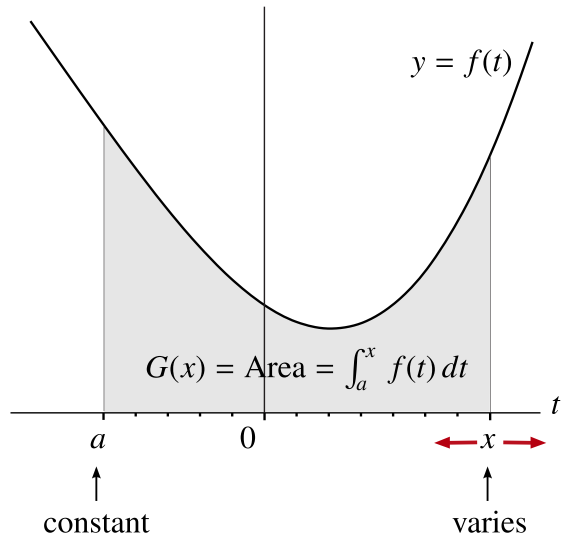
The goal in this section is to introduce the *Fundamental Theorem of Calculus* which is attributed independently to Sir Issac Newton and to Gottfried Leibniz. As the name suggests, this is perhaps the most important theorem in Calculus and many would argue, one of the most important discoveries in the history of mathematics. Despite this lofty claim, the Fundamental Theorem is at its heart a *simple rule of differentiation*. However, from this simple rule, we can derive a method that will allow us to evaluate many types of integrals without having to appeal to the complicated process involving Riemann sums. Consequently, the Fundamental Theorem of Calculus enables us to *link together differential calculus and integral calculus* in a very profound way.

Let's begin by assuming that the function f is continuous on an interval $[a, b]$.

Let's also define the *integral function*

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

What does this integral function do? If $f \geq 0$, then $G(x)$ is the *function* that calculates the area under the graph of $y = f(t)$ as x *varies* over an interval $[a, b]$ starting from a.



The objective is to determine *the rate of change in the area $G(x)$ as x changes*. In other words, to *find the derivative of the integral function $G(x)$*.

Before we consider the general case, let's look at a simple example.

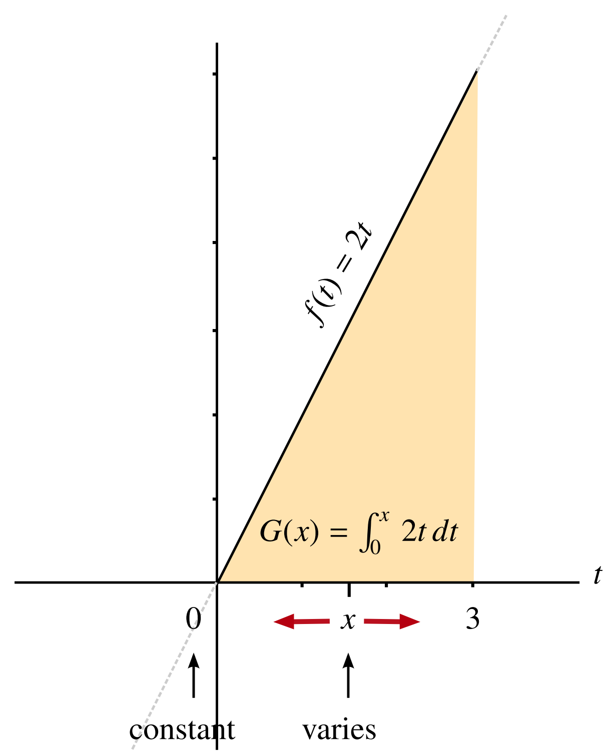
EXAMPLE 9 Let $f(t) = 2t$ on the interval $[0, 3]$. Find a formula for $G(x)$.

Recall that the integral function is defined by

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

Then the integral function for $f(t) = 2t$ starting at $a = 0$ is

$$G(x) = \int_0^x 2t\,dt.$$



In other words, $G(x)$ is the integral function that calculates the area under the curve of $f(t) = 2t$ as x *varies* over the interval $[0, 3]$ starting at $t = 0$.

Let's try to calculate the area under $f(t)$ as t varies from $x = 0$, $x = 1$, $x = 2$ and $x = 3$. We will use these values to see if we can determine $G(x)$.

Case $x = 0$:

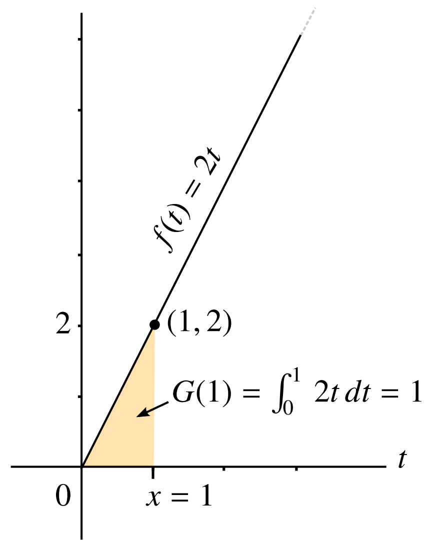
If $x = 0$, we have that $G(0) = \int_0^0 2t\,dt = 0$ since the limits of integration are identical. (There is no area to calculate.) Thus we have the area under $f(t)$ on the interval $[0, 0]$ is 0 and $G(0) = 0$.

Case $x = 1$:

If $x = 1$, we have that $G(1) = \int_0^1 2t\,dt$ and $G(1)$ is the area under the graph of $f(t) = 2t$ on the interval $[0, 1]$. We can calculate this using geometry since the area is a triangle.

$$\begin{aligned} \text{Area} = G(1) &= \int_0^1 2t\,dt \\ &= \frac{1}{2} \times \text{base} \times \text{height} \\ &= \frac{1}{2}(1)(2(1)) \\ &= 1 \end{aligned}$$

Thus we have the area under $f(t)$ on the interval $[0, 1]$ is 1 and $G(1) = 1$.

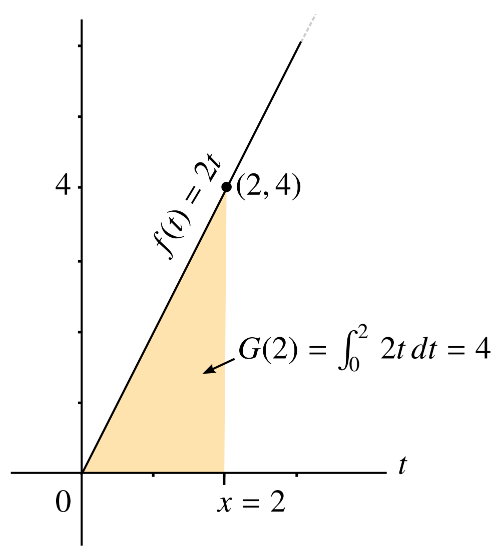


Case $x = 2$:

If $x = 2$, we have that $G(2) = \int_0^2 2t\,dt$ and $G(2)$ is the area under the graph of $f(t) = 2t$ on the interval $[0, 2]$. We can again calculate this using geometry since the area is a triangle.

$$\begin{aligned} \text{Area} = G(2) &= \int_0^2 2t\,dt \\ &= \frac{1}{2} \times \text{base} \times \text{height} \\ &= \frac{1}{2}(2)(2(2)) \\ &= 4 \end{aligned}$$

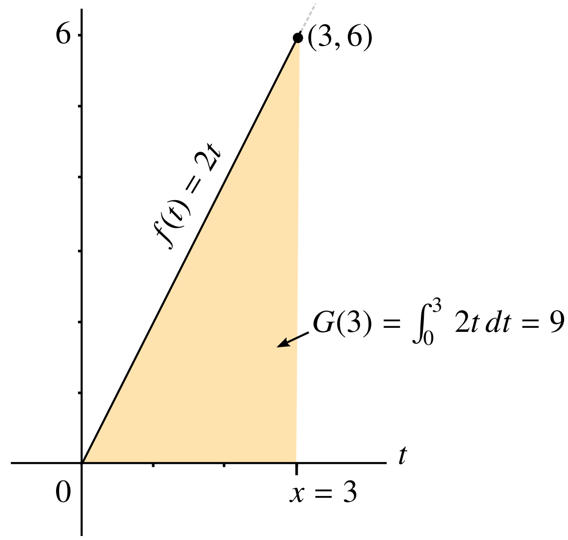
Thus we have the area under $f(t)$ on the interval $[0, 2]$ is 4 and $G(2) = 4$.



Case $x = 3$**:**

If $x = 3$, we have that $G(3) = \int_0^3 2t\,dt$ and $G(3)$ is the area under the graph of $f(t) = 2t$ on the interval $[0, 3]$. Once more we can calculate this using geometry since the area is a triangle.

$$\begin{aligned} \text{Area} = G(3) &= \int_0^3 2t\,dt \\ &= \frac{1}{2} \times \text{base} \times \text{height} \\ &= \frac{1}{2}(3)(2(3)) \\ &= 9 \end{aligned}$$



Thus we have the area under $f(t)$ on the interval $[0, 3]$ is 9 and $G(3) = 9$.

Though we could continue to allow x to vary and perform this calculation, let's consider the results for $G(x)$. They are summarized in the following table.

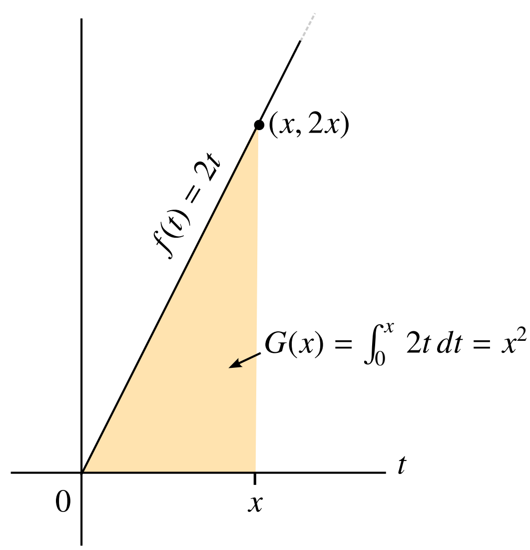
x	$G(x)$
1	1
2	4
3	9
$\vdots$	$\vdots$
x	x^2

Notice that a pattern is forming. It appears that as "x" varies, $G(x)$ takes on the value of "x^2." In fact, this is indeed the case.

Case: Generic x**:**

If $x > 0$, we have that $G(x) = \int_0^x 2t\,dt$ and $G(x)$ is the area under the graph of $f(t) = 2t$ on the interval $[0, x]$. We are still able to calculate this area using geometry since the region is a triangle.

$$\begin{aligned} \text{Area} = G(x) &= \int_0^x 2t\,dt \\ &= \frac{1}{2} \times \text{base} \times \text{height} \\ &= \frac{1}{2}(x)(2(x)) \\ &= x^2 \end{aligned}$$



Thus we have the area under $f(t)$ on the interval $[0, x]$ is x^2 and $G(x) = x^2$.

Important Observation: Notice that $G(x) = x^2$ and the derivative of G is

$G'(x) = 2x$. In other words, we have just seen that

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

This means that

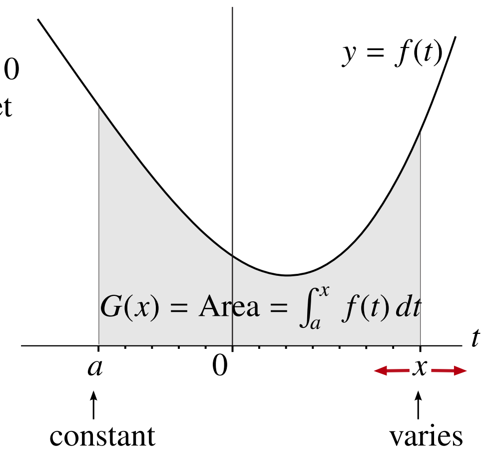
$$G'(x) = \frac{d}{dx}\int_0^x f(t)\,dt = f(x).$$

In the previous example, we were able to calculate the area geometrically because f was a linear function and the region under the graph of f was always triangular. Normally we will not have an integrand that has its area calculated so easily. We will now discuss the case where f is a generic function.

Again we begin by assuming that $f(t) \geq 0$ is continuous on the interval $[a, b]$ and let the integral function be defined by

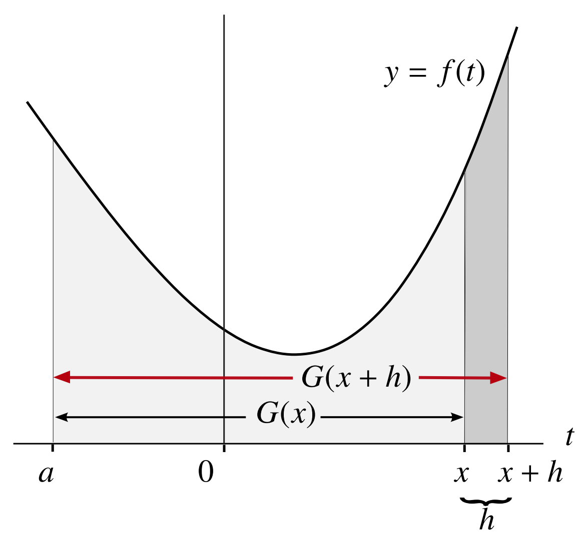
$$G(x) = \int_a^x f(t)\,dt.$$

In this case, $G(x)$ represents the area bounded by the graph of $f(t)$, the t-axis, and the lines $t = a$ and $t = x$.



The objective is to *determine the rate of change in the area $G(x)$ as x changes*. In other words, to *find the derivative $G'(x)$ of the integral function $G(x)$*.

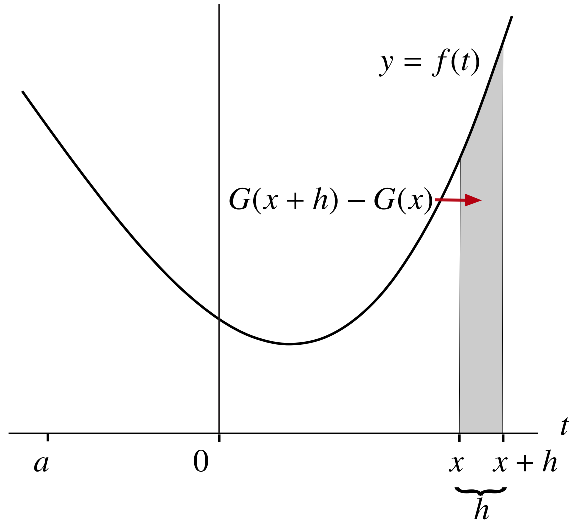
First we increment x by adding a small amount denoted by h. Then $G(x + h)$ is the area obtained by adding the first region $G(x)$ with the shaded area between the lines $t = x$ and $t = x + h$.



Next, consider $G(x+h) - G(x)$. This difference is exactly the shaded area.

It is important to remember that the area of this region can also be expressed as an integral, namely

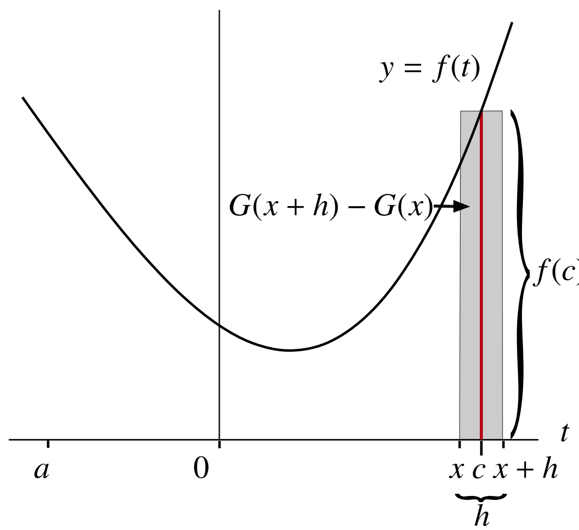
$$\int_x^{x+h} f(t)\,dt.$$



Moreover, the *Average Value Theorem* tells us that there is a c with $x < c < x+h$ such that

$$\begin{aligned}\int_x^{x+h} f(t)dt &= f(c)((x+h) - x) \\ &= f(c)h\end{aligned}$$

So the area of this rectangle is $f(c)h$.



As the diagram suggests, when h is small this means that there exists a c with $x < c < x+h$ such that

$$G(x+h) - G(x) = f(c)h$$

and hence

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

However, if h is very small, then c must also be very close to x. Since f is continuous, this means that $f(c) = \dfrac{G(x+h) - G(x)}{h}$ must be very close to $f(x)$. In fact, if we let h approach 0 from the right, we get that c must also approach x from the right. All of this together with the assumption that f is continuous gives us

$$\lim_{h\to 0^+} \frac{G(x+h) - G(x)}{h} = \lim_{c\to x^+} f(c) = f(x).$$

This is one side of the limit that defines $G'(x)$. A similar argument shows that

$$\lim_{h\to 0^-} \frac{G(x+h) - G(x)}{h} = f(x)$$

and hence that

$$\lim_{h\to 0} \frac{G(x+h)-G(x)}{h} = f(x).$$

That is,

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

Note: When we calculated the one-sided limit we made a number of assumptions. First, we assumed that f was continuous. This was used in two places. First, so that we could apply the *Average Value Theorem* to get that

$$\frac{G(x+h)-G(x)}{h} = f(c)$$

and second, to conclude that

$$\lim_{c\to x^+} f(c) = f(x).$$

The other assumptions were that $f(t) \geq 0$ and that the increment h was positive.

The assumption that f be continuous is essential, but the other two assumptions were only for our convenience and they can actually be omitted. This gives us a very simple rule of differentiation for integral functions, though a rule with a profound impact.

THEOREM 5

Fundamental Theorem of Calculus (Part 1) [FTC1]

Assume that f is continuous on an open interval I containing a point a. Let

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

Then $G(x)$ is differentiable at each $x \in I$ and

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

Equivalently,

$$G'(x) = \frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

PROOF

Assume that $G(x) = \int_a^x f(t)dt$ and that f is continuous at $x_0 \in I$. Let $\epsilon > 0$. Then there exists a $\delta > 0$ so that if $0 < |c - x_0| < \delta$, then

$$|f(c) - f(x_0)| < \epsilon.$$

Now let $0 < |x - x_0| < \delta$. Then

$$\begin{aligned}\frac{G(x) - G(x_0)}{x - x_0} &= \frac{\int_a^x f(t)dt - \int_a^{x_0} f(t)dt}{x - x_0} \\ &= \frac{1}{x - x_0} \int_{x_0}^{x} f(t)dt\end{aligned}$$

But then there exists a c between x and x_0 with

$$f(c) = \frac{1}{x - x_0} \int_{x_0}^{x} f(t)dt$$

by the Average Value Theorem.

This means that if $0 < |x - x_0| < \delta$, then since $0 < |c - x_0| < \delta$ as well

$$\left| \frac{G(x) - G(x_0)}{x - x_0} - f(x_0) \right| = |\, f(c) - f(x_0)\,| < \epsilon.$$

By the definition of a limit we get that

$$\begin{aligned}G'(x_0) &= \lim_{x \to x_0} \frac{G(x) - G(x_0)}{x - x_0} \\ &= f(x_0).\end{aligned}$$

■

NOTE

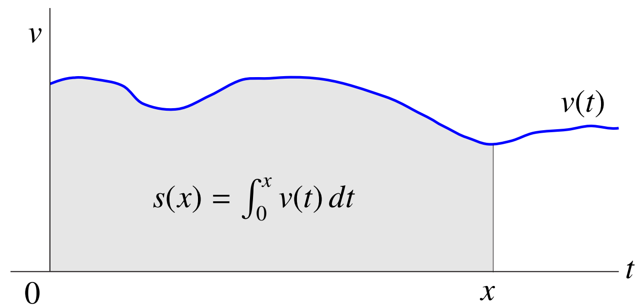
If we use Leibniz notation for derivatives, the Fundamental Theorem of Calculus (Part 1) can be written as

$$\frac{d}{dx} \int_a^x f(t)\,dt = f(x)$$

This equation roughly states that if you first integrate f and then differentiate the result, you will return back to the original function f. ◀

In the following example, a physical interpretation of the Fundamental Theorem is presented.

EXAMPLE 10 Assume that a vehicle travels forward along a straight road with a velocity at time t given by the function $v(t)$. If we fix a starting point at $t = 0$, then we saw from the section about Riemann sums that the displacement $s(x)$ up to time $t = x$ is the area under the velocity graph. That is, $s(x) = \int_0^x v(t)\,dt$.



We can assume that velocity is a continuous function of time. Consequently, the Fundamental Theorem of Calculus applies to the function $s(x)$. Moreover, the theorem tells us that $s(x)$ is differentiable and that the derivative of displacement is velocity

$$s'(x) = v(x)$$

exactly as we would expect!

EXAMPLE 11 (a) Find $F'(x)$ if $F(x) = \displaystyle\int_3^x e^{t^2}\,dt$.

Since $f(t) = e^{t^2}$ is a continuous function, the Fundamental Theorem of Calculus applies. Therefore to find $F'(x)$ we can simply replace t by x in $f(t)$ to get

$$F'(x) = e^{x^2}.$$

(b) Let's modify the previous question. Let $G(x) = \displaystyle\int_3^{x^2} e^{t^2}\,dt$. Find $G'(x)$.

This is not quite the same as the previous example. In fact, in order to find $G'(x)$ we note that

$$G(x) = F(x^2)$$

where $F(x) = \displaystyle\int_3^x e^{t^2}\,dt$. But this means we can use the Chain Rule to get that

$$G'(x) = F'(x^2)\frac{d}{dx}(x^2).$$

But to find $F'(x^2)$ we replace t by x^2 in e^{t^2}. That is

$$F'(x^2) = e^{(x^2)^2} = e^{x^4}$$

and $\frac{d}{dx}(x^2) = 2x$. It follows that if $G(x) = \displaystyle\int_3^{x^2} e^{t^2}\,dt$, then

$$G'(x) = 2xe^{x^4}.$$

There is an alternate method to solve this problem. First let $u = x^2$ since the goal is to get an integral in the form $\int_a^x$. Then

$$\begin{aligned}
G'(x) &= \frac{d}{dx}\int_3^{x^2} e^{t^2}\,dt \\
&= \frac{d}{dx}\int_3^{u} e^{t^2}\,dt \qquad \text{(substituting } u = x^2\text{)} \\
&= \frac{d}{du}\left(\int_3^{u} e^{t^2}\,dt\right)\frac{du}{dx} \qquad \text{(by the Chain Rule)} \\
&= e^{u^2}\frac{du}{dx} \qquad \text{(by FTCI)} \\
&= e^{x^4}\cdot 2x
\end{aligned}$$

which is the same answer we calculated by the previous method.

In the statement of the Fundamental Theorem of Calculus, the lower limit of the integral was always fixed. That is, it did not vary with x. We can now make our example even more complicated by letting the lower limit of the integral vary as a function of x. Let

$$H(x) = \int_{\cos(x)}^{x^2} e^{t^2}\,dt.$$

How would we find $H'(x)$?

We can cleverly use the properties of the integral. In fact, we can write

$$H(x) = \int_{\cos(x)}^{x^2} e^{t^2}\,dt = \int_{\cos(x)}^{3} e^{t^2}\,dt + \int_{3}^{x^2} e^{t^2}\,dt.$$

Furthermore, we know that

$$\int_{\cos(x)}^{3} e^{t^2}\,dt = -\int_{3}^{\cos(x)} e^{t^2}\,dt$$

and this integral is in the form where we can use the Fundamental Theorem. Therefore, we have that

$$H(x) = \int_{\cos(x)}^{x^2} e^{t^2}\,dt = \int_{3}^{x^2} e^{t^2}dt - \int_{3}^{\cos(x)} e^{t^2}dt.$$

If we now let

$$H_1(x) = \int_{3}^{\cos(x)} e^{t^2}\,dt$$

then

$$H(x) = G(x) - H_1(x)$$

where $G(x)$ is defined as before. But then

$$H'(x) = G'(x) - H_1{}'(x)$$

and we already know that
$$G'(x) = 2xe^{x^4}.$$

This means that we only need to find $H_1{}'(x)$. To accomplish this we do exactly what we did to find $G'(x)$. We note that

$$H_1(x) = F(\cos(x))$$

so

$$\begin{aligned} H_1{}'(x) &= F'(\cos(x))\frac{d}{dx}(\cos(x)) \\ &= -\sin(x)e^{(\cos(x))^2} \end{aligned}$$

Combining all of this together gives us that

$$\begin{aligned} H'(x) &= G'(x) - H_1{}'(x) \\ &= 2xe^{x^4} - (-\sin(x)e^{(\cos(x))^2}) \\ &= 2xe^{x^4} + \sin(x)e^{(\cos(x))^2} \end{aligned}$$

◀

The previous example leads us to an extended version of the Fundamental Theorem of Calculus.

THEOREM 6 **Extended Version of the Fundamental Theorem of Calculus**

Assume that f is continuous and that g and h are differentiable. Let

$$H(x) = \int_{g(x)}^{h(x)} f(t)\,dt.$$

Then $H(x)$ is differentiable and

$$H'(x) = f(h(x))h'(x) - f(g(x))g'(x).$$

1.6 The Fundamental Theorem of Calculus (Part 2)

We have seen that the Fundamental Theorem of Calculus provides us with a *simple rule for differentiating integral functions* and so it provides the key link between differential and integral calculus. However, we will soon see it also provides us with a *powerful tool for evaluating integrals*. First we must briefly review the topic of antiderivatives from your study of differential calculus.

1.6.1 Antiderivatives

We know a number of techniques for calculating derivatives. In this section, we will review how we can sometimes "undo" differentiation. That is, given a function f, we will look for a new function F with the property that $F'(x) = f(x)$.

DEFINITION

Antiderivative

Given a function f, an *antiderivative* is a function F such that

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

If $F'(x) = f(x)$ for all x in an interval I, we say that F is an antiderivative for f on I.

EXAMPLE 12 Let $f(x) = x^3$. Let $F(x) = \frac{x^4}{4}$. Then

$$F'(x) = \frac{4x^{4-1}}{4} = x^3 = f(x),$$

so $F(x) = \frac{x^4}{4}$ is an antiderivative of $f(x) = x^3$. ◀

While the derivative of a function is always unique, this is *not* true of antiderivatives. In the previous example, if we let $G(x) = \frac{x^4}{4} + 2$, then $G'(x) = x^3$. Therefore, both $F(x) = \frac{x^4}{4}$ and $G(x) = \frac{x^4}{4} + 2$ are antiderivatives of the same function $f(x) = x^3$.

This holds in greater generality: if F is an antiderivative of a given function f, then so is $G(x) = F(x) + C$ for every $C \in \mathbb{R}$. A question naturally arises–are these all of the antiderivatives of f?

To answer this question, we appeal to the Mean Value Theorem. Assume that F and G are both antiderivatives of a given function f. Let

$$H(x) = G(x) - F(x).$$

Then

$$\begin{aligned} H'(x) &= G'(x) - F'(x) \\ &= f(x) - f(x) \\ &= 0 \end{aligned}$$

for every x.

The Mean Value Theorem showed that there exists a constant C such that

$$H(x) = G(x) - F(x) = C.$$

But this means that

$$G(x) = F(x) + C.$$

It follows that once we have one antiderivative F of a function f, we can find all of the antiderivatives by considering all functions of the form

$$G(x) = F(x) + C.$$

EXAMPLE 13 Let $f(x) = x^3$. Find all of the antiderivatives of f.

We have already seen that $F(x) = \frac{x^4}{4}$ is an antiderivative of x^3. It follows that the family of all antiderivatives consists of functions of the form

$$G(x) = \frac{x^4}{4} + C$$

for $C \in \mathbb{R}$. ◀

Notation: We will denote the *family of antiderivatives* of a function f by

$$\int f(x)\,dx.$$

For example,

$$\int x^3\,dx = \frac{x^4}{4} + C.$$

The symbol

$$\int f(x)\,dx$$

is called the ***in**definite integral of* f. The function f is called the *integrand*.

We will be content to find the antiderivatives of many of the basic functions that we will use in this course. The next theorem tells us how to find the antiderivatives of one of the most important classes of functions, the powers of x.

THEOREM 7 **Power Rule for Antiderivatives**

If $\alpha \neq -1$, then

$$\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + C.$$

To see that this theorem is correct we need only differentiate. Since

$$\frac{d}{dx}\left(\frac{x^{\alpha+1}}{\alpha+1} + C\right) = x^\alpha,$$

we have found all of the antiderivatives.

The following table lists the antiderivatives of several basic functions. You can use differentiation to verify each antiderivative.

Integrand	Antiderivative
$f(x) = x^n \quad$ where $n \neq -1$	$\int x^n\,dx = \dfrac{x^{n+1}}{n+1} + C$
$f(x) = \dfrac{1}{x}$	$\int \dfrac{1}{x}\,dx = \ln(\mid x \mid) + C$
$f(x) = e^x$	$\int e^x\,dx = e^x + C$
$f(x) = \sin(x)$	$\int \sin(x)\,dx = -\cos(x) + C$
$f(x) = \cos(x)$	$\int \cos(x)\,dx = \sin(x) + C$
$f(x) = \sec^2(x)$	$\int \sec^2(x)\,dx = \tan(x) + C$
$f(x) = \dfrac{1}{1+x^2}$	$\int \dfrac{1}{1+x^2}\,dx = \arctan(x) + C$
$f(x) = \dfrac{1}{\sqrt{1-x^2}}$	$\int \dfrac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x) + C$
$f(x) = \dfrac{-1}{\sqrt{1-x^2}}$	$\int \dfrac{-1}{\sqrt{1-x^2}}\,dx = \arccos(x) + C$
$f(x) = \sec(x)\tan(x)$	$\int \sec(x)\tan(x)\,dx = \sec(x) + C$
$f(x) = a^x \quad$ where $a > 0$ and $a \neq 1$	$\int a^x\,dx = \dfrac{a^x}{\ln(a)} + C$

1.6.2 Evaluating Definite Integrals

Suppose that we want to evaluate $\int_0^2 t^3\,dt$. At this point we would have to resort to using Riemann sums, a process that we have seen is very tedious and is best avoided if possible. Instead let's define

$$G(x) = \int_0^x t^3\,dt$$

and note that

$$G(2) = \int_0^2 t^3\,dt.$$

Why does this help us? The Fundamental Theorem of Calculus shows that

$$G'(x) = x^3.$$

That is, $G(x)$ is an *antiderivative* of x^3. However, we know from the Mean Value Theorem and by the power rule for antiderivatives that if F is any antiderivative of x^3 then

$$F(x) = \frac{x^4}{4} + C$$

where C is some unknown constant. This means that

$$G(x) = \int_0^x t^3\,dt = \frac{x^4}{4} + C_1$$

for some constant C_1. If we knew C_1 we would be done.

To determine C_1 we know that

$$G(x) = \int_0^x t^3\,dt = \frac{x^4}{4} + C_1$$

and

$$0 = \int_0^0 t^3\,dt = G(0) = \frac{0^4}{4} + C_1 = C_1$$

so

$$G(x) = \int_0^x t^3\,dt = \frac{x^4}{4}.$$

Finally,

$$\int_0^2 t^3\,dt = G(2) = \frac{2^4}{4} = 4.$$

Question: Did we really need to find C_1?

To answer this question we will make the following very important observation.

Key Observation: Let F and G be any two antiderivatives of the same function f. Then

$$G(x) = F(x) + C$$

Let $a, b \in \mathbb{R}$. Then

$$\begin{aligned} G(b) - G(a) &= (F(b) + C) - (F(a) + C) \\ &= F(b) - F(a) \end{aligned}$$

Assume that f is continuous. We want to calculate

$$\int_a^b f(t)\,dt.$$

Let

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

Then the Fundamental Theorem of Calculus (Part1) shows that G is an antiderivative of f. Moreover, if F is any other antiderivative of f, then

$$\begin{aligned} \int_a^b f(t)\,dt &= G(b) \\ &= G(b) - G(a) \quad (\text{since } G(a) = \int_a^a f(t)\,dt = 0) \\ &= F(b) - F(a) \end{aligned}$$

For example, to evaluate

$$\int_0^2 t^3\,dt,$$

we know that

$$F(x) = \frac{x^4}{4}$$

is an antiderivative for $f(x) = x^3$. It follows from the previous observation that

$$\begin{aligned}\int_0^2 t^3\,dt &= F(2) - F(0)\\ &= \frac{2^4}{4} - \frac{0^4}{4}\\ &= 4\end{aligned}$$

This example shows us how we can now use antiderivatives to help us evaluate an integral, which is further evidence that the two branches of calculus–differentiation and integration–are intimately linked. Moreover, the observation we have just made establishes a procedure for evaluating definite integrals that works in general because of the Fundamental Theorem of Calculus (Part 1). This is summarized in the following theorem. Because this procedure is essentially a consequence of the Fundamental Theorem of Calculus (Part 1), this result is called the Fundamental Theorem of Calculus (Part 2).

THEOREM 8

Fundamental Theorem of Calculus (Part 2) [FTC2]

Assume that f is continuous and that F is any antiderivative of f.

Then

$$\int_a^b f(t)dt = F(b) - F(a).$$

Going forward, it is no longer necessary to use Riemann sums to calculate integrals; we can use the Fundamental Theorem of Calculus (Part 2) instead.

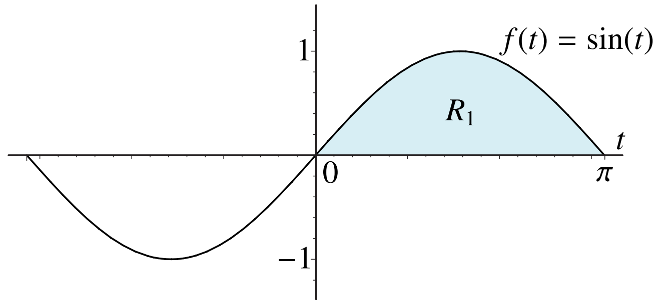
We will now introduce the following notation to use in evaluating integrals. We write

$$F(x)\Big|_a^b = F(b) - F(a)$$

to indicate that the value of the antiderivative F evaluated at b minus the value of the antiderivative F evaluated at a.

EXAMPLE 14 Evaluate $\int_0^\pi \sin(t)\,dt$.

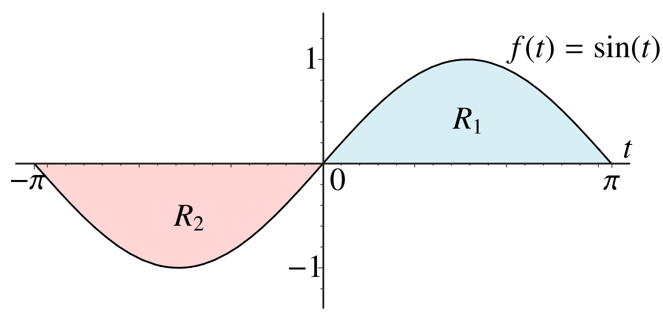
This is the area of the region R_1 under the graph of $\sin(t)$ between $t = 0$ and $t = \pi$. The value for the area is not a number that we can guess since the region is not a familiar shape.



However, $f(t) = \sin(t)$ is continuous and $F(t) = -\cos(t)$ is an antiderivative of f. The Fundamental Theorem of Calculus II tells us that

$$\begin{aligned}\int_0^\pi \sin(t)\,dt &= F(\pi) - F(0)\\ &= (-\cos(\pi)) - (-\cos(0))\\ &= -(-1) - (-1)\\ &= 1 + 1\\ &= 2\end{aligned}$$

Next let's evaluate $\int_{-\pi}^{\pi} \sin(t)dt$.



Using a geometric argument, the value of this integral should be the area of region R_1 minus the area of region R_2. But since $\sin(x)$ is an *odd* function, the symmetry of the graph shows that R_1 and R_2 should have the same area. This means the integral should be 0. To confirm this result we can again use the Fundamental Theorem of Calculus to get

$$\begin{aligned}\int_{-\pi}^{\pi} \sin(t)\,dt &= (-\cos(t))\,\Big|_{-\pi}^{\pi} \\ &= (-\cos(\pi)) - (-\cos(-\pi)) \\ &= (-(-1)) - (-(-1)) \\ &= 1 - 1 \\ &= 0\end{aligned}$$

as expected. ◀

Before we end this section, it is important that we emphasize the difference between the meaning of

$$\int_a^b f(t)\,dt \qquad \text{and} \qquad \int f(t)dt$$

The first expression,

$$\int_a^b f(t)\,dt$$

is called a *definite* integral. It represents a *number* that is defined as a limit of Riemann sums.

The second expression,

$$\int f(t)dt$$

is called an ***in****definite* integral. It represents the *family of all functions that are antiderivatives* of the given function f.

The use of similar notation for these very distinct objects is a direct consequence of the Fundamental Theorem of Calculus.

1.7 Change of Variables

While the Fundamental Theorem of Calculus is a very powerful tool for evaluating definite integrals, the ability to use this tool is limited by our ability to identify antiderivatives.

Finding antiderivatives is essentially **"undoing differentiation."** While we have antiderivative rules for polynomials and for some of the trigonometric and exponential functions, unfortunately it is generally much more difficult to find antiderivatives than it is to differentiate. For example, it is actually possible to prove (using sophisticated algebra that is well beyond this course) that the function

$$f(x) = e^{x^2}$$

does not have an antiderivative that we can state in terms of any functions with which we are familiar. This is a serious flaw in our process since, for example, integrals involving such functions are required for statistical analysis. However, in the next section a method is presented that can undo the most complex rule of differentiation—the Chain Rule. By using this technique, you will be able to evaluate many more types of integrals.

1.7.1 Change of Variables for the Indefinite Integral

Assume that we have two functions $f(u)$ and $h(u)$ with

$$h'(u) = f(u).$$

Then using the notation of antiderivatives we have

$$\int f(u)\,du = h(u) + C.$$

Now let $u = g(x)$ be a function of x. The Chain Rule says that if $H(x) = h(g(x))$ then

$$\begin{aligned} H'(x) &= h'(g(x))g'(x) \\ &= f(g(x))g'(x) \qquad \text{(since } h'(u) = f(u)\text{)} \end{aligned}$$

Integrating both sides we get

$$\begin{aligned} \int f(g(x))g'(x)\,dx &= H(x) + C \\ &= h(g(x)) + C \qquad \text{(since } H(x) = h(g(x))\text{)} \end{aligned}$$

However, this shows that

$$\begin{aligned} \int f(g(x))g'(x)\,dx &= h(g(x)) + C \\ &= h(u)\Big|_{u=g(x)} + C \\ &= \int f(u)\,du\Big|_{u=g(x)} \end{aligned}$$

where the symbol $h(u)\Big|_{u=g(x)}$ means replace u by $g(x)$ in the formula for $h(u)$ and the symbol $\int f(u)\,du\Big|_{u=g(x)}$ means replace u by $g(x)$ once the antiderivative has been found.

Let's see how this works in practice.

EXAMPLE 15 Evaluate $\displaystyle\int 2xe^{x^2}\,dx$.

In this case, note that if we let $u = g(x) = x^2$, then $g'(x) = 2x$. If we also let $f(u) = e^u$, then

$$\int 2x\, e^{x^2}\, dx = \int f(g(x))\, g'(x)\, dx$$

We get that

$$\begin{aligned}
\int 2xe^{x^2}\, dx &= \int f(g(x))g'(x)\, dx \\
&= \int f(u)\, du \Big|_{u=g(x)} \\
&= \int e^u\, du \Big|_{u=x^2} \\
&= e^u \Big|_{u=x^2} + C \\
&= e^{x^2} + C
\end{aligned}$$

To verify that

$$\int 2xe^{x^2}\, dx = e^{x^2} + C$$

we can check the answer by differentiating. Using the Chain Rule, we see that

$$\frac{d}{dx}(e^{x^2} + C) = 2xe^{x^2}$$

which is the integrand in the original question, exactly as we expected.

The method just outlined is called *Change of Variables*. It is often also called *Substitution*, since we **"substitute** $g(x)$ **for** u**."**

There is a notational trick that can help you to remember the process. Start with

$$\int f(g(x))g'(x)\, dx.$$

We want to make the substitution

$$u = g(x).$$

Differentiating both sides gives us

$$\frac{du}{dx} = g'(x).$$

If we treat du and dx as if they were "numbers", then

$$du = g'(x)\, dx.$$

We can now substitute u for $g(x)$ and du for $g'(x)dx$ to get

$$\int f(g(x))\, g'(x)\, dx = \int f(u)\, du \,\Big|_{u = g(x)}$$

which is the Change of Variables formula.

It is important to note that the expression

$$du = g'(x)\, dx$$

does *not* really have any mathematical meaning. None the less, this trick works and it is how integrals are actually computed in practice.

EXAMPLE 16 Evaluate $\displaystyle\int \frac{2x}{1+x^2}\, dx$ by making the substitution $u = 1 + x^2$.

If

$$u = 1 + x^2,$$

then

$$du = 2x\, dx.$$

Substituting $u = 1 + x^2$ and $du = 2x\, dx$ into the original integral gives us

$$\int \frac{2x}{1+x^2}\, dx = \int \frac{1}{u}\, du \Big|_{u=1+x^2}$$

but

$$\int \frac{1}{u}\, du = \ln(|\, u \,|) + C.$$

Hence

$$\begin{aligned}
\int \frac{2x}{1+x^2}\, dx &= \int \frac{1}{u}\, du \Big|_{u=1+x^2} \\
&= \ln(|\, u \,|)\Big|_{u=1+x^2} + C \\
&= \ln(|\, 1 + x^2 \,|) + C \\
&= \ln(1 + x^2) + C
\end{aligned}$$

where the last equality holds since $1 + x^2 > 0$.

Until you become comfortable with this technique it is always a good idea to **check your answer by differentiating**. (In this case, the answer is $f(x) = \ln(1 + x^2) + C$. Differentiating we get $f'(x) = \frac{1}{(1+x^2)} \frac{d}{dx}(1 + x^2) = \frac{2x}{(1+x^2)}$ which is the original integrand for the question . . . so the answer is correct!) ◄

EXAMPLE 17 Evaluate $\int x\cos(x^2)\,dx$.

In this example, using the substitution $u = x^2$ causes us to be out by a constant factor. However, this will not be a problem by proceeding as follows.

The method of substitution suggests letting $u = x^2$ and $f(u) = \cos(u)$. In this case we have

$$du = 2x\,dx \qquad (*)$$

However, it appears that we really wanted

$$du = x\,dx.$$

We can rearrange equation $(*)$ to get that

$$\frac{1}{2}\,du = x\,dx.$$

Then substituting $\frac{1}{2}\,du$ for $x\,dx$ and u for x^2 we get

$$\begin{aligned}
\int x\cos(x^2)\,dx &= \int \frac{1}{2}\cos(u)\,du\Big|_{u=x^2} \\
&= \frac{1}{2}\int \cos(u)\,du\Big|_{u=x^2} \\
&= \frac{1}{2}\sin(u)\Big|_{u=x^2} + C \\
&= \frac{1}{2}\sin(x^2) + C
\end{aligned}$$

You should check that this answer is correct by differentiating $\frac{1}{2}\sin(x^2) + C$ to ensure it equals the integrand. ◀

The next unusual substitution allows us to evaluate an important trigonometric integral.

EXAMPLE 18 Evaluate $\int \sec(\theta)\,d\theta$.

At first glance this does not appear to be an integral that can be evaluated by using substitution. However, we can make the following clever observation:

$$\begin{aligned}
\sec(\theta) &= \sec(\theta)\left(\frac{\sec(\theta)+\tan(\theta)}{\sec(\theta)+\tan(\theta)}\right) \\
&= \frac{\sec^2(\theta)+\sec(\theta)\tan(\theta)}{\sec(\theta)+\tan(\theta)}
\end{aligned}$$

Let $u = \sec(\theta) + \tan(\theta)$. Then

$$du = \sec(\theta)\tan(\theta) + \sec^2(\theta)\,d\theta.$$

It follows that

$$\begin{aligned}\int \sec(\theta)\,d\theta &= \int \frac{\sec^2(\theta) + \sec(\theta)\tan(\theta)}{\sec(\theta) + \tan(\theta)}\,d\theta \\ &= \int \frac{du}{u}\end{aligned}$$

The antiderivative rules give us that

$$\int \frac{du}{u} = \ln(|u|) + C.$$

Letting $u = \sec(\theta) + \tan(\theta)$, we get

$$\int \sec(\theta)\,d\theta = \ln(|\sec(\theta) + \tan(\theta)|) + C.$$

◀

1.7.2 Change of Variables for the Definite Integral

We can also use the Change of Variables technique for definite integrals. However, for the case of definite integrals, we must *be careful with the limits of integration*.

Suppose that we want to evaluate

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

where f and g' are continuous functions. We have just seen that if $h(u)$ is an antiderivative of $f(u)$, then

$$H(x) = h(g(x))$$

is an antiderivative of

$$f(g(x))g'(x).$$

This means that we can apply the Fundamental Theorem of Calculus to get

$$\begin{aligned}\int_a^b f(g(x))g'(x)\,dx &= H(b) - H(a) \\ &= h(g(b)) - h(g(a))\end{aligned}$$

However, since $h'(u) = f(u)$, the Fundamental Theorem of Calculus also shows us that

$$\int_{g(a)}^{g(b)} f(u)\,du = h(g(b)) - h(g(a)).$$

Combining these two results shows that

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

THEOREM 9

Change of Variables

Assume that $g'(x)$ is continuous on $[a,b]$ and $f(u)$ is continuous on $g([a,b])$,

then

$$\int_{x=a}^{x=b} f(g(x))g'(x)\,dx = \int_{u=g(a)}^{u=g(b)} f(u)\,du.$$

Let's see what this theorem implies by looking at some examples. Notice that you must give special attention to the limits of integration.

EXAMPLE 19 Evaluate $\int_2^4 (5x-6)^3\,dx$.

Let $u = g(x) = 5x - 6$. Then $f(u) = u^3$. Since

$$du = g'(x)\,dx = 5\,dx,$$

we have

$$\frac{1}{5}\,du = dx.$$

The Change of Variables Theorem shows us that

$$\begin{aligned}\int_2^4 (5x-6)^3\,dx &= \int_{u=g(a)}^{u=g(b)} f(u)\,du \\ &= \int_{u=g(2)}^{u=g(4)} u^3\,\frac{1}{5}du\end{aligned}$$

Now since $g(a) = g(2) = 5(2) - 6 = 4$ and $g(b) = g(4) = 5(4) - 6 = 14$ we have

$$
\begin{aligned}
&= \frac{1}{5}\int_{4}^{14} u^3\,du \\
&= \frac{1}{5}\left(\frac{1}{4}u^4\right)\Big|_{4}^{14} \\
&= \frac{1}{20}(14^4 - 4^4) \\
&= 1908
\end{aligned}
$$

◀

EXAMPLE 20 Evaluate $\displaystyle\int_0^1 \frac{x\,dx}{\sqrt{x^2+1}}$.

Let $u = g(x) = x^2 + 1$. Then

$$du = g'(x)\,dx = 2x\,dx,$$

so

$$\frac{1}{2}\,du = x\,dx.$$

We also have,

$$f(u) = \frac{1}{\sqrt{u}} = u^{-\frac{1}{2}}.$$

The Change of Variables Theorem shows us that

$$
\begin{aligned}
\int_0^1 \frac{x\,dx}{\sqrt{x^2+1}} &= \int_{u=g(a)}^{u=g(b)} f(u)\,du \\
&= \int_{u=g(0)}^{u=g(1)} (u^{-\frac{1}{2}})\,(\frac{1}{2}du)
\end{aligned}
$$

Now since $g(a) = g(0) = 0^2 + 1 = 1$ and $g(b) = g(1) = 1^2 + 1 = 2$ we have

$$
\begin{aligned}
&= \frac{1}{2}\int_1^2 u^{-\frac{1}{2}}\,du \\
&= \frac{1}{2}\left(2\,u^{\frac{1}{2}}\right)\Big|_1^2 \\
&= 2^{\frac{1}{2}} - 1^{\frac{1}{2}} \\
&= \sqrt{2} - 1 \\
&\cong 0.41
\end{aligned}
$$

◀

EXAMPLE 21 Evaluate $\int_0^{\frac{\pi}{4}} -2\cos^2(2x)\ \sin(2x)\,dx$.

Let $u = g(x) = \cos(2x)$ and $f(u) = u^2$. Then since

$$du = g'(x)\,dx = -2\sin(2x)\,dx$$

the Change of Variables Theorem shows us that

$$\begin{aligned}
\int_0^{\frac{\pi}{4}} -2\cos^2(2x)\sin(2x)\,dx &= \int_{u=g(a)}^{u=g(b)} f(u)\,du \\
&= \int_{\cos(2(0))}^{\cos(2(\frac{\pi}{4}))} u^2\,du \\
&= \int_{\cos(0)}^{\cos(\frac{\pi}{2})} u^2\,du \\
&= \int_1^0 u^2\,du \\
&= \left.\frac{u^3}{3}\right|_1^0 \\
&= \frac{0^3}{3} - \frac{1^3}{3} \\
&= \frac{-1}{3}
\end{aligned}$$

◀

Chapter 2

Techniques of Integration

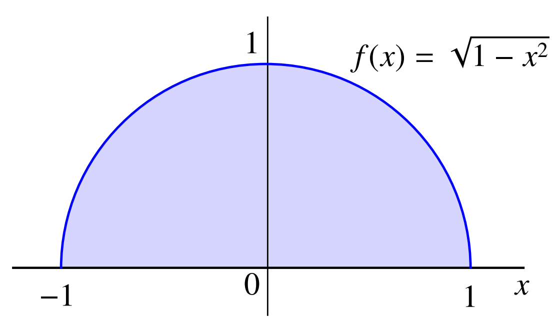
The Change of Variables Theorem gave us a method for evaluating integrals when the antiderivative of the integrand was not obvious. The underlying method involved "substitution" of one variable for another. In this chapter, we continue using the Change of Variables Theorem to calculate integrals. However, the substitutions will involve trigonometric functions.

2.1 Inverse Trigonometric Substitutions

From the geometric interpretation of the integral we have seen that

$$\int_{-1}^{1} \sqrt{1-x^2}\, dx = \frac{\pi}{2}.$$

However, we did not explicitly calculate this integral because we did not know many integration techniques. Instead, we found the area by deducing that $\sqrt{1-x^2}$ was the top half of a circle with radius 1 and using the high school formula for the area of a circle, *area* $= \pi r^2$.



We will now explicitly calculate this integral using a trigonometric substitution.

EXAMPLE 1 Evaluate

$$\int_{-1}^{1} \sqrt{1-x^2}\, dx = \frac{\pi}{2}.$$

We will require the use of the Pythagorean Identity for trigonometric functions to evaluate this integral:

$$\sin^2(x) + \cos^2(x) = 1.$$

Let

$$x = \sin(u)$$

for $\dfrac{-\pi}{2} \leq u \leq \dfrac{\pi}{2}$.

This might seem like an unusual suggestion since the substitution rule usually asks us to make a substitution of the form

$$u = g(x).$$

However, we have actually done this! To see why this is the case, recall that on the interval $[\frac{-\pi}{2}, \frac{\pi}{2}]$, the function $x = \sin(u)$ has a unique inverse given by $u = \arcsin(x)$. In fact, we are really making the substitution $u = \arcsin(x)$ and this method is called *inverse* trigonometric substitution.

Using the substitution $x = \sin(u)$, the integrand

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

becomes

$$\begin{aligned} \sqrt{1 - \sin^2(u)} &= \sqrt{\cos^2(u)} \\ &= |\cos(u)| \\ &= \cos(u) \end{aligned}$$

where the last equality holds since $\cos(u) \geq 0$ on the interval $[\frac{-\pi}{2}, \frac{\pi}{2}]$.

We can differentiate

$$x = \sin(u)$$

to get

$$dx = \cos(u)\, du.$$

This means that

$$\begin{aligned} \int_{-1}^{1} \sqrt{1 - x^2}\, dx &= \int_{\arcsin(-1)}^{\arcsin(1)} \cos(u)\ \cos(u)\, du \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(u)\, du \end{aligned}$$

Note: Since $u = g(x) = \arcsin(x)$, we have that the new limits of integration are $u = g(-1) = \arcsin(-1) = -\frac{\pi}{2}$ and $u = g(1) = \arcsin(1) = \frac{\pi}{2}$.

To finish the integral calculation we need to use the following trigonometric identity:

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

This means that

$$\begin{aligned}
\int_{-1}^{1} \sqrt{1-x^2}\,dx &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(u)\,du \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1+\cos(2u)}{2}\,du \\
&= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{1}{2}\,du + \int_{\frac{-\pi}{2}}^{\frac{\pi}{2}} \frac{\cos(2u)}{2}\,du \\
&= \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1\,du + \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(2u)\,du
\end{aligned}$$

The first integral is

$$\begin{aligned}
\frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1\,du &= \frac{1}{2}u\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \\
&= \frac{1}{2}\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) \\
&= \frac{\pi}{2}
\end{aligned}$$

To evaluate the second integral

$$\frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(2u)\,du$$

use the substitution

$$v = 2u$$

with

$$dv = 2\,du.$$

If $u = \frac{\pi}{2}$, then $v = \pi$ and if $u = -\frac{\pi}{2}$, then $v = -\pi$. We can apply the Change of Variable Formula to get that

$$\begin{aligned}
\frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(2u)\,du &= \frac{1}{2}\int_{-\pi}^{\pi} \cos(v)\,\frac{dv}{2} \\
&= \frac{1}{4}\int_{-\pi}^{\pi} \cos(v)\,dv \\
&= \frac{1}{4}\sin(v)\Big|_{-\pi}^{\pi} \\
&= \frac{1}{4}(\sin(\pi) - \sin(-\pi))
\end{aligned}$$

$$= \frac{1}{4}(0-0)$$

$$= 0$$

We have then shown that

$$\begin{aligned}
\int_{-1}^{1} \sqrt{1-x^2}\,dx &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2(u)\,du \\
&= \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1\,du + \frac{1}{2}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(2u)\,du \\
&= \frac{\pi}{2} + 0 \\
&= \frac{\pi}{2}
\end{aligned}$$

which agrees with our previous argument.

◀

REMARK

In general, there are *three* main classes of inverse trigonometric substitutions. The first class are integrals with integrands of the form

$$\sqrt{a^2 - b^2x^2}.$$

The substitution, based on the pythagorean identity $\sin^2(x) + \cos^2(x) = 1$, is

$$bx = a\sin(u).$$

The previous example demonstrated this class.

The second class of trigonometric substitution covers integrands of the form

$$\sqrt{a^2 + b^2x^2}.$$

The substitution is based on the identity $\sec^2(x) - 1 = \tan^2(x)$ and is given by

$$bx = a\tan(u).$$

The third type of substitution covers integrands of the form

$$\sqrt{b^2x^2 - a^2}.$$

The substitution is again based on the identity $\sec^2(x) - 1 = \tan^2(x)$ and is given by

$$bx = a\sec(u).$$

◀

The next example illustrates the third class of trigonometric substitution.

EXAMPLE 2 Evaluate $\displaystyle\int_{\sqrt{3}}^{3} \frac{\sqrt{4x^2-9}}{x}\,dx$.

Consider the expression $\sqrt{4x^2-9}$. We could try to substitute $u = 4x^2 - 9$ giving us $du = 8x\,dx$ or $dx = \frac{du}{8x}$. This would have worked if the integral had been

$$\int_{\sqrt{3}}^{3} x\sqrt{4x^2-9}\,dx$$

because then the x's would cancel. (You should verify this statement.) However, the substitution $u = 4x^2 - 9$ does not help in this example.

Since the numerator of the integrand takes the form $\sqrt{b^2x^2-a^2}$ where $b = 2$ and $a = 3$, the correct substitution is

$$2x = 3\sec(u)$$

or

$$u = \operatorname{arcsec}\left(\frac{2x}{3}\right)$$

where $0 \leq u < \frac{\pi}{2}$.

Since $4x^2 = (2x)^2$, we have $(3\sec(u))^2 = 9\sec^2(u)$. Therefore,

$$\begin{aligned}
\sqrt{4x^2-9} &= \sqrt{9\sec^2(u)-9} \\
&= \sqrt{9(\sec^2(u)-1)} \\
&= 3\sqrt{\sec^2(u)-1} \\
&= 3\sqrt{\tan^2(u)} \\
&= 3\,|\tan(u)\,| \\
&= 3\tan(u)
\end{aligned}$$

with the last equality holding since if $0 \leq u < \frac{\pi}{2}$, then $\tan(u) \geq 0$.

Since $2x = 3\sec(u)$ we also have that

$$2dx = 3\sec(u)\tan(u)\,du$$

or

$$dx = \frac{3}{2}\sec(u)\tan(u)\,du.$$

Next, we must find the new limits of integration for this integral. When $x = \sqrt{3}$, we have

$$\begin{aligned}
\sec(u) &= \frac{2\sqrt{3}}{3} \\
&= \frac{2}{\sqrt{3}}
\end{aligned}$$

This means that

$$\begin{aligned}\cos(u) &= \frac{1}{\sec(u)}\\ &= \frac{\sqrt{3}}{2}\end{aligned}$$

The only angle u with $0 \leq u < \frac{\pi}{2}$ and $\cos(u) = \frac{\sqrt{3}}{2}$, is $u = \frac{\pi}{6}$. (Verify this fact by considering the Unit Circle.)

Similarly, if $x = 3$, then $\sec(u) = 2$ so $\cos(u) = \frac{1}{2}$. This means that $u = \frac{\pi}{3}$.

Finally, we must also substitute for the x in the denominator of the integrand

$$\frac{\sqrt{4x^2-9}}{x}$$

since it has not cancelled in the substitution. However, we know that $x = \frac{3}{2}\sec(u)$.

Combining everything we know gives us

$$\begin{aligned}\int_{\sqrt{3}}^{3} \frac{\sqrt{4x^2-9}}{x}dx &= \int_{\frac{\pi}{6}}^{\frac{\pi}{3}} \frac{3\tan(u)}{\frac{3}{2}\sec(u)}\left(\frac{3}{2}\right)\sec(u)\tan(u)\,du\\ &= 3\int_{\frac{\pi}{6}}^{\frac{\pi}{3}} \tan^2(u)\,du\\ &= 3\int_{\frac{\pi}{6}}^{\frac{\pi}{3}} (\sec^2(u)-1)\,du\\ &= 3(\tan(u)-u)\Big|_{\frac{\pi}{6}}^{\frac{\pi}{3}}\\ &= 3\left(\tan\left(\frac{\pi}{3}\right)-\frac{\pi}{3}\right)-3\left(\tan\left(\frac{\pi}{6}\right)-\frac{\pi}{6}\right)\\ &= \left(3\sqrt{3}-\pi\right)-\left(\frac{3}{\sqrt{3}}-\frac{\pi}{2}\right)\\ &= 2\sqrt{3}-\frac{\pi}{2}\end{aligned}$$

◀

In general, when presented with an integral that you are unsure of how to solve, be aware of the following classes of trigonometric substitutions to see if they can help you evaluate the integral.

Summary of Inverse Trigonometric Substitutions

Class of Integrand	Integral	Trig Substitution	Trig Identity
$\sqrt{a^2 - b^2x^2}$	$\int \sqrt{a^2 - b^2x^2}\, dx$	$bx = a\sin(u)$	$\sin^2(x) + \cos^2(x) = 1$
$\sqrt{a^2 + b^2x^2}$	$\int \sqrt{a^2 + b^2x^2}\, dx$	$bx = a\tan(u)$	$\sec^2(x) - 1 = \tan^2(x)$
$\sqrt{b^2x^2 - a^2}$	$\int \sqrt{b^2x^2 - a^2}\, dx$	$bx = a\sec(u)$	$\sec^2(x) - 1 = \tan^2(x)$

2.2 Integration by Parts

Suppose that we want to calculate

$$\int x\sin(x)\, dx.$$

There is no obvious substitution that will help. Fortunately, there is another method that will work for this integral called *Integration by Parts*.

While the method of integration by substitution was based on trying to *undo* the Chain Rule, Integration by Parts is derived from the Product Rule. If f and g are differentiable, then the Product Rule states that

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

Since the antiderivative of a derivative is just the original function up to a constant, we have

$$\begin{aligned} f(x)g(x) &= \int \frac{d}{dx}(f(x)g(x))\, dx \\ &= \int (f'(x)g(x) + f(x)g'(x))\, dx \\ &= \int f'(x)g(x)\, dx + \int f(x)g'(x)\, dx \end{aligned}$$

Rearranging this equation leads to the following formula:

DEFINITION

The Integration by Parts Formula

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx.$$

EXAMPLE 3 Use integration by parts to evaluate $\int x\sin(x)\,dx$.

The task is to choose the functions f and g' in such a way that the integral $\int x\sin(x)\,dx$ has the form $\int f(x)g'(x)\,dx$ and the expression

$$f(x)g(x) - \int f'(x)g(x)\,dx$$

can be easily evaluated. The key is to view $x\sin(x)$ as a product of the functions x and $\sin(x)$ and to note that differentiating x produces the constant 1. This will leave us with only a simple trigonometric function to integrate. Therefore, we let $f(x) = x$ and let $g'(x) = \sin(x)$.

The next step is to determine f' and g.

Since $f(x) = x$, we have that $f'(x) = 1$. We can choose any antiderivative of $\sin(x)$ to play the role of $g(x)$ so we choose $g(x) = -\cos(x)$. Substituting into the Integration by Parts Formula

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

gives

$$\int x\sin(x)\,dx = x(-\cos(x)) - \int (1)(-\cos(x))\,dx$$

or

$$\int x\sin(x)\,dx = -x\cos(x) + \int \cos(x)\,dx.$$

Since

$$\int \cos(x)\,dx = \sin(x) + C$$

we get

$$\int x\sin(x)\,dx = -x\cos(x) + \sin(x) + C.$$

We can verify this solution by differentiation:

$$\begin{aligned}\frac{d}{dx}(-x\cos(x) + \sin(x) + C) &= -\cos(x) + x\sin(x) + cos(x) + 0 \\ &= x\sin(x)\end{aligned}$$

which is the original integrand as we expected.

◀

The next example shows that we might have to combine Integration by Parts with substitution to calculate an integral.

EXAMPLE 4 Evaluate $\int x\cos(2x)\,dx$.

The strategy for this integral is again to use Integration by Parts to eliminate the x from the integrand so that we are left with a simple trigonometric function to integrate. Therefore, we let

$$f(x) = x \qquad \text{and} \qquad g'(x) = \cos(2x).$$

We must now find $f'(x)$ and $g(x)$. Since $f(x) = x$, then $f'(x) = 1$. To find $g(x)$, we evaluate

$$\int \cos(2x)\,dx.$$

Let $u = 2x$, then $du = 2\,dx$ so that $dx = \frac{du}{2}$. Substitute these into the integral to get

$$\begin{aligned}\int \cos(2x)\,dx &= \int \cos(u)\,\frac{du}{2}\\ &= \frac{1}{2}\int \cos(u)\,du\\ &= \frac{\sin(u)}{2} + C\\ &= \frac{\sin(2x)}{2} + C\end{aligned}$$

We can choose any antiderivative of $\cos(2x)$ for $g(x)$, so let $C = 0$ to get

$$\begin{array}{ll} f(x) = x & g'(x) = \cos(2x)\\ f'(x) = 1 & g(x) = \frac{\sin(2x)}{2}\end{array}$$

Applying the Integration by Parts formula gives

$$\begin{aligned}\int x\cos(2x)\,dx &= \frac{x\sin(2x)}{2} - \int (1)\frac{\sin(2x)}{2}\,dx\\ &= \frac{x\sin(2x)}{2} - \frac{1}{2}\int \sin(2x)\,dx\end{aligned}$$

To evaluate $\int \sin(2x)\,dx$, we can again use the substitution $u = 2x$ so that $du = 2dx$

and $dx = \frac{du}{2}$. This shows that

$$\begin{aligned}\int \sin(2x)\,dx &= \int \sin(u)\,\frac{du}{2}\\ &= \frac{1}{2}\int \sin(u)\,du\\ &= \frac{-\cos(u)}{2}+C\\ &= \frac{-\cos(2x)}{2}+C\end{aligned}$$

Therefore,

$$\begin{aligned}\int x\cos(2x)\,dx &= \frac{x\sin(2x)}{2}-\frac{1}{2}\int \sin(2x)\,dx\\ &= \frac{x\sin(2x)}{2}-\frac{1}{2}\left(\frac{-\cos(2x)}{2}\right)+C\\ &= \frac{x\sin(2x)}{2}+\frac{\cos(2x)}{4}+C\end{aligned}$$

NOTE

Since C is an arbitray constant we did not need to multiply it by $\frac{1}{2}$ when we substituted for $\int \sin(2x)\,dx$ in this calculation. ◀

EXAMPLE 5 Evaluate $\int x^2e^x\,dx$.

Once again there is no obvious substitution so we will try Integration by Parts. We will use differentiation to eliminate the polynomial x^2 so that only a simple exponential function is left to integrate. However, this time we will need to apply the process twice.

We begin with

$$\begin{array}{ll} f(x) = x^2 & g'(x) = e^x \\ f'(x) = 2x & g(x) = e^x \end{array}$$

Applying the Integration by Parts formula gives

$$\begin{aligned}\int x^2 e^x\,dx &= x^2e^x - \int (2x)e^x\,dx\\ &= x^2e^x - 2\int xe^x\,dx\end{aligned}$$

We are left to evaluate $\int xe^x\,dx$. This integral is again an ideal candidate for Integration by Parts. Let

$$\begin{array}{ll} f(x) = x & g'(x) = e^x \\ f'(x) = 1 & g(x) = e^x \end{array}$$

to get

$$\begin{aligned}\int xe^x\,dx &= xe^x - \int (1)e^x\,dx\\ &= xe^x - \int e^x\,dx\\ &= xe^x - e^x + C\end{aligned}$$

We can now substitute for $\int xe^x\,dx$ to get,

$$\begin{aligned}\int x^2e^x\,dx &= x^2e^x - 2\int xe^x\,dx\\ &= x^2e^x - 2(xe^x - e^x) + C\\ &= x^2e^x - 2xe^x + 2e^x + C\end{aligned}$$

◀

REMARK

You might guess from the previous example that the integral

$$\int x^3e^x\,dx$$

could be evaluated with *three* applications of Integration by Parts since we would have to *differentiate three times to eliminate the polynomial* x^3. ◀

The next example illustrates another class of functions that are ideally suited to Intergation by Parts.

EXAMPLE 6 Evaluate $\int e^x \sin(x)\,dx$.

This example presents a different type of problem than any of the previous examples. It is not clear which function should be f and which should be g' since no amount of differentiation will eliminate either e^x or $\sin(x)$. In this case, we will simply choose g' to be the easiest function to integrate. For this example, this means that $g'(x) = e^x$. Therefore, we have

$$\begin{array}{ll} f(x) = \sin(x) & g'(x) = e^x \\ f'(x) = \cos(x) & g(x) = e^x \end{array}$$

The Integration by Parts Formula gives

$$\int e^x \sin(x)\,dx = e^x \sin(x) - \int e^x \cos(x)\,dx.$$

This result may appear somewhat discouraging because there is no reason to believe that the integral

$$\int e^x \cos(x)\,dx$$

is any easier to evaluate than the original integral $\int e^x \sin(x)\,dx$.

The key is to apply the formula again to the integral $\int e^x \cos(x)\,dx$ with

$$\begin{array}{ll} f(x) = \cos(x) & g'(x) = e^x \\ f'(x) = -\sin(x) & g(x) = e^x \end{array}$$

to get

$$\begin{aligned} \int e^x \cos(x)\,dx &= e^x \cos(x) - \int e^x(-\sin(x))\,dx \\ &= e^x \cos(x) + \int e^x(\sin(x))\,dx \end{aligned}$$

We seem to be left with having to evaluate $\int e^x(\sin(x))\,dx$ **which is exactly where we started!** However, if we substitute $e^x \cos(x) + \int e^x(\sin(x))\,dx$ for $\int e^x \cos(x)\,dx$ in the original equation something interesting happens. We get

$$\begin{aligned} \int e^x \sin(x)\,dx &= e^x \sin(x) - \int e^x \cos(x)\,dx \\ &= e^x \sin(x) - (e^x \cos(x) + \int e^x(\sin(x))\,dx) \\ &= e^x \sin(x) - e^x \cos(x) - \int e^x(\sin(x))\,dx \end{aligned}$$

You will notice that $\int e^x(\sin(x))\,dx$ appears on both sides of our equation but with opposite signs. We can treat this expression as some unknown variable and then gather like terms as we would in basic algebra. This means adding $\int e^x(\sin(x))\,dx$ to both sides of the expression to get

$$2\left(\int e^x(\sin(x))\,dx\right) = e^x\sin(x) - e^x\cos(x).$$

Divide by 2 so that

$$\int e^x(\sin(x))\,dx = \frac{e^x\sin(x) - e^x\cos(x)}{2}.$$

At this point, you might notice that the constant of integration is missing in this expression yet all general antiderivatives must include a constant. In fact, this is due to the way that the Integration by Parts formula handles these constants (the constants are always there implicitly even if they are not explicitly written). We have identified just one possible antiderivative for the function $e^x\sin(x)$. To state all of the antiderivatives we know that we simply add an arbitrary constant so that

$$\int e^x(\sin(x))\,dx = \frac{e^x\sin(x) - e^x\cos(x)}{2} + C.$$

◀

Observation: The method we outlined in the previous example works because of the cyclic nature of the derivatives for $\sin(x)$ and e^x. This statement is also true of both $\cos(x)$ and e^x. Therefore, it should not be surprising to discover that $\int e^x\cos(x)\,dx$ can be evaluated in the same manner.

Important Note: In summary, the Integration by Parts formula is ideally suited to evaluating integrals of the following types:

- $\int x^n\cos(x)\,dx$
- $\int x^n\sin(x)\,dx$
- $\int x^n e^x\,dx$
- $\int e^x\cos(x)\,dx$
- $\int e^x\sin(x)\,dx$

◀

However, there are other more unusual examples of integrals that are also suitable for Integration by Parts.

EXAMPLE 7 Evaluate

$$\int \arctan(x)\,dx.$$

At first glance this integral does not seem to be of the form

$$\int f(x)g'(x)\,dx$$

since there is no product in the integrand. However, the key is to rewrite the integrand as

$$\arctan(x) = (1)\arctan(x).$$

Since $\arctan(x)$ is easy to differentiate and 1 is easily integrated, we can now try Integration by Parts with $f(x) = \arctan(x)$ and $g'(x) = 1$. This leads to

$$\begin{array}{ll} f(x) = \arctan(x) & g'(x) = 1 \\ f'(x) = \frac{1}{1+x^2} & g(x) = x \end{array}$$

Applying the Integration by Parts formula gives

$$\begin{aligned} \int \arctan(x)\,dx &= x\arctan(x) - \int x\left(\frac{1}{1+x^2}\right)dx \\ &= x\arctan(x) - \int \frac{x}{1+x^2}\,dx \end{aligned}$$

The integral $\int \frac{x}{1+x^2}\,dx$ can be handled by substitution. Let $u = 1 + x^2$ so $du = 2xdx$ and $dx = \frac{du}{2x}$. Substitution gives

$$\begin{aligned} \int \frac{x}{1+x^2}\,dx &= \int \frac{x}{u}\frac{du}{2x} \\ &= \frac{1}{2}\int \frac{1}{u}\,du \\ &= \frac{\ln(|\,u\,|)}{2} + C \\ &= \frac{\ln(|\,1+x^2\,|)}{2} + C \\ &= \frac{\ln(1+x^2)}{2} + C \end{aligned}$$

since $1 + x^2 > 0$.

Returning to the original equation, we have

$$\begin{aligned}\int \arctan(x)\,dx &= x\arctan(x) - \int \frac{x}{1+x^2}\,dx \\ &= x\arctan(x) - \left(\frac{\ln(1+x^2)}{2}\right) + C\end{aligned}$$

We can check this answer by differentiating:

$$\begin{aligned}\frac{d}{dx}\left(x\arctan(x) - \frac{\ln(1+x^2)}{2} + C\right) &= \arctan(x) + x\left(\frac{1}{1+x^2}\right) - \frac{1}{2}\left(\frac{1}{1+x^2}\right)(2x) \\ &= \arctan(x) + \frac{x}{1+x^2} - \frac{x}{1+x^2} \\ &= \arctan(x)\end{aligned}$$

exactly as expected. ◀

EXAMPLE 8 Evaluate $\int \ln(x)\,dx$.

Notice

$$\ln(x) = \underset{\substack{\uparrow \\ g'(x)}}{1} \cdot \underset{\substack{\uparrow \\ f(x)}}{\ln(x)}$$

This gives

$$\begin{array}{ll} f(x) = \ln(x) & g'(x) = 1 \\ f'(x) = \frac{1}{x} & g(x) = x \end{array}$$

Applying the Integration by Parts Formula

$$\begin{aligned}\int \ln(x)\,dx &= x\ln(x) - \int x\left(\frac{1}{x}\right)dx \\ &= x\ln(x) - \int 1\,dx \\ &= x\ln(x) - x + C\end{aligned}$$

◀

It is important to be cautious when trying to evaluate an integral. For example, the integral

$$\int xe^{x^2}\,dx$$

appears to be a candidate for Integration by Parts. However, this is not the case. Instead, it can be evaluated by using the substitution $u = x^2$.

The Integration by Parts Formula can also be applied to definite integrals. The following theorem is a direct consequence of combining the Integration by Parts formula with the Fundamental Theorem of Calculus.

THEOREM 1 **Integration by Parts**

Assume that f and g are such that both f' and g' are continuous on an interval containing a and b. Then

$$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b f'(x)g(x)\,dx.$$

EXAMPLE 9 Evaluate $\int_0^1 xe^x\,dx$.

Let

$$f(x) = x \qquad g'(x) = e^x$$

$$f'(x) = 1 \qquad g(x) = e^x$$

Integration by Parts shows that

$$\begin{aligned}\int_0^1 xe^x\,dx &= xe^x\Big|_0^1 - \int_0^1 e^x\,dx \\ &= [e-0] - e^x\Big|_0^1 \\ &= e - [e-1] \\ &= 1\end{aligned}$$

◀

2.3 Partial Fractions

Recall that a *rational function* is a function of the form

$$f(x) = \frac{p(x)}{q(x)}$$

where p and q are polynomials. In this section, we will discuss a method for integrating rational functions called *Partial Fractions*. We will illustrate this method with an example.

EXAMPLE 10 Evaluate $\int \frac{1}{x^2-1} dx$.

Step 1:

First factor the denominator to get that

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

Step 2:

Find constants A and B so that

$$\frac{1}{(x-1)(x+1)} = \frac{A}{x-1} + \frac{B}{x+1} \qquad (*)$$

To find A and B, we multiply both sides of the identity $(*)$ by $(x-1)(x+1)$ to get

$$1 = A(x+1) + B(x-1) \qquad (**)$$

The two roots of the denominator were $x = 1$ and $x = -1$. If we substitute $x = 1$ into equation $(**)$, we have

$$1 = A(1+1) + B(1-1)$$

or

$$1 = 2A.$$

Therefore,

$$A = \frac{1}{2}.$$

If we then substitute $x = -1$ into equation $(**)$, we get

$$1 = A(-1+1) + B(-1-1)$$

or

$$1 = -2B$$

and

$$B = -\frac{1}{2}.$$

Using these values of A and B we have

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

Therefore

$$\int \frac{1}{x^2-1} dx = \frac{1}{2}\int \frac{1}{x-1} dx - \frac{1}{2}\int \frac{1}{x+1} dx.$$

Recall that

$$\int \frac{1}{x-a} dx = \ln(|\, x-a\,|) + C.$$

Hence

$$\begin{aligned}
\int \frac{1}{x^2-1}\,dx &= \frac{1}{2}\int \frac{1}{x-1}\,dx - \frac{1}{2}\int \frac{1}{(x+1)}\,dx \\
&= \frac{1}{2}\ln(|\,x-1\,|) - \frac{1}{2}\ln(|\,x+1\,|) + C \\
&= \frac{1}{2}\ln\left(\frac{|\,x-1\,|}{|\,x+1\,|}\right) + C
\end{aligned}$$

since $\ln(b) - \ln(a) = \ln\left(\frac{b}{a}\right)$. ◀

The method we have just outlined required us to separate the function $f(x) = \frac{1}{x^2-1}$ into rational functions with *first degree* denominators. This is called a *partial fraction decomposition of f*.

DEFINITION

Type I Partial Fraction Decomposition

Assume that

$$f(x) = \frac{p(x)}{q(x)}$$

where p and q are polynomials such that

1. $degree(p(x)) < degree(q(x)) = k$,
2. $q(x)$ can be factored into the product of linear terms each with distinct roots. That is
$$q(x) = a(x-a_1)(x-a_2)(x-a_3)\cdots(x-a_k)$$
where the a_i's are unique and none of the a_i's are roots of $p(x)$.

Then there exists constants $A_1, A_2, A_3, \cdots, A_k$ such that

$$f(x) = \frac{1}{a}\left[\frac{A_1}{x-a_1} + \frac{A_2}{x-a_2} + \frac{A_3}{x-a_3} + \cdots + \frac{A_k}{x-a_k}\right]$$

we say that f admits a *Type I Partial Fraction Decomposition*.

Key Observation: The existence of the constants $A_1, A_2, \cdots, A_k$ follows from some basic algebra. However, what is important to us is that if the rational function $f(x) = \frac{p(x)}{q(x)}$ has a Type I Decomposition, then it is easy to find its integral.

THEOREM 2 **Integration of Type I Partial Fractions**

Assume that $f(x) = \frac{p(x)}{q(x)}$ admits a Type I Partial Fraction Decomposition of the form

$$f(x) = \frac{1}{a}\left[\frac{A_1}{x-a_1} + \frac{A_2}{x-a_2} + \cdots + \frac{A_k}{x-a_k}\right].$$

Then

$$\begin{aligned}\int f(x)\,dx &= \frac{1}{a}\left[\int \frac{A_1}{x-a_1}\,dx + \int \frac{A_2}{x-a_2}\,dx + \cdots + \int \frac{A_k}{x-a_k}\,dx\right] \\ &= \frac{1}{a}[A_1 \ln(|\ x-a_1\ |) + A_2 \ln(|\ x-a_2\ |)) + \cdots + A_k \ln(|\ x-a_k\ |)] + C\end{aligned}$$

EXAMPLE 11 Evaluate $\int \frac{x+2}{x(x-1)(x-3)}\,dx.$

In this case, $p(x) = x+2$ and $q(x) = x(x-1)(x-3)$. Notice that $1 = degree(p(x)) < degree(q(x)) = 3$. Since $q(x)$ has the distinct roots 0, 1 and 3, we have a Type I Decomposition. Therefore, there are constants A, B, and C such that

$$\frac{x+2}{x(x-1)(x-3)} = \frac{A}{x} + \frac{B}{x-1} + \frac{C}{x-3}.$$

Step 1: Cross-multiply to get

$$x+2 = A(x-1)(x-3) + B(x)(x-3) + C(x)(x-1) \qquad (*)$$

Step 2: To find the constants we will substitute each of the roots into the identity (*). If $x = 0$, then

$$0+2 = A(0-1)(0-3) + B(0)(0-3) + C(0)(0-1)$$

or

$$2 = 3A.$$

Hence

$$A = \frac{2}{3}.$$

Let $x = 1$ to get

$$1+2 = A(1-1)(1-3) + B(1)(1-3) + C(1)(1-1).$$

Therefore,

$$3 = -2B$$

so that

$$B = -\frac{3}{2}.$$

Finally, with $x = 3$

$$3 + 2 = A(3-1)(3-3) + B(3)(3-3) + C(3)(3-1)$$

so

$$5 = 6C$$

and

$$C = \frac{5}{6}.$$

Notice that when we substitute the root a_j into identity (*), we get back the coefficient A_j corresponding to this root.

This means that

$$\frac{x+2}{x(x-1)(x-3)} = \frac{\frac{2}{3}}{x} + \frac{-\frac{3}{2}}{x-1} + \frac{\frac{5}{6}}{x-3}.$$

Therefore

$$\begin{aligned}\int \frac{x+2}{x(x-1)(x-3)}\,dx &= \frac{2}{3}\int \frac{1}{x}\,dx - \frac{3}{2}\int \frac{1}{x-1}\,dx + \frac{5}{6}\int \frac{1}{x-3}\,dx \\ &= \frac{2}{3}\ln(|\,x\,|) - \frac{3}{2}\ln(|\,x-1\,|) + \frac{5}{6}\ln(|\,x-3\,|) + c\end{aligned}$$

◀

DEFINITION

Type II Partial Fraction Decomposition

Assume that

$$f(x) = \frac{p(x)}{q(x)}$$

where $p(x)$ and $q(x)$ are polynomials such that

1. $degree(p(x)) < degree(q(x)) = k$,
2. $q(x)$ can be factored into the product of linear terms with non-distinct roots. That is

$$q(x) = a(x-a_1)^{m_1}(x-a_2)^{m_2}(x-a_3)^{m_3}\cdots(x-a_l)^{m_l}$$

   where at least one of the m_j's is greater than 1.

We say that f admits a *Type II Partial Fraction Decomposition*.

In this case, the partial fraction decomposition can be built as follows.

Each expression $(x - a_j)^{m_j}$ in the factorization of $q(x)$ will contribute m_j terms to the decomposition, one for each power of $x - a_j$ from 1 to m_j, which when combined will be of the form

$$\frac{p(x)}{q(x)} = \sum_{j=1}^{l} \frac{A_{j,1}}{x - a_j} + \frac{A_{j,2}}{(x - a_j)^2} + \frac{A_{j,3}}{(x - a_j)^3} + \cdots + \frac{A_{j,m_j}}{(x - a_j)^{m_j}}.$$

The number m_j is called the *multiplicity* of the root a_j.

EXAMPLE 12 Evaluate $\displaystyle\int \frac{1}{x^2(x-1)}\,dx$.

In this case, $p(x) = 1$ and $q(x) = x^2(x - 1)$. The roots of $q(x)$ are 0 and 1 and since the root 0 has multiplicity 2, this is a Type II Partial Fraction. Therefore, we can find constants A, B and C such that

$$\frac{1}{x^2(x-1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x-1}.$$

To find the constants A, B and C, we follow a similar procedure.

Step 1: Cross-multiply to get

$$1 = Ax(x - 1) + B(x - 1) + Cx^2 \qquad (*)$$

Step 2: Substitute the roots $x = 0$ and $x = 1$ into identity (*).

If $x = 0$, then

$$1 = A(0)(0 - 1) + B(0 - 1) + C(0^2)$$

or

$$1 = -B$$

so

$$B = -1.$$

Notice that substituting $x = 0$ only gave us the coefficient of the term with the *highest power of x* in the decomposition.

Next, let $x = 1$. Then

$$1 = A(1)(1 - 1) + B(1 - 1) + C(1^2)$$

and hence

$$1 = C.$$

Step 3: We have not yet found the coefficient A. There are a number of methods we could use to find A. We could, for example, substitute into the identity (*) any value other than 0 and 1 and use the fact that we already know B and C to solve for A. For example, if we let $x = 2$, we have

$$1 = A(2)(2-1) + B(2-1) + C(2^2)$$

or

$$1 = 2A + B + 4C.$$

Substituting $B = -1$ and $C = 1$ gives

$$1 = 2A - 1 + 4$$

or

$$-2 = 2A.$$

Hence

$$A = -1.$$

Perhaps an easier method is to *compare coefficients*. If we expand the expressions on the right-hand side of identity (*) we get

$$1 = A(x^2 - x) + B(x-1) + C(x^2) = (A+C)x^2 + (B-A)x - B.$$

We can rewrite this as

$$0x^2 + 0x + 1 = (A+C)x^2 + (B-A)x - B.$$

Since the two sides must agree for all x, they must both be the same polynomial. This means that the coefficients must be equal. In particular, the coefficients of x^2 must agree so that

$$0 = A + C$$

or

$$A = -C.$$

Since $C = 1$, this tells us that

$$A = -1.$$

Therefore,

$$\begin{aligned}
\int \frac{1}{x^2(x-1)}\,dx &= \int \frac{A}{x}\,dx + \int \frac{B}{x^2}\,dx + \int \frac{C}{x-1}\,dx \\
&= \int \frac{-1}{x}\,dx + \int \frac{-1}{x^2}\,dx + \int \frac{1}{x-1}\,dx \\
&= -\int \frac{1}{x}\,dx - \int \frac{1}{x^2}\,dx + \int \frac{1}{x-1}\,dx \\
&= -\ln(|\,x\,|) + \frac{1}{x} + \ln(|\,x-1\,|) + c
\end{aligned}$$

since

$$\begin{aligned}\int \frac{1}{x^2}\,dx &= \int x^{-2}\,dx \\ &= \frac{x^{-1}}{(-1)} + c \\ &= -\frac{1}{x} + c\end{aligned}$$

◀

Unfortunately, not all polynomials factor over the real numbers into products of linear terms. For example, the polynomial $x^2 + 1$ cannot be factored any further. This is an example of an *irreducible quadratic*. In fact, a quadratic $ax^2 + bx + c$ is irreducible if its discriminant $b^2 - 4ac < 0$.

However, the *Fundamental Theorem of Algebra* shows that every polynomial $q(x)$ factors in the form

$$q(x) = a(p_1(x))^{m_1}(p_2(x))^{m_2}(p_3(x))^{m_3}\cdots(p_k(x))^{m_k}$$

where each $p_i(x)$ is either of the form $(x - a)$ or it is an irreducible quadratic of the form $x^2 + bx + c$.

DEFINITION

Type III Partial Fraction Decomposition

Let $f(x) = \frac{p(x)}{q(x)}$ be a rational function with $degree(p(x)) < degree(q(x))$, but $q(x)$ does not factor into linear terms. We say that f admits a *Type III Partial Fraction Decomposition*.

In this case, the partial fraction decomposition can be built as follows:

Suppose that $q(x)$ has an irreducible factor $x^2 + bx + c$ with multiplicity m. Then this factor will contribute terms of the form

$$\frac{B_1x + C_1}{x^2 + bx + c} + \frac{B_2x + C_2}{(x^2 + bx + c)^2} + \cdots + \frac{B_mx + C_m}{(x^2 + bx + c)^m}$$

to the decomposition.

The linear terms are handled exactly as they were in the previous cases.

Note: We will not consider the case where $m > 1$ for some irreducible quadratic in evaluating integrals of rational functions with Type III Partial Fraction Decompositions. ◀

The method for finding the constants in a Type III Partical Fraction Decomposition is very similar to that of the first two types. We illustrate this with an example.

EXAMPLE 13 Evaluate $\int \frac{1}{x^3+x}\,dx$.

First observe that

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

so there will be constants A, B and C such that

$$f(x) = \frac{1}{x^3+x} = \frac{1}{x(x^2+1)} = \frac{A}{x} + \frac{Bx+C}{x^2+1}.$$

Step 1: To find the constants, we begin by cross-multiplying to obtain the identity

$$1 = A(x^2+1) + (Bx+C)x \qquad (*)$$

Step 2: Substitute $x = 0$, the only Real root, to find the coefficient A. This gives

$$1 = A(0^2+1) + (B(0)+C)(0)$$

or

$$A = 1.$$

Step 3: The remaining constants are found by comparing coefficients. Expanding the identity (*) gives

$$1 = (A+B)x^2 + Cx + A.$$

Comparing the coefficients of x^2 gives

$$0 = A + B$$

or

$$B = -A$$

Since $A = 1$, we get

$$B = -1.$$

Comparing the coefficients of x gives

$$C = 0.$$

Therefore,

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

and

$$\begin{aligned}
\int \frac{1}{x^3+x}\,dx &= \int \frac{1}{x(x^2+1)}\,dx \\
&= \int \frac{1}{x}\,dx - \int \frac{x}{x^2+1}\,dx \\
&= \ln(|\,x\,|) - \int \frac{x}{x^2+1}\,dx
\end{aligned}$$

To finish the calculation, we use the substitution $u = x^2 + 1$, so $du = 2x\,dx$ and $dx = \frac{du}{2x}$ to get

$$\begin{aligned}
\int \frac{x}{x^2+1}\,dx &= \int \frac{x}{u}\frac{du}{2x} \\
&= \frac{1}{2}\int \frac{1}{u}\,du \\
&= \frac{1}{2}\ln(|\,u\,|) + c \\
&= \frac{1}{2}\ln(x^2+1) + c
\end{aligned}$$

Putting this all together we get that

$$\int \frac{1}{x^3+x}\,dx = \ln(|\,x\,|) - \frac{1}{2}\ln(x^2+1) + c.$$

◀

NOTE

1. Since $C = 0$ in the previous example, the integral was straightforward to evaluate. Unfortunately, this is not the case for most Type III Partial Fractions.
2. All of the Partial Fraction Decompositions required the rational function to satisfy $degree(p(x)) < degree(q(x))$. If this is not the case, then we need to use division of polynomials to find new polynomials $r(x)$ and $p_1(x)$ such that

$$\frac{p(x)}{q(x)} = r(x) + \frac{p_1(x)}{q(x)}$$

with $degree(p_1(x)) < degree(q(x))$.

◀

2.4 Introduction to Improper Integrals

The Riemann Integral was defined for certain bounded functions on closed intervals $[a, b]$. In many applications, most notably for statistical and data analysis, we want to be able to integrate functions over intervals of *infinite* length. However, at this point

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

has no meaning.

If $f(x) \geq 0$ and $a < b$, then

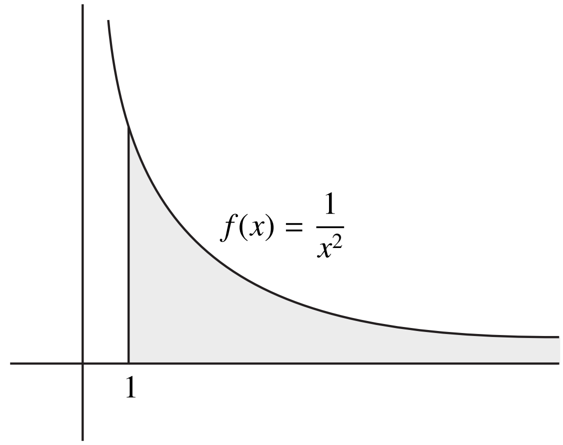
$$\int_a^b f(x)\,dx$$

can be interpreted geometrically as the area bounded by the graph of $y = f(x)$, the x-axis and the vertical lines $x = a$ and $x = b$. We could use this to guide us in defining the integral over an infinite interval.

For example, for the function $f(x) = \frac{1}{x^2}$, we might want

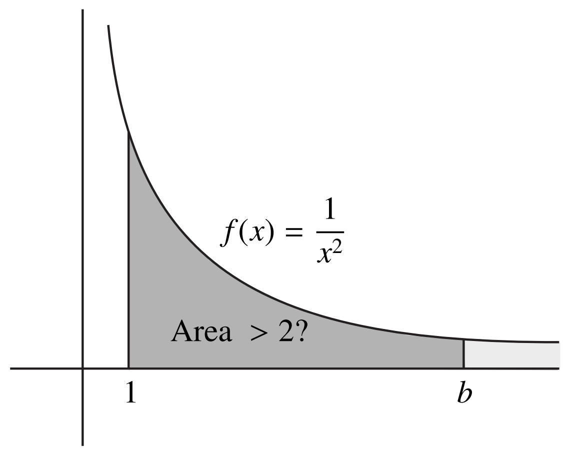
$$\int_1^\infty \frac{1}{x^2}\,dx$$

to represent the area bounded by the graph of $y = f(x)$, the x-axis and the vertical line $x = 1$.



Since this region is *unbounded* it might seem likely that this area should be infinite.

However, if this was true then it should be the case that by choosing b large enough we should be able to make the area bounded by the graph of $y = f(x)$, the x-axis and the vertical lines $x = 1$ and $x = b$ at least as large as 2.



However, this area is given by $\int_1^b \frac{1}{x^2}\,dx$ and

$$\begin{aligned}
\int_1^b \frac{1}{x^2}\,dx &= \int_1^b x^{-2}\,dx \\
&= -\frac{1}{x}\Big|_1^b \\
&= -\frac{1}{b} + 1 \\
&= 1 - \frac{1}{b} \\
&< 2
\end{aligned}$$

This shows that no matter how large b is the area bounded by the graph of $y = f(x)$, the x-axis and the vertical lines $x = 1$ and $x = b$ will always be less than 2. In fact, it is always less than 1. This suggests that the original region should have *finite area* despite the fact that it is unbounded.

We are now in a position similar to when we first defined the integral. While we have an intuitive idea of what *area* means, we do not have a formal definition of area that applies to such unbounded regions. One method to avoid this problem would be to define the area of the unbounded region as the limit of the bounded areas as b goes to ∞. That is

$$\text{Area} = \lim_{b\to\infty} \int_1^b \frac{1}{x^2}\,dx.$$

In particular, we would have

$$\begin{aligned}\text{Area} &= \lim_{b\to\infty} \int_1^b \frac{1}{x^2}\,dx \\ &= \lim_{b\to\infty} 1 - \frac{1}{b} \\ &= 1\end{aligned}$$

In this case, we would like to have

$$\int_1^\infty \frac{1}{x^2}\,dx = 1.$$

This motivates the following definition:

DEFINITION

Type I Improper Integral

1) Let f be integrable on $[a, b]$ for each $a \le b$. We say that the *Type I Improper Integral*

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

converges if

$$\lim_{b\to\infty} \int_a^b f(x)\,dx$$

exists. In this case, we write

$$\int_a^\infty f(x)\,dx = \lim_{b\to\infty} \int_a^b f(x)\,dx.$$

Otherwise, we say that $\int_a^\infty f(x)\,dx$ *diverges*.

2) Let f be integrable on $[b, a]$ for each $b \leq a$. We say that the *Type I Improper Integral*

$$\int_{-\infty}^{a} f(x)\,dx$$

converges if

$$\lim_{b\to-\infty} \int_{b}^{a} f(x)\,dx$$

exists. In this case, we write

$$\int_{-\infty}^{a} f(x)\,dx = \lim_{b\to-\infty} \int_{b}^{a} f(x)\,dx.$$

Otherwise, we say that $\int_{-\infty}^{a} f(x)\,dx$ diverges.

3) Assume that f is integrable on $[a, b]$ for each $a, b \in \mathbb{R}$ with $a < b$. We say that the *Type I Improper Integral*

$$\int_{-\infty}^{\infty} f(x)\,dx$$

converges if both $\int_{-\infty}^{c} f(x)\,dx$ and $\int_{c}^{\infty} f(x)\,dx$ converge for some $c \in \mathbb{R}$. In this case, we write

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{c} f(x)\,dx + \int_{c}^{\infty} f(x)\,dx$$

Otherwise, we say that $\int_{-\infty}^{\infty} f(x)\,dx$ diverges.

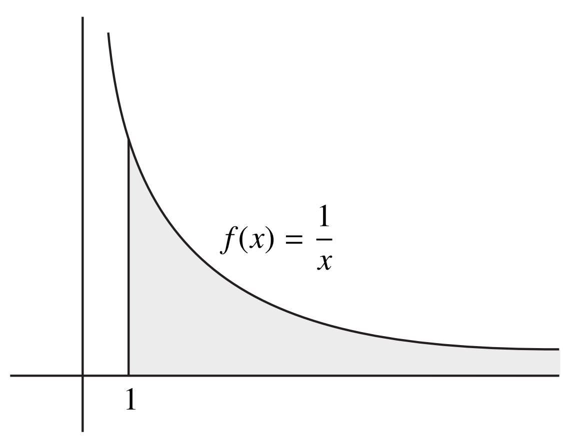
Note: In general, we will focus our attention on Type I improper integrals of the form

$$\int_{a}^{\infty} f(x)\,dx$$

◀

EXAMPLE 14 Show that $\int_{1}^{\infty} \frac{1}{x}\,dx$ diverges.

$\int_{1}^{\infty} \frac{1}{x}\,dx$ represents the area of the region bounded by the graph of $f(x) = \frac{1}{x}$, the x-axis and the vertical line $x = 1$.



This unbounded region looks very similar to the region discussed previously. However, this time

$$\begin{aligned}
\int_1^\infty \frac{1}{x}\,dx &= \lim_{b\to\infty}\int_1^b \frac{1}{x}\,dx \\
&= \lim_{b\to\infty} \ln(x)\Big|_1^b \\
&= \lim_{b\to\infty}(\ln(b)-\ln(1)) \\
&= \lim_{b\to\infty}\ln(b) \\
&= \infty
\end{aligned}$$

This shows that $\int_1^\infty \frac{1}{x}\,dx$ diverges to ∞ and hence that the area of the region bounded by the graph of $f(x)=\frac{1}{x}$, the x-axis and the vertical line $x=1$ is also infinite. ◀

We have seen that $\int_1^\infty \frac{1}{x^2}\,dx$ converges while $\int_1^\infty \frac{1}{x}\,dx$ diverges. More generally, we have the following natural question.

Question: For which p does

$$\int_1^\infty \frac{1}{x^p}\,dx$$

converge? In fact, the answer to this question will be crucial to our study of series. Since we already know what happens if $p=1$, we can assume that $p\neq 1$.

To answer this question we require the following facts. If $\alpha>0$, then

$$\lim_{b\to\infty} b^\alpha = \infty$$

and if $\alpha<0$, then

$$\lim_{b\to\infty} b^\alpha = 0.$$

For any $b>1$,

$$\begin{aligned}
\int_1^b \frac{1}{x^p}\,dx &= \int_1^b x^{-p}\,dx \\
&= \frac{x^{-p+1}}{-p+1}\Big|_1^b \\
&= \frac{b^{1-p}}{1-p}-\frac{1}{-p+1} \\
&= \frac{b^{1-p}}{1-p}+\frac{1}{p-1}
\end{aligned}$$

If $p < 1$, then $1 - p > 0$. Therefore,

$$\lim_{b\to\infty} \frac{b^{1-p}}{1-p} = \infty$$

since the *exponent is positive*. This means that

$$\lim_{b\to\infty} \frac{b^{1-p}}{1-p} + \frac{1}{p-1} = \infty$$

and hence that $\int_1^\infty \frac{1}{x^p}\,dx$ diverges.

However, if $p > 1$, then $1 - p < 0$. This time since the *exponent is negative*,

$$\lim_{b\to\infty} \frac{b^{1-p}}{1-p} = 0$$

and hence

$$\lim_{b\to\infty} \frac{b^{1-p}}{1-p} + \frac{1}{p-1} = \frac{1}{p-1}.$$

Therefore, if $p > 1$, $\int_1^\infty \frac{1}{x^p}\,dx$ converges and

$$\int_1^\infty \frac{1}{x^p}\,dx = \frac{1}{p-1}.$$

This is summarized in the following important theorem.

THEOREM 3

p-Test for Type I Improper Integrals

The improper integral

$$\int_1^\infty \frac{1}{x^p}\,dx$$

converges if and only if $p > 1$.

If $p > 1$, then

$$\int_1^\infty \frac{1}{x^p}\,dx = \frac{1}{p-1}.$$

We end this section with one more important example of a convergent improper integral.

EXAMPLE 15 Evaluate $\int_0^\infty e^{-x}\,dx$.

We have

$$\begin{aligned}\int_0^\infty e^{-x}\,dx &= \lim_{b\to\infty}\int_0^b e^{-x}\,dx \\ &= \lim_{b\to\infty} -e^{-x}\Big|_0^b \\ &= \lim_{b\to\infty}(-e^{-b}+e^0) \\ &= \lim_{b\to\infty}(-e^{-b}+1) \\ &= 1\end{aligned}$$

since $\lim\limits_{b\to\infty} -e^{-b} = 0$. ◀

2.4.1 Properties of Type I Improper Integrals

Since the evaluation of an improper integral results from taking limits at ∞, it makes sense that improper integrals should inherit many of the properties of these types of limits.

THEOREM 4 **Properties of Type I Improper Integrals**

Assume that $\int_a^\infty f(x)\,dx$ and $\int_a^\infty g(x)\,dx$ both converge.

1. $\int_a^\infty cf(x)\,dx$ converges for each $c \in \mathbb{R}$ and
$$\int_a^\infty cf(x)\,dx = c\int_a^\infty f(x)\,dx.$$
2. $\int_a^\infty (f(x)+g(x))\,dx$ converges and
$$\int_a^\infty (f(x)+g(x))\,dx = \int_a^\infty f(x)\,dx + \int_a^\infty g(x)\,dx.$$
3. If $f(x) \le g(x)$ for all $a \le x$, then
$$\int_a^\infty f(x)\,dx \le \int_a^\infty g(x)\,dx.$$
4. If $a < c < \infty$, then $\int_c^\infty f(x)\,dx$ converges and
$$\int_a^\infty f(x)\,dx = \int_a^c f(x)\,dx + \int_c^\infty f(x)\,dx.$$

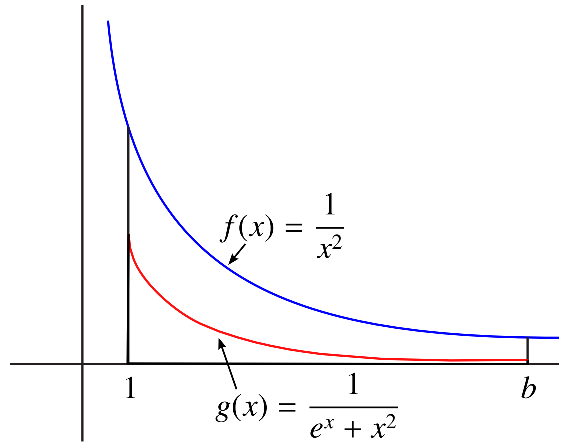
Note: Unfortunately, it is usually not possible to explicitly evaluate an integral $\int_a^b f(x)\,dx$ for every $b > 1$. Therefore, it may be difficult to apply the definition to determine if an improper integral converges. For example, it is not obvious how to evaluate the integral

$$\int_1^b \frac{1}{e^x + x^2}\,dx.$$

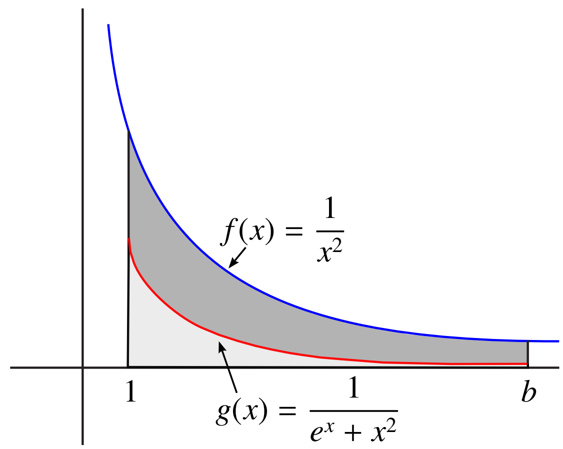
However, we do know that

$$\int_1^b \frac{1}{e^x + x^2}\,dx \le \int_1^b \frac{1}{x^2}\,dx$$

for each $b \ge 1$ since $e^x > 0$ for each $x \ge 1$.



From this we can immediately conclude that the area under the graph of $g(x) = \frac{1}{e^x + x^2}$ between $x = 1$ and $x = b$ should be less than the area under the graph of $f(x) = \frac{1}{x^2}$ between $x = 1$ and $x = b$.



We also know that $\int_1^\infty \frac{1}{x^2}\,dx$ is finite and

$$\int_1^b \frac{1}{x^2}\,dx < \int_1^\infty \frac{1}{x^2}\,dx$$

for each $b \ge 1$. It follows that for every $b \ge 1$

$$\int_1^b \frac{1}{e^x + x^2}\,dx < \int_1^\infty \frac{1}{x^2}\,dx.$$

This suggests that the area of the region under the graph of $g(x) = \frac{1}{e^x + x^2}$ which lies above the x-axis and to the right of the line $x = 1$ should be finite.

For functions g with $g(x) > 0$ for all $x \ge a$, we want to interpret the integral $\int_a^\infty g(x)\,dx$ as the area of the region bounded by the graph of $y = g(x)$, the x-axis and the vertical line $x = a$. To say that the $\int_a^\infty g(x)\,dx$ converges should be equivalent to the area being finite. As such we should be able to conclude that $\displaystyle\int_1^{\infty} \frac{1}{e^x + x^2}\, dx$ converges.

Unfortunately, we have not explicitly shown that

$$\lim_{b\to\infty} \int_1^b \frac{1}{e^x + x^2}\, dx$$

actually exists because we cannot evaluate $\displaystyle\int_1^b \frac{1}{e^x + x^2}\, dx$ with any of the techniques we have developed.

We can however make the following observation: If we let $G(b) = \displaystyle\int_1^b \frac{1}{e^x + x^2}\, dx$, then when viewed as a function on the interval $[1, \infty)$, G is increasing and

$$G(b) \leq \int_1^{\infty} \frac{1}{x^2}\, dx$$

for all $b \in [1, \infty)$. In the next section, we will see that this is enough to show that

$$\lim_{b\to\infty} G(b) = \lim_{b\to\infty} \int_1^b \frac{1}{e^x + x^2}\, dx$$

exists and hence that the improper integral $\displaystyle\int_1^{\infty} \frac{1}{e^x + x^2}\, dx$ does in fact converge.

2.4.2 Comparison Test for Type I Improper Integrals

Suppose that

$$0 \leq g(x) \leq f(x)$$

on $[a, \infty)$. Assume also that $\int_a^{\infty} f(x)\, dx$ converges. Then the area under the graph of f from $x = a$ to ∞ is finite. But $0 \leq g(x) \leq f(x)$ so the area under the graph of g from $x = a$ to ∞ should be less than the area under the graph of f, and hence it should also be finite. This should imply that $\int_a^{\infty} g(x)\, dx$ converges.

On the other hand, if $\int_a^{\infty} g(x)\, dx$ diverges, then the area under the graph of g from $x = a$ to ∞ is infinite. Since $f(x)$ is larger than $g(x)$, it should be true that area under the graph of f from $x = a$ to ∞ is infinite. That is, $\int_a^{\infty} f(x)\, dx$ diverges.

In summary, if

$$0 \leq g(x) \leq f(x)$$

on $[a, \infty)$ and the integral of the larger function converges, so does the integral of the smaller function. However, if the smaller function has an integral that diverges to infinity, so should the larger function. To see why this is the case, we can use an analogue of the Monotone Convergence Theorem for functions.

Recall that the *Monotone Convergence Theorem* tells us that a **non-decreasing** sequence $\{a_n\}$ converges if and only if it is bounded above and that if it does converge, then $\lim\limits_{n\to\infty} a_n = lub(\{a_n\})$. We can now prove a similar result for functions.

THEOREM 5 **The Monotone Convergence Theorem for Functions**

Assume that f is non-decreasing on $[a, \infty)$.

1. If $\{f(x) \mid x \in [a, \infty)\}$ is bounded above, then $\lim\limits_{x\to\infty} f(x)$ exists and

$$\lim_{x\to\infty} f(x) = L = lub(\{f(x) \mid x \in [a, \infty)\}).$$

2. If $\{f(x) \mid x \in [a, \infty)\}$ is not bounded above, then $\lim\limits_{x\to\infty} f(x) = \infty$.

PROOF

The proof of this theorem is very similar to that of the Monotone Convergence Theorem for sequences.

1. Assume that $\{f(x) \mid x \in [a, \infty)\}$ is bounded above, and let

$$L = lub(\{f(x) \mid x \in [a, \infty)\}).$$

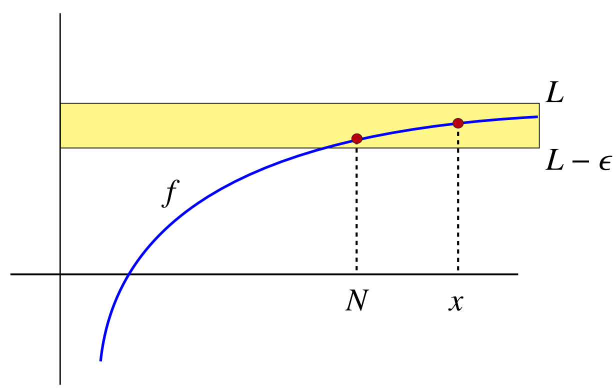
Let $\epsilon > 0$. Then $L - \epsilon$ is not an upper bound for $\{f(x) \mid x \in [a, \infty)\}$. Therefore, there exists an $N \in [a, \infty)\}$ so that

$$L - \epsilon < f(N) \le L.$$

But if $x \ge N$ we would have that

$$L - \epsilon < f(N) \le f(x) \le L.$$

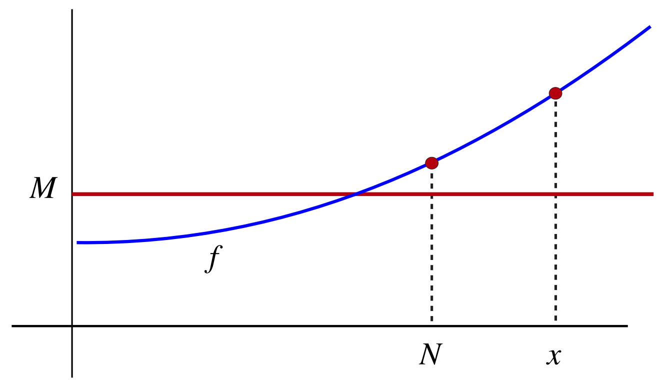
This means that if $x \ge N$, then $|f(x) - L| < \epsilon$ so that $\lim\limits_{x\to\infty} f(x) = L$ as claimed.



2. Assume $\{f(x) \mid x \in [a, \infty)\}$ is not bounded above. Let $M > 0$. Since M is not an upper bound for $\{f(x) \mid x \in [a, \infty)\}$, there exists an $N \in [a, \infty)\}$ so that $M < f(N)$. But if $x \ge N$, we have

$$M < f(N) \le f(x)$$

which shows that $\lim\limits_{x\to\infty} f(x) = \infty$.



■

We can now establish *one of the most important tools* for determining the convergence or divergence of improper integrals.

THEOREM 6 **Comparison Test for Type I Improper Integrals**

Assume that $0 \leq g(x) \leq f(x)$ for all $x \geq a$ and that both f and g are continuous on $[a, \infty)$.

1. If $\int_a^\infty f(x)\,dx$ converges, then so does $\int_a^\infty g(x)\,dx$.
2. If $\int_a^\infty g(x)\,dx$ diverges, then so does $\int_a^\infty f(x)\,dx$.

PROOF

The two statements are logically equivalent (why?). This means that we only have to prove the first statement and then the second statement follows. As such, assume that $\int_a^\infty f(x)\,dx$ converges. Next let

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

Since $f(x) \geq 0$, we have that F is non-decreasing on $[a, \infty)$ and that

$$\lim_{t\to\infty} F(t) = \lim_{t\to\infty} \int_a^t f(x)\,dx = \int_a^\infty f(x)\,dx < \infty.$$

Let

$$G(t) = \int_a^t g(x)\,dx.$$

This time since $0 \leq g(x) \leq f(x)$ we have that G is non-decreasing on $[a, \infty)$ and that $G(t) \leq F(t)$ for any $t \in [a, \infty)$. But then

$$G(t) \leq \int_a^\infty f(x)\,dx < \infty$$

for all $t \in [a, \infty)$. This shows that $\{G(t) \mid t \in [a, \infty)\}$ is bounded above. Finally, the Monotone Convergence Theorem for Functions tells us that

$$\lim_{t\to\infty} G(t) = \lim_{t\to\infty} \int_a^t g(x)\,dx$$

exists. This proves statement 1. ■

EXAMPLE 16 Show that $\int_1^\infty \frac{1}{e^x+x^2}\,dx$ converges.

We have already seen that

$$0 < \frac{1}{e^x + x^2} < \frac{1}{x^2}$$

for all $x \geq 1$ and that $\int_1^\infty \frac{1}{x^2}\,dx$ converges. It follows immediately from the Comparison Theorem that $\int_1^\infty \frac{1}{e^x+x^2}\,dx$ also converges. ◀

EXAMPLE 17 Does $\int_1^\infty \dfrac{1}{x + \sqrt{x}}\,dx$ converge or diverge?

We know that

$$0 < \frac{1}{x + \sqrt{x}} < \frac{1}{x}$$

for all $x \geq 1$. However, $\int_1^\infty \frac{1}{x}\,dx$ diverges so the Comparison Test does not apply since *we cannot say anything about the smaller integral if the larger one diverges*.

The key observation is that $x + \sqrt{x} \leq x + x = 2x$ for $x \geq 1$. Therefore,

$$0 < \frac{1}{2x} \leq \frac{1}{x + \sqrt{x}}$$

for $x \geq 1$. Moreover, since $\int_1^\infty \frac{1}{x}\,dx$ diverges, so does

$$\int_1^\infty \frac{1}{2x}\,dx.$$

This time we can use the Comparison Test to conclude that $\int_1^\infty \frac{1}{x+\sqrt{x}}\,dx$ diverges. ◀

So far we have dealt almost exclusively with improper integrals involving positive functions. In particular, the Comparison Test applies to positive functions.

Moreover the Monotone Convergence Theorem for Functions allows us to establish the following important fact.

Fact

If f is integrable on $[a, b)$ for every $b \geq a$ and if $f(x) \geq 0$ on $[a, \infty)$, then $\int_a^\infty f(x)\,dx$ converges if and only if there exists an M such that

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

for all $b > a$.

Note: The previous statement is not true without the assumption that $f(x) \geq 0$ as the following example illustrates. ◀

EXAMPLE 18 Show that $\int_0^\infty \cos(x)\,dx$ diverges.

By definition,

$$\begin{aligned}
\int_0^\infty \cos(x)\,dx &= \lim_{b\to\infty} \int_0^b \cos(x)\,dx \\
&= \lim_{b\to\infty} \sin(x)\Big|_0^b \\
&= \lim_{b\to\infty} (\sin(b) - \sin(0)) \\
&= \lim_{b\to\infty} \sin(b)
\end{aligned}$$

However, $\sin(b)$ oscillates between -1 and 1 as $b \to \infty$ so that $\lim\limits_{b\to\infty} \sin(b)$ does not exist. Therefore, $\int_0^\infty \cos(x)\,dx$ diverges despite the fact that

$$-1 \leq \int_0^b \cos(x)\,dx = \sin(b) \leq 1$$

for all $b > 0$.

◀

However, we can still often use our tools for improper integrals of positive functions to determine convergence of some improper integrals of more general functions. To do so we introduce the notion of *absolute convergence* which is an analog of absolute convergence for sequences.

DEFINITION

Absolute Convergence for Type I Improper Integrals

Let f be integrable on $[a, b)$ for all $b \geq a$. We say that the improper integral $\int_a^\infty f(x)\,dx$ *converges absolutely* if

$$\int_a^\infty |\ f(x)\ |\ dx$$

converges.

Similar to the case for sequences we will now see that absolute convergence implies convergence for improper integrals.

THEOREM 7

Absolute Convergence Theorem for Improper Integrals

Let f be integrable on $[a, b]$ for all $b > a$. Then $|f|$ is also integrable on $[a, b]$ for all $b > a$. Moreover, if we assume that

$$\int_a^\infty |\ f(x)\ |\ dx$$

converges, then so does

$$\int_a^\infty f(x)\,dx.$$

In particular, if $0 \leq |f(x)| \leq g(x)$ for all $x \geq a$, both f and g are integrable on $[a, b]$ for all $b \geq a$, and if $\int_a^\infty g(x)\,dx$ converges, then so does

$$\int_a^\infty f(x)\,dx.$$

PROOF

Assume that $\int_a^\infty |\ f(x)\ |\ dx$ converges. Then so does

$$\int_a^\infty 2\,|\ f(x)\ |\ dx.$$

We also have that

$$0 \leq f(x) + |f(x)| \leq 2|f(x)|$$

so by the Comparison Theorem

$$\int_a^\infty f(x) + |\ f(x)\ |\ dx$$

converges. Finally, since

$$f(x) = [f(x) + |f(x)|] - |f(x)|$$

we get that $\int_a^\infty f(x)\,dx$ converges with

$$\int_a^\infty f(x)\,dx = \int_a^\infty f(x) + |\,f(x)\,|\;dx - \int_a^\infty |\,f(x)\,|\;dx.$$

To prove the second statement, assume that $\int_a^\infty g(x)\,dx$ converges. The Comparison Test shows that

$$\int_a^\infty |\,f(x)\,|\;dx$$

also converges. We can now apply the first statement to conclude that

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

converges. ■

EXAMPLE 19 Show that $\int_3^\infty \frac{\cos(x)}{x^2+2x+1}\,dx$ converges.

We know

$$\begin{aligned} \left|\frac{\cos(x)}{x^2+2x+1}\right| &\leq \frac{1}{x^2+2x+1} \\ &\leq \frac{1}{x^2} \end{aligned}$$

for all $x \geq 3$.

The p-Test shows that $\int_3^\infty \frac{1}{x^2}\,dx$ converges.

Therefore, by the Comparison Test,

$$\int_3^\infty \left|\frac{\cos(x)}{x^2+2x+1}\right| dx$$

converges.

The Absolute Convergence Theorem shows that $\int_3^\infty \frac{\cos(x)}{x^2+2x+1}\,dx$ converges.

◀

2.4.3 The Gamma Function

An important class of examples of improper integrals arise as values of a function called the *Gamma function*.

DEFINITION

The Gamma Function

For each $x \in \mathbb{R}$, define
$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$
The function Γ is called the Gamma function.

Observation: In order to properly define the Γ function we should really show that the improper integral that arises from each choice of x is actually convergent. We have actually already seen this to be true for $x = 1$ as the example below reminds us. Later we will provide strong evidence of why this is so in general, but the verification of convergence is left as an exercise.

EXAMPLE 20 Calculate
$$\Gamma(1) = \int_0^\infty t^0 e^{-t}dt = \int_0^\infty e^{-t}dt.$$

By definition
$$\begin{aligned}
\Gamma(1) &= \int_0^\infty e^{-t}dt \\
&= \lim_{b\to\infty}\int_0^b e^{-t}dt \\
&= \lim_{b\to\infty} -e^{-t}\Big|_0^b \\
&= \lim_{b\to\infty}(-e^{-b}) - (-e^{-0}) \\
&= 1
\end{aligned}$$

◀

Note: To see why the integrals involved in the definition of the Gamma function always converge we first note that by modifying the previous example we can show that for any $M > 0$, the improper integral
$$\int_M^\infty e^{-\frac{t}{2}}dt$$
also converges. Next we observe that the Fundamental Log Limit shows that for any $x \in \mathbb{R}$ we have that
$$\lim_{t\to\infty} t^{x-1}e^{-\frac{t}{2}} = 0.$$
Combining these two observations and using the Comparison Test for Improper Integrals we can show that $\int_0^\infty t^{x-1}e^{-t}dt$ is always convergent. ◀

Observation: If we apply the Integration by Parts formula we get the following interesting result:

$$\int t^x e^{-t} dt = -t^x e^{-t} + x \cdot \int t^{x-1} e^{-t} dt.$$

It follows that

$$\begin{aligned}
\Gamma(x+1) &= \lim_{b\to\infty} \int_0^b t^x e^{-t} dt \\
&= \lim_{b\to\infty} -t^x e^{-t}\Big|_0^b + x \cdot \int_0^b t^{x-1} e^{-t} dt \\
&= \lim_{b\to\infty} -b^x e^{-b} + x \cdot \lim_{b\to\infty} \int_0^b t^{x-1} e^{-t} dt \\
&= 0 + x \cdot \int_0^\infty t^{x-1} e^{-t} dt \\
&= x \cdot \Gamma(x)
\end{aligned}$$

As an immediate application of this observation we can show that for any $n \in \mathbb{N}$, we have that

$$\Gamma(n) = (n-1)!$$

In fact, we know that $\Gamma(1) = 1 = (1-0)!$. Now suppose that $\Gamma(k) = (k-1)!$. Then

$$\begin{aligned}
\Gamma(k+1) &= k \cdot \Gamma(k) \\
&= k \cdot (k-1)! \\
&= k!
\end{aligned}$$

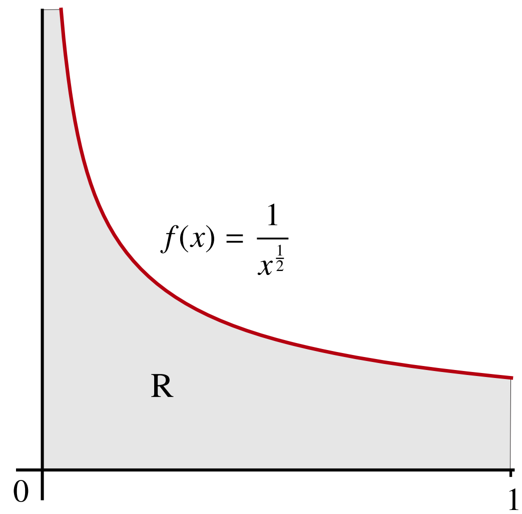
so we can deduce that $\Gamma(n) = (n-1)!$ by using Mathematical Induction. For this reason the Γ function is viewed as a means of generating *factorial* values for non-natural numbers.

Note: The Γ function also has important applications in statistics and probability theory.

2.4.4 Type II Improper Integrals

So far in considering improper integrals we have only considered the case where the interval over which we are integrating is unbounded. There is a second type of improper integral which we call a Type II Improper Integral. In the case of a Type II Improper Integral the assumption will be that the integrand f has a vertical asymptote at some point $a \in \mathbb{R}$. To illustrate what we mean by a Type II Improper Integral we consider the function $f(x) = \frac{1}{x^{\frac{1}{2}}}$ on the interval $(0, 1]$. We might ask:

Problem: What is the area under the graph of f above the x-axis and between the lines $x = 0$ and $x = 1$?
We note that f has a vertical asymptote at $x = 0$ and as a consequence this region (which we denote by R) is also *unbounded*.



However, if this was the case, then for any $M > 0$ we should be able to find a $b \in (0.1]$ so that the area under the graph of f above the x-axis and between the lines $x = b$ and $x = 1$ should be at least M. We know that this latter area is

$$\begin{aligned}\int_b^1 \frac{1}{x^{\frac{1}{2}}}\,dx &= 2x^{\frac{1}{2}}\Big|_b^1 \\ &= 2(1-\sqrt{b}) \\ &< 2\end{aligned}$$

Just as we deduced that the unbounded area under the graph of $f(x) = \frac{1}{x^2}$ on the interval $[1, \infty)$ could be finite, we again see that the area of this unbounded region R could be finite. In fact, similar to Improper Integrals of Type I, we could define the area A of region R to be

$$A = \lim_{b\to 0^+} \int_b^1 \frac{1}{x^{\frac{1}{2}}}\,dx = \lim_{b\to 0^+} 2x^{\frac{1}{2}}\Big|_b^1 = \lim_{b\to 0^+} 2(1-\sqrt{b}) = 2.$$

DEFINITION

Type II Improper Integral

1) Let f be integrable on $[t, b]$ for every $t \in (a, b]$ with either $\lim\limits_{x\to a^+} f(x) = \infty$ or $\lim\limits_{x\to a^+} f(x) = -\infty$. We say that the *Type II Improper Integral*

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

converges if

$$\lim_{t\to a^+} \int_t^b f(x)\,dx$$

exists. In this case, we write

$$\int_a^b f(x)\,dx = \lim_{t\to a^+} \int_t^b f(x)\,dx.$$

Otherwise, we say that $\int_a^b f(x)\,dx$ *diverges*.

2) Let f be integrable on $[a,t]$ for every $t \in [a,b)$ with either $\lim\limits_{x\to b^-} f(x) = \infty$ or $\lim\limits_{x\to b^-} f(x) = -\infty$. We say that the *Type II Improper Integral*

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

converges if

$$\lim_{t\to b^-} \int_a^t f(x)\,dx$$

exists. In this case, we write

$$\int_a^b f(x)\,dx = \lim_{t\to b^-} \int_a^t f(x)\,dx.$$

Otherwise, we say that $\int_a^b f(x)\,dx$ *diverges*.

3) If f has an infinite discontinuity at $x = c$ where $a < c < b$, then we say that the *Type II Improper Integral*

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

converges if both $\displaystyle\int_a^c f(x)\,dx$ and $\displaystyle\int_c^b f(x)\,dx$ converge. In this case, we write

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

If one or both of these integrals diverge, then we say that $\displaystyle\int_a^b f(x)\,dx$ *diverges*.

Remark: In determining the convergence or divergence of a Type I Improper Integral the p-test was an important tool. There is a natural analog of the p-test for Type II Improper Integrals.

THEOREM 8 **p-Test for Type II Improper Integrals**

The improper integral

$$\int_0^1 \frac{1}{x^p}\,dx$$

converges if and only if $p < 1$.

If $p < 1$, then

$$\int_0^1 \frac{1}{x^p}\,dx = \frac{1}{1-p}.$$

PROOF

First assume that $p \neq 1$. By definition

$$\begin{aligned}\int_0^1 \frac{1}{x^p}\,dx &= \lim_{t\to 0^+}\int_t^1 \frac{1}{x^p}\,dx \\ &= \lim_{t\to 0^+} \frac{1}{1-p}x^{1-p}\Big|_t^1 \\ &= \lim_{t\to 0^+} \frac{1}{1-p} - \frac{1}{1-p}t^{1-p}\end{aligned}$$

Now if $p < 1$, then $1 - p > 0$ so

$$\lim_{t\to 0^+} \frac{1}{1-p} - \frac{1}{1-p}t^{1-p} = \frac{1}{1-p}.$$

Now if $p > 1$, then $1 - p < 0$ so

$$\lim_{t\to 0^+} \frac{1}{1-p} - \frac{1}{1-p}t^{1-p} = \infty.$$

Finally, if $p = 1$, then

$$\begin{aligned}\int_0^1 \frac{1}{x}\,dx &= \lim_{t\to 0^+}\int_t^1 \frac{1}{x}\,dx \\ &= \lim_{t\to 0^+} \ln(x)\Big|_t^1 \\ &= \lim_{t\to 0^+} \ln(1) - \ln(t) \\ &= \infty\end{aligned}$$

It follows that $\int_0^1 \frac{1}{x^p}\,dx$ converges precisely when $p < 1$ and that in this case

$$\int_0^1 \frac{1}{x^p}\,dx = \frac{1}{1-p}$$

as claimed. ∎

Chapter 3

Applications of Integration

In this chapter, we will consider four types of calculations that use integration: areas between curves, volumes using the disk method, volumes using the shell method, and arc length.

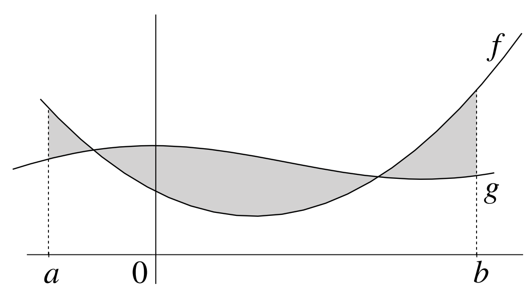
3.1 Areas Between Curves

We have already seen that there is a strong relationship between integration and area. In particular, if the continuous function f is positive on $[a, b]$, then we interpreted $\int_a^b f(t)\,dt$ to be the area under the graph of f that is above the t-axis and bounded by the lines $t = a$ and $t = b$.

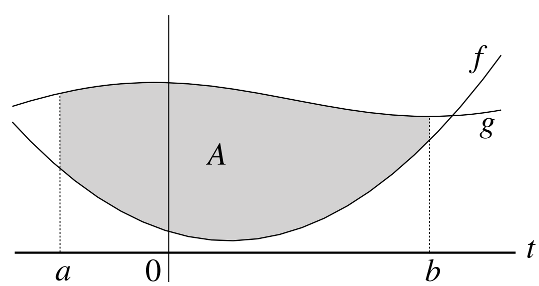
In this section, integration is used to answer a more general problem—that of calculating areas *between* curves (rather than between the curve and the x-axis).

Problem

Let f and g be continuous on an interval $[a, b]$. Find the area of the region bounded by the graphs of the two functions, f and g, and the lines $t = a$ and $t = b$.



Let's begin with a simple case. Assume that $f(t) \leq g(t)$ for all $t \in [a, b]$. The task is to find the area A of the region in the diagram.

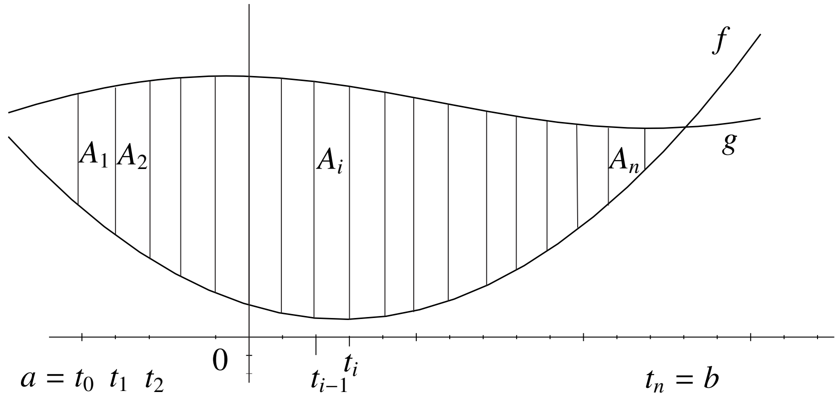


We will use a method similar to our previous area calculation under a curve with Riemann sums. That is, we begin by constructing a regular n-partition

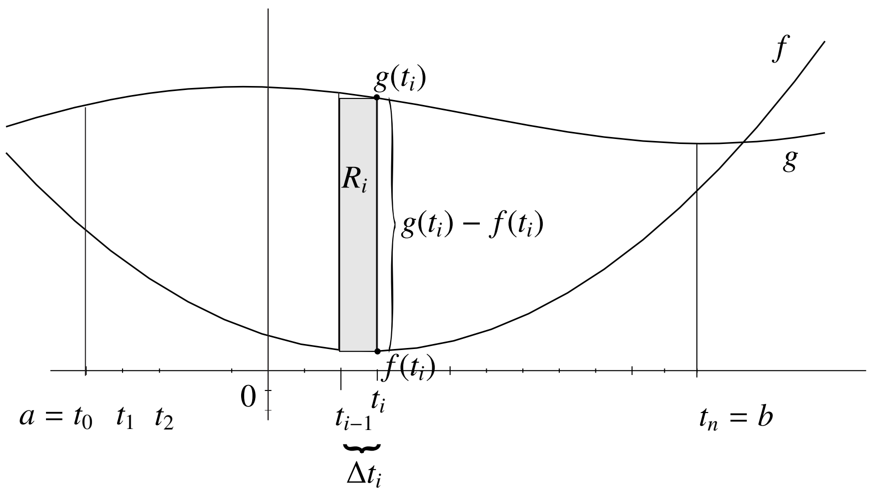
$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

with $\Delta t_i = \frac{b-a}{n}$ and $t_i = a + \frac{i(b-a)}{n}$.

This partition divides A into n subregions which we label as $A_1, A_2, \cdots, A_n$ where A_i is the region bounded by the graphs f and g, and the lines $t = t_{i-1}$ and $t = t_i$.



A rectangle R_i can now be used to estimate the area A_i as follows:



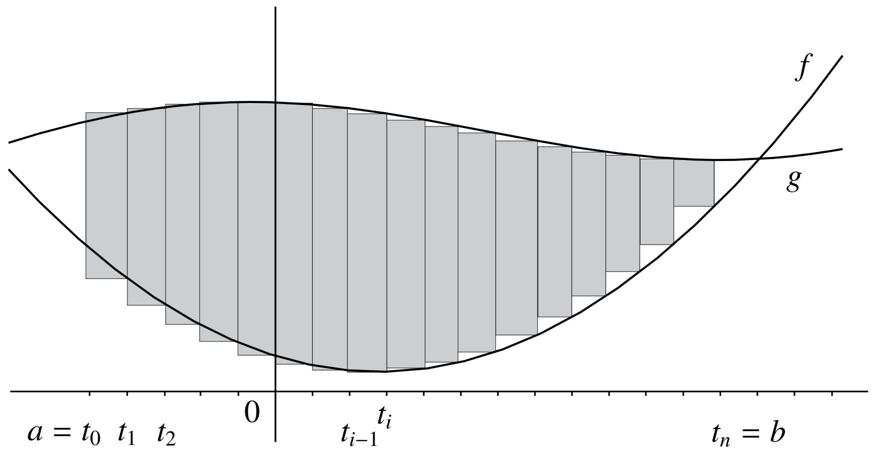
The height of the rectangle R_i is $h = g(t_i) - f(t_i)$ and its width is $\Delta t_i = \frac{b-a}{n}$, so the area A_i is estimated by

$$A_i \cong (g(t_i) - f(t_i))\,\Delta t_i.$$

Thus

$$\begin{aligned} A &= \sum_{i=1}^{n} A_i \\ &\cong \sum_{i=1}^{n} (g(t_i) - f(t_i))\Delta t_i \\ &\cong \sum_{i=1}^{n} (g(t_i) - f(t_i))\frac{b-a}{n} \end{aligned}$$

with the latter sum equal to a right-hand Riemann sum for the function $g - f$ on $[a, b]$.



Let $n \to \infty$. Then

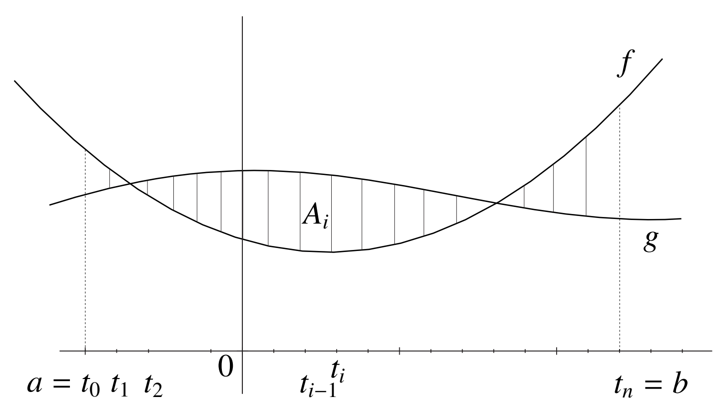
$$\begin{aligned} A &= \lim_{n\to\infty} \sum_{i=1}^{n} (g(t_i) - f(t_i))\Delta t_i \\ &= \lim_{n\to\infty} \sum_{i=1}^{n} (g(t_i) - f(t_i))\frac{b-a}{n} \\ &= \int_a^b (g(t) - f(t))\, dt \end{aligned}$$

The general case where f and g may cross at one or more locations on the interval $[a, b]$ is similar. We again construct a regular n-partition

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

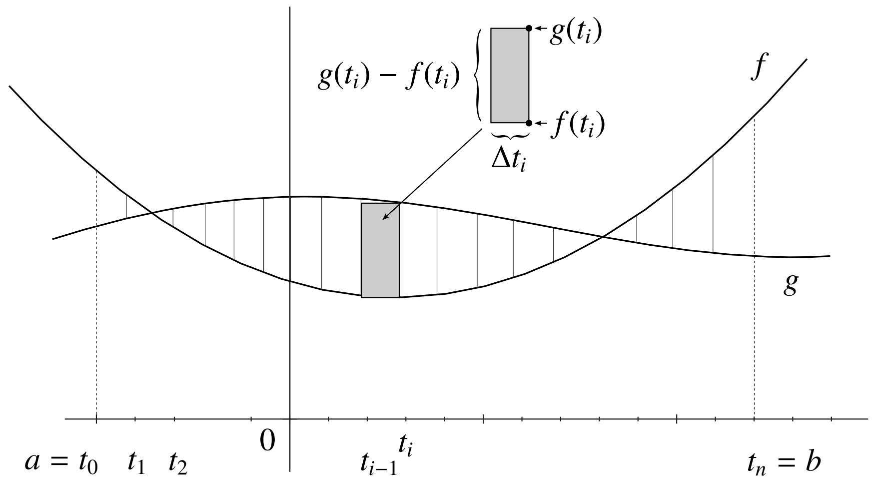
with $\Delta t_i = \frac{b-a}{n}$ and $t_i = a + \frac{i(b-a)}{n}$. This divides A into n subregions $A_1, A_2, \cdots, A_n$ where A_i is the region bounded by the graphs f and g, and the lines $t = t_{i-1}$ and $t = t_i$.

Moreover, we can again estimate the area A_i by constructing rectangle R_i. However, this time we must be concerned with whether $f(t_i) \leq g(t_i)$ or whether $g(t_i) \leq f(t_i)$.



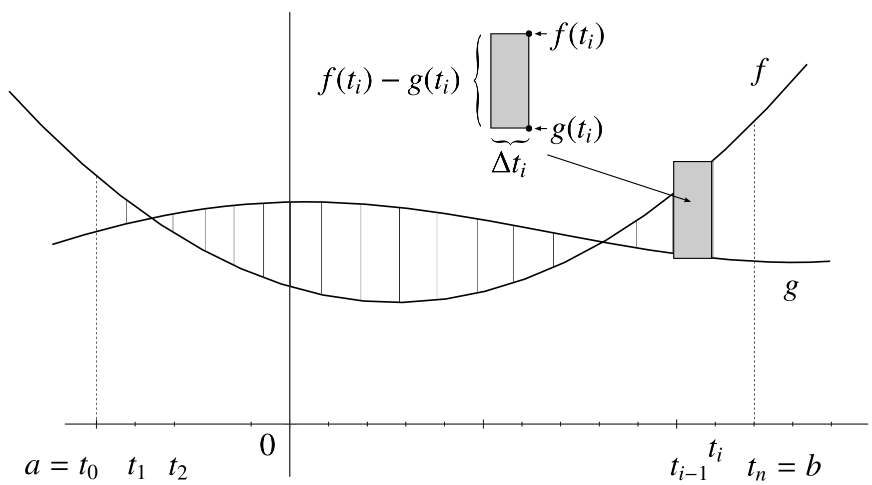
If $f(t_i) \leq g(t_i)$, then the height of the rectangle R_i is $h_i = g(t_i) - f(t_i)$ and its width is $\Delta t_i = \frac{b-a}{n}$. That is

$$A_i \cong (g(t_i) - f(t_i))\, \Delta t_i$$



However, if $g(t_i) \leq f(t_i)$, then the height of the rectangle is now $h_i = f(t_i) - g(t_i)$. The width remains as $\Delta t_i = \frac{b-a}{n}$, so

$$A_i \cong (f(t_i) - g(t_i))\,\Delta t_i$$

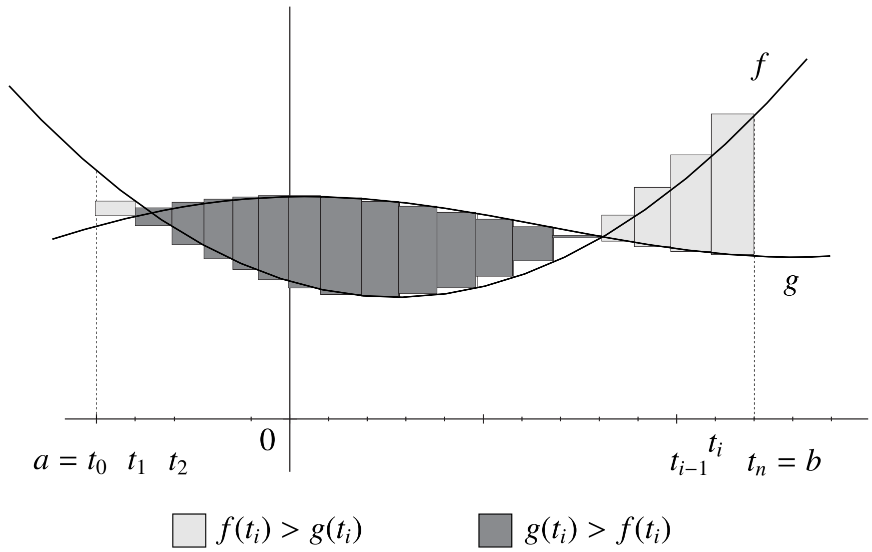


To summarize, we have that

$$A_i \cong h_i\,\Delta t_i$$

where

$$h_i = \begin{cases} g(t_i) - f(t_i) & \text{if} \quad g(t_i) - f(t_i) \geq 0 \\ f(t_i) - g(t_i) & \text{if} \quad g(t_i) - f(t_i) < 0 \end{cases}$$



Once more, the estimate for the area between f and g on $[a, b]$ is

$$\begin{aligned} A &= \sum_{i=1}^{n} A_i \\ &\cong \sum_{i=1}^{n} h_i \, \Delta t_i \end{aligned}$$

However, since

$$h_i = \begin{cases} g(t_i) - f(t_i) & \text{if} \quad g(t_i) - f(t_i) \geq 0 \\ f(t_i) - g(t_i) & \text{if} \quad g(t_i) - f(t_i) < 0 \end{cases}$$

then h_i is equivalent to

$$h_i = \mid g(t_i) - f(t_i) \mid$$

so

$$\begin{aligned} A &= \sum_{i=1}^{n} A_i \\ &\cong \sum_{i=1}^{n} \mid g(t_i) - f(t_i) \mid \Delta t_i \end{aligned}$$

and the latter sum is a right-hand Riemann sum for the function $\mid g - f \mid$ on $[a, b]$.

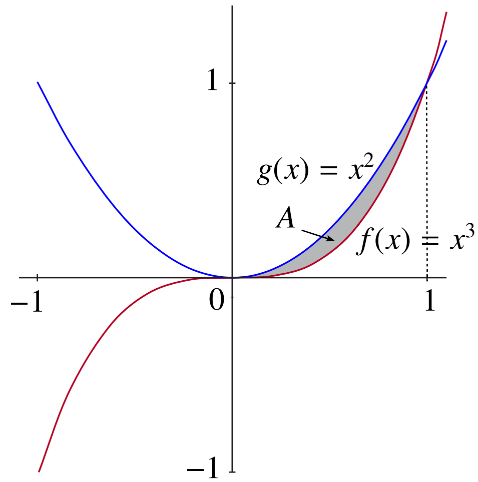
Finally, letting $n \to \infty$ gives us

$$\begin{aligned} A &= \lim_{n\to\infty} \sum_{i=1}^{n} \mid g(t_i) - f(t_i) \mid \Delta t_i \\ &= \int_a^b \mid g(t) - f(t) \mid dt \end{aligned}$$

Area Between Curves

Let f and g be continuous on $[a, b]$. Let A be the region bounded by the graphs of f and g, the line $t = a$ and the line $t = b$. Then the area of region A is given by

$$A = \int_a^b |\, g(t) - f(t)\, |\ dt.$$



EXAMPLE 1 Find the area A of the closed region bounded by the graphs of the functions $g(x) = x^2$ and $f(x) = x^3$. This area is the shaded region in the diagram.

The graphs cross when $x^3 = x^2$ or equivalently when

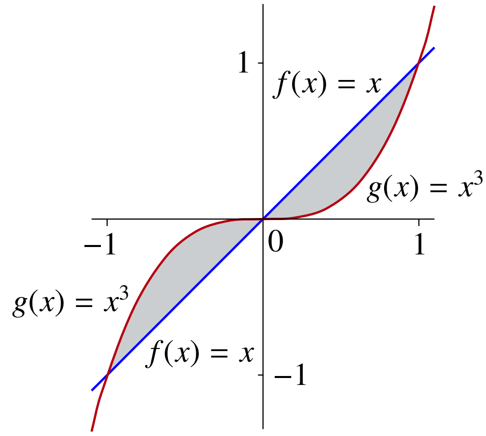
$$\begin{aligned} & 0 = x^3 - x^2 \\ \Rightarrow \quad & 0 = x^2(x-1) \end{aligned}$$

This occurs when $x = 0$ and $x = 1$. It follows that we are looking for the area bounded by the functions $g(x) = x^2$ and $f(x) = x^3$ between the lines $x = 0$ and $x = 1$. Moreover, on the interval $[0, 1]$ notice that $x^2 \geq x^3$. This means that the area is

$$\begin{aligned} A &= \int_0^1 \left(x^2 - x^3\right) dx \\ &= \left(\frac{x^3}{3} - \frac{x^4}{4}\right)\Bigg|_0^1 \\ &= \left(\frac{1}{3} - \frac{1}{4}\right) - (0 - 0) \\ &= \frac{1}{12} \end{aligned}$$

◀

EXAMPLE 2 Find the total area A of the closed regions bounded by the graphs of the functions $f(x) = x$ and $g(x) = x^3$. The shaded regions in the diagram represent A.



First we must locate the points where the graphs intersect. That is, where $x^3 = x$ or equivalently where

$$\begin{aligned} & 0 = x^3 - x \\ \Rightarrow \quad & 0 = x(x^2 - 1) \\ \Rightarrow \quad & 0 = x(x+1)(x-1) \end{aligned}$$

The solutions are $x = -1$, $x = 0$, and $x = 1$. This means that the left-hand bound is $x = -1$ and the right-hand bound is $x = 1$. Using this information, we know that the area is given by

$$\int_{-1}^{1} |\, x^3 - x \,|\ dx.$$

However, we **cannot** apply the Fundamental Theorem of Calculus directly to $|\, x^3 - x \,|$ to finish the calculation since f and g intersect on the interval $[-1, 1]$. Instead, we must consider the area in two parts, $A1$ and $A2$.

Area of A1

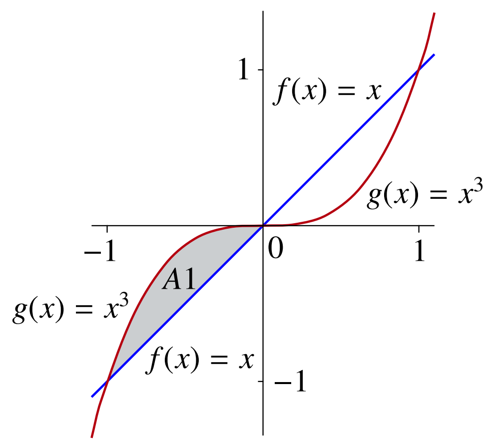
On the interval $[-1, 0]$ we have

$$x^3 \geq x$$

It follows that

$$\int_{-1}^{0} |\, x^3 - x \,|\ dx = \int_{-1}^{0} (x^3 - x)\, dx$$

This integral represents $A1$, the shaded area in the diagram.



Area of A2

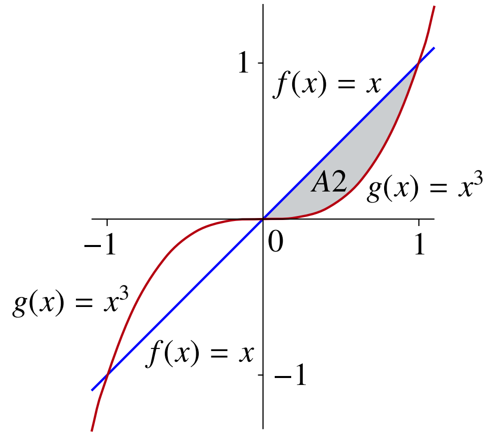
On the interval $[0, 1]$ we have

$$x \geq x^3.$$

It follows that

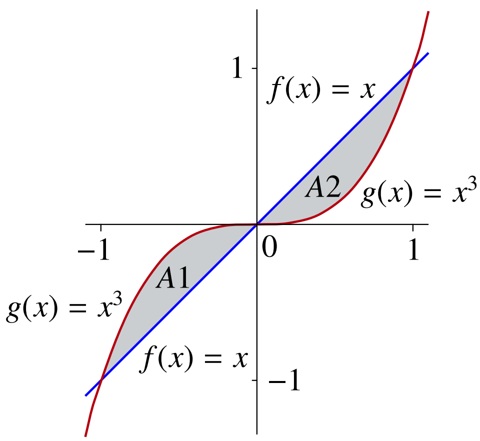
$$\int_0^1 |\, x^3 - x \,|\; dx = \int_0^1 (x - x^3)\, dx$$

This integral represents $A2$, the shaded area in the diagram.



Total Area Between the Curves

The total area A between the curves $f(x) = x$ and $g(x) = x^3$ on the interval $[-1, 1]$ is $A = A1 + A2$. Thus the total area is:



$$\begin{aligned}
A &= \int_{-1}^{1} |\, x^3 - x \,|\; dx \\
&= A1 + A2 \\
&= \int_{-1}^{0} |\, x^3 - x \,|\; dx + \int_0^1 |\, x^3 - x \,|\; dx \\
&= \int_{-1}^{0} (x^3 - x)\, dx + \int_0^1 (x - x^3)\, dx \\
&= \left(\frac{x^4}{4} - \frac{x^2}{2}\right)\Bigg|_{-1}^{0} + \left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Bigg|_{0}^{1} \\
&= \left((0-0) - \left(\frac{1}{4} - \frac{1}{2}\right)\right) + \left(\left(\frac{1}{2} - \frac{1}{4}\right) - (0-0)\right) \\
&= \frac{1}{4} + \frac{1}{4} \\
&= \frac{1}{2}
\end{aligned}$$

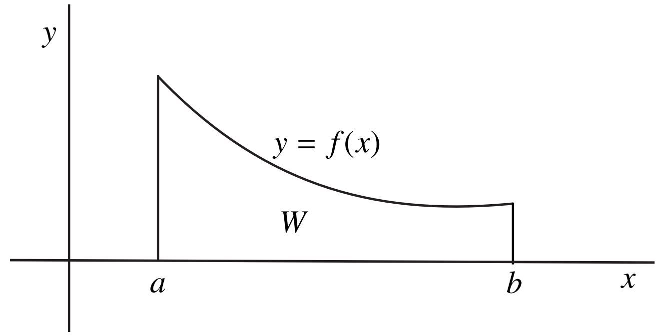
◀

3.2 Volumes of Revolution: Disk Method

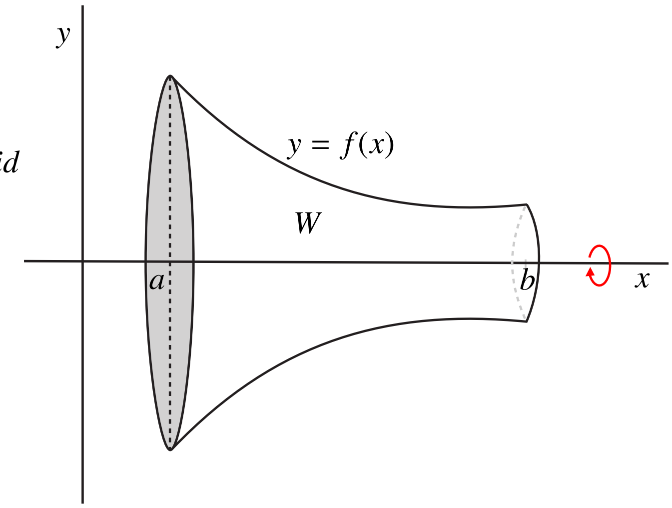
In this section we will use integration to calculate the volume of various types of solids obtained by rotating a region in the plane around a fixed line.

Problem 1:

Assume that f is continuous on $[a,b]$ and that $f(x) \geq 0$ on $[a,b]$. Let W be the region bounded by the graph of f, the lines $x=a$ and $x=b$ and the line $y=0$.



If region W is revolved around the x-axis an object called a *solid of revolution* is generated with the property that each vertical cross section of the solid is a circle with radius equal to the value of the function at the location of the slice.

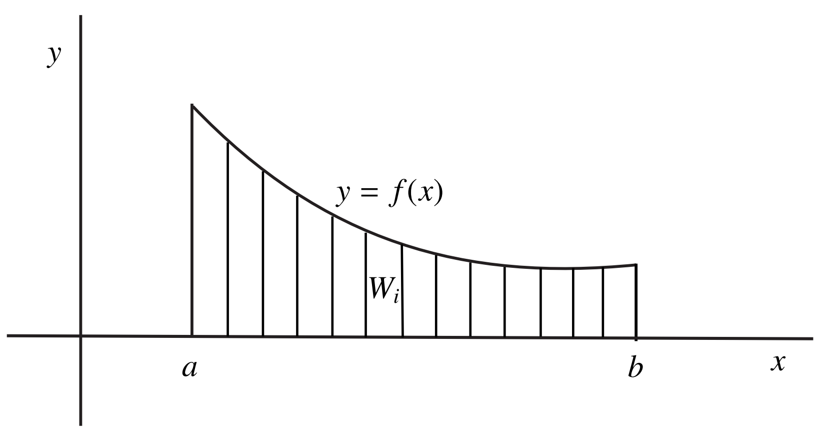


Our goal is to determine the volume V of this solid.

We will use integration to solve this problem and begin with a regular n-partition

$$a = t_0 < t_1 < t_2 < \cdots < t_{i-1} < t_i < \cdots < t_{n-1} < t_n = b$$

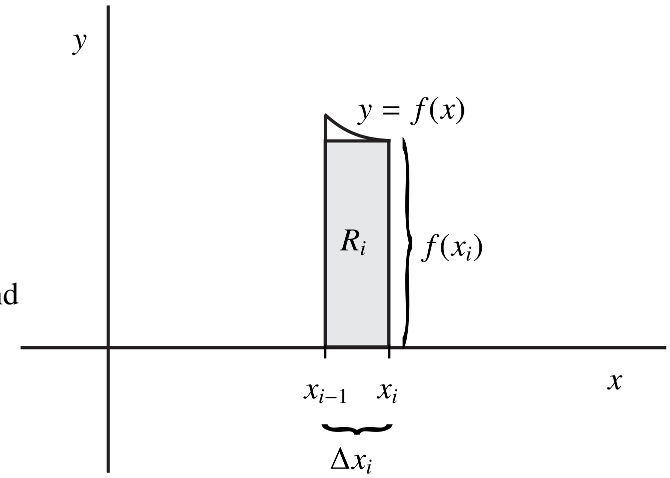
of $[a,b]$ with $\Delta t_i = \frac{b-a}{n}$ and $t_i = a + \frac{i(b-a)}{n}$. This partition subdivides the region W into n subregions. Let W_i denote the subregion of W in the interval $[x_{i-1}, x_i]$.



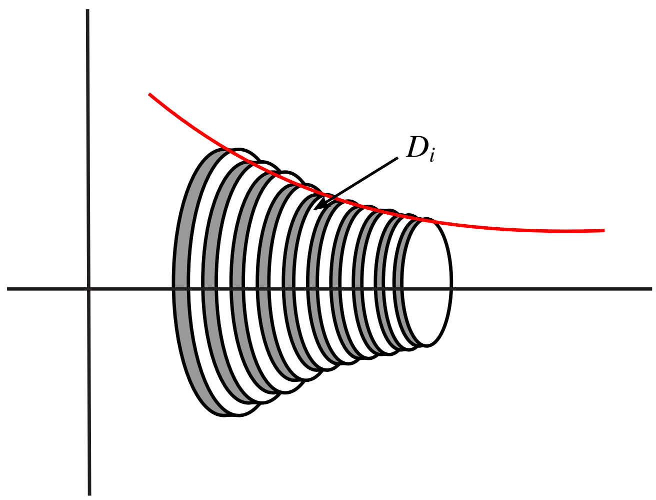
If we let V_i be the volume obtained by rotating W_i around the axis, then

$$V = \sum_{i=1}^{n} V_i.$$

We will use the same idea that was used to calculate areas to estimate the volume V_i. In particular, replace W_i by the rectangle R_i with height $f(x_i)$ and base on the interval $[x_{i-1}, x_i]$.



If $\triangle x_i$ is small, then V_i is approximately equal to the volume obtained by rotating R_i around the x-axis. Rotating each R_i generates a thin cylindrical disk D_i. Therefore, the solid is approximated by a series of thin disks.



For this reason, this method to find the volume of revolution is often called the *Disk Method*.

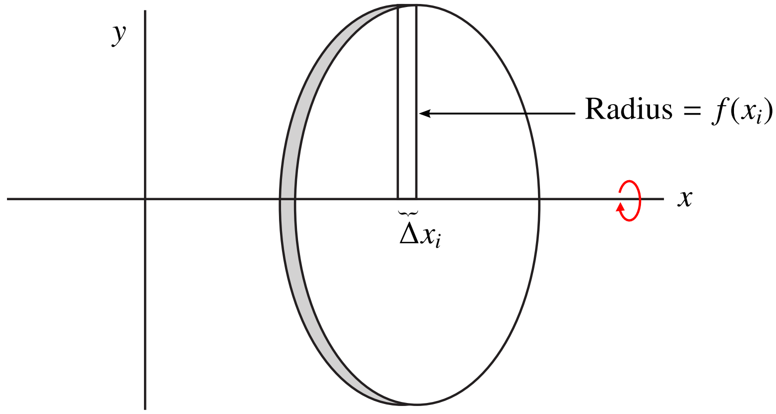
The next step is to determine the volume V_i^* of the disk D_i.

However, a close look at this disk shows that it has radius equal to the value of the function at x_i and its thickness is $\triangle x_i$. Therefore, since the volume of a cylindrical disk is

$$\pi \times (\text{radius})^2 \times (\text{thickness})$$

we get that

$$V_i^* = \pi f(x_i)^2 \triangle x_i.$$



Volume = $\pi f(x_i)^2 \Delta x_i$

Then the approximation for the total volume of the solid of revolution is:

$$\begin{aligned} V &= \sum_{i=1}^{n} V_i \\ &\cong \sum_{i=1}^{n} V_i^* \\ &= \sum_{i=1}^{n} \pi f(x_i)^2 \triangle x_i \end{aligned}$$

It follows that

$$V \cong \sum_{i=1}^{n} \pi f(x_i)^2 \triangle x_i$$

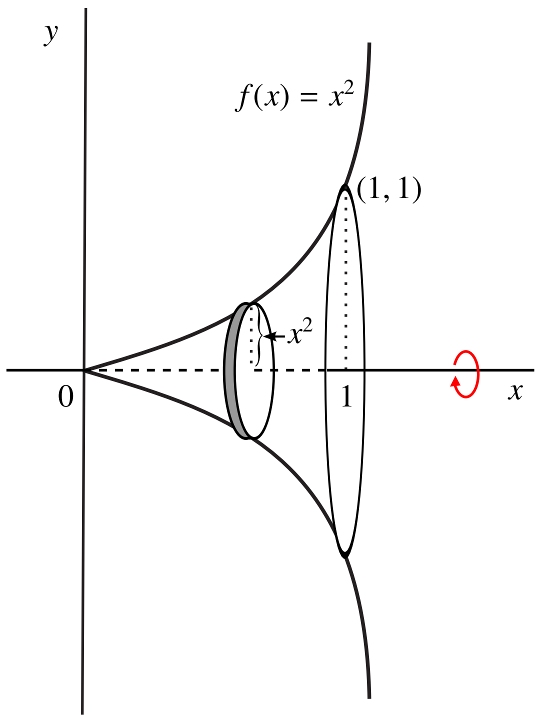
and this is a Riemann sum for the function $\pi f(x)^2$ over the interval $[a, b]$. Therefore, letting $n \to \infty$, we achieve the formula for the volume of revolution.

Volumes of Revolution: The Disk Method I

Let f be continuous on $[a, b]$ with $f(x) \geq 0$ for all $x \in [a, b]$. Let W be the region bounded by the graphs of f, the x-axis and the lines $x = a$ and $x = b$. Then the volume V of the solid of revolution obtained by rotating the region W around the x-axis is given by

$$V = \int_a^b \pi f(x)^2 dx.$$

EXAMPLE 3 Find the volume of the solid of revolution obtained by rotating the region bounded by the graph of the function $f(x) = x^2$, the x-axis, and the lines $x = 0$ and $x = 1$, around the x-axis.



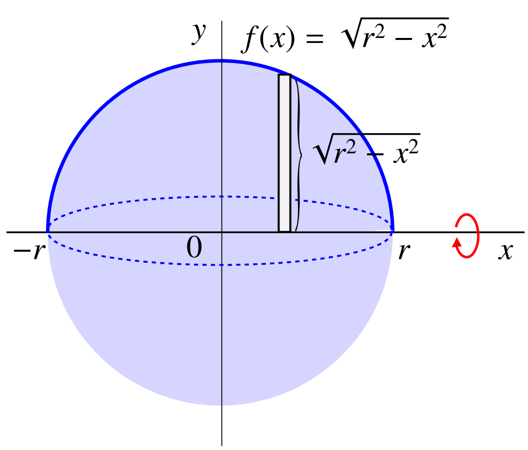
Using the formula, the volume is

$$\begin{aligned} V &= \int_0^1 \pi f(x)^2\,dx \\ &= \int_0^1 \pi (x^2)^2\,dx \\ &= \pi \int_0^1 x^4\,dx \\ &= \pi \frac{x^5}{5}\Big|_0^1 \\ &= \frac{\pi}{5} \end{aligned}$$

In the next example, we will use what we have learned about volumes of revolution to derive the formula for the volume of a sphere.

EXAMPLE 4 Find the volume of the sphere of radius r obtained by rotating the semi-circular region bounded by the graph of $f(x) = \sqrt{r^2 - x^2}$, the lines $x = -r$, $x = r$ and $y = 0$ around the x-axis.

Applying the disk method we get



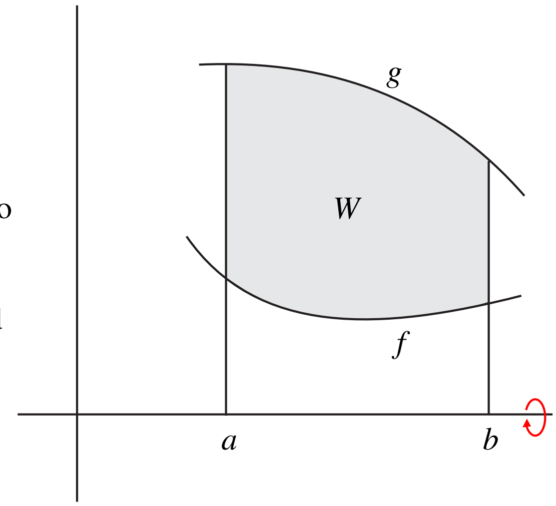
$$\begin{aligned} V &= \int_{-r}^{r} \pi f(x)^2 dx \\ &= \int_{-r}^{r} \pi \left(\sqrt{r^2 - x^2}\right)^2 dx \\ &= \int_{-r}^{r} \pi (r^2 - x^2)\,dx \\ &= \pi \left(r^2 x - \frac{x^3}{3}\right)\Big|_{-r}^{r} \\ &= \pi \left(\left(r^3 - \frac{r^3}{3}\right) - \left(-r^3 - \frac{(-r)^3}{3}\right)\right) \\ &= \frac{4}{3}\pi r^3 \end{aligned}$$

which is the general formula for the volume of a sphere.

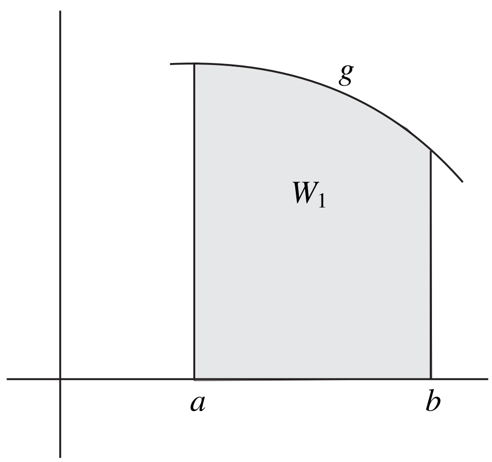
Until now we have looked at volume problems that involved a region that was bounded by a function f and the x-axis. Next we will look at a more general problem where the region that is revolved is bounded by two functions.

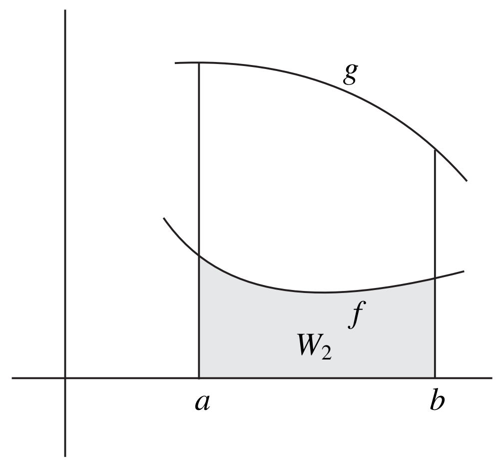
Problem 2:

Suppose that $0 \leq f(x) \leq g(x)$. We want to find the volume V of the solid formed by revolving the region W bounded by the graphs of f and g and the lines $x = a$ and $x = b$ around the x-axis.



Observe that if we let W_1 denote the region bounded by the graph of g, the x-axis, and the lines $x = a$ and $x = b$ and we let W_2 denote region bounded by the graph f, the x-axis, and the lines $x = a$ and $x = b$,





then W is the region that remains when we remove W_2 from W_1. It follows that the solid generated by revolving W around the x-axis is the same as the solid we would get by revolving W_1 around the x-axis and then removing the portion that would correspond to the solid obtained by revolving W_2 around the x-axis.

If we let V_1 be the volume of the solid obtained by rotating W_1 and V_2 be the volume of the solid obtained by rotating W_2, then we have

$$V = V_1 - V_2.$$

However, the formula for volumes of revolution tells us that

$$V_1 = \int_a^b \pi g(x)^2 \, dx$$

and

$$V_2 = \int_a^b \pi f(x)^2\, dx.$$

Therefore,

$$\begin{aligned} V &= V_1 - V_2 \\ &= \int_a^b \pi g(x)^2\, dx - \int_a^b \pi f(x)^2\, dx \\ &= \int_a^b \pi(g(x)^2 - f(x)^2)\, dx \end{aligned}$$

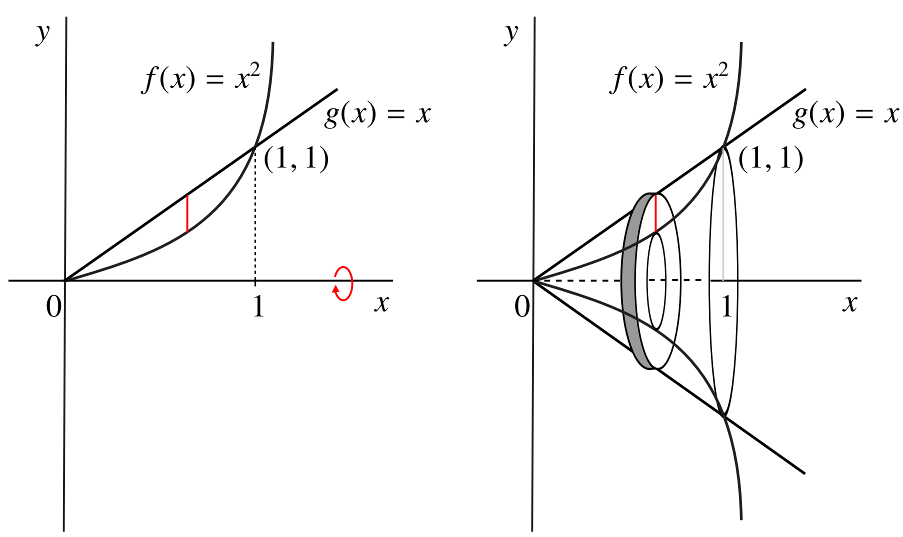
Volumes of Revolution: The Disk Method II

Let f and g be continuous on $[a, b]$ with $0 \le f(x) \le g(x)$ for all $x \in [a, b]$. Let W be the region bounded by the graphs of f and g, and the lines $x = a$ and $x = b$. Then the volume V of the solid of revolution obtained by rotating the region W around the x-axis is given by

$$V = \int_a^b \pi(g(x)^2 - f(x)^2)\, dx.$$

EXAMPLE 5 Find the volume V of the solid obtained by revolving the closed region bounded by the graphs of $g(x) = x$ and $f(x) = x^2$ around the x-axis.

Since we have not been given the interval over which we will integrate, we must find the x-coordinates of the points where the graphs intersect. But this means that we must solve $x = x^2$ so $x^2 - x = x(x - 1) = 0$. Hence, the points of intersection are located at $x = 0$ or $x = 1$. Moreover, on $[0, 1]$, we have $0 \le x^2 \le x$ (i.e., the graph of x^2 lies below the graph of x on this interval). Therefore, the region appears as follows:



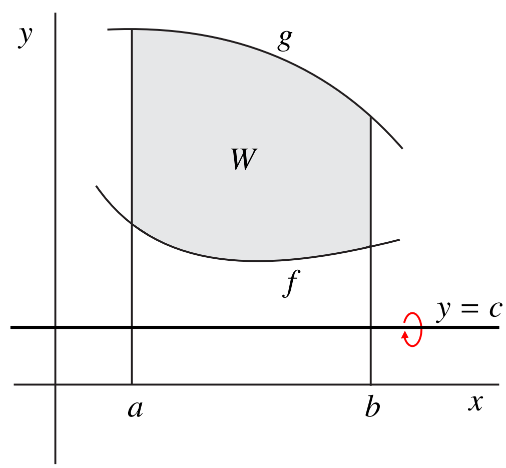
Then the volume is

$$\begin{aligned}
V &= \int_0^1 \pi\left(g(x)^2 - f(x)^2\right)dx \\
&= \int_0^1 \pi\left((x)^2 - (x^2)^2\right)dx \\
&= \int_0^1 \pi\left(x^2 - x^4\right)dx \\
&= \pi\left(\frac{x^3}{3} - \frac{x^5}{5}\right)\Bigg|_0^1 \\
&= \pi\left(\frac{1}{3} - \frac{1}{5}\right) \\
&= \frac{2\pi}{15}
\end{aligned}$$

Exercise:

Suppose that f and g are continuous on $[a,b]$ with $c \le f(x) \le g(x)$ for all $x \in [a,b]$. Let W be the region bounded by the graphs of f and g, and the lines $x = a$ and $x = b$. What is the volume V of the solid of revolution obtained by revolving the region W around the line $y = c$? The previous analysis still applies. Therefore, as an exercise, verify that the volume in this case is

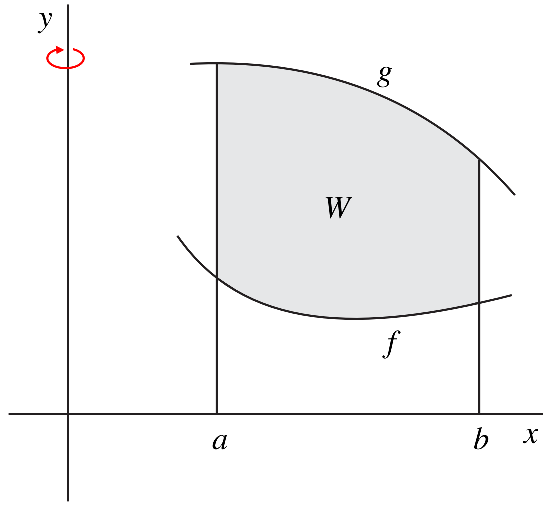
$$V = \int_a^b \pi((g(x) - c)^2 - (f(x) - c)^2)\,dx.$$



3.3 Volumes of Revolution: Shell Method

Sometimes using the Disk Method to find the volume of a solid of revolution can be onerous due to the algebra involved. Additionally, the Disk Method is sometimes difficult to use if the region is revolved around the y-axis instead of the x-axis. There is an alternate method called the *Shell Method* that may be easier to implement in such cases.

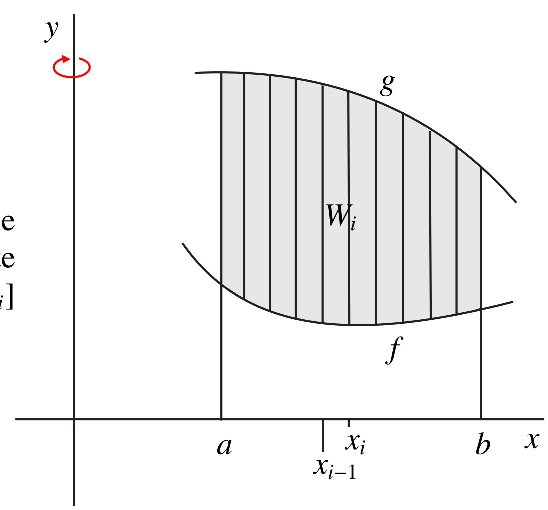
Problem: Assume that f and g are continuous on $[a,b]$, with $a \geq 0$ and $f(x) \leq g(x)$ on $[a,b]$. Let W be the region bounded by the graphs of f and g and the lines $x = a$ and $x = b$. Find the volume V of the solid obtained by rotating the region W around the y-axis.



We can proceed in a manner similar to the development of the Disk Method by constructing a regular n-partition of $[a,b]$ with

$$a = x_0 < x_1 < \cdots < x_{i-1} < x_i < \cdots < x_n = b.$$

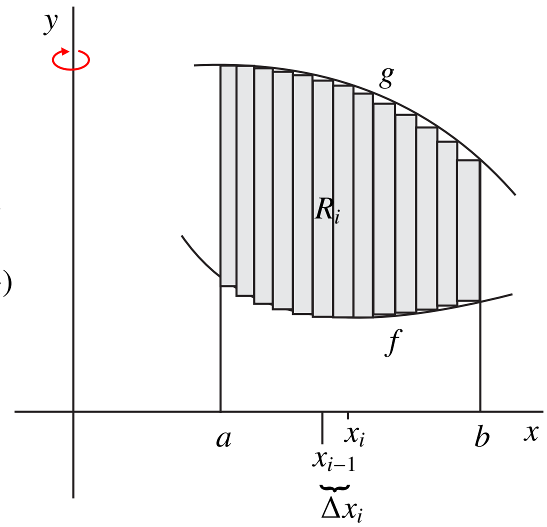
Once again, this partition subdivides the region W into n subregions. Let W_i denote the subregion of W on the interval $[x_{i-1}, x_i]$



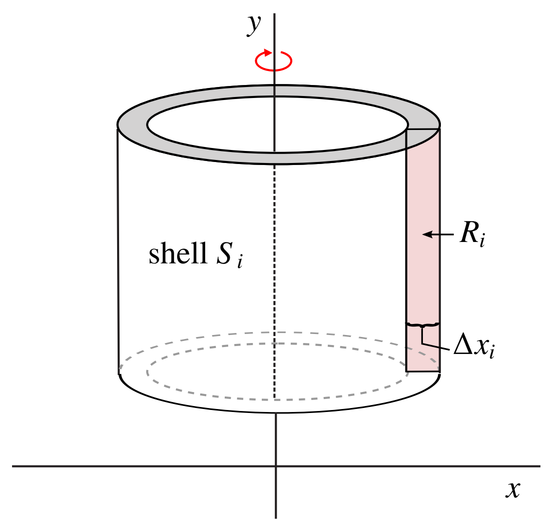
Let V_i be the volume obtained by rotating W_i around the y-axis so that

$$V = \sum_{i=1}^{n} V_i.$$

Next, approximate W_i by the rectangle R_i with height $g(x_i) - f(x_i)$, and base on the line $y = f(x_i)$ and top on the line $y = g(x_i)$ in the interval $[x_{i-1}, x_i]$.



It follows that if $\triangle x_i$ is small, then V_i is approximately equal to the volume obtained by rotating R_i around the y-axis. This time rotating R_i generates a thin cylindrical shell S_i.



For this reason, this method for finding volumes is called the *Shell Method* (or *Cylindrical Shell Method*).

The volume V_i^* of the shell generated by R_i is

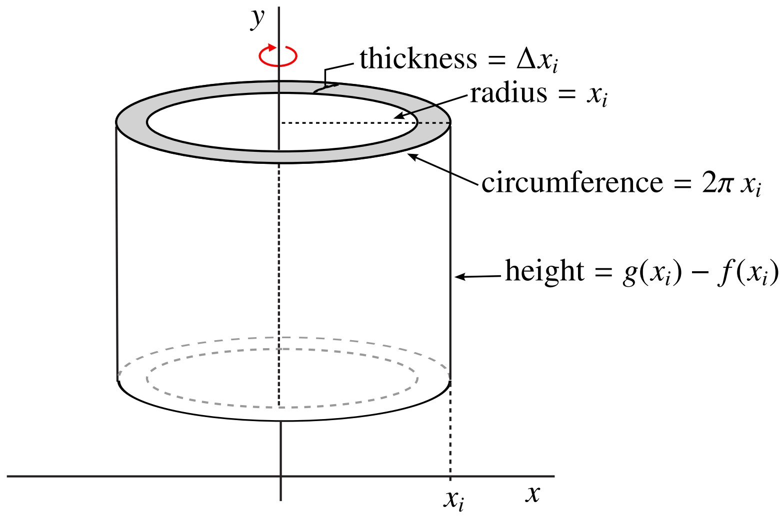
$$\text{(circumference)} \times \text{(height)} \times \text{(thickness)}$$

which is the same as

$$2\pi \times \text{(radius)} \times \text{(height)} \times \text{(thickness)}.$$

The height of the shell is $g(x_i) - f(x_i)$, its thickness is $\triangle x_i$, and the radius of revolution is x_i (the distance from the y-axis). Therefore, the volume V_i^* of S_i is

$$2\pi x_i(g(x_i) - f(x_i))\triangle x_i.$$



$$\text{Volume } = 2\pi\, x_i\,(g(x_i) - f(x_i))\,\Delta x_i$$

It follows that

$$\begin{aligned} V &= \sum_{i=1}^{n} V_i \\ &\cong \sum_{i=1}^{n} V_i^* \\ &= \sum_{i=1}^{n} 2\pi x_i(g(x_i) - f(x_i))\triangle x_i \end{aligned}$$

Letting $n \to \infty$, we get

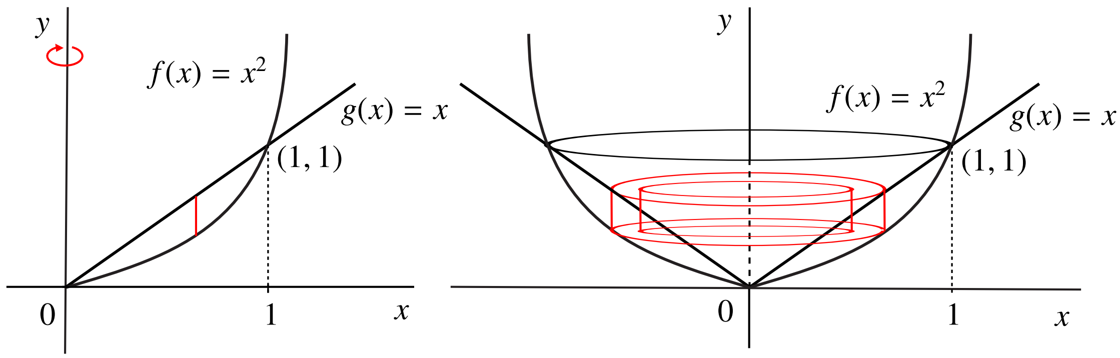
$$V = \int_a^b 2\pi x(g(x) - f(x))\,dx.$$

Volumes of Revolution: The Shell Method

Let $a \geq 0$. Let f and g be continuous on $[a, b]$ with $f(x) \leq g(x)$ for all $x \in [a, b]$. Let W be the region bounded by the graphs of f and g, and the lines $x = a$ and $x = b$. Then the volume V of the solid of revolution obtained by rotating the region W around the y-axis is given by

$$V = \int_a^b 2\pi x(g(x) - f(x))\,dx.$$

EXAMPLE 6 Find the volume of the solid obtained by revolving the closed region in the first quadrant bounded by the graphs of $g(x) = x$ and $f(x) = x^2$ around the y-axis.



As we have seen from a previous example, the graphs intersect in the first quadrant when $x = 0$ and $x = 1$ on the interval $[0, 1]$ with $f(x) \leq g(x)$. Thus

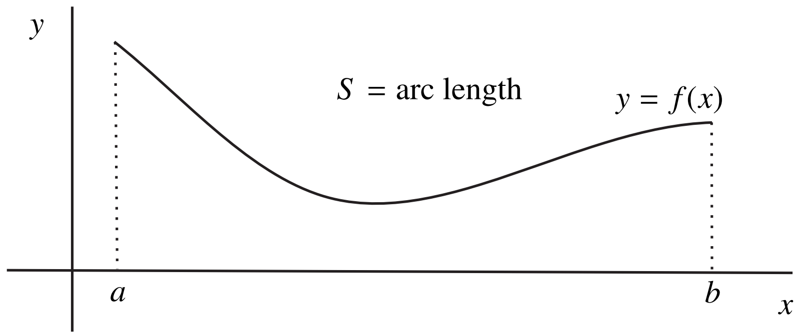
$$\begin{aligned} V &= \int_0^1 2\pi x(g(x) - f(x))dx \\ &= \int_0^1 2\pi x(x - x^2)dx \\ &= \int_0^1 2\pi(x^2 - x^3)dx = \frac{\pi}{6} \end{aligned}$$

Observe that previously we calculated the volume obtained by rotating this same region around the x-axis to equal $\frac{2\pi}{15}$ (which is less than $\frac{\pi}{6}$). This should not be surprising since the region is closer to the x-axis than to the y-axis and the further away from the axis of revolution, the larger the volume. ◀

3.4 Arc Length

The next application of integration that we will develop is a method for finding the length of the graph of a function over an interval $[a, b]$. The calculation of arc length has many important applications though most are beyond the scope of this course.

Problem: Let f be continuously differentiable on $[a, b]$. What is the arc length S of the graph of f on the interval $[a, b]$?

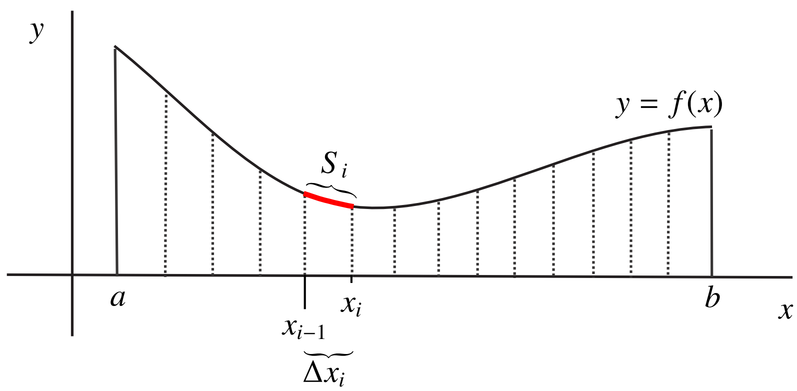


Let

$$a = x_0 < x_1 < \cdots < x_{i-1} < x_i < \cdots < x_n = b$$

be a regular n-partition of $[a, b]$.

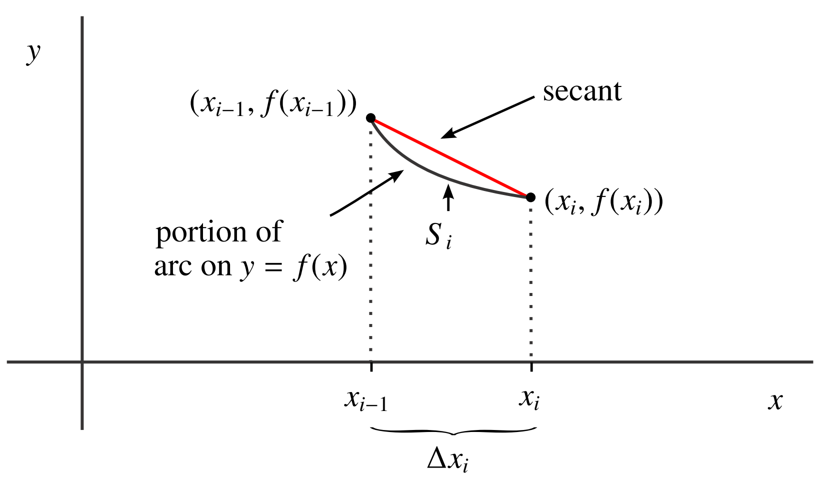
Let S_i denote the length of the arc joining $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$.



Then the length of the graph of f on the interval $[a, b]$ is

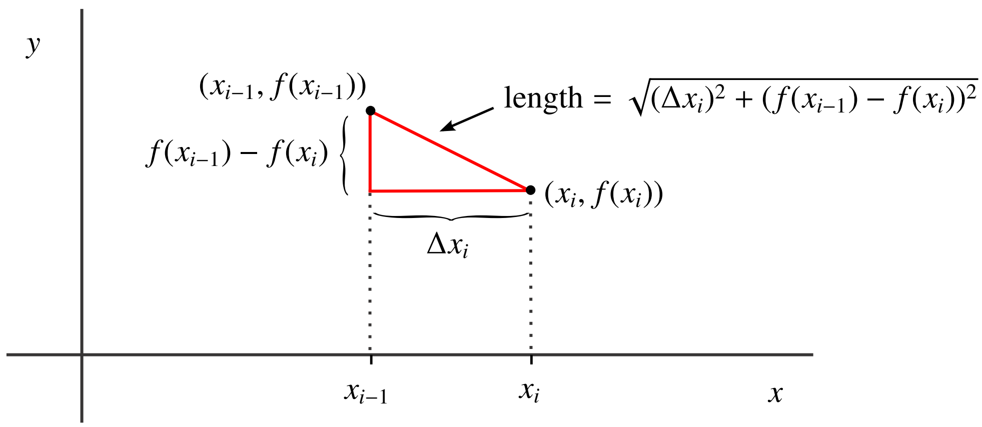
$$S = \sum_{i=1}^{n} S_i.$$

Observe that if $\triangle x_i$ is small, then S_i is approximately equal to the length of the secant line joining $(x_{i-1}, f(x_{i-1}))$ and $(x_i, f(x_i))$.



It follows that

$$\begin{aligned} S_i &\cong \sqrt{(\triangle x_i)^2 + (\triangle y_i)^2} \\ &= \sqrt{(\triangle x_i)^2 + (f(x_i) - f(x_{i-1}))^2} \end{aligned}$$



Next, applying the Mean Value Theorem guarantees a $c_i \in (x_{i-1}, x_i)$ such that

$$f(x_i) - f(x_{i-1}) = f'(c_i)\triangle x_i.$$

Therefore,

$$\begin{aligned} S_i &\cong \sqrt{(\triangle x_i)^2 + (f(x_i) - f(x_{i-1}))^2} \\ &= \sqrt{(\triangle x_i)^2 + (f'(c_i)\triangle x_i)^2} \\ &= \sqrt{(\triangle x_i)^2 + (f'(c_i))^2(\triangle x_i)^2} \\ &= \sqrt{(\triangle x_i)^2(1 + (f'(c_i))^2)} \\ &= \sqrt{1 + (f'(c_i))^2}\,\triangle x_i. \end{aligned}$$

Hence

$$\begin{aligned} S &= \sum_{i=1}^{n} S_i \\ &\cong \sum_{i=1}^{n} \sqrt{1 + (f'(c_i))^2}\,\triangle x_i \end{aligned}$$

This is a Riemann sum for the function $\sqrt{1 + (f'(x))^2}$ over the interval $[a, b]$. Therefore, letting $n \to \infty$, we get

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

Arc Length

Let f be continuously differentiable on $[a,b]$. Then the arc length S of the graph of f over the interval $[a,b]$ is given by

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

REMARK

The derivation of the arc length formula has many important applications that are beyond the scope of this course. Unfortunately, due to the square root in the integrand of the formula, there are very few functions for which we can calculate the arc length explicitly. Even calculating the arc length of the graph of $f(x) = x^3$ over the interval $[0,1]$ is beyond our current ability. However, there are a few examples that we can evaluate explicitly. ◀

EXAMPLE 7 Find the length S of the portion of the graph of the function $f(x) = \frac{2x^{\frac{3}{2}}}{3}$ between $x = 1$ and $x = 2$.

In this case, $f'(x) = x^{\frac{1}{2}}$. Hence

$$\begin{aligned}
S &= \int_1^2 \sqrt{1+(f'(x))^2}\,dx \\
&= \int_1^2 \sqrt{1+(x^{\frac{1}{2}})^2}\,dx \\
&= \int_1^2 \sqrt{1+x}\,dx \\
&= \left.\frac{2(1+x)^{\frac{3}{2}}}{3}\right|_1^2 \\
&= \frac{2(3)^{\frac{3}{2}}}{3} - \frac{2(2)^{\frac{3}{2}}}{3} \\
&= \frac{2}{3}(3^{\frac{3}{2}} - 2^{\frac{3}{2}}) \\
&\cong 1.578
\end{aligned}$$

◀

Chapter 4

Differential Equations

In this chapter we introduce and study *differential equations*. Differential equations are equations involving functions and their derivatives. These equations are used to model problems in the physical, biological, and social sciences.

4.1 Introduction to Differential Equations

Differential equations (DEs) often arise from studying real world problems. For example, if we let

$$P(t)$$

denote the population of a colony of bacteria at time t, then empirical evidence suggests that in an environment with unlimited resources the population will grow at a rate that is proportional to its size. This makes sense since the more bacteria that are present, the more “offspring” they will produce. Mathematically, this gives rise to the *differential equation*

$$P'(t) = kP(t)$$

where k is the constant of proportionality. If a function satisfying this equation can be determined, it would be helpful in predicting how the population will evolve.

The goal of this section is to introduce differential equations and to see how to find solutions for some basic examples.

DEFINITION

Differential Equation

A *differential equation* is an equation involving an independent variable such as x, a function $y = y(x)$ and various derivatives of y. In general, we will write

$$F(x, y, y', y'', \cdots, y^{(n)}) = 0.$$

A *solution* to the differential equation is a function φ such that

$$F(x, \varphi(x), \varphi'(x), \cdots, \varphi^{(n)}(x)) = 0.$$

The highest order of a derivative appearing in the equation is called the *order* of the differential equation.

EXAMPLE 1 Consider the equation

$$F(x, y, y'') = (\cos(x))y + y'' = 0.$$

This is an example of a differential equation of order 2. The constant function $\varphi(x) = 0$ is a solution to this equation since

$$\cos(x)\varphi + \varphi'' = \cos(x) \cdot 0 + 0 = 0.$$

However, at this point we have no tools to find any other solutions should they exist.

◀

NOTE

1) In this course, we will typically consider only first-order differential equations. Such DEs can be written in the form

$$y' = f(x, y).$$

A solution for a first-order differentiable equation is a function φ for which

$$\varphi'(x) = f(x, \varphi(x)).$$

2) The simplest first-order DE is the equation

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

Hence $y = y(x)$ is a solution if and only if y is an antiderivative of f. Therefore, the solutions to this equation are given by

$$\int f(x)\,dx = F(x) + C$$

where F is any antiderivative of f and $C \in \mathbb{R}$ is an arbitrary constant. This shows that differential equations do not need to have unique solutions. In particular, each different choice of C results in a new solution. The constant C is called a *parameter* and the collection of solutions $\{F(x) + C \mid C \in \mathbb{R}\}$ is called a *one parameter family.* ◀

EXAMPLE 2 Solve the differential equation $y' = \cos(x)$.

We have that

$$y = \int \cos(x)\,dx = \sin(x) + C$$

where C is an arbitrary constant. ◀

4.2 Separable Differential Equations

In this section we consider an important class of first-order differential equations and outline a technique for finding their solutions.

DEFINITION

Separable Differential Equation

A first-order differentiable equation is *separable* if there exists functions $f = f(x)$ and $g = g(y)$ such that the differentiable equation can be written in the form

$$y' = f(x)g(y).$$

EXAMPLE 3 Consider the following differentiable equations:

i) $y' = xy^2$ is separable. In this case, $f(x) = x$ and $g(y) = y^2$.

ii) $y' = y$ is separable. In this case, $f(x) = 1$ and $g(y) = y$.

iii) $y' = \cos(xy)$ is not separable since it can not be written in the form $y' = f(x)g(y)$. ◀

Solving Separable Differential Equations

There is a simple process to follow to find the solutions to a separable differential equation $y' = f(x)g(y)$. The steps are:

Step 1: Identify $f(x)$ and $g(y)$

Step 2: Find all constant (equilibrium) solutions

Step 3: Find the implicit solution

Step 4: Find the explicit solutions

We will illustrate this process with a series of examples.

Step 1: Identify $f(x)$ and $g(y)$

Often when you are presented with a differential equation, it will not be obvious whether the DE is separable. You may have to factor the differential equation in order to identify $f(x)$ and $g(y)$.

EXAMPLE 4 Let $y' = xy^2 + x$. Determine if this DE is separable.

In this case, the DE must be factored first in order to identify $f(x)$ and $g(y)$. Note that

$$\begin{aligned} y' &= xy^2 + x \\ &= x(y^2 + 1) \end{aligned}$$

so $f(x) = x$ and $g(y) = y^2 + 1$. Since this differential equation can be rewritten in the form $y' = f(x)g(y)$, it is a separable DE. ◀

Step 2: Find all constant (equilibrium) solutions

Let

$$y' = f(x)g(y)$$

be a separable DE.

Suppose that $g(y_0) = 0$ for some y_0. Then the constant function

$$y = \varphi(x) = y_0$$

is a solution to the separable differential equation since

$$\varphi'(x) = 0 = f(x)g(y_0) = f(x)g(\varphi(x))$$

for every x.

DEFINITION

Constant (Equilibrium) Solution to a Separable Differential Equation

If

$$y' = f(x)g(y)$$

is a separable differential equation and if $y_0 \in \mathbb{R}$ is such that $g(y_0) = 0$, then

$$\phi(x) = y_0$$

is called a *constant* or *equilibrium solution* to the differential equation.

EXAMPLE 5 Consider the separable differential equation

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

Then $f(x) = 1$ and $g(y) = y(1 - y)$. Moreover, $g(y) = 0$ if $y = 0$ or $y = 1$.

Let $\varphi(x) = 0$ for each x. Then, since φ is constant $\varphi'(x) = 0$,

$$0 = \varphi'(x) = \varphi(x)(1 - \varphi(x))$$

and hence $\varphi(x) = 0$ is a constant (equilibrium) solution.

If $\psi(x) = 1$ for all x, then $\psi'(x) = 0$ and $(1 - \psi(x)) = 0$ for each x. Therefore,

$$0 = \psi'(x) = \psi(x)(1 - \psi(x))$$

for each x. Hence $\psi(x) = 1$ is also a constant solution of the differential equation. These are the only two constant solutions to this separable differential equation.

◀

Step 3: Find the implicit solution

If $y' = f(x)g(y)$ is a separable differential equation, when $g(y) \neq 0$ we can divide by $g(y)$ to get

$$\frac{y'}{g(y)} = f(x).$$

Integrating both sides with respect to x gives

$$\int \frac{y'}{g(y)}\, dx = \int f(x)\, dx.$$

However, if we note that $y = y(x)$, we can apply the Change of Variables theorem to the left-hand integral to get

$$\begin{aligned} \int \frac{y'}{g(y)}\, dx &= \int \frac{1}{g(y(x))} y'(x)\, dx \\ &= \int \frac{1}{g(y)} dy \end{aligned}$$

This gives us the formula

$$\boxed{\int \frac{1}{g(y)} dy = \int f(x)\, dx}$$

Evaluating these integrals gives us an *implicit solution* to the differential equation of the form

$$G(y) = F(x) + C$$

where C is an arbitrary real constant.

This step will be successful only if we are able to evaluate $\int \frac{1}{g(y)} dy$ and $\int f(x)\, dx$.

Step 4: Find the explicit solutions

Try to solve the implicit equation

$$G(y) = F(x) + C$$

for y in terms of x. This will be the explicit solution to the differential equation. Unfortunately, it is not always easy to solve this equation for y in terms of x.

The next example illustrates all four of the steps required to solve a separable differential equation.

EXAMPLE 6 Solve the separable differential equation

$$y' = x(y^2 + 1).$$

Step 1: Identify $f(x)$ and $g(y)$

Since this separable differential equation is already in factored form, we can see that $f(x) = x$ and $g(y) = y^2 + 1$.

Step 2: Find all constant (equilibrium) solutions

We must find all values y_0 such that $g(y_0) = (y_0^2 + 1) = 0$. Since $y^2 + 1 > 0$ for all y, there are no such y_0's. Thus we conclude that there are no constant solutions.

Step 3: Find the implicit solution

Using the formula $\int \frac{1}{g(y)} dy = \int f(x)\, dx$, we get

$$\int \frac{1}{y^2 + 1}\, dy = \int x\, dx.$$

Since

$$\int \frac{1}{y^2 + 1}\, dy = \arctan(y) + C1$$

and

$$\int x\, dx = \frac{x^2}{2} + C2$$

evaluating the integrals gives the implicit solution

$$\arctan(y) = \frac{x^2}{2} + C \qquad (*)$$

Note: We only need to include the constant *once* since it is arbitrary.

Step 4: Find the explicit solutions

To find the explicit solutions, we must try to solve equation $(*)$ for y. Observe that

$$\tan(\arctan(y)) = y$$

so that we can apply the *tangent* function to both sides of the implicit solution $(*)$ to get

$$y = \tan(\arctan(y)) = \tan\left(\frac{x^2}{2} + C\right).$$

Therefore, the explicit solutions to the separable differential equation

$$y' = x(y^2 + 1)$$

are all functions of the form

$$y = \tan\left(\frac{x^2}{2} + C\right)$$

where C is an arbitrary constant. There were no constant solutions.

Check your work: You can verify that these are the correct solutions by differentiating.

If $y = \tan\left(\frac{x^2}{2} + C\right)$, then

$$\begin{aligned} y' &= \sec^2\left(\frac{x^2}{2} + C\right)\left(\frac{2x}{2}\right) \\ &= x\sec^2\left(\frac{x^2}{2} + C\right) \end{aligned}$$

The trigonometric identity $\sec^2(\theta) = 1 + \tan^2(\theta)$ give us that

$$\begin{aligned} y' &= x\sec^2\left(\frac{x^2}{2} + C\right) \\ &= x\left(1 + \tan^2\left(\frac{x^2}{2} + C\right)\right) \\ &= x(1 + y^2) \end{aligned}$$

which was the original separable differential equation, exactly as expected. ◀

EXAMPLE 7 Solve the differentiable equation

$$y' = xy + y.$$

Step 1: Identify $f(x)$ and $g(y)$ (if possible) to determine if the DE is separable

This DE factors as

$$y' = (x + 1)y$$

so $f(x) = x + 1$ and $g(y) = y$, so this DE is separable.

Step 2: Find all the constant (equilibrium) solutions

Since $g(y) = y$, the only y_0 such that $g(y_0) = 0$ is $y_0 = 0$. Therefore,

$$y = 0$$

is the only constant solution.

Step 3: Find the implicit solution

If $y \neq 0$, using the formula $\int \frac{1}{g(y)} dy = \int f(x)\, dx$, we get

$$\int \frac{1}{y}\, dy = \int (x+1)\, dx$$

so

$$\ln(|\, y\,|) = \frac{x^2}{2} + x + C \qquad (*)$$

where C is an arbitrary constant.

Step 4: Find the explicit solutions

To solve the implicit equation $(*)$ for y, we first exponentiate both sides to get

$$e^{\ln(|y|)} \;=\; e^{\frac{x^2}{2}+x+C}$$

so

$$\begin{aligned} |\, y\,| &= e^C e^{\frac{x^2}{2}+x} \\ &= C_1 e^{\frac{x^2}{2}+x} \end{aligned}$$

where $C_1 = e^C > 0$.

However,

$$|\, y\,| = C_1 e^{\frac{x^2}{2}+x}$$

means that

$$\begin{aligned} y &= \pm C_1 e^{\frac{x^2}{2}+x} \\ &= C_2 e^{\frac{x^2}{2}+x} \end{aligned}$$

where $C_2 = \pm C_1 \neq 0$.

Therefore, the solutions are

$$y = 0 \qquad \textit{(equilibrium solution)}$$

or

$$y = C_2 e^{\frac{x^2}{2}+x} \qquad \textit{(explicit solutions)}$$

where $C_2 \neq 0$.

Finally, since

$$y = 0 = 0e^{\frac{x^2}{2}+x}$$

we actually have that all of the solutions are of the form

$$y = C_3 e^{\frac{x^2}{2}+x}$$

where C_3 is an arbitrary constant. ◀

Strategy [Solving Separable Differential Equations]

Solving the separable differential equation

$$y' = f(x)g(y)$$

consists of 4 steps.

Step 1: Determine whether the DE is separable. You may have to factor the DE to identify $f(x)$ and $g(y)$.

Step 2: Determine the *constant solution(s)* by finding all the values y_0 such that $g(y_0) = 0$. For each such y_0, the constant function

$$y = y(x) = y_0$$

is a solution.

Step 3: If $g(y) \neq 0$, integrate both sides of the following equation

$$\int \frac{1}{g(y)}\,dy = \int f(x)\,dx$$

to solve the differential equation implicitly.

Step 4: Solve the implicit equation from Step 3 explicitly for y in terms of x.

Step 5: **[Optional]** Check your solution by differentiating y to deterimine if this derivative is equal to the original DE y'.

NOTE

Each of these steps could be difficult or even impossible to complete! For this reason, it is often necessary to find qualitative solutions or numerical solutions for the differential equation. We will discuss qualitative solutions later in the chapter. ◀

4.3 First-Order Linear Differential Equations

Linear differential equations form one of the most important classes of differential equations. There is a very well developed theory for dealing with these equations both algebraically and numerically. Furthermore, a common strategy for handling many differential equations that arise in real world problems is to use approximation techniques to replace the given equation by a linear one. In this section, we will develop an algorithm for solving first-order linear differential equations that provides a rather simple formula for determining all solutions to this class of equations.

DEFINITION **First-Order Linear Differentiable Equations [FOLDE]**

A first-order differential equation is said to be *linear* if it can be written in the form

$$y' = f(x)y + g(x).$$

EXAMPLE 8 Consider the following differential equations:

i) The separable differential equation

$$y' = 3x(y-1)$$

may be rewritten as

$$y' = 3xy - 3x$$

so it is also linear.

ii) The differentiable equation

$$y' = x^2y^3$$

is not linear since the term y^3 is of third degree.

◀

The next example will introduce a method for solving first-order linear differential equations.

EXAMPLE 9 Solve the first-order linear differential equation

$$y' = 3xy - 3x.$$

The first step is to rewrite the differential equation so that "$g(x)$" is alone on the right-hand side of the equation,

$$y' - 3xy = -3x.$$

The next step is to multiply both sides of the equation by a *nonzero* function $I = I(x)$ to get

$$Iy' - 3xIy = -3xI \qquad (1)$$

The goal is to find the nonzero function $I = I(x)$ such that if we differentiate $I(x)y(x)$ we will get the left-hand side of equation (1). That is,

$$\frac{d}{dx}(I(x)y(x)) = Iy' - 3xIy$$

Using the Product Rule we see that

$$\frac{d}{dx}(I(x)y(x)) = Iy' + I'y$$

so we require that

$$Iy' + I'y = Iy' - 3xIy \qquad (2)$$

A close look at equation (2) shows us that we need

$$I' = -3xI \qquad (3)$$

Equation (3) is a separable differential equation which we know how to solve.

Since the only constant solution is $I = I(x) = 0$ and we require a nonzero function, we proceed to Step 2 of the algorithm for solving separable equations.

Write

$$\int \frac{1}{I}\, dI = \int (-3x)\, dx.$$

This gives

$$\ln(|\, I\,|) = -\frac{3}{2}x^2 + C.$$

Exponentiating shows that

$$|\, I\,| = C_1 e^{-\frac{3}{2}x^2}$$

where $C_1 = e^C > 0$ and hence that

$$I = C_2 e^{-\frac{3}{2}x^2}$$

with $C_2 \neq 0$.

We only require one such function, so choose $C_2 = 1$. Then

$$I = I(x) = e^{-\frac{3}{2}x^2}.$$

With this choice of I we now have an equation of the form

$$\frac{d}{dx}(I(x)y(x)) = -3xI$$

where $I = I(x) = e^{\frac{-3}{2}x^2}$.

Integrating both sides of this equation gives us

$$\begin{aligned} I(x)y &= \int \left(\frac{d}{dx}(I(x)y(x))\right) dx \\ &= \int -3xI(x)\, dx \\ &= \int -3xe^{-\frac{3}{2}x^2}\, dx \end{aligned}$$

Let $u = \frac{-3}{2}x^2$ to get that $du = -3xdx$ so $dx = \frac{du}{-3x}$ which gives

$$\begin{aligned} \int (-3xI(x))\, dx &= \int -3xe^{\frac{-3}{2}x^2}\, dx \\ &= \int e^u\, du \\ &= e^u + C \\ &= e^{\frac{-3}{2}x^2} + C \end{aligned}$$

This means

$$I(x)y = e^{\frac{-3}{2}x^2} + C.$$

Solving for y gives us

$$\begin{aligned} y &= \frac{e^{\frac{-3}{2}x^2} + C}{I(x)} \\ &= \frac{e^{\frac{-3}{2}x^2} + C}{e^{\frac{-3}{2}x^2}} \\ &= 1 + Ce^{\frac{3}{2}x^2} \end{aligned}$$

where C is an arbitrary constant.

Finally, we can verify this answer by differentiating y to get

$$\begin{aligned} y' &= Ce^{\frac{3}{2}x^2}(3x) \\ &= (3x)(y-1) \\ &= 3xy - 3x \end{aligned}$$

which, as we expected, is the original DE that we were trying to solve. ◀

The function $I = I(x)$ in the previous example is called the *integrating factor*. The reasons for introducing such a function may look a little mysterious. However, in general, the use of $I(x)$ works for solving all first-order linear differential equations.

Strategy [Solving First-Order Linear Differential Equations]

Solving the first-order linear differential equation

$$y' = f(x)y + g(x)$$

consists of 3 steps.

Step 1: Determine whether the DE is linear. Write the equation in the form

$$y' - f(x)y = g(x)$$

and identify $f(x)$ and $g(x)$.

Step 2: Calculate the integrating factor $I(x)$ with $I(x) \neq 0$. Solve for I by using

$$I = e^{-\int f(x)\,dx}$$

Step 3: Since $I(x) \neq 0$, the solution is

$$y = \frac{\int g(x)I(x)\,dx}{I(x)}$$

Step 4: **[Optional]** Check your solution by differentiating y.

We can now state the following theorem which summarizes what we have learned about solving first-order linear differential equations.

THEOREM 1 **Solving First-Order Linear Differential Equations**

Let f and g be continuous and let

$$y' = f(x)y + g(x)$$

be a first-order linear differential equation. Then the solutions to this equation are of the form

$$y = \frac{\int g(x)I(x)\,dx}{I(x)}$$

where $I(x) = e^{-\int f(x)\,dx}$.

Note: In theory, the method we have just outlined provides us with a means of solving all first-order linear differential equations. However, in practice this only works provided that we can perform the required integrations.

EXAMPLE 10 Solve the first-order linear differential equation

$$y' = x - y.$$

First we rewrite the DE in the form $y' - f(x)y = g(x)$ to get

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

so this DE is linear and we have $f(x) = -1$ and $g(x) = x$.

The integrating factor is

$$I(x) = e^{-\int (-1)\,dx} = e^x.$$

The general solution can be found by using

$$\begin{aligned} y &= \frac{\int g(x)I(x)\,dx}{I(x)} \\ &= \frac{\int xe^x\,dx}{e^x} \end{aligned}$$

Integration by parts can be used to show

$$\int xe^x\,dx = xe^x - e^x + C.$$

It then follows that

$$\begin{aligned} y &= \frac{\int xe^x\,dx}{e^x} \\ &= \frac{xe^x - e^x + C}{e^x} \\ &= x - 1 + Ce^{-x} \end{aligned}$$

The general solution is $y = x - 1 + Ce^{-x}$.

Check your work. You can verify that this is the correct solution by differentiating y. Since

$$y = x - 1 + Ce^{-x}$$

then

$$y' = 1 - Ce^{-x}.$$

We get

$$\begin{aligned} y' &= 1 - Ce^{-x} \\ &= x - (x - 1 + Ce^{-x}) \\ &= x - y \end{aligned}$$

which is the original FOLDE. This verifies that $y = x - 1 + Ce^{-x}$ is the correct solution.

◀

4.4 Initial Value Problems

We have seen that a first-order differential equation

$$y' = f(x, y)$$

generally produces infinitely many solutions.

However, there are times when we are looking for a particular solution. This happens, for example, when a real world problem dictates that the solution must take particular values at a set of predetermined points. That is, we must satisfy constraints such as

$$\begin{aligned} y(x_0) &= y_0 \\ y(x_1) &= y_1 \\ y(x_2) &= y_2 \\ y(x_3) &= y_3 \\ &\vdots \end{aligned}$$

These constraints are called *initial values* or *initial conditions* and a differential equation specified with *initial values* is called an *initial value problem*.

The use of initial values are often important in real world problems. For example, we have seen that with unlimited resources, we can expect the population $P(t)$ of a bacteria colony to satisfy the differential equation

$$P' = kP$$

for some k. It is then easy to verify that the general solution to this equation is

$$P(t) = Ce^{kt}$$

where C is arbitrary and k is potentially unknown.

In this form, it is impossible to use the solutions to derive information about the population at a specific time. However, suppose we knew that the initial population at time $t = 0$ was P_0. This specifies the initial condition

$$P(0) = P_0.$$

Substituting $t = 0$ into the general solution gives us

$$P_0 = P(0) = Ce^{k(0)} = Ce^0 = C.$$

Therefore, we now know that

$$P(t) = P_0 e^{kt}.$$

If we also knew that the population at $t = 1$ was P_1 we would have

$$P_1 = P_0 e^{k(1)} = P_0 e^k.$$

This gives us that

$$e^k = \frac{P_1}{P_0}$$

and hence that

$$k = \ln\left(\frac{P_1}{P_0}\right).$$

The population function has now been completely determined. (In fact, knowing the population at any two distinct times completely determines the population function).

We have just seen how initial values can help us to determine which of the potentially infinitely many solutions to a differential equation arising from a real world problem gives the desired solution. For linear equations, this fact is illustrated by the next theorem.

THEOREM 2 **Existence and Uniqueness Theorem for First-Order Linear Differential Equations**

Assume that f and g are continuous functions on an interval I. Then for each $x_0 \in I$ and for all $y_0 \in \mathbb{R}$, the initial value problem

$$\begin{aligned} y' &= f(x)y + g(x) \\ y(x_0) &= y_0 \end{aligned}$$

has exactly *one* solution $y = \varphi(x)$ on the interval I.

EXAMPLE 11 Solve the initial value problem

$$y' = xy$$

with $y(0) = 1$.

Observe that this differential equation is linear since it takes the form $y' = f(x)y + g(x)$ where $f(x) = x$ and $g(x) = 0$, so the previous theorem tells us that there will be a unique solution. However, this differential equation is also separable since it can be written in the form $y' = f(x)g(y)$ with $f(x) = x$ and $g(y) = y$, so we can use the method developed for separable equations to find the solution.

The only constant solution is $y = y(x) = 0$ which does not satisfy the initial conditions. Hence we have

$$\int \frac{1}{y}\,dy = \int x\,dx.$$

This shows that

$$\ln(|\,y\,|) = \frac{x^2}{2} + C$$

so

$$y = C_1 e^{\frac{x^2}{2}}.$$

We also have that

$$1 = y(0) = C_1 e^{\frac{0^2}{2}} = C_1 e^0 = C_1.$$

Therefore $y = e^{\frac{x^2}{2}}$ is the unique solution to this initial value problem. ◀

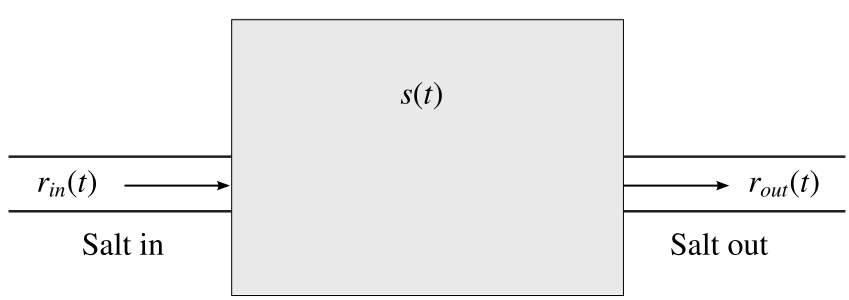
EXAMPLE 12 **A Mixing Problem**

Assume that a brine containing $30g$ of salt per litre of water is pumped into a $1000L$ tank at a rate of 1 litre per second. The tank initially contains $1000L$ of fresh water. It also contains a device that thoroughly mixes its contents. The resulting solution is simultaneously drained from the tank at a rate of 1 litre per second.

Problem: How much salt will be in the tank at any given time?

Let $s(t)$ denote the amount of salt in the tank at time t. Then $s'(t)$ is the difference between the rate at which salt is entering the tank (in the brine) and the rate at which salt is leaving the tank (in the discharge). Label these $r_{\text{in}}(t)$ and $r_{\text{out}}(t)$, respectively. That is,

$$s'(t) = r_{\text{in}}(t) - r_{\text{out}}(t).$$



To find $r_{\text{in}}(t)$ we note that the concentration of salt in the brine entering the tank is constant at $30g$ per litre. The flow rate is $1L$ per second and the rate at which the salt is entering the tank is the product of the concentration and the flow rate. Hence

$$r_{\text{in}}(t) = 30\frac{g}{L} \times 1\frac{L}{s} = 30\frac{g}{s}$$

and so the rate at which salt is entering the tank is 30 grams per second.

Calculating $r_{\text{out}}(t)$ is similar. It is the concentration of the discharge times the rate of flow. The rate of flow is again $1L$ per second but this time the concentration is not constant. In fact the concentration of the discharge is the same as that of the tank. Since the concentration of salt in the tank is $\frac{s(t)}{1000}$, we get

$$r_{\text{out}}(t) = \frac{s(t)}{1000} \times 1 = \frac{s(t)}{1000}$$

grams per second. It follows that

$$s'(t) = 30 - \frac{s(t)}{1000}.$$

This is a first-order linear differential equation with $f(t) = -\frac{1}{1000}$ and $g(t) = 30$. (Note: It is also a separable DE). To solve the equation as a FOLDE, the integrating factor $I(t)$ is

$$I(t) = e^{-\int \frac{-1}{1000}\,dt} = e^{\frac{t}{1000}}.$$

Using the FOLDE formula

$$y = \frac{\int g(t)I(t)dt}{I(t)}$$

gives us

$$\begin{aligned} s(t) &= \frac{\int 30e^{\frac{t}{1000}}\,dt}{e^{\frac{t}{1000}}} \\ &= \frac{30000e^{\frac{t}{1000}} + C}{e^{\frac{t}{1000}}} \\ &= 30000 + Ce^{-\frac{t}{1000}} \end{aligned}$$

Since $s(0) = 0$, we get

$$0 = 30000 + Ce^{0} = 30000 + C$$

and hence

$$C = -30000.$$

Therefore, at any given time

$$s(t) = 30000 - 30000e^{-\frac{t}{1000}}$$

grams.

Finally, since $\lim_{x\to\infty} e^{-x} \to 0$, observe that

$$\lim_{t\to\infty} s(t) = \lim_{t\to\infty} 30000 - 30000e^{-\frac{t}{1000}} = 30000$$

grams. This means that if the system was allowed to continue indefinitely, the amount of salt in the tank would approach 30000 grams. At that level, the concentration in the 1000L tank would be 30 grams per litre, which would be the same as the inflow rate. Therefore, the system is moving towards a stable equilibrium.

◀

Note: In general, initial value problems need not have any solutions or may not have unique solutions. For example, to see that the solutions need not be unique consider the initial value problem

$$y' = y^{\frac{1}{3}}$$

with $y(1) = 0$. Then the constant solution

$$y = y(x) = 0$$

satisfies the initial condition. However so do both

$$y = \pm[\frac{2}{3}(x-1)]^{\frac{3}{2}}.$$

◀

4.5 Graphical and Numerical Solutions to Differential Equations

It is often is the case that an explicit formula for the solution to a differential equation cannot be determined. When this occurs, we can still learn about the possible solutions to a differential equation through a graphical analysis (direction fields) or a numerical analysis (Euler's Method).

4.5.1 Direction Fields

Most differential equations cannot be solved by obtaining an explicit formula for the solution. However, we can construct local approximations to solutions by looking at short segments of their tangent lines at a number of points (x, y), with the slope of these tangent lines determined by the differential equation. This set of tangent line segments form a *direction field* and the direction field helps to visualize the *solution curve* that passes through any point that sits on a solution to the differential equation.

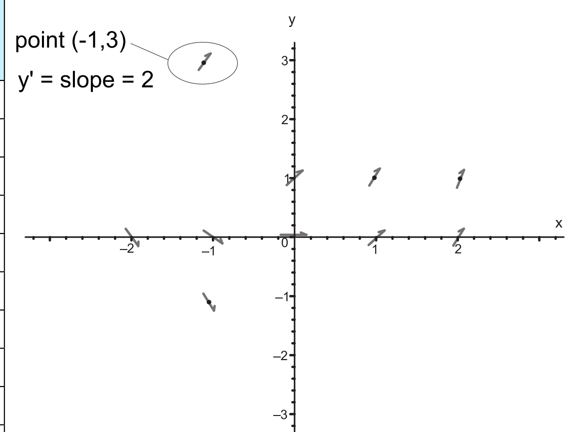
For example, consider the differential equation

$$y' = x + y.$$

This differential equation tells us that the derivative y' (or slope of the tangent line) of any solution whose graph contains the point (x, y) is the sum of the components of the points, $x + y$.

Let's consider a set of points (x, y) chosen at random. For each pair (x, y) the corresponding value of y' is calculated. This information is listed in the following table. The tangent line segments through (x, y) with the given slopes y' are then plotted.

x	y	tangent line slope from DE $y' = x + y$
-2	0	$y' = -2 + 0 = -2$
-1	0	$y' = -1 + 0 = -1$
0	0	$y' = 0 + 0 = 0$
1	0	$y' = 1 + 0 = 1$
2	0	$y' = 2 + 0 = 2$
0	1	$y' = 0 + 1 = 1$
1	1	$y' = 1 + 1 = 2$
2	1	$y' = 2 + 1 = 3$
-1	3	$y' = -1 + 3 = 2$
-1	-1	$y' = -1 + -1 = -2$



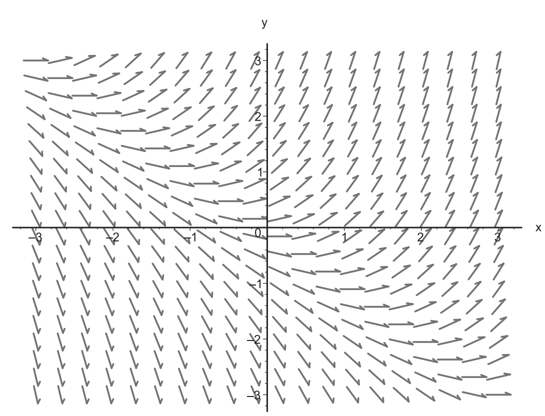
For example, at the origin $(0, 0)$, we have $y' = 0 + 0 = 0$, so the tangent line has slope 0 at the origin which gives us a *horizontal* line segment there. Similarly, at $(1, 1)$, the tangent line has slope 2 so a line segment *rising* to the right is drawn. At $(-1, -1)$, the tangent line has slope -2 so a line segment *falling* to the right is drawn.

The more tangent line segments that are drawn in the direction field, the easier it is to visualize the solution curves to this differential equation. However, this exercise can become tedious if done by hand. Instead, a mathematical software program is normally used to render the direction field.

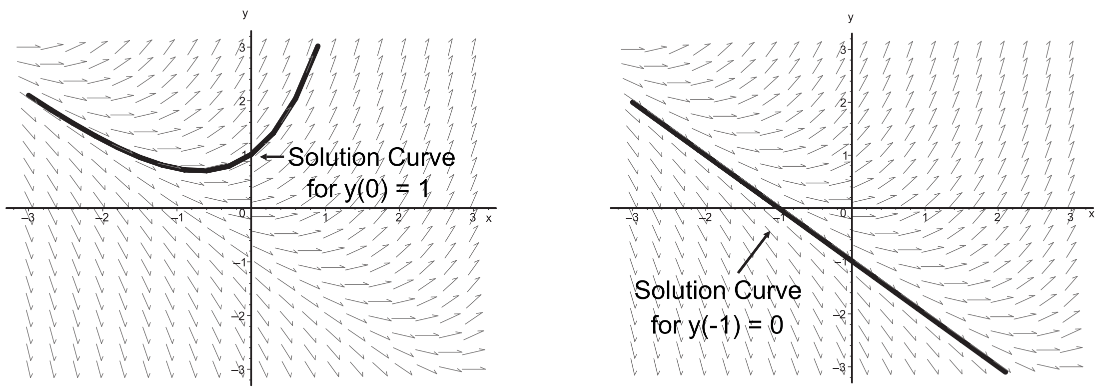
The completed direction field for

$$y' = x + y$$

is shown.



Once we can view the direction field, specific solutions can be sketched by drawing along the tangent line segments. For example, the following diagrams show the hand sketch of the solution to $y' = x + y$ for $y(0) = 1$ and for $y(-1) = 0$ suggested by the direction field. Notice that the solution curve for $y(-1) = 0$ is *linear*! By studying the shape of these solution curve sketches, we can better understand the nature of the solution set of the differential equation even though we may not know the explicit solutions.



Note: Since this linear differential equation $y' = x + y$ can be solved explicitly with the initial values $y(0) = 1$ or $y(-1) = 0$, it is a worthwhile exercise to compare the explicit solutions with the solution curves we obtained from the direction field. ◀

4.5.2 Euler's Method

In the previous section, we used linear approximation to construct direction fields to graphically approximate solutions to differential equations. We will explore this idea further and describe an algorithm known as *Euler's Method* for numerically building an approximate solution to a differential equation $y' = f(x, y)$ on a closed interval $[a, b]$.

Suppose the function $y = y(x)$ is known to be a solution of a differential equation

$$y' = f(x, y)$$

and that $y(x_0) = y_0$. This means that the point (x_0, y_0) is located on the graph of $y = y(x)$. Moreover,

$$y'(x_0) = f(x_0, y_0)$$

is the slope of the tangent line to the graph of $y = y(x)$ through the point (x_0, y_0). We know *near* x_0 that $y(x)$ can be approximated by its linear approximation:

$$\begin{aligned} L_{x_0}(x) &= y(x_0) + y'(x_0)(x - x_0) \\ &= y_0 + f(x_0, y_0)(x - x_0) \end{aligned}$$

In the following algorithm, the key idea will be the fact that if x is close to x_0, then

$$y(x) \cong L_{x_0}(x) = y(x_0) + y'(x_0)(x - x_0).$$

Euler's Method

The first step in this algorithm is to choose a partition

$$P = \{a = x_0 < x_1 < x_2 < \cdots < x_n = b\}$$

of the closed interval $[a, b]$.

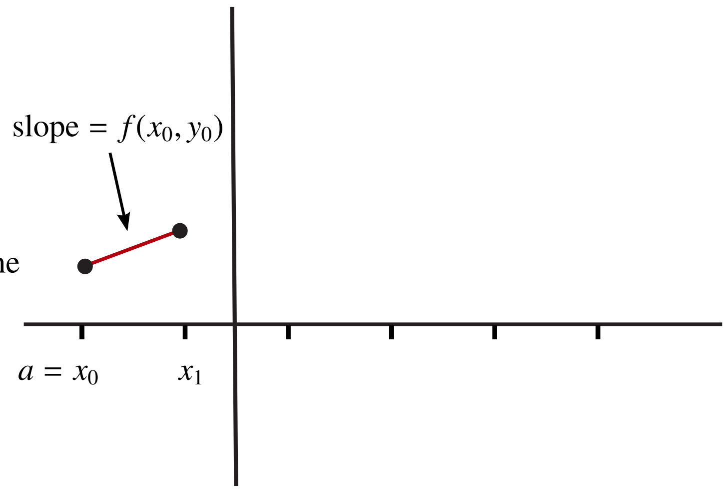
Let

$$\varphi(x_0) = \varphi(a) = y_0.$$

On the interval $[x_0, x_1]$, we define $\varphi(x)$ to be the function

$$L_{x_0}(x) = y_0 + f(x_0, y_0)(x - x_0).$$

Since this is the tangent line approximation to y' at the point (x_0, y_0), the graph of $L_{x_0}(x)$ is a line through the point (x_0, y_0) with slope equal to $f(x_0, y_0)$.



The next step is to calculate the value of this linear approximation at x_1 to determine the y-coordinate of the right-hand endpoint of the tangent line approximation. This is given by

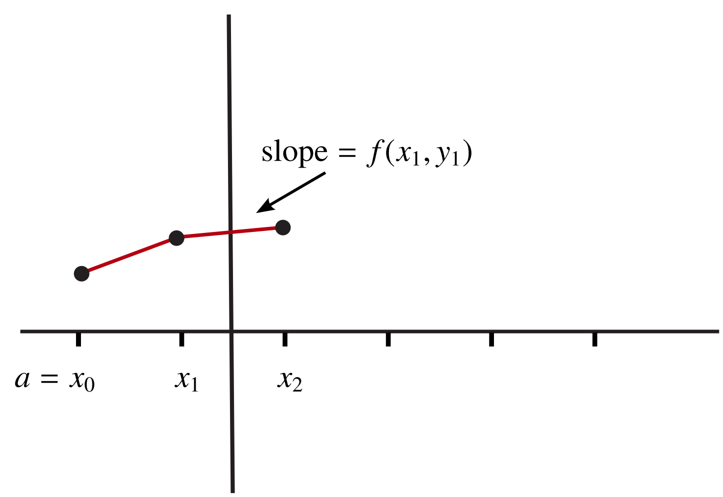
$$y_1 = L_{x_0}(x_1) = y_0 + f(x_0, y_0)(x_1 - x_0).$$

We can now find the linear approximation to the solution function through the new point (x_1, y_1). The approximate solution will then be defined as

$$L_{x_1}(x) = y_1 + f(x_1, y_1)(x - x_1)$$

on the interval $[x_1, x_2]$.

The graph is again a line through the point (x_1, y_1) with slope equal to $f(x_1, y_1)$.



We then find y_2, the y-coordinate of the right-hand endpoint of the new tangent line approximation. We have

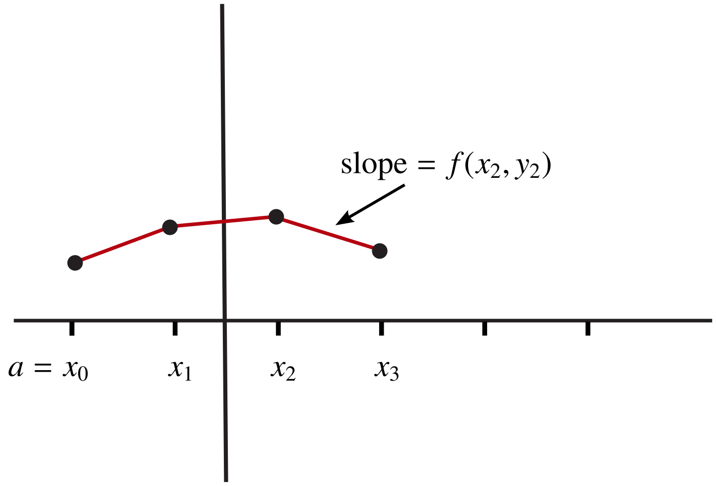
$$y_2 = L_{x_1}(x_2) = y_1 + f(x_1, y_1)(x_2 - x_1).$$

In a similar manner we find the linear approximation to the solution function through the new point (x_2, y_2). This new approximation is

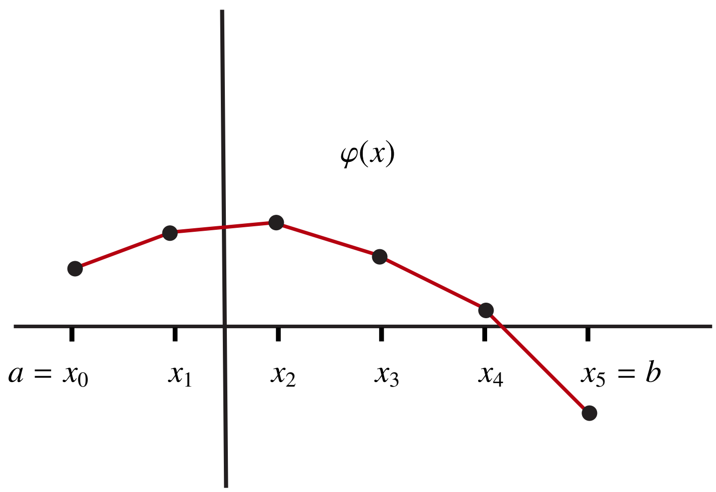
$$L_{x_2}(x) = y_2 + f(x_2, y_2)(x - x_2)$$

and its graph is the line through the point (x_2, y_2) with slope equal to $f(x_2, y_2)$.

Once again we let the approximate solution agree with this linear approximation on the interval $[x_2, x_3]$.



We proceed in this manner moving left to right until we have defined $\varphi(x)$ on the entire interval $[a, b]$.



In summary, Euler's method begins at the left-hand endpoint $x_0 = a$ of an interval and a short tangent line is created in the direction indicated by the direction field. After proceeding a short distance along this tangent line, stop at the next location in the partition $x = x_1$, adjust the slope of the tangent line, and proceed in this new direction. Repeat this process according to the direction field to find an approximate solution $\varphi(x)$. It is important to emphasize that Euler's method will not produce an exact solution to an initial value problem. However, the larger the number of terms in the partition and the closer together we choose successive points in the partition, the closer $\varphi(x)$ will be to a true solution to the differential equation.

4.6 Exponential Growth and Decay

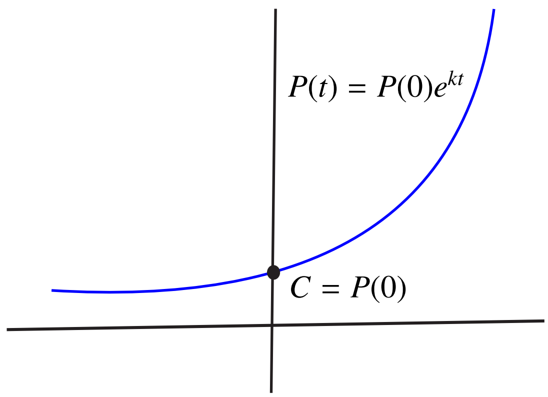
It is known that a population of bacteria in an environment with unlimited resources grows at a rate that is proportional to the size of the population. Therefore, if $P(t)$ represents the size of the population at time t, there is a constant k such that

$$P' = kP.$$

The general solution to this differential equation is given by

$$P(t) = Ce^{kt}$$

where $C = P(0)$ represents the initial population.



Exponential Growth

From the shape of the graph, it makes sense when we say that the bacteria population exhibits *exponential growth*.

Physical considerations generally limit the possible solutions to the equation. In the case of the bacteria population we will see that if we know the initial population as well as the size of the population at a one other fixed time, then the exact population function can be determined.

EXAMPLE 13 At time $t = 0$, a bacteria colony's population is estimated to be 7.5×10^5. One hour later, at $t = 1$, the population has doubled to 1.5×10^6. How long will it take until the population reaches 10^7?

Let $P(t)$ represent the size of the population at time t. We know that there is a constant k such that

$$P' = kP$$

so

$$P(t) = Ce^{kt}$$

and $C = P(0) = 7.5 \times 10^5$.

We also know that

$$1.5 \times 10^6 = P(1) = 7.5 \times 10^5 e^{k(1)}.$$

Therefore

$$e^k = \frac{1.5 \times 10^6}{7.5 \times 10^5} = 2.$$

To find k, take the natural logarithm of both sides of the equation to get

$$k = \ln(2).$$

This tells us that the population function is

$$P(t) = 7.5 \times 10^5 e^{(\ln(2))t}.$$

Now that we know the general formula for $P(t)$, to answer the original question we need to find t_0 such that

$$P(t_0) = 7.5 \times 10^5 e^{(\ln(2))t_0} = 10^7.$$

Therefore,

$$e^{(\ln(2))t_0} = \frac{10^7}{7.5 \times 10^5}$$

so

$$(\ln(2))t_0 = \ln\left(\frac{10^7}{7.5 \times 10^5}\right)$$

and

$$t_0 = \frac{\ln\left(\frac{10^7}{7.5\times 10^5}\right)}{\ln(2)} \cong 3.74\,\text{hours}.$$

◀

There are many other real world phenomena that behave in a manner similar to the growth of a bacteria population. In other cases, rather than exponential growth, we have exponential *decay*. For example, the rate at which radioactive material breaks down is proportional to the mass of material present.

Let $m(t)$ denote the mass of a certain radioactive material at time t. Then there is a constant k such that

$$\frac{dm}{dt} = m' = km.$$

We have

$$m(t) = Ce^{kt}$$

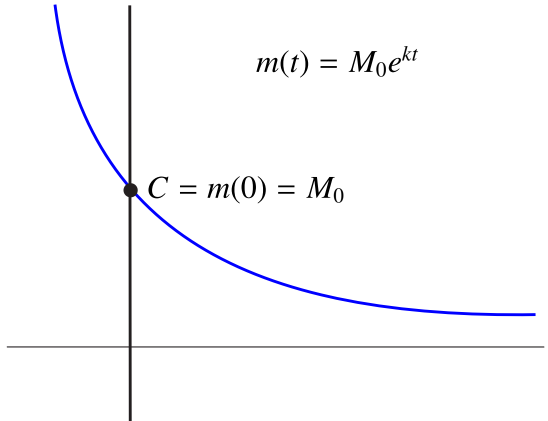
where $C = m(0) = M_0$ is the initial mass of the material. Therefore,

$$m(t) = M_0 e^{kt}.$$

Since the amount of material is decreasing, $m'(t) < 0$. But

$$m'(t) = km(t)$$

and $m(t) > 0$ so it follows that $k < 0$. Therefore, the graph of $m(t)$ appears as follows:



Exponential Decay

In particular, notice that

$$\lim_{t\to\infty} m(t) = \lim_{t\to\infty} M_0e^{kt} = 0$$

since $k < 0$.

We call such a process *exponential decay*.

All radioactive materials have associated with them a quantity t_h known as the *half-life* of the material. This is the amount of time it would take for one-half of the material to decay. The half-life is a fundamental characteristic of the material.

Mathematically, if

$$m(t) = M_0e^{kt}$$

then t_h is the time at which

$$m(t_h) = M_0e^{kt_h} = \frac{M_0}{2}.$$

Dividing by M_0 shows that

$$e^{kt_h} = \frac{1}{2}$$

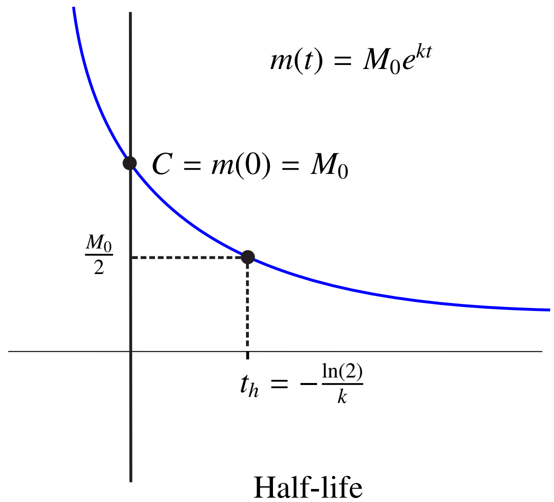
and hence that

$$kt_h = \ln\left(\frac{1}{2}\right) = -\ln(2).$$

Therefore, the half-life is given by the formula

$$t_h = \frac{-\ln(2)}{k}.$$

In particular, this shows that the half-life of a material is independent of the original mass.



Half-life

EXAMPLE 14 **Carbon Dating**

All living organisms contain a small amount of radioactive carbon-14. Moreover, each type of organism has a particular equilibrium ratio of carbon-14 compared to the stable isotope carbon-12.

When an organism dies the equilibrium is no longer maintained since the radioactive carbon-14 slowly breaks down into carbon-12. It is also known that carbon-14 breaks down at a rate of 1 part in 8000 per year. This means that after 1 year an initial quantity of 8000 particles will be reduced to 7999. Hence

$$7999 = m(1) = 8000e^{k(1)}$$

so that

$$k = \ln\left(\frac{7999}{8000}\right).$$

Problem 1: Find the half-life of carbon-14.

From the previous discussion, we know that

$$\begin{aligned} t_h &= \frac{-\ln(2)}{k} \\ &= \frac{-\ln(2)}{\ln\left(\frac{7999}{8000}\right)} \\ &\cong 5544.83 \text{ years} \end{aligned}$$

Problem 2: After a fossil was found research showed that the amount of carbon-14 was 23% of the amount that would have been present at the time of death. How old was the fossil?

Let M_0 be the expected amount of carbon-14 in the fossil and let t_o be the age of the fossil. Then the research shows that

$$(0.23)M_0 = m(t_0) = M_0 e^{kt_0}.$$

We must solve this equation for t_0. The first step is to recognize that

$$e^{kt_0} = \frac{(0.23)M_0}{M_0} = 0.23$$

This shows that we did not need to find the quantity M_0 explicitly to solve this question.

Taking the natural logarithm of both sides of the equation gives

$$kt_0 = \ln(0.23)$$

and hence that

$$\begin{aligned} t_0 &= \frac{\ln(0.23)}{k} \\ &= \frac{\ln(0.23)}{\ln\left(\frac{7999}{8000}\right)} \\ &= 11\,756 \text{ years} \end{aligned}$$

◀

4.7 Newton's Law of Cooling

Newton's law of cooling states that an object will cool (or warm) at a rate that is proportional to the difference between the temperature of the object and the ambient temperature T_a of its surroundings. Therefore, if $T(t)$ denotes the temperature of an object at time t, then there is a constant k such that

$$T' = k(T - T_a).$$

If $D = D(t) = T(t) - T_a$, then

$$D' = T' = kD$$

so D satisfies the equation for exponential growth (or decay). We know

$$D = Ce^{kt}.$$

It follows that

$$T(t) = Ce^{kt} + T_a$$

where $C = D(0) = T_0 - T_a$ and $T_0 = T(0)$.

Therefore,

$$T(t) = (T_0 - T_a)e^{kt} + T_a.$$

There are three possible cases.

1. $T_0 > T_a$.

   Physically, this means that the object is originally at a temperature that is greater than the ambient temperature. This means that the object will be *cooling*.

   Since $T(t)$ is decreasing

   $$T' = k(T - T_a) < 0.$$

   However, $T > T_a$, so that $k < 0$.

2. $T_0 < T_a$.

   In this case, the object is originally at a temperature that is lower than the ambient temperature. Therefore, the object will be *warming*.

   This time $T(t)$ is increasing so

   $$T' = k(T - T_a) > 0.$$

   Since $T < T_a$, it follows again that $k < 0$.

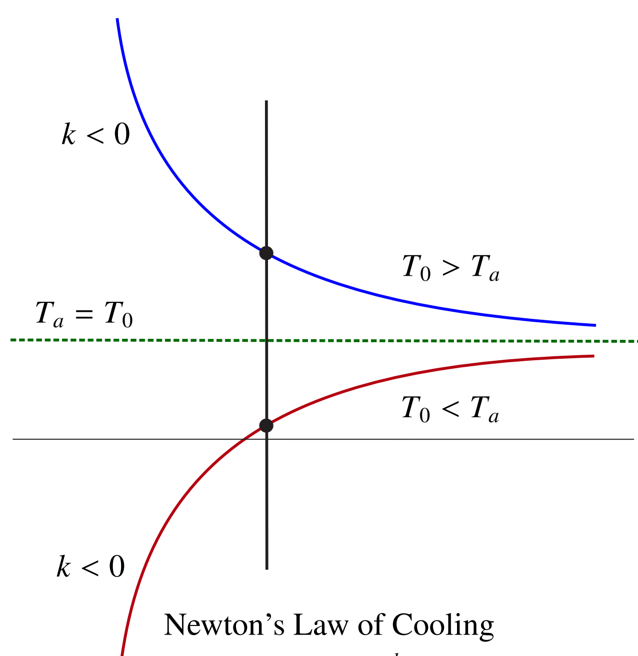
3. $T_0 = T_a$.

   Then

   $$T' = k(T - T_a) = 0$$

   so the temperature remains constant. We call this the *equilibrium state*.

The diagram summarizes the possible graphs of the temperature function.



Newton's Law of Cooling
$T(t) = (T_0 - T_a)e^{kt} + T_a$

Notice that in all three cases,

$$\lim_{t\to\infty} T(t) = T_a.$$

Regardless of the initial starting point, if a process always moves towards a particular equilibrium value, we call this value a *stable equilibrium*.

EXAMPLE 15 A cup of boiling water at $100°C$ is allowed to cool in a room where the ambient temperature is $20°C$. If after 10 minutes the water has cooled to $70°C$, what will be the temperature after the water has cooled for 25 minutes?

Let $T(t)$ denote the temperature of the water at time t minutes after cooling commences. The initial temperature is $T_0 = 100°C$ and the ambient temperature is $T_a = 20°C$. Newton's Law of Cooling shows that there is a constant $k < 0$ such that

$$\begin{aligned} T(t) &= (T_0 - T_a)e^{kt} + T_a \\ &= (100 - 20)e^{kt} + 20 \\ &= 80e^{kt} + 20 \end{aligned}$$

The next step is to determine k. Note that

$$70 = T(10) = 80e^{k(10)} + 20$$

so

$$50 = 80e^{10k}.$$

Hence,

$$10k = \ln\left(\frac{50}{80}\right)$$

and

$$\begin{aligned} k &= \frac{\ln\left(\frac{50}{80}\right)}{10} \\ &= -0.047 \end{aligned}$$

We can now evaluate $T(25)$ to get that the temperature after 25 minutes is

$$\begin{aligned} T(25) &= 80e^{-0.047(25)} + 20 \\ &= 44.71 \end{aligned}$$

degrees Celsius.

◀

4.8 Logistic Growth

We have seen that a population with unlimited resources grows at a rate that is proportional to its size. This leads to the differential equation

$$P' = kP.$$

However, the assumption that resources will be unlimited is usually unrealistic. More likely, there is a maximum population M that the surrounding environment can sustain. This means that as the population $P(t)$ approaches M, resources will become more scarce and the growth rate of the population will slow. On the other hand, when the population is small in comparison to the maximum population possible, the growth rate will be similar to that of the unrestricted case since there will be little resource pressure. It is known that such a population satisfies a differential equation of the form

$$P' = kP(M - P).$$

This equation means that the rate of growth is proportional to the product of the current population and the difference from the maximum sustainable population.

Populations of this type are said to satisfy *logistic growth* and the differential equation

$$y' = ky(M - y)$$

is called the *logistic equation*.

The logistic equation need not only model a population. However, in the special case where we are trying to describe the behavior of a population, we have the additional constraint that $P(t) > 0$.

Let $P_0 = P(0)$ be the initial population at the beginning of a study.

Observe that if the initial population is smaller than M, then the population will be growing. This means that we would have

$$0 < P' = kP(M - P)$$

since both P and $M - P$ are positive. As such, we would expect that $k > 0$.

However, if the initial population exceeds the maximum sustainable population, then the population would decrease so

$$0 > P' = kP(M - P)$$

and again we would have $k > 0$ since $P > 0$ and $M - P < 0$.

A third possible case occurs when the initial population is already at the maximum. In this case,

$$P' = kP(M - P) = 0$$

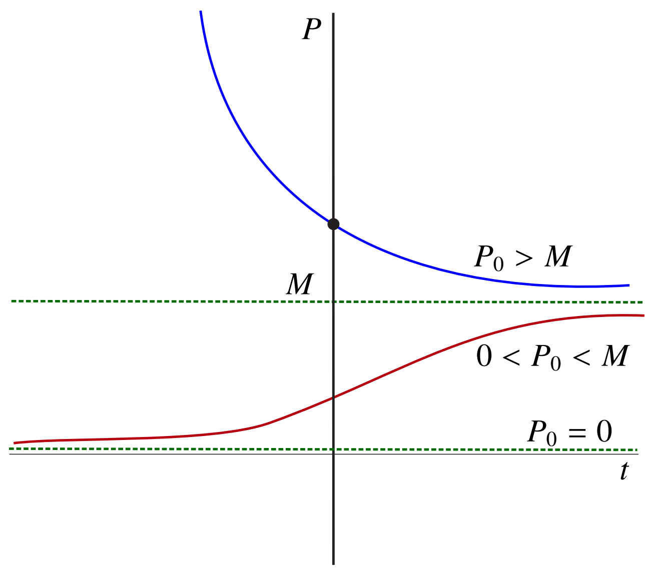
so the population would remain constant. This shows that $P(t) = M$ is an equilibrium solution.

The last case we will consider occurs when $P_0 = 0$. In this case, we have that

$$P' = kP(M - P) = 0$$

which makes sense since there are no parents to produce offspring. Therefore, $P(t) = 0$ is also an equilibrium, but its nature is quite different than that of the equilibrium at $P(t) = M$.

It follows that in all cases, we may assume that $P(t) > 0$ for all t so that the possible solutions look as follows:



Logistic Growth

You will notice that as long as $P_0 \neq 0$ we have

$$\lim_{t \to \infty} P(t) = M.$$

This means that $P(t) = M$ is a *stable* equilibrium. However, since we will never move towards an equilibrium of $P(t) = 0$ once there is a nonzero population, $P(t) = 0$ is called an ***un****stable equilibrium*.

So far, we have presented a qualitative solution to the logistic growth problem. However, since the equation is separable, we can try to solve it algebraically. We have already observed that $P(x) = 0$ and $P(x) = M$ are the constant solutions. We can then try to solve

$$\int \frac{1}{P(M - P)}\, dP = \int k\, dt = kt + C_1$$

To evaluate $\int \frac{1}{P(M-P)}\, dP$ we use partial fractions.

The constants A and B are such that

$$\frac{1}{P(M - P)} = \frac{A}{P} + \frac{B}{M - P}$$

or

$$1 = A(M - P) + B(P).$$

Letting $P = 0$ gives

$$1 = A(M)$$

so

$$A = \frac{1}{M}.$$

Letting $P = M$, we get

$$1 = B(M)$$

and again

$$B = \frac{1}{M}.$$

Therefore

$$\frac{1}{P(M-P)} = \frac{1}{M}\left[\frac{1}{P} + \frac{1}{M-P}\right].$$

It follows that

$$\begin{aligned}
\int \frac{1}{P(M-P)}\,dP &= \frac{1}{M}\left[\int \frac{1}{P}\,dP + \int \frac{1}{M-P}\,dP\right] \\
&= \frac{1}{M}[\ln(\mid P \mid) - \ln(\mid M-P \mid)] + C_2 \\
&= \frac{1}{M}\ln\left(\frac{\mid P \mid}{\mid M-P \mid}\right) + C_2
\end{aligned}$$

We now have that

$$\frac{1}{M}\ln\left(\frac{\mid P(t) \mid}{\mid M-P(t) \mid}\right) + C_2 = kt + C_1.$$

Therefore,

$$\ln\left(\frac{\mid P(t) \mid}{\mid M-P(t) \mid}\right) = Mkt + C_3$$

where C_3 is arbitrary.

This shows that

$$\frac{\mid P(t) \mid}{\mid M-P(t) \mid} = Ce^{Mkt}$$

where $C = e^{C_3} > 0$.

There are two cases to consider.

Case 1: Assume that $0 < P(t) < M$. Then

$$\frac{\mid P(t) \mid}{\mid M-P(t) \mid} = \frac{P(t)}{M-P(t)} = Ce^{Mkt}.$$

Solving for $P(t)$ would give

$$\begin{aligned} P(t) &= (M - P(t))Ce^{Mkt} \\ &= MCe^{Mkt} - P(t)Ce^{Mkt} \end{aligned}$$

so that

$$P(t) + P(t)Ce^{Mkt} = MCe^{Mkt}.$$

We then have

$$P(t)(1 + Ce^{Mkt}) = MCe^{Mkt}$$

and finally that

$$\begin{aligned} P(t) &= \frac{MCe^{Mkt}}{1 + Ce^{Mkt}} \\ &= M\frac{Ce^{Mkt}}{1 + Ce^{Mkt}} \end{aligned}$$

There are two important observations we can make about this solution.

(a) Since $C > 0$, the denominator is never 0 so the function $P(t)$ is continuous and

$$0 < \frac{Ce^{Mkt}}{1 + Ce^{Mkt}} < 1$$

so that

$$0 < P(t) < M$$

which agrees with our assumption.

(b) Since $k > 0$, we have that

$$\begin{aligned} \lim_{t\to\infty} P(t) &= \lim_{t\to\infty} M\frac{Ce^{Mkt}}{1 + Ce^{Mkt}} \\ &= M \lim_{t\to\infty} \frac{Ce^{Mkt}}{1 + Ce^{Mkt}} \\ &= M \end{aligned}$$

and

$$\lim_{t\to-\infty} P(t) = \lim_{t\to-\infty} M\frac{Ce^{Mkt}}{1 + Ce^{Mkt}} = 0.$$

This shows that the population would eventually approach the maximum population M but if you went back in time far enough, the population would be near 0. Both of these limits are consistent with our expectations.

If $t = 0$, then

$$P_0 = P(0) = M\frac{Ce^0}{1+Ce^0} = M\frac{C}{1+C}.$$

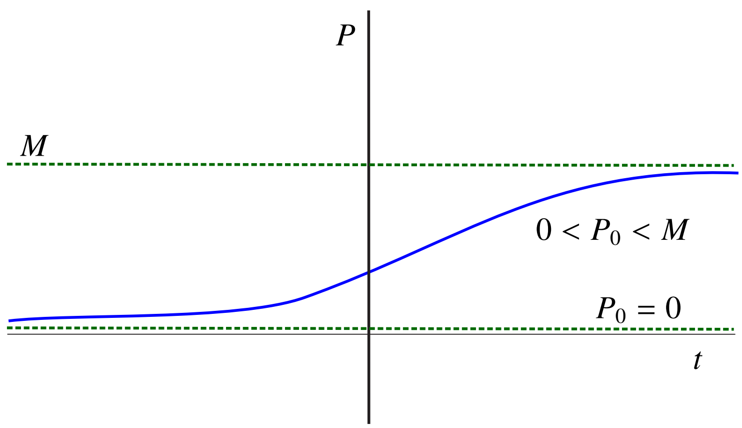
Solving for C yields

$$P_0(1+C) = MC$$
$$P_0 + P_0C = MC$$
$$P_0 = (M - P_0)C$$

and finally that

$$C = \frac{P_0}{M - P_0}.$$

The graph of the function $P(t) = M\frac{Ce^{Mkt}}{1+Ce^{Mkt}}$ looks as follows:



Logistic Growth

Case 2: If $P(0) > M$, then

$$\frac{|\,P(t)\,|}{|\,M - P(t)\,|} = -\frac{P(t)}{M - P(t)} = \frac{P}{P - M} = Ce^{Mkt}.$$

Proceeding in a manner similar to the previous case, we get that there exists a positive constant C such that

$$P(t) = M\frac{Ce^{Mkt}}{Ce^{Mkt} - 1}.$$

Notice that this function has a vertical asymptote when the denominator

$$Ce^{Mkt} - 1 = 0.$$

Moreover, the function is only positive if

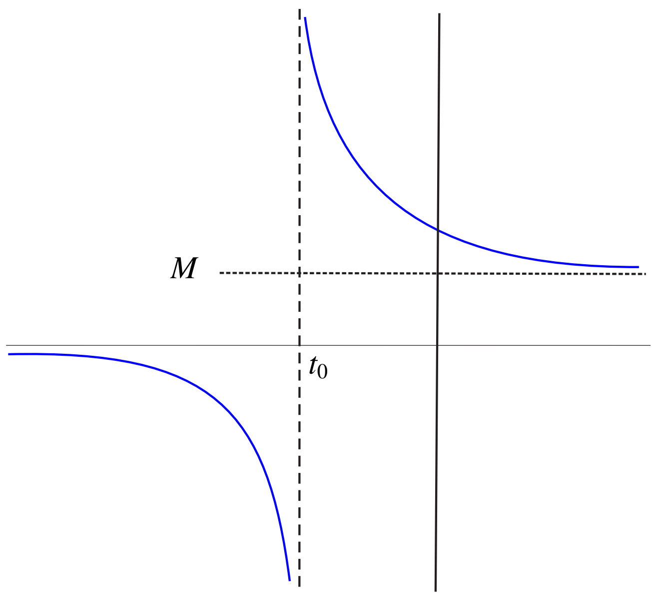
$$Ce^{Mkt} > 1$$

or equivalently if

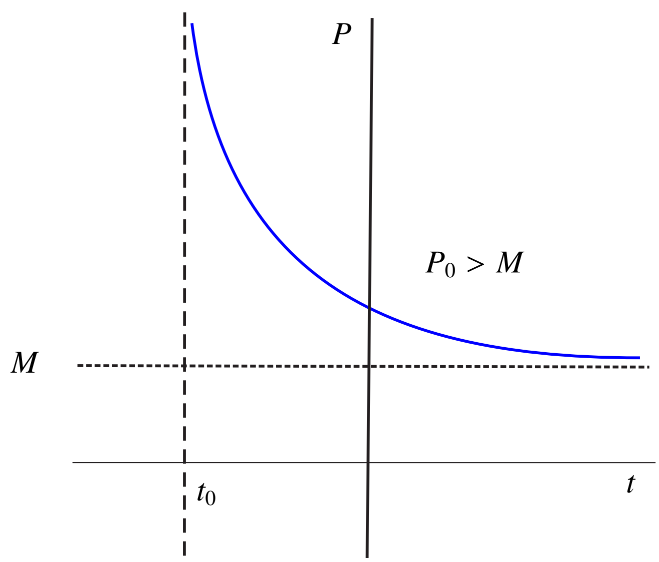
$$e^{Mkt} > \frac{1}{C}.$$

The use of some algebra shows that this happens if and only if $t > \dfrac{\ln\left(\frac{1}{C}\right)}{Mk} = t_0$.

If we ignore the fact that population must be positive, the graph of the solution function $P(t) = M\dfrac{Ce^{Mkt}}{Ce^{Mkt} - 1}$ appears as follows:



Since we are looking for a population function and so we require $P(t) \geq 0$, we will only consider values of t which exceed t_0. Therefore, the graph of the population function is:



It is still true that

$$\lim_{t\to\infty} P(t) = \lim_{t\to\infty} M\frac{Ce^{Mkt}}{Ce^{Mkt}-1} = M.$$

and

$$\lim_{t\to t_0^+} P(t) = \infty.$$

EXAMPLE 16 A game reserve can support at most 800 elephants. An initial population of 50 elephants is introduced in the park. After 5 years the population has grown to 120 elephants. Assuming that the population satisfies a logistic growth model, how large will the population be 25 years after this introduction?

Let $P(t)$ denote the elephant population t years after they are introduced to the park. We know that there are positive constants C and k such that the population of elephants is given by

$$P(t) = 800\frac{Ce^{800kt}}{1+Ce^{800kt}}.$$

Recall that if $P_0 = P(0)$, then

$$C = \frac{P_0}{M-P_0}.$$

We are given that $P_0 = P(0) = 50$ and $M = 800$. Then

$$C = \frac{50}{800-50} = \frac{50}{750} = \frac{1}{15}.$$

Therefore,

$$P(t) = 800\frac{\frac{1}{15}e^{800kt}}{1+\frac{1}{15}e^{800kt}}.$$

To find k, we use the fact that

$$120 = P(5) = 800\frac{\frac{1}{15}e^{800k(5)}}{1+\frac{1}{15}e^{800k(5)}}.$$

Hence

$$\frac{120}{800} = \frac{3}{20} = \frac{\frac{1}{15}e^{800k(5)}}{1+\frac{1}{15}e^{800k(5)}}$$

and thus

$$\frac{9}{4} = \frac{e^{4000k}}{1+\frac{1}{15}e^{4000k}}.$$

Cross-multiplying gives

$$\frac{9}{4}(1+\frac{1}{15}e^{4000k}) = e^{4000k}$$

and

$$\frac{9}{4} + \frac{3}{20}e^{4000k} = e^{4000k}$$

so

$$\frac{9}{4} = \frac{17}{20}e^{4000k}.$$

This means

$$\frac{45}{17} = e^{4000k}$$

and finally that

$$k = \frac{\ln\left(\frac{45}{17}\right)}{4000}.$$

Substituting k back into the population model and evaluating at $t = 25$ we get

$$\begin{aligned} P(25) &= 800\frac{\frac{1}{15}e^{800\frac{\ln\left(\frac{45}{17}\right)}{4000}(25)}}{1+\frac{1}{15}e^{800\frac{\ln\left(\frac{45}{17}\right)}{4000}(25)}} \\ &= 800\frac{\frac{1}{15}e^{5\ln\left(\frac{45}{17}\right)}}{1+\frac{1}{15}e^{5\ln\left(\frac{45}{17}\right)}} \\ &= 717 \text{ elephants} \end{aligned}$$

It follows that after 25 years the population has very nearly reached its maximum (800 elephants).

◀

Logistic growth also applies to many other situations.

EXAMPLE 17 A rumor is circulating around a university campus. A survey revealed that at one point only 5% of the students in the school were aware of the rumor. However, since news on campus spreads quickly, after 10 hours the rumor is known by 10% of the student body. How long will it take until 30% of the students are aware of the rumor?

Let $r(t)$ be the fraction of the student body at time t that have heard this rumor. Then $0 \leq r(t) \leq 1$.

Experiments have shown that the rate at which a rumor spreads through a population is proportional to the product of the fraction of the population that have heard the rumor and the fraction that have not. Therefore, there is a constant k such that

$$r' = kr(1-r)$$

and so this is a logistic growth model with $M = 1$. It follows that there is a positive constant C such that

$$r(t) = \frac{Ce^{kt}}{1+Ce^{kt}}.$$

We know that at $r(0) = 0.05$ so

$$C = \frac{r(0)}{1 - r(0)} = \frac{0.05}{0.95} = 0.0526315$$

and hence that

$$\begin{aligned} 0.1 &= r(10) \\ &= \frac{\frac{0.05}{0.95}e^{10k}}{1 + \frac{0.05}{0.95}e^{10k}} \end{aligned}$$

Therefore

$$0.1 + \frac{0.005}{0.95}e^{10k} = \frac{0.05}{0.95}e^{10k}$$

and

$$0.1 = \frac{0.045}{0.95}e^{10k}.$$

This gives

$$e^{10k} = \frac{0.095}{0.045}$$

and

$$k = \frac{\ln\left(\frac{0.095}{0.045}\right)}{10} = 0.07472$$

Finally, we want to find t_0 such that

$$0.3 = \frac{Ce^{kt_0}}{1 + Ce^{kt}}.$$

Therefore,

$$0.3(1 + Ce^{kt}) = Ce^{kt_0}$$

so

$$0.3 = 0.7Ce^{kt_0}$$

and

$$e^{kt_0} = \frac{0.3}{0.7C}.$$

This shows that

$$t_0 = \frac{\ln\left(\frac{0.3}{0.7C}\right)}{k} = \frac{\ln\left(\frac{0.3}{0.7(0.0526315)}\right)}{0.07472} = 28.07$$

hours.

After 28.07 hours, 30% of the student population had heard the rumor.

◀

There are many other important examples of logistic models that are similar to the previous example. For example, the spread of disease through a population also behaves like the spread of a rumor and as such can be studied with a logistic growth model.

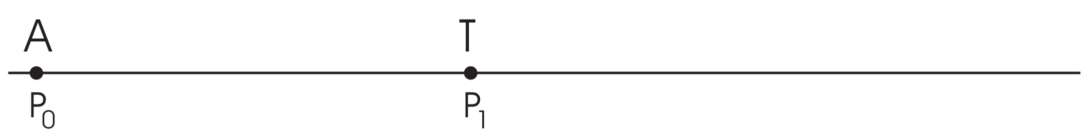
Chapter 5

Numerical Series

The main topic in this chapter is infinite series. You will learn that a series is just a sum of infinitely many terms. One of the main problems that you will encounter is to try to determine what it means to add infinitely many terms . We will accomplish this task by defining the *sequence of partial sums* and then studying the convergence of the series.

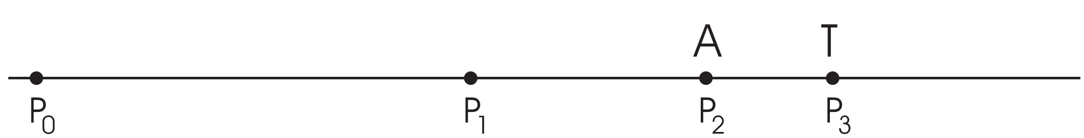
5.1 Introduction to Series

The Greek philosopher Zeno, who lived from 490-425 BC, proposed many paradoxes. The most famous of these is the *Paradox of Achilles and the Tortoise*. In this paradox, Achilles is supposed to race a tortoise. To make the race fair, Achilles (A) gives the tortoise (T) a substantial head start.



Zeno would argue that before Achilles could catch the tortoise, he must first go from his starting point at P_0 to that of the tortoise at P_1. However, by this time the tortoise has moved forward to P_2.

This time, before Achilles could catch the tortoise, he must first go from P_1 to where the tortoise was at P_2. However, by the time Achilles completes this task, the tortoise has moved forward to P_3.



Each time Achilles reaches the position that the tortoise had been, the tortoise has moved further ahead.

This process of Achilles trying to reach where the tortoise was ad infinitum led Zeno to suggest that Achilles could *never catch the tortoise*.

Zeno's argument seems to be supported by the following observation:

Let t_1 denote the time it would take for Achilles to get from his starting point P_0 to P_1. Let t_2 denote the time it would take for Achilles to get from P_1 to P_2, and let t_3 denote the time it would take for Achilles to get from P_2 to P_3. More generally, let t_n denote the time it would take for Achilles to get from P_{n-1} to P_n. Then the time it would take to catch the tortoise would be at least as large as the sum

$$t_1 + t_2 + t_3 + t_4 + \cdots + t_n + \cdots$$

of all of these infinitely many time periods.

Since each $t_n > 0$, Achilles is being asked to complete infinitely many tasks (each of which takes a positive amount of time) in a finite amount of time. It may seem that this is impossible. However, this is certainly a paradox because we know from our own experience that someone as swift as Achilles will eventually catch and even pass the tortoise. Hence, the sum

$$t_1 + t_2 + t_3 + t_4 + \cdots + t_n + \cdots$$

must be *finite*.

This statement brings into question the following very fundamental problem.

Problem: Given an *infinite* sequence $\{a_n\}$ of real numbers, what do we mean by the sum

$$a_1 + a_2 + a_3 + a_4 + \cdots + a_n + \cdots ?$$

To see why this is an issue, consider the following example.

EXAMPLE 1 Let $a_n = (-1)^{n-1}$. Consider

$$1 + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + \cdots .$$

If we want to find this sum, we could try to use the associative property of finite sums and group the terms as follows:

$$[1 + (-1)] + [1 + (-1)] + [1 + (-1)] + [1 + (-1)] + \cdots .$$

This would give

$$0 + 0 + 0 + 0 + \cdots$$

which must be 0. Therefore, we might expect that

$$1 + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + \cdots = 0.$$

This makes sense since *there appears to be the same number of* 1*'s and* -1*'s*, so cancellation should make the sum 0.

However, if we choose to group the terms the differently,

$$1 + [(-1) + 1] + [(-1) + 1] + [(-1) + 1] + [(-1) + 1] + \cdots$$

then we get

$$1 + 0 + 0 + 0 + 0 + \cdots = 1.$$

Both methods seem to be equally valid so we cannot be sure of the real sum. It seems that *the usual rules of arithmetic do not hold for infinite sums*. We must look for an alternate approach.

◀

Since finite sums behave very well, we might try adding all of the terms up to a certain cut-off k and then see if a pattern develops as k gets very large. This is in fact how we will proceed.

DEFINITION

Series

Given a sequence $\{a_n\}$, the *formal* sum

$$a_1 + a_2 + a_3 + a_4 + \cdots + a_n + \cdots$$

is called a *series*. The series is called *formal* because we have not yet given it a meaning numerically.

The a_n's are called the *terms* of the series. For each term a_n, the *index* of the term is n.

We will denote the series by

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

Note that all of the series we have listed so far have started with the first term indexed by 1. This is not necessary. In fact, it is quite common for a series to begin with the initial index being 0. In fact, the series can start at any initial point.

$$\overbrace{\sum_{\underbrace{n=j}_{\text{initial index}}}^{\infty}}^{\text{final index}} a_n = a_j + a_{j+1} + a_{j+2} + a_{j+3} + \ldots$$

DEFINITION

Convergence of a Series

Given a series

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

for each $k \in \mathbb{N}$, we define the *k-th partial sum* S_k by

$$S_k = \sum_{n=1}^{k} a_n.$$

We say that the series $\sum\limits_{n=1}^{\infty} a_n$ *converges* if the sequence $\{S_k\}$ of partial sums converges. In this case, if $L = \lim\limits_{k\to\infty} S_k$, then we write

$$\sum_{n=1}^{\infty} a_n = L$$

and assign the sum this value. Otherwise, we say that the series $\sum\limits_{n=1}^{\infty} a_n$ *diverges*.

We can apply these definitions to the series that we considered earlier in this section.

EXAMPLE 2 Let $a_n = (-1)^{n-1}$. Consider the sequence of partial sums:

$$\begin{array}{lllllllll}
S_1 & = & a_1 & & & & & = 1 \\
S_2 & = & a_1 + a_2 & = & S_1 + a_2 & = & S_1 - 1 & = 0 \\
S_3 & = & a_1 + a_2 + a_3 & = & S_2 + a_3 & = & S_2 + 1 & = 1 \\
S_4 & = & a_1 + a_2 + a_3 + a_4 & = & S_3 + a_4 & = & S_3 - 1 & = 0 \\
S_5 & = & a_1 + a_2 + a_3 + a_4 + a_5 & = & S_4 + a_5 & = & S_4 + 1 & = 1 \\
 & \vdots & & & & & &
\end{array}$$

Therefore,

$$S_k = \begin{cases} 0 & \text{if } k \text{ is even} \\ 1 & \text{if } k \text{ is odd} \end{cases}$$

This shows that the sequence of partial sums $\{S_k\}$ diverges, and hence so does $\sum\limits_{n=1}^{\infty}(-1)^{n-1}$. ◀

EXAMPLE 3 Determine if the series

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

converges or diverges.

Observe that

$$a_n=\frac{1}{n^2+n}=\frac{1}{n(n+1)}.$$

Moreover, we can write

$$a_n=\frac{1}{n(n+1)}=\frac{1}{n}-\frac{1}{n+1}.$$

Therefore the series becomes

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

To calculate S_k note that

$$\begin{aligned} S_k &= \sum_{n=1}^{k}\left(\frac{1}{n}-\frac{1}{n+1}\right)\\ &= (1-\frac{1}{2})+(\frac{1}{2}-\frac{1}{3})+(\frac{1}{3}-\frac{1}{4})+(\frac{1}{4}-\frac{1}{5})+\cdots+(\frac{1}{k}-\frac{1}{k+1}) \end{aligned}$$

If we regroup these terms, we get

$$\begin{aligned} S_k &= (1-\frac{1}{2})+(\frac{1}{2}-\frac{1}{3})+(\frac{1}{3}-\frac{1}{4})+(\frac{1}{4}-\frac{1}{5})+\cdots+(\frac{1}{k}-\frac{1}{k+1})\\ &= 1-(\frac{1}{2}-\frac{1}{2})-(\frac{1}{3}-\frac{1}{3})-(\frac{1}{4}-\frac{1}{4})-(\frac{1}{5}-\frac{1}{5})-\cdots-(\frac{1}{k}-\frac{1}{k})-\frac{1}{k+1}\\ &= 1-0-0-0-0-\cdots-0-\frac{1}{k+1}\\ &= 1-\frac{1}{k+1} \end{aligned}$$

Then

$$\lim_{k\to\infty}S_k=\lim_{k\to\infty}\left(1-\frac{1}{k+1}\right)=1.$$

Since the sequence of partial sums $\{S_k\}$ converges to 1, the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^2+n}$ converges and

$$\sum_{n=1}^{\infty} \frac{1}{n^2+n} = 1.$$

What is remarkable about the previous series is not that we were able to show that it converges, but rather that we could find its sum so easily. Generally, this will not be the case. In fact, even if we know a series converges, it may be very difficult or even impossible to determine the exact value of its sum. In most cases, we will have to be content with either showing that a series converges or that it diverges and, in the case of a convergent series, *estimating* its sum.

The next section deals with an important class of series known as *geometric series*. Not only can we determine if such a series converges, but we can easily find the sum.

5.2 Geometric Series

Perhaps the most important type of series are the geometric series.

DEFINITION

Geometric Series

A *geometric series* is a series of the form

$$\sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + r^4 + \cdots$$

The number r is called the *ratio* of the series.

If $r = (-1)$, the series is

$$\sum_{n=0}^{\infty} (-1)^n = 1 + (-1) + 1 + (-1) + 1 + (-1) + 1 + (-1) + \cdots$$

which we have already seen diverges.

If $r = 1$, the series is

$$\sum_{n=0}^{\infty} 1^n = 1 + 1 + 1 + 1 + 1 + \cdots$$

which again diverges since $S_k = \sum\limits_{n=0}^{k} 1^n = k + 1$ diverges to ∞.

Question: Which if any of the geometric series converge?

Assume that $r \neq 1$. Let

$$S_k = 1 + r + r^2 + r^3 + r^4 + \cdots + r^k.$$

Then

$$\begin{aligned} rS_k &= r(1 + r + r^2 + r^3 + r^4 + \cdots + r^k) \\ &= r + r^2 + r^3 + r^4 + \cdots + r^{k+1} \end{aligned}$$

Therefore

$$\begin{aligned} S_k - rS_k &= (1 + r + r^2 + r^3 + r^4 + \cdots + r^k) - (r + r^2 + r^3 + r^4 + \cdots + r^k + r^{k+1}) \\ &= 1 - r^{k+1} \end{aligned}$$

Hence

$$(1 - r)S_k = S_k - rS_k = 1 - r^{k+1}$$

and since $r \neq 1$,

$$S_k = \frac{1 - r^{k+1}}{1 - r}.$$

The only term in this expression that depends on k is r^{k+1}, so $\lim\limits_{k\to\infty} S_k$ exists if and only if $\lim\limits_{k\to\infty} r^{k+1}$ exists. However, if $| r |< 1$, then r^{k+1} becomes very small for large k. That is $\lim\limits_{k\to\infty} r^{k+1} = 0$.

If $| r |> 1$, then $| r^{k+1} |$ becomes very large as k grows. That is, $\lim\limits_{k\to\infty} | r^{k+1} |= \infty$. Hence, $\lim\limits_{k\to\infty} r^{k+1}$ does not exist.

Finally, if $r = -1$, then r^{k+1} alternates between 1 and -1, so $\lim\limits_{k\to\infty} r^{k+1}$ again diverges.

This shows that r^{k+1}, and hence the series $\sum\limits_{n=0}^{\infty} r^n$, converges if and only if $| r |< 1$. Moreover, in this case,

$$\begin{aligned} \lim_{k\to\infty} S_k &= \lim_{k\to\infty} \frac{1 - r^{k+1}}{1 - r} \\ &= \frac{1 - \lim\limits_{k\to\infty} r^{k+1}}{1 - r} \\ &= \frac{1}{1 - r} \end{aligned}$$

THEOREM 1 **Geometric Series Test**

The geometric series $\sum_{n=0}^{\infty} r^n$ converges if $|r| < 1$ and diverges otherwise.

If $|r| < 1$, then

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}.$$

EXAMPLE 4 Evaluate $\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$.

SOLUTION This is a geometric series with ratio $r = \frac{1}{2}$. Since $0 < \frac{1}{2} < 1$, the Geometric Series Test shows that $\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n$ converges. Moreover,

$$\begin{aligned} \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n &= \frac{1}{1-\frac{1}{2}} \\ &= 2 \end{aligned}$$

◀

5.3 Divergence Test

It makes sense that if we are to add together infinitely many positive numbers and get something finite, then the terms must eventually be small. We will now see that this statement holds for any convergent series.

THEOREM 2 **Divergence Test**

Assume that $\sum_{n=1}^{\infty} a_n$ converges. Then

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

Equivalently, if $\lim_{n\to\infty} a_n \neq 0$ or if $\lim_{n\to\infty} a_n$ does not exist, then $\sum_{n=1}^{\infty} a_n$ diverges.

The Divergence Test gets its name because it can identify certain series as being divergent, but **it cannot show that a series converges**.

EXAMPLE 5 Consider the geometric series $\sum_{n=0}^{\infty} r^n$ with $|r| \geq 1$. Then $\lim_{n\to\infty} r^n = 1$ if $r = 1$ and it does not exist for all other r with $|r| \geq 1$ (in other words, if $r = -1$ or if $|r| > 1$). The Divergence Test shows that if $|r| \geq 1$, then $\sum_{n=0}^{\infty} r^n$ diverges. ◀

The Divergence Test works for the following reason. Assume that $\sum\limits_{n=1}^{\infty} a_n$ converges to L. This is equivalent to saying that

$$\lim_{k\to\infty} S_k = L.$$

By the basic properties of convergent sequences, we get that

$$\lim_{k\to\infty} S_{k-1} = L$$

as well.

However, for $k \geq 2$,

$$\begin{aligned} S_k - S_{k-1} &= \sum_{n=1}^{k} a_n - \sum_{n=1}^{k-1} a_n \\ &= (a_1 + a_2 + a_3 + a_4 + \cdots + a_{k-1} + a_k) - (a_1 + a_2 + a_3 + a_4 + \cdots + a_{k-1}) \\ &= a_k \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{k\to\infty} a_k &= \lim_{k\to\infty}(S_k - S_{k-1}) \\ &= \lim_{k\to\infty} S_k - \lim_{k\to\infty} S_{k-1} \\ &= L - L \\ &= 0. \end{aligned}$$

EXAMPLES

1. Consider the sequence $\{\frac{n}{n+1}\}$. Then

   $$\lim_{n\to\infty} \frac{n}{n+1} = 1.$$

   Therefore, the Divergence Test shows that

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

   diverges.

2. While it is difficult to do so, it is possible to show that

   $$\lim_{n\to\infty} \sin(n)$$

   does not exist. Therefore, the Divergence Test shows that the series

   $$\sum_{n=1}^{\infty} \sin(n)$$

   diverges.

3. The Divergence Test shows that if either $\lim\limits_{n\to\infty} a_n \neq 0$ or if $\lim\limits_{n\to\infty} a_n$ does not exist, then $\sum\limits_{n=1}^{\infty} a_n$ diverges. It would seem natural to ask if the *converse* statement holds. That is:

   Question: If $\lim\limits_{n\to\infty} a_n = 0$, does this mean that $\sum\limits_{n=1}^{\infty} a_n$ converges?

   We will see that the answer to the question above is: **No**, the fact that $\lim\limits_{n\to\infty} a_n = 0$, does not mean that $\sum\limits_{n=1}^{\infty} a_n$ converges.

   Let $a_n = \frac{1}{n}$ and

$$S_k = \sum_{n=1}^{k} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{k}.$$

   Then

$$\begin{array}{rcl}
S_1 & = & 1 \\
S_2 & = & 1 + \frac{1}{2} \\
S_4 & = & 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} \\
 & = & 1 + \frac{1}{2} + (\frac{1}{3} + \frac{1}{4}) \\
 & > & 1 + \frac{1}{2} + (\frac{1}{4} + \frac{1}{4}) \\
 & = & 1 + \frac{1}{2} + \frac{1}{2} \\
 & = & 1 + \frac{2}{2} \\
S_8 & = & 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \\
 & = & 1 + \frac{1}{2} + (\frac{1}{3} + \frac{1}{4}) + (\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}) \\
 & > & 1 + \frac{1}{2} + (\frac{1}{4} + \frac{1}{4}) + (\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}) \\
 & = & 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} \\
 & = & 1 + \frac{3}{2} \\
 & \vdots &
\end{array}$$

   We have seen that

$$\begin{array}{rcccl}
S_1 & = & S_{2^0} & = & 1 + \frac{0}{2} \\
S_2 & = & S_{2^1} & = & 1 + \frac{1}{2} \\
S_4 & = & S_{2^2} & > & 1 + \frac{2}{2} \\
S_8 & = & S_{2^3} & > & 1 + \frac{3}{2} \\
 & & \vdots & &
\end{array}$$

A pattern has emerged. In general, we can show that for any m

$$S_{2^m} \geq 1 + \frac{m}{2}.$$

However, the sequence $1 + \frac{m}{2}$ grows without bounds. It follows that the partial sums of the form S_{2^m} also grow without bound. This shows that the series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges to ∞ as well. ◀

This example shows that even if $\lim\limits_{n\to\infty} a_n = 0$, it is still possible for $\sum\limits_{n=1}^{\infty} a_n$ to **diverge!!!**

Note:

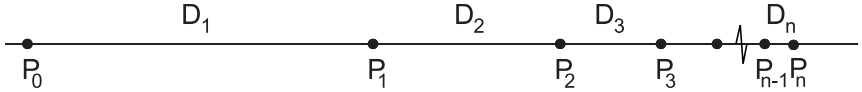
1. The sequence $\{\frac{1}{n}\}$ was first studied in detail by Pythagoras who felt that these ratios represented *musical harmony*. For this reason the sequence $\{\frac{1}{n}\}$ is called *the Harmonic Progression* and the series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ is called *the Harmonic Series*.

   We have just shown that the Harmonic Series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges to ∞. However, the argument to do this was quite clever. Instead, we might ask if we could use a computer to add up the first k terms for some large k and show that the sums are getting large? In this regard, we may want to know how many terms it would take so that

   $$S_k = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \cdots + \frac{1}{k} > 100\ ?$$

   The answer to this question is very surprising. It can be shown that k must be at least 10^{30}, which is an enormous number. No modern computer could ever perform this many additions!!!

2. Recall that in Zeno's paradox, Achilles had to travel infinitely many distances in a finite amount of time to catch the tortoise. If D_n represents the distance between points P_{n-1} (where Achilles is after $n-1$ steps) and P_n (where the tortoise is currently located), then the D_n's are becoming progressively smaller.

   If t_n is the time it takes Achilles to cover the distance D_n, then the t_n's are also becoming progressively smaller. In fact, they are so small that $\lim\limits_{n\to\infty} t_n = 0$ and indeed it is reasonable to assume that

   $$\sum_{n=1}^{\infty} t_n$$

   converges! This is how we can resolve Zeno's paradox.

5.4 Arithmetic of Series

Since convergent series can be viewed as the limit of their sequences of partial sums, the arithmetic rules for sequences can be applied whenever they are appropriate. With this in mind, we get:

THEOREM 3 **Arithmetic for Series I**

Assume that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ both converge.

1. The series $\sum_{n=1}^{\infty} ca_n$ converges for every $c \in \mathbb{R}$ and
$$\sum_{n=1}^{\infty} ca_n = c\sum_{n=1}^{\infty} a_n.$$

2. The series $\sum_{n=1}^{\infty} (a_n + b_n)$ converges and
$$\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n.$$

These rules should not be surprising. They follow immediately from the corresponding rules for sequences.

There is one other rule that we will need that does not have an analog for sequences.

Given a series $\sum_{n=1}^{\infty} a_n$, let $j \in \mathbb{N}$. Let

$$\sum_{n=j}^{\infty} a_n = a_j + a_{j+1} + a_{j+2} + a_{j+3} + \cdots .$$

We say that $\sum_{n=j}^{\infty} a_n$ *converges* if

$$\lim_{k\to\infty} T_k$$

exists, where

$$T_k = \sum_{n=j}^{j+k-1} a_n = a_j + a_{j+1} + a_{j+2} + a_{j+3} + \cdots + a_{j+k-1}.$$

The following theorem relates the convergence of the series $\sum_{n=j}^{\infty} a_n$ with that of the original series $\sum_{n=1}^{\infty} a_n$.

THEOREM 4 **Arithmetic for Series II**

1. If $\sum\limits_{n=1}^{\infty} a_n$ converges, then $\sum\limits_{n=j}^{\infty} a_n$ also converges for **each** j.

2. If $\sum\limits_{n=j}^{\infty} a_n$ converges for **some** j, then $\sum\limits_{n=1}^{\infty} a_n$ converges.

In either of these two cases,

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + \cdots + a_{j-1} + \sum_{n=j}^{\infty} a_n.$$

Observation: As a consequence of the previous theorem, we make the following very important observation. Given a sequence

$$a_1, a_2, a_3, \cdots, a_n, \cdots$$

and the first $j-1$ terms are changed to create a new sequence

$$b_1, b_2, b_3, \cdots, b_n, \cdots$$

where $b_n = a_n$ and if $n \geq j$, then the series $\sum\limits_{n=j}^{\infty} a_n$ and $\sum\limits_{n=j}^{\infty} b_n$ are identical. Hence, they either both converge or both diverge. The previous theorem can now be used to show that **either both $\sum\limits_{n=1}^{\infty} a_n$ and $\sum\limits_{n=1}^{\infty} b_n$ converge or they both diverge!!!** Therefore, **convergence of a series depends completely on the *tail* of the sequence of terms. Changing finitely many terms will not affect convergence, though in the case of a convergent series it may affect the final sum.**

EXAMPLE 6 A ball is launched straight up from the ground to a height of 30m. When the ball returns to the ground it will bounce to a height that is exactly $\frac{1}{3}$ of its previous height. Assuming that the ball continues to bounce each time it returns to the ground, how far does the ball travel before coming to rest?

Prior to returning to the ground for the first time, the ball travels 30m on its way up and then 30m down for a total of $2(30) = 60$m.

On the first bounce, the ball will travel upwards $\frac{30}{3}$m and down again the same distance for a total of $2\left(\frac{30}{3}\right)$m.

On the second bounce, the ball will travel upwards one third the distance of the first bounce or

$$\frac{1}{3}\left(\frac{30}{3}\right) = \frac{30}{3^2}\text{m}.$$

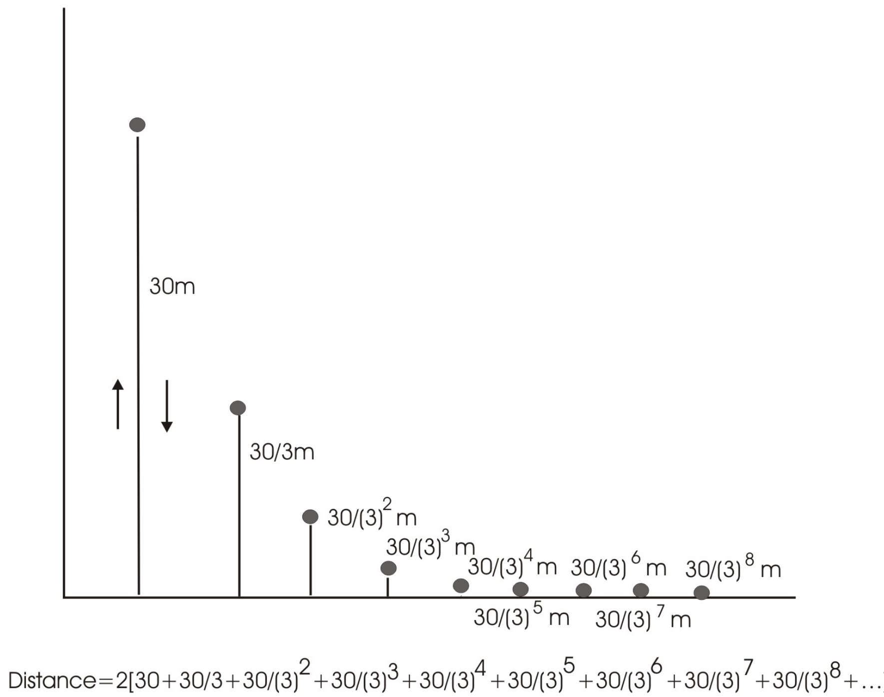
It will also travel down the same distance for a total of $2\left(\frac{30}{3^2}\right)$m.

On the third bounce, the ball will again travel upwards one third the distance it traveled on the second bounce or

$$\frac{1}{3}\left(\frac{30}{3^2}\right) = \left(\frac{30}{3^3}\right)\text{m}.$$

With the downward trip, the third bounce covers a distance of $2\left(\frac{30}{3^3}\right)$m.

Note the pattern that has formed. On the n-th bounce, the ball will travel a distance of $2\left(\frac{30}{3^n}\right)$m. The total distance D the ball travels will be the sum of each of these distances.



Therefore, using our rules of arithmetic and what we know about geometric series, we get

$$\begin{aligned}
D &= 2(30) + 2\left(\frac{30}{3}\right) + 2\left(\frac{30}{3^2}\right) + 2\left(\frac{30}{3^3}\right) + \cdots + 2\left(\frac{30}{3^n}\right) + \cdots \\
&= 2(30)[1 + \frac{1}{3} + \frac{1}{3^2} + \frac{1}{3^3} + \cdots + \frac{1}{3^n} + \cdots] \\
&= 60\sum_{n=0}^{\infty}\left(\frac{1}{3}\right)^n \\
&= \frac{60}{1-\frac{1}{3}} \\
&= 90 \text{ meters.}
\end{aligned}$$

Since we are assuming that the ball will bounce infinitely often, we might expect that this process would continue forever. However, this is not the case. In fact, the reasoning is very similar to that of the resolution of Zeno's paradox since the amount of time it takes for the ball to complete each bounce decreases very rapidly. Indeed, by using some basic physics, we can actually calculate the total time it would take for the ball to complete its travels.

It is known from physics that if a ball is dropped, it will fall a distance

$$S = \frac{1}{2}gt^2$$

meters in t seconds, where $g = 9.81$ m/(sec)^2 is the acceleration due to gravity. Therefore, if a ball is dropped from a height h we can determine how long it will take to reach the ground. We have

$$h = \frac{1}{2}gt^2$$

so that

$$t^2 = \frac{2h}{g}$$

or

$$t = \sqrt{\frac{2h}{g}}.$$

In our case, the ball will take the same amount of time to make the upward trip as the downward trip. Therefore, the total time it will take to complete the n-th bounce (the 0-th bounce is the original trip) will be

$$t_n = 2\sqrt{\frac{2h_n}{g}}$$

where h_n is the height of the n-th bounce. But

$$h_n = \frac{30}{3^n}$$

so

$$t_n = 2\sqrt{\frac{2(30)}{g(3^n)}} = \left(2\sqrt{\frac{60}{g}}\right)\left(\frac{1}{\sqrt{3}}\right)^n$$

It follows that the total time it will take for the ball to complete all of the bounces is

$$\begin{aligned}
\sum_{n=0}^{\infty} t_n &= \sum_{n=0}^{\infty}\left(2\sqrt{\frac{60}{g}}\right)\left(\frac{1}{\sqrt{3}}\right)^n \\
&= 2\sqrt{\frac{60}{g}}\sum_{n=0}^{\infty}\left(\frac{1}{\sqrt{3}}\right)^n \\
&= 2\sqrt{\frac{60}{g}}\left(\frac{1}{1-\frac{1}{\sqrt{3}}}\right) \\
&= 11.708 \text{ seconds.}
\end{aligned}$$

◀

5.5 Positive Series

We will soon see that the Monotone Convergence Theorem can help us determine the convergence of series, particularly those series with positive terms. Recall the following definition:

DEFINITION

Monotonic Sequences

Given a sequence $\{a_n\}$, we say that the sequence is

i) *non-decreasing* if $a_{n+1} \geq a_n$ for every $n \in \mathbb{N}$.

ii) *increasing* if $a_{n+1} > a_n$ for every $n \in \mathbb{N}$.

iii) *non-increasing* if $a_{n+1} \leq a_n$ for every $n \in \mathbb{N}$.

iv) *decreasing* if $a_{n+1} < a_n$ for every $n \in \mathbb{N}$.

We say that $\{a_n\}$ is *monotonic* if it satisfies one of these four conditions.

The Monotone Convergence Theorem gives us a simple criterion for determining the convergence or divergence of a monotonic sequence.

THEOREM 5

Monotone Convergence Theorem (MCT)

Let $\{a_n\}$ be a non-decreasing sequence.

1. If $\{a_n\}$ is bounded above, then $\{a_n\}$ converges to $L = lub(\{a_n\})$.
2. If $\{a_n\}$ is not bounded above, then $\{a_n\}$ diverges to ∞.

In particular, $\{a_n\}$ converges if and only if it is bounded above.

Note: A similar statement can be made about non-increasing sequences by replacing the least upper bound with the greatest lower bound and ∞ by $-\infty$. ◀

DEFINITION

Positive Series

We call a series $\sum\limits_{n=1}^{\infty} a_n$ *positive* if the terms $a_n \geq 0$ for all $n \in \mathbb{N}$.

Assume that $a_n \geq 0$ for all n. Let

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

be the k-th partial sum of the series with terms $\{a_n\}$. Then

$$\begin{aligned} S_{k+1} - S_k &= \sum_{n=1}^{k+1} a_n - \sum_{n=1}^{k} a_n \\ &= (a_1 + a_2 + \cdots + a_k + a_{k+1}) - (a_1 + a_2 + \cdots + a_k) \\ &= a_{k+1} \\ &\geq 0 \end{aligned}$$

This shows that

$$S_{k+1} \geq S_k$$

so $\{S_k\}$ is a non-decreasing sequence. The Monotone Convergence Theorem tells us that there are two possibilities for the sequence $\{S_k\}$:

1. $\{S_k\}$ is bounded and therefore, by the MCT it converges.
2. $\{S_k\}$ diverges to ∞.

In terms of the series this is equivalent to

1. $\sum\limits_{n=1}^{\infty} a_n$ converges.
2. $\sum\limits_{n=1}^{\infty} a_n$ diverges to ∞.

Key Observation: Therefore, for positive series, the convergence of the series essentially depends only on how large are the terms a_n. Generally speaking, the larger the a_n's, the more likely it is that a series will diverge to ∞ and the smaller the a_n's, the more likely it is that a series will converge.

To make this statement more precise, assume that we have two series,

$$\sum_{n=1}^{\infty} a_n \quad \text{and} \quad \sum_{n=1}^{\infty} b_n$$

with $0 \leq a_n \leq b_n$ for all $n \in \mathbb{N}$. Assume also that the series $\sum\limits_{n=1}^{\infty} b_n$ with the larger terms converges to some number L. Since $a_n \leq b_n$ for all $n \in \mathbb{N}$, we would *not* expect $\sum\limits_{n=1}^{\infty} a_n$ to diverge to ∞. In fact, if we let

$$S_k = \sum_{n=1}^{k} a_n$$

and

$$T_k = \sum_{n=1}^{k} b_n,$$

then

$$\begin{aligned} S_k &= a_1 + a_2 + \cdots + a_k \\ &\leq b_1 + b_2 + \cdots + b_k \\ &= T_k \\ &\leq L \end{aligned}$$

since $L = \lim\limits_{k\to\infty} T_k$.

We have shown that for each k,

$$S_k \leq L = \sum_{n=1}^{\infty} b_n.$$

However, the sequence $\{S_k\}$ is increasing and we have just shown that it is bounded above by L. The Monotone Convergence Theorem shows that $\{S_k\}$ converges to some M with $M \leq L$. In other words,

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

On the other hand, if the series $\sum\limits_{n=1}^{\infty} a_n$ with the smaller terms diverges to infinity, then we can make the partial sum S_k as large as we like. But, $S_k \leq T_k$, so that we can make the T_k's as large as we like. This shows that

$$\lim_{k\to\infty} T_k = \infty$$

so that $\sum\limits_{n=1}^{\infty} b_n$ diverges to ∞.

This leads us to one of the **most important tools** we will have for determining the convergence or divergence of positive series.

5.5.1 Comparison Test

THEOREM 6 **Comparison Test for Series**

Assume that $0 \leq a_n \leq b_n$ for each $n \in \mathbb{N}$.

1. If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.
2. If $\sum_{n=1}^{\infty} a_n$ diverges, then $\sum_{n=1}^{\infty} b_n$ diverges.

Remark: We must make three important observations concerning the Comparison Test.

1. If $\sum_{n=1}^{\infty} a_n$ converges, then we cannot say anything about $\sum_{n=1}^{\infty} b_n$.
2. If $\sum_{n=1}^{\infty} b_n$ diverges, then we cannot say anything about $\sum_{n=1}^{\infty} a_n$.
3. Since the first few terms do not affect whether or not a series diverges, for the Comparison Test to hold, we really only need that

$$0 \leq a_n \leq b_n$$

for each $n \geq K$, where $K \in \mathbb{N}$. That is, the conditions of the theorem need only be satisfied by the elements of the tails of the two sequences.

EXAMPLE 7 We have seen that the Harmonic Series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges. We also know that

$$0 < \frac{1}{n} \leq \frac{1}{\sqrt{n}}$$

for all $n \in \mathbb{N}$. Let $a_n = \frac{1}{n}$ and $b_n = \frac{1}{\sqrt{n}}$. Then the Comparison Test shows that the series $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ also diverges. ◀

EXAMPLE 8 Consider the series

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

If we use a method similar to that of the Partial Fraction decomposition, we get that

$$\frac{1}{n(n-1)} = \frac{1}{n-1} - \frac{1}{n}.$$

Let

$$\begin{aligned}
S_k &= \sum_{n=2}^{k} \frac{1}{n^2-n} \\
&= \sum_{n=2}^{k} \frac{1}{n-1} - \frac{1}{n} \\
&= \left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{k-2}-\frac{1}{k-1}\right)+\left(\frac{1}{k-1}-\frac{1}{k}\right) \\
&= 1-\left(\frac{1}{2}-\frac{1}{2}\right)-\left(\frac{1}{3}-\frac{1}{3}\right)-\left(\frac{1}{4}-\frac{1}{4}\right)-\cdots-\left(\frac{1}{k-2}-\frac{1}{k-2}\right)-\left(\frac{1}{k-1}-\frac{1}{k-1}\right)-\frac{1}{k} \\
&= 1-\frac{1}{k}
\end{aligned}$$

Since $\lim\limits_{k\to\infty} S_k = \lim\limits_{k\to\infty} 1-\frac{1}{k} = 1$, the series $\sum\limits_{n=2}^{\infty} \frac{1}{n^2-n}$ converges with

$$\sum_{n=2}^{\infty} \frac{1}{n^2-n} = 1.$$

Next, observe that for $n \geq 2$, we have $n^2 > n^2-n$ and hence that

$$0 < \frac{1}{n^2} < \frac{1}{n^2-n}.$$

However, we have just shown that $\sum\limits_{n=2}^{\infty} \frac{1}{n^2-n}$ converges to 1. We can use the Comparison Test to conclude that $\sum\limits_{n=2}^{\infty} \frac{1}{n^2}$ converges and that

$$\sum_{n=2}^{\infty} \frac{1}{n^2} \leq 1.$$

We can now immediately conclude that the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ also converges and

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{1^2} + \sum_{n=2}^{\infty} \frac{1}{n^2} \leq 1+1 = 2.$$

In fact, using techniques that are beyond the scope of this course, it can be shown that $\sum\limits_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \cong 1.64493.$ ◀

EXAMPLE 9 Let $p \in \mathbb{R}$ with $p \geq 2$. Then for each $n \in \mathbb{N}$, $n^2 \leq n^p$. Hence

$$0 \leq \frac{1}{n^p} \leq \frac{1}{n^2}.$$

Since $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ converges, the Comparison Test shows that if $p \geq 2$, $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ also converges.

If $p \leq 1$, then $n^p \leq n$ for each $n \in \mathbb{N}$. It follows that

$$0 \leq \frac{1}{n} \leq \frac{1}{n^p}.$$

This time, we know that the series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges. The Comparison Test shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ also diverges.

Therefore, $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p \geq 2$ and diverges if $p \leq 1$. It would be natural to ask about what happens when $1 < p < 2$. For example, does $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converge?

Unfortunately, we have

$$\frac{1}{n^2} \leq \frac{1}{n^{\frac{3}{2}}} \leq \frac{1}{n}.$$

Since $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ converges and $\frac{1}{n^2} \leq \frac{1}{n^{\frac{3}{2}}}$, this tells us nothing about $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$. Similarly, since $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges and $\frac{1}{n^{\frac{3}{2}}} \leq \frac{1}{n}$, this tells us nothing about $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$.

By making use of what we know about Improper Integrals, we will see later that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ actually converges. ◀

EXAMPLE 10 Recall that

$$n! = 1 \cdot 2 \cdot 3 \cdots n$$

for $n \geq 1$ and that

$$0! = 1.$$

Problem: Show that $\sum\limits_{n=0}^{\infty} \frac{1}{n!}$ converges and that

$$2 < \sum_{n=0}^{\infty} \frac{1}{n!} < 3.$$

Note: In fact $\sum\limits_{n=0}^{\infty} \frac{1}{n!} = e$.

First consider that

$$\begin{aligned}\sum_{n=0}^{\infty}\frac{1}{n!} &= \frac{1}{0!}+\frac{1}{1!}+\frac{1}{2!}+\cdots \\ &= 1+\sum_{n=1}^{\infty}\frac{1}{n!}\end{aligned}$$

We will show that $\sum\limits_{n=1}^{\infty}\frac{1}{n!}$ converges and that

$$1<\sum_{n=1}^{\infty}\frac{1}{n!}<2.$$

Let $a_n=\frac{1}{n!}$ and let $b_n=\frac{1}{2^{n-1}}$.

Then

$$\begin{aligned}a_1 &= \frac{1}{1}=\frac{1}{2^0}=\frac{1}{2^{1-1}}=b_1 \\ a_2 &= \frac{1}{1\cdot 2}=\frac{1}{2}=b_2 \\ a_3 &= \frac{1}{1\cdot 2\cdot 3}<\frac{1}{1\cdot 2\cdot 2}=\frac{1}{2^2}=b_3 \\ a_4 &= \frac{1}{1\cdot 2\cdot 3\cdot 4}<\frac{1}{1\cdot 2\cdot 2\cdot 2}=\frac{1}{2^3}=b_4 \\ &\vdots \\ a_n &= \frac{1}{1\cdot 2\cdot 3\cdot 4\cdots n}<\frac{1}{1\cdot 2\cdot 2\cdot 2\cdots 2}=\frac{1}{2^{n-1}}=b_n\end{aligned}$$

This shows that $0<a_n\leq b_n$ for each $n\in\mathbb{N}$.

Now

$$\begin{aligned}\sum_{n=1}^{\infty}b_n &= \sum_{n=1}^{\infty}\frac{1}{2^{n-1}} \\ &= \frac{1}{2^0}+\frac{1}{2^1}+\frac{1}{2^2}+\cdots \\ &= \sum_{j=0}^{\infty}\frac{1}{2^j} \\ &= \frac{1}{1-\frac{1}{2}} \\ &= 2\end{aligned}$$

Since $0 < a_n \leq b_n$ for each $n \in \mathbb{N}$ and $\sum\limits_{n=1}^{\infty} b_n$ converges, the Comparison Test shows that $\sum\limits_{n=1}^{\infty} a_n = \sum\limits_{n=1}^{\infty} \frac{1}{n!}$ also converges and that

$$0 < \sum_{n=1}^{\infty} \frac{1}{n!} \leq \sum_{n=1}^{\infty} b_n = 2.$$

However, since $a_n < b_n$ for each $n \geq 3$, we have that

$$0 < \sum_{n=1}^{\infty} \frac{1}{n!} < \sum_{n=1}^{\infty} b_n = 2.$$

But $\sum\limits_{n=1}^{\infty} \frac{1}{n!} = \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots > 1$. Therefore,

$$1 < \sum_{n=1}^{\infty} \frac{1}{n!} < 2.$$

Finally, since

$$\sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \sum_{n=1}^{\infty} \frac{1}{n!},$$

we get that

$$2 = 1 + 1 < \sum_{n=0}^{\infty} \frac{1}{n!} < 1 + 2 = 3.$$

◀

We have seen that the Comparison Test can help determine whether certain series converge. We will now present a variation of the Comparison Test that will work for a significant collection of series, including all of those series where the terms are ratios of polynomials in n. We begin with such an example.

EXAMPLE 11 Let $a_n = \frac{2n}{n^3-n+1}$. Notice that if n is very large, then $(-n+1)$ is negligible in comparison to n^3. Therefore, we could say that for large n

$$n^3 - n + 1 \cong n^3$$

and so

$$a_n = \frac{2n}{n^3 - n + 1} \cong \frac{2n}{n^3} = \frac{2}{n^2}.$$

In other words, for very large n the terms $a_n = \frac{2n}{n^3-n+1}$ are roughly comparable in size to $\frac{2}{n^2}$.

We know that $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ converges and hence so does $\sum\limits_{n=1}^{\infty} \frac{2}{n^2}$. Since

$$a_n = \frac{2n}{n^3 - n + 1} \cong \frac{2}{n^2}$$

we might guess that $\sum\limits_{n=1}^{\infty} \frac{2n}{n^3-n+1}$ also converges. Unfortunately,

$$0 < \frac{2}{n^2} \leq \frac{2n}{n^3 - n + 1}$$

for all n. This means that we cannot immediately apply the Comparison Test to establish the convergence. However, we will be able to show that the next theorem will work in this case. It is essentially an *upgraded* version of the Comparison Test.

◀

5.5.2 Limit Comparison Test

THEOREM 7

Limit Comparison Test (LCT)

Assume that $a_n > 0$ and $b_n > 0$ for each $n \in \mathbb{N}$. Assume also that

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

where either $L \in \mathbb{R}$ or $L = \infty$.

1. If $0 < L < \infty$, then $\sum\limits_{n=1}^{\infty} a_n$ converges if and only if $\sum\limits_{n=1}^{\infty} b_n$ converges.
2. If $L = 0$ and $\sum\limits_{n=1}^{\infty} b_n$ converges, then $\sum\limits_{n=1}^{\infty} a_n$ converges. Equivalently, if $\sum\limits_{n=1}^{\infty} a_n$ diverges, then so does $\sum\limits_{n=1}^{\infty} b_n$.
3. If $L = \infty$ and $\sum\limits_{n=1}^{\infty} a_n$ converges, then $\sum\limits_{n=1}^{\infty} b_n$ converges. Equivalently, if $\sum\limits_{n=1}^{\infty} b_n$ diverges, then so does $\sum\limits_{n=1}^{\infty} a_n$.

PROOF

Assume that

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

1) If $0 < L < \infty$, the interval $(\frac{L}{2}, 2L)$ is an open interval containing L. It follows that we can find a cutoff $N \in \mathbb{N}$ so that if $n \geq N$, then
$$\frac{L}{2} < \frac{a_n}{b_n} < 2L$$
or equivalently that
$$\frac{L}{2} \cdot b_n < a_n < 2Lb_n$$
Now if $\sum\limits_{n=1}^{\infty} a_n$ converges, then the Comparison Test shows that
$$\sum_{n=1}^{\infty} \frac{L}{2} \cdot b_n$$
converges and hence so does
$$\sum_{n=1}^{\infty} b_n.$$
If $\sum\limits_{n=1}^{\infty} b_n$ converges, then so does
$$\sum_{n=1}^{\infty} 2L \cdot b_n.$$
But again we can use the Comparison Test to show that
$$\sum_{n=1}^{\infty} a_n$$
converges.

2) If $L = 0$, then we can find a cut off $N \in \mathbb{N}$ so that if $n \geq N$, then
$$0 < \frac{a_n}{b_n} < 1$$
or equivalently that
$$0 < a_n < b_n.$$
In this case, if $\sum\limits_{n=1}^{\infty} b_n$ converges, then $\sum\limits_{n=1}^{\infty} a_n$ converges as well by the Comparison Test. Equivalently, if $\sum\limits_{n=1}^{\infty} a_n$ diverges, then so does $\sum\limits_{n=1}^{\infty} b_n$.

3) If $L = \infty$, then we can find a cut off $N \in \mathbb{N}$ so that if $n \geq N$, then
$$\frac{a_n}{b_n} > 1$$
or equivalently that
$$b_n < a_n.$$
If $\sum\limits_{n=1}^{\infty} a_n$ converges, then $\sum\limits_{n=1}^{\infty} b_n$ converges by the Comparison Test. Equivalently, if $\sum\limits_{n=1}^{\infty} b_n$ diverges, then so does $\sum\limits_{n=1}^{\infty} a_n$.

■

Remarks:

We can informally summarize why the Limit Comparison Test works.

If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = L$ where $0 < L < \infty$, then for large n we have

$$\frac{a_n}{b_n} \cong L$$

or

$$a_n \cong Lb_n.$$

This suggests that $\sum\limits_{n=1}^{\infty} a_n$ converges if and only if $\sum\limits_{n=1}^{\infty} Lb_n$ converges. However, the properties of convergent series show that if $0 < L < \infty$, then $\sum\limits_{n=1}^{\infty} Lb_n$ converges if and only if $\sum\limits_{n=1}^{\infty} b_n$ converges. Combining these statements gives us that $\sum\limits_{n=1}^{\infty} a_n$ converges if and only if $\sum\limits_{n=1}^{\infty} b_n$ converges.

When $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = L$ where $0 < L < \infty$, we say that *a_n and b_n have the same order of magnitude*. We write

$$a_n \approx b_n.$$

The Limit Comparison Test says that **two positive series with terms of the same order of magnitude will have the same convergence properties.**

If $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = 0$, then b_n must eventually be much larger than a_n. In this case, we write $a_n \ll b_n$ and we say that *the order of magnitude of a_n is smaller than the order of magnitude of b_n*.

In this case, if the smaller series $\sum\limits_{n=1}^{\infty} a_n$ diverges to ∞, it would make sense that $\sum\limits_{n=1}^{\infty} b_n$ also diverges to ∞.

Finally, if $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \infty$, then a_n must eventually be much larger than b_n. That is $b_n \ll a_n$. This time, if the larger series $\sum\limits_{n=1}^{\infty} a_n$ converges, it would make sense that $\sum\limits_{n=1}^{\infty} b_n$ would converge as well. ◀

EXAMPLE 12 Let $a_n = \frac{2n}{n^3-n+1}$ and $b_n = \frac{1}{n^2}$. Then

$$\begin{aligned}\frac{a_n}{b_n} &= \frac{\frac{2n}{n^3-n+1}}{\frac{1}{n^2}} \\ &= \frac{2n^3}{n^3-n+1} \\ &= \frac{n^3}{n^3}\left(\frac{2}{1-\frac{1}{n^2}+\frac{1}{n^3}}\right) \\ &= \frac{2}{1-\frac{1}{n^2}+\frac{1}{n^3}}\end{aligned}$$

Therefore,

$$\lim_{n\to\infty}\frac{a_n}{b_n} = \lim_{n\to\infty}\frac{2}{1-\frac{1}{n^2}+\frac{1}{n^3}} = \frac{2}{1} = 2.$$

This confirms for large n that $a_n = \frac{2n}{n^3-n+1} \cong 2b_n = \frac{2}{n^2}$.

Since $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}$ converges, the first statement in the Limit Comparison Test shows that $\sum\limits_{n=1}^{\infty}\frac{2n}{n^3-n+1}$ converges as we expected. ◀

EXAMPLE 13 Show that $\sum\limits_{n=1}^{\infty}\sin\left(\frac{1}{n}\right)$ diverges.

It can be shown that for any $0 < x \leq 1$ that $0 < \sin(x) < x$ and hence that

$$0 < \sin\left(\frac{1}{n}\right) < \frac{1}{n}$$

for each $n \in \mathbb{N}$. However, since $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ diverges, we cannot use the Comparison Test directly to show that $\sum\limits_{n=1}^{\infty}\sin\left(\frac{1}{n}\right)$ diverges. (Why?)

Recall that the Fundamental Trigonometric Limit states that as $x \to 0$, $\frac{\sin(x)}{x} \to 1$. As $n \to \infty$, we have $\frac{1}{n} \to 0$. Therefore, by the Sequential Characterization of Limits

$$\lim_{n\to\infty}\frac{\sin\left(\frac{1}{n}\right)}{\frac{1}{n}} = 1.$$

If we let $a_n = \sin\left(\frac{1}{n}\right)$ and $b_n = \frac{1}{n}$, then we have just shown that

$$\lim_{n\to\infty}\frac{a_n}{b_n} = 1.$$

Since $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ diverges, the Limit Comparison Test shows that $\sum\limits_{n=1}^{\infty}\sin\left(\frac{1}{n}\right)$ diverges. ◀

5.6 Integral Test for Convergence of Series

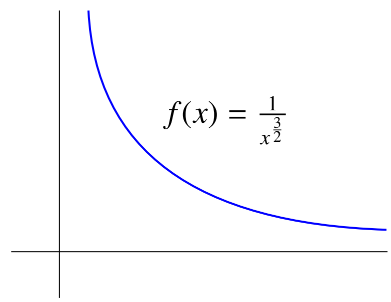
We have seen that the series $\sum_{n=1}^{\infty} \frac{1}{n}$ diverges while the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges. The Comparison Test can then be used to show that if $p \geq 2$ the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges, while if $p \leq 1$ the series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ diverges. We are not yet able to determine what happens when $1 < p < 2$ since the Comparison Test fails in this case since $\frac{1}{n^2} < \frac{1}{n^p} < \frac{1}{n}$ for $n > 1$, and the convergence of a series with smaller terms or the divergence of a series with larger terms does not help us determine whether a particular series converges or diverges.

It turns out that we can use improper integrals to establish the convergence or divergence of this remaining case. To see how we do this, we will consider the series $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$.

EXAMPLE 14 Show that $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges.

We begin by noting that the series $\sum_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges if and only if the series $\sum_{n=2}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges. Next we will consider the function $f(x) = \frac{1}{x^{\frac{3}{2}}}$. This function is continuous on $[1, \infty)$ and is decreasing on this interval. You can verify the last statement by noting that the derivative is $f'(x) = -\frac{3}{2}x^{\frac{-5}{2}}$, which is negative if $x > 0$.

The function appears as follows:

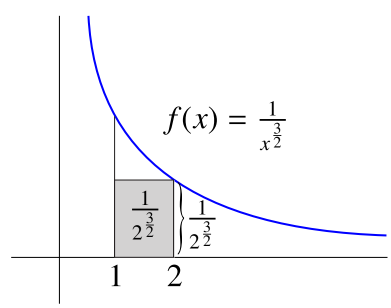


Observe that on the interval $[1, 2]$, f has a minimum value of $\frac{1}{2^{\frac{3}{2}}}$ at the right-endpoint $x = 2$. Therefore,

$$f(x) \geq \frac{1}{2^{\frac{3}{2}}}$$

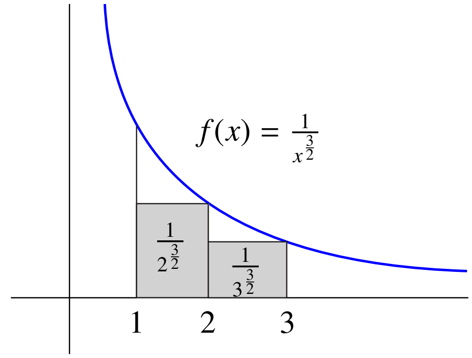
for all $x \in [1, 2]$.

From this fact and the following diagram, we see that the area of the rectangle with height $\frac{1}{2^{\frac{3}{2}}}$ and width 1 is less than the integral $\int_1^2 \frac{1}{x^{\frac{3}{2}}}\, dx$.



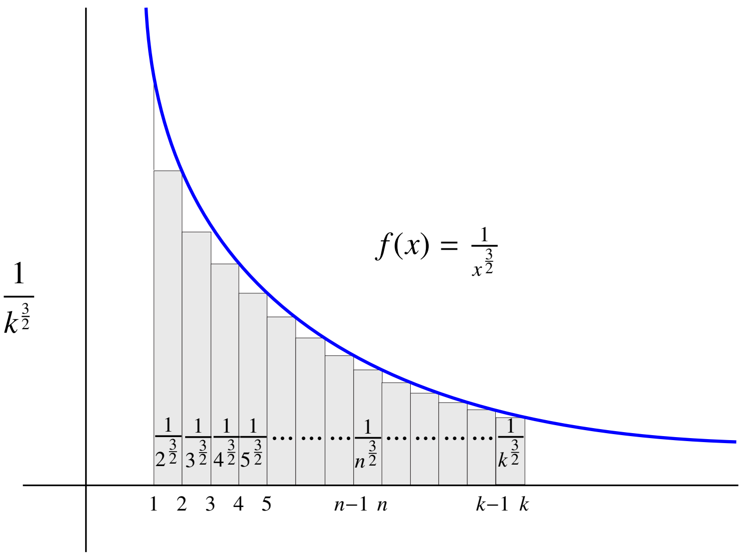
Similarly, we see that

$$\frac{1}{2^{\frac{3}{2}}} + \frac{1}{3^{\frac{3}{2}}} \le \int_1^3 \frac{1}{x^{\frac{3}{2}}}\,dx.$$



Continuing on we get that for any $k \in \mathbb{N}$, $k \ge 2$,

$$\begin{aligned} S_k &= \sum_{n=2}^{k} \frac{1}{n^{\frac{3}{2}}} \\ &= \frac{1}{2^{\frac{3}{2}}} + \frac{1}{3^{\frac{3}{2}}} + \cdots + \frac{1}{k^{\frac{3}{2}}} \\ &\le \int_1^k \frac{1}{x^{\frac{3}{2}}}\,dx \end{aligned}$$



We know that

$$\begin{aligned} \int_1^\infty \frac{1}{x^{\frac{3}{2}}}\,dx &= \lim_{b\to\infty} \int_1^b \frac{1}{x^{\frac{3}{2}}}\,dx \\ &= \lim_{b\to\infty} -2x^{-\frac{1}{2}}\Big|_1^b \\ &= \lim_{b\to\infty} \left[\frac{-2}{\sqrt{b}} + \frac{2}{\sqrt{1}}\right] \\ &= 2 \end{aligned}$$

This shows that for every k,

$$S_k = \sum_{n=2}^{k} \frac{1}{n^{\frac{3}{2}}} \le \int_1^k \frac{1}{x^{\frac{3}{2}}}\,dx < 2.$$

Therefore, $\{S_k\}$ is an increasing sequence that we have just seen is bounded. The Monotone Convergence Theorem tells us that it must converge. Therefore, $\sum\limits_{n=2}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges. Finally, this shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ also converges as we expected. ◄

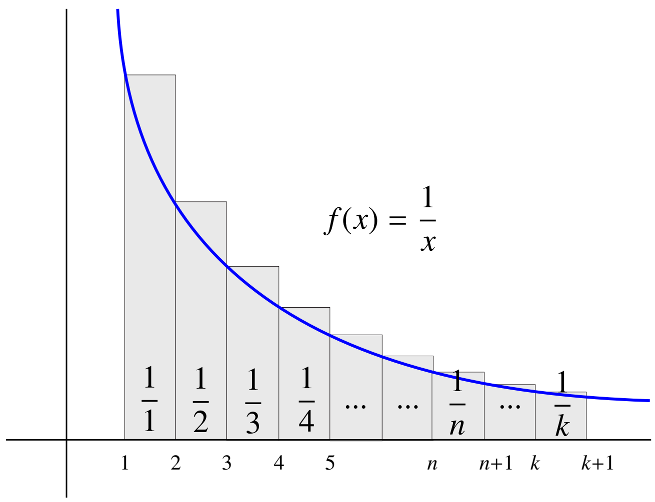
EXAMPLE 15 We can use a similar argument to provide us another way to show that $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges.

Let $f(x) = \frac{1}{x}$. Then we know that

$$\begin{aligned}
\int_1^{\infty} \frac{1}{x}\,dx &= \lim_{b\to\infty} \int_1^{b} \frac{1}{x}\,dx \\
&= \lim_{b\to\infty} \ln(x)\Big|_1^b \\
&= \lim_{b\to\infty} [\ln(b) - \ln(1)] \\
&= \infty
\end{aligned}$$

so that $\int_1^{\infty} \frac{1}{x}\,dx$ diverges.

Since $f(x) = \frac{1}{x}$ is decreasing, the maximum value for the function on an interval of the form $[n, n+1]$ occurs at the left-hand endpoint n with $f(n) = \frac{1}{n}$. It follows that $\int_n^{n+1} \frac{1}{x}\,dx$ is smaller than the area of the rectangle with height $f(n) = \frac{1}{n}$ and base of width 1 between n and $n+1$.



Moreover, as the diagram illustrates, for each $k \in \mathbb{N}$

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k} > \int_1^{k+1} \frac{1}{x}\,dx.$$

However, because $\int_1^{\infty} \frac{1}{x}\,dx$ diverges, $\int_1^{k+1} \frac{1}{x}\,dx \to \infty$ as $k \to \infty$. This means that

$$S_k = \sum_{n=1}^{k} \frac{1}{n} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{k}$$

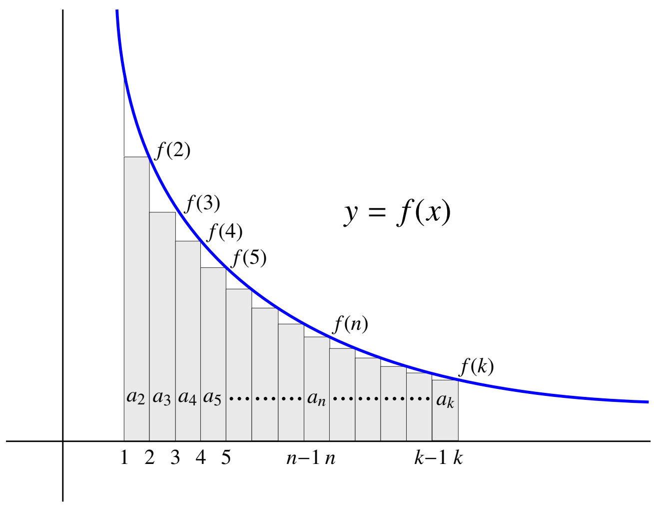
must also grow toward ∞ as k gets large. Therefore, $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges. ◀

Remark: It turns out that the process we have used in the last two examples provides us with powerful tools for studying the convergence and divergence of many important series. In general, we will assume that

1. f is continuous on $[1, \infty)$.
2. $f(x) > 0$ on $[1, \infty)$.
3. f is decreasing on $[1, \infty)$.

Let $a_n = f(n)$. Then just as was the case for $f(x) = \frac{1}{x^{\frac{3}{2}}}$, for any $n \in \mathbb{N}$, $n \geq 2$, the minimum value for the function $f(x)$ on the interval $[n-1, n]$ is at the right-hand endpoint and as such is $f(n) = a_n$. Again, as in the case of $f(x) = \frac{1}{x^{\frac{3}{2}}}$, for any $k \in \mathbb{N}$, $k \geq 2$, we have

$$\sum_{n=2}^{k} a_n = a_2 + a_3 + \cdots + a_k \leq \int_1^k f(x)\,dx.$$



Therefore, if $\int_1^\infty f(x)\,dx$ converges, then

$$\sum_{n=2}^{k} a_n \leq \int_1^\infty f(x)\,dx < \infty$$

for each k. Using the Monotone Convergence Theorem, this shows that if $\int_1^\infty f(x)\,dx$ converges, then so does $\sum\limits_{n=2}^{\infty} a_n$. Finally, we get that $\sum\limits_{n=1}^{\infty} a_n$ also converges.

As we have just seen $\sum\limits_{n=2}^{k} a_n = a_2 + a_3 + \cdots + a_k \leq \int_1^k f(x)\,dx$. From this observation we get that

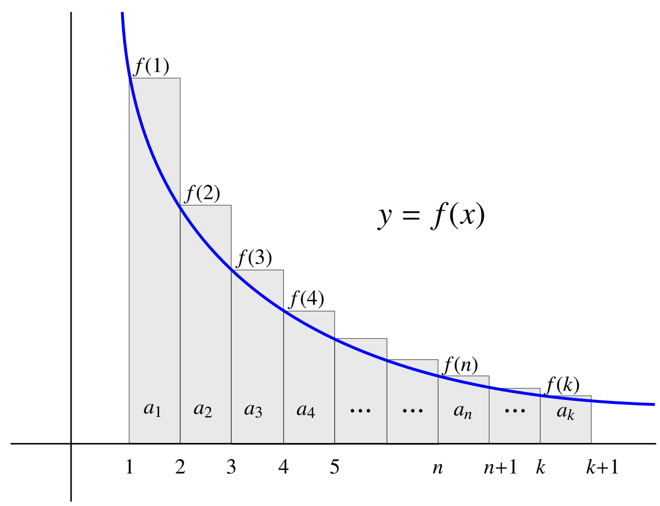
$$\begin{aligned} \sum_{n=1}^{k} a_n &= a_1 + \sum_{n=2}^{k} a_n \\ &\leq a_1 + \int_1^k f(x)\,dx. \end{aligned}$$

Allowing k to approach ∞, we see that if $\int_1^\infty f(x)\,dx < \infty$ converges, then

$$\sum_{n=1}^{\infty} a_n \le a_1 + \int_1^\infty f(x)\,dx.$$

Again, just as was the case for $f(x) = \frac{1}{x}$, for any $n \in \mathbb{N}$, the maximum value for the function $f(x)$ on the interval $[n, n+1]$ is at the left-hand endpoint and as such is $f(n) = a_n$. It follows that for any $k \in \mathbb{N}$, we have

$$\int_1^{k+1} f(x)\,dx \le a_1 + a_2 + a_3 + \cdots + a_k = \sum_{n=1}^{k} a_n.$$



This means that if $\int_1^\infty f(x)\,dx$ diverges to ∞, then $\sum\limits_{n=1}^{\infty} a_n$ must also diverge or equivalently that if $\sum\limits_{n=1}^{\infty} a_n$ converges, then so does $\int_1^\infty f(x)\,dx$.

Combining what we have done so far we get that $\sum\limits_{n=1}^{\infty} a_n$ converges if and only if $\int_1^\infty f(x)\,dx$ converges!

Note: We have just seen how improper integrals can help us analyze the growth rates of the partial sums of a series. In fact, so long as the series arises from a function f with the stated properties we have that for each $k \in \mathbb{N}$ that

$$\int_1^{k+1} f(x)\,dx \le \sum_{n=1}^{k} a_n \le a_1 + \int_1^{k} f(x)\,dx.$$

Assume now that $\sum\limits_{n=1}^{\infty} a_n$ converges to S. Then by allowing k to approach ∞ in the previous inequality we get

$$\int_1^\infty f(x)\,dx \le S \le \int_1^\infty f(x)\,dx + a_1.$$

Unfortunately, if a_1 is large, this estimate for S is rather crude. The good news is that we can do better!

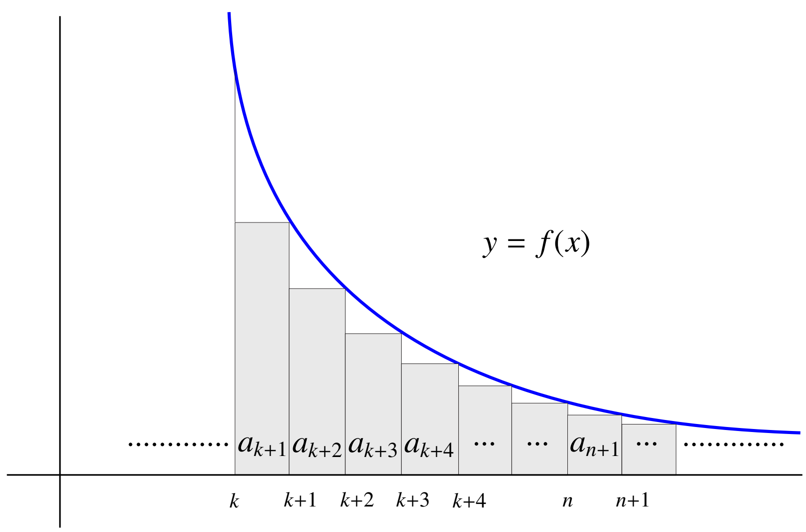
Observe that since the terms in the series are positive, we have

$$\begin{aligned} 0 &\le S - S_k \\ &= \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{k} a_n \\ &= \sum_{n=k+1}^{\infty} a_n \end{aligned}$$

This means that estimating how close the partial sum is to the final limit is equivalent to estimating how large is the sum of the **tail** of the series.

However, as the following diagram shows

$$0 \le S - S_k = \sum_{n=k+1}^{\infty} a_n \le \int_k^{\infty} f(x)\, dx.$$



The previous discussion leads us to the following important test for convergence.

THEOREM 8 **Integral Test for Convergence**

Assume that

1. f is continuous on $[1, \infty)$,
2. $f(x) > 0$ on $[1, \infty)$,
3. f is decreasing on $[1, \infty)$, and
4. $a_k = f(k)$.

For each $n \in \mathbb{N}$, let $S_n = \sum\limits_{k=1}^{n} a_k$. Then

i) For all $n \in \mathbb{N}$,
$$\int_1^{n+1} f(x)\,dx \leq S_n \leq a_1 + \int_1^n f(x)\,dx.$$

ii) $\sum\limits_{k=1}^{\infty} a_k$ converges if and only if $\int_1^\infty f(x)\,dx$ converges.

iii) In the case that $\sum\limits_{k=1}^{\infty} a_k$ converges, then
$$\int_1^\infty f(x)\,dx \leq \sum_{k=1}^{\infty} a_k \leq a_1 + \int_1^\infty f(x)\,dx$$
and
$$\int_{n+1}^\infty f(x)\,dx \leq S - S_n \leq \int_n^\infty f(x)\,dx,$$
where $S = \sum\limits_{k=1}^{\infty} a_k$. (Note that by (ii), $\int_n^\infty f(x)\,dx$ exists.)

Note: In the case where the series and the improper integral converge, they do **not** have to converge to the same value. Furthermore, all of the conditions are important, particularly that f is eventually decreasing (or at least non-increasing). Otherwise, the series and the improper integral could converge or diverge independent of one another.

The conditions of the Integral Test **do not have to hold on all of** $[1, \infty)$ for this analysis to be useful. What is really important is that **they hold from some point onward**. In fact, if these three conditions hold on the interval $[m, \infty)$ for some positive integer m, then we can conclude that
$$\sum_{n=m}^{\infty} a_n \text{ converges if and only if } \int_m^\infty f(x)\,dx \text{ converges.}$$

We have essentially used the Integral Test to show that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges and that $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges. Since $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges, the Comparison Test shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ converges for any $p \geq \frac{3}{2}$ and we know from before that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ diverges if $p \leq 1$. We still don't know what happens if $1 < p < \frac{3}{2}$. The Integral Test can help us fill in this gap.

THEOREM 9 ***p*-Series Test**

The series $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$.

PROOF

We already know that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ converges if $p \geq 2$ and it diverges for $p \leq 1$. We can use the Integral Test to address the missing interval.

Consider, the function $f(x) = \frac{1}{x^p}$. It is easy to see that if $p > 0$, then $f(x)$ satisfies the three hypotheses of the Integral Test. If we let $a_n = f(n) = \frac{1}{n^p}$, then for any $p > 0$, the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ will converge if and only if the improper integral $\int_1^{\infty} \frac{1}{x^p}\, dx$ converges. However, we know that

$$\int_1^{\infty} \frac{1}{x^p}\, dx$$

converges if and only if $p > 1$. Since we know that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ diverges if $p \leq 0$, this tells us that $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ will converge if and only if $p > 1$. ■

EXAMPLE 16 Show that

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

converges.

We will use the Limit Comparison Test and the p-Series Test.

A close look at the terms of the series shows that for large n,

$$\frac{1}{n^{\frac{3}{2}} - n + 1} \cong \frac{1}{n^{\frac{3}{2}}}.$$

In fact

$$\begin{aligned} \lim_{n\to\infty} \frac{\frac{1}{n^{\frac{3}{2}}}}{\frac{1}{n^{\frac{3}{2}}-n+1}} &= \lim_{n\to\infty} \frac{n^{\frac{3}{2}} - n + 1}{n^{\frac{3}{2}}} \\ &= \lim_{n\to\infty} (1 - \frac{1}{n^{\frac{1}{2}}} + \frac{1}{n^{\frac{3}{2}}}) \\ &= 1 \end{aligned}$$

Therefore, the Limit Comparison Test shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}-n+1}$ converges if and only if $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges. However, the p-Series test shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}}$ converges. We can conclude that $\sum\limits_{n=1}^{\infty} \frac{1}{n^{\frac{3}{2}}-n+1}$ converges as well. ◀

EXAMPLE 17 Determine whether the series $\sum\limits_{n=2}^{\infty} \frac{1}{n\ln(n)}$ converges.

Consider the function $f(x) = \frac{1}{x\ln(x)}$. Then

1. $f(x)$ is continuous on $[2, \infty)$.
2. $f(x) > 0$ on $[2, \infty)$.
3. $f(x)$ is decreasing on $[2, \infty)$.

While all of these claims are easy to verify, we will explicitly show that condition 3 holds. To do this, note that from the quotient rule it follows that

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

if $x \geq 2$.

This shows that $f'(x) < 0$ for every $x \in [2, \infty)$. Hence, $f(x)$ is decreasing on $[2, \infty)$. Alternatively, we could have observed that the function $x\ln(x)$ is increasing on $[2, \infty)$ and as such its reciprocal $\frac{1}{x\ln(x)}$ must be decreasing.

We can apply the Integral Test to see that $\sum\limits_{n=2}^{\infty} \frac{1}{n\ln(n)}$ converges if and only if $\int_2^{\infty} \frac{1}{x\ln(x)}\,dx$ converges.

Now

$$\int_2^{\infty} \frac{1}{x\ln(x)}\,dx = \lim_{b\to\infty} \int_2^{b} \frac{1}{x\ln(x)}\,dx.$$

To evaluate $\int_2^{b} \frac{1}{x\ln(x)}\,dx$ use the substitution $u = \ln(x)$, $du = \frac{dx}{x}$ to get

$$\begin{aligned} \int_2^{b} \frac{1}{x\ln(x)}\,dx &= \int_{\ln(2)}^{\ln(b)} \frac{1}{u}\,du \\ &= \ln(u)\Big|_{\ln(2)}^{\ln(b)} \\ &= \ln(\ln(b)) - \ln(\ln(2)) \end{aligned}$$

Therefore,

$$\begin{aligned}\int_2^\infty \frac{1}{x\ln(x)}\,dx &= \lim_{b\to\infty}\int_2^b \frac{1}{x\ln(x)}\,dx \\ &= \lim_{b\to\infty} \ln(\ln(b)) - \ln(\ln(2)) \\ &= \infty\end{aligned}$$

Since $\int_2^\infty \frac{1}{x\ln(x)}\,dx$ diverges, the Integral Test shows that $\sum\limits_{n=2}^{\infty}\frac{1}{n\ln(n)}$ also diverges. ◀

EXAMPLE 18 Show that $\sum\limits_{n=2}^{\infty}\frac{1}{n(\ln(n))^2}$ converges.

Let $f(x) = \frac{1}{x(\ln(x))^2}$. It is easy to verify that

1. f is continuous on $[2,\infty)$.
2. $f(x) > 0$ on $[2,\infty)$.

Moreover, since $x(\ln(x))^2$ is increasing for $x \geq 2$, it follows that $f(x) = \frac{1}{x(\ln(x))^2}$ is decreasing. Alternatively, $f'(x) = -\frac{2+\ln(x)}{x^2(\ln(x))^3}$ which is negative for $x \geq 2$. Therefore, the Integral Test can be used to conclude that $\sum\limits_{n=2}^{\infty}\frac{1}{n(\ln(n))^2}$ converges if and only if $\int_2^\infty \frac{1}{x(\ln(x))^2}\,dx$ converges.

To evaluate $\int_2^b \frac{1}{x(\ln(x))^2}\,dx$, use the substitution $u = \ln(x)$, $du = \frac{dx}{x}$ to get

$$\begin{aligned}\int_2^b \frac{1}{x(\ln(x))^2}\,dx &= \int_{\ln(2)}^{\ln(b)} \frac{1}{u^2}\,du \\ &= -\frac{1}{u}\Big|_{\ln(2)}^{\ln(b)} \\ &= \frac{1}{\ln(2)} - \frac{1}{\ln(b)}\end{aligned}$$

Therefore,

$$\begin{aligned}\int_2^\infty \frac{1}{x(\ln(x))^2}\,dx &= \lim_{b\to\infty}\int_2^b \frac{1}{x(\ln(x))^2}\,dx \\ &= \lim_{b\to\infty} \frac{1}{\ln(2)} - \frac{1}{\ln(b)} \\ &= \frac{1}{\ln(2)}\end{aligned}$$

Since $\int_2^\infty \frac{1}{x(\ln(x))^2}\,dx$ converges, so does $\sum\limits_{n=2}^{\infty} \frac{1}{n(\ln(n))^2}$. ◀

5.6.1 Integral Test and Estimation of Sums and Errors

The Integral Test is a powerful tool for determining the convergence or divergence of many important series. However, much more can be said. In fact, we have seen that if

1. f is continuous on $[1,\infty)$,
2. $f(x) > 0$ on $[1,\infty)$,
3. $f(x)$ is decreasing on $[1,\infty)$,
4. $a_n = f(n)$, and
5. $S_k = \sum\limits_{n=1}^{k} a_n$,

then

$$\int_1^{k+1} f(x)\,dx \le S_k \le \int_1^k f(x)\,dx + a_1.$$

Therefore, we can use integration to estimate the value of the partial sum S_k of the series $\sum\limits_{n=1}^{\infty} a_n$.

EXAMPLE 19 How large is the k-th partial sum S_k of the harmonic series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$?

In this case, $f(x) = \frac{1}{x}$ and $a_1 = 1$. But we know that

$$\int_1^{k+1} \frac{1}{x}\,dx \le S_k = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{k} \le \int_1^k \frac{1}{x}\,dx + a_1$$

Since

$$\int_1^c \frac{1}{x}\,dx = \ln(x)\Big|_1^c = \ln(c) - \ln(1) = \ln(c)$$

for any value of $c > 0$, this inequality becomes

$$\ln(k+1) \le \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{k} \le \ln(k) + 1$$

If we let $k = 1000$, we see that $\ln(1001) = 6.908754779$ and $\ln(1000) = 6.907755279$. This tells us that

$$6.908754779 \le \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{1000} \le 7.907755279$$

If $k = 10^9$, then the inequality would show us that

$$20.72326584 \le \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{10^9} \le 21.72326584$$

What is remarkable about these estimates is not their accuracy in estimating a particular partial sum, but that while we know the harmonic series will diverge to ∞, after 1000 terms the sum has not yet exceeded 8, and after 1 billion terms, the sum is still less than 22! This shows us that we would not have been able to guess that this series diverged by testing some partial sums on a computer.

To emphasize home this last comment consider how many terms one would have to add together so that

$$S_k = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{k} > 1000.$$

We know that

$$S_k \cong \ln(k).$$

As such to have $S_k > 1000$ we would need enough terms so that

$$1000 < \ln(k).$$

This means that we should have

$$k > e^{1000} \cong 10^{434}.$$

This is an enormous number. ◀

EXAMPLE 20 The p-Series Test shows that the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^4}$ converges. Let

$$S_k = \sum_{n=1}^{k} \frac{1}{n^4} \quad \text{and} \quad S = \sum_{n=1}^{\infty} \frac{1}{n^4}.$$

Estimate the error in using the first 100 terms in the series to approximate S. That is, estimate $|S - S_{100}|$.

The first observation we make is that since all the terms are positive we have that $S - S_{100} > 0$ and hence that

$$|S - S_{100}| = S - S_{100}.$$

But we also know from the Integral Test that

$$\int_{101}^{\infty} \frac{1}{x^4}\,dx \le S - S_{100} \le \int_{100}^{\infty} \frac{1}{x^4}\,dx.$$

But for any $m \in \mathbb{N}$, we have that

$$\begin{aligned}
\int_m^\infty \frac{1}{x^4}\,dx &= \lim_{b\to\infty} \int_m^b \frac{1}{x^4}\,dx \\
&= \lim_{b\to\infty} -\frac{1}{3x^3}\Big|_m^b \\
&= \lim_{b\to\infty} \frac{1}{3m^3} - \frac{1}{3b^3} \\
&= \frac{1}{3m^3}
\end{aligned}$$

Substituting $m = 101$ and $m = 100$ respectively in $\frac{1}{3m^3}$ we get that

$$\frac{1}{3(101)^3} \leq S - S_{100} \leq \frac{1}{3(100)^3}$$

or

$$3.2353 \times 10^{-7} \leq S - S_{100} \leq 3.3333 \times 10^{-7}.$$

Now if we calculate S_{100} we get $S_{100} = 1.082322905$ (up to 9 decimal places) and hence our prediction would be that

$$1.082323229 \leq S \leq 1.082323238$$

In fact, it is actually known that

$$S = \sum_{n=1}^\infty \frac{1}{n^4} = \frac{\pi^4}{90} = 1.082323234$$

which does indeed lie within our range. ◀

EXAMPLE 21 The Integral Test tells us that the series $\sum\limits_{n=2}^\infty \frac{1}{n(\ln(n))^2}$ converges. But it can also show us that the series converges very slowly. For example, suppose that $S = \sum\limits_{n=2}^\infty \frac{1}{n(\ln(n))^2}$ and $S_k = \sum\limits_{n=2}^k \frac{1}{n(\ln(n))^2}$. Then we know that

$$S - S_k \cong \int_k^\infty \frac{1}{x(\ln(x))^2}\,dx.$$

But

$$\begin{aligned}
\int_k^b \frac{1}{x(\ln(x))^2}\,dx &= \int_{\ln(k)}^{\ln(b)} \frac{1}{u^2}\,du \\
&= -\frac{1}{u}\Big|_{\ln(k)}^{ln(b)}
\end{aligned}$$

$$= \frac{1}{\ln(k)} - \frac{1}{\ln(b)}$$

Therefore,

$$\begin{aligned}
\int_k^\infty \frac{1}{x(\ln(x))^2}\,dx &= \lim_{b\to\infty} \int_k^b \frac{1}{x(\ln(x))^2}\,dx \\
&= \lim_{b\to\infty} \frac{1}{\ln(k)} - \frac{1}{\ln(b)} \\
&= \frac{1}{\ln(k)}
\end{aligned}$$

This means if we want

$$S - S_k < \frac{1}{100}$$

we would choose k so that

$$\frac{1}{\ln(k)} < \frac{1}{100}$$

or equivalently that

$$100 < \ln(k) \iff k > e^{100}.$$

Therefore, to have the partial sum S_k approximate the final sum to within only $\frac{1}{100}$ we would need roughly $e^{100} \cong 10^{43}$ terms. ◀

5.7 Alternating Series

We have seen that for a series $\sum\limits_{n=1}^{\infty} a_n$ with positive terms (in other words, $a_n \geq 0$ for all n) that $\sum\limits_{n=1}^{\infty} a_n$ will either converge if the terms are *small enough* or it will diverge to ∞.

Without the assumption that $a_n \geq 0$ for all n, the situation can become much more complicated. In this section, we will look at one more class of series whose behavior is particularly nice.

DEFINITION

Alternating Series

A series of the form

$$\sum_{n=1}^{\infty} (-1)^{n-1} a_n = a_1 - a_2 + a_3 - a_4 + \cdots$$

or of the form

$$\sum_{n=1}^{\infty} (-1)^{n} a_n = -a_1 + a_2 - a_3 + a_4 - \cdots$$

is said to be *alternating* provided that $a_n > 0$ for all n.

EXAMPLE 22 The most important example of an alternating series is

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

Problem: Does the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ converge?

For positive series, we saw that two series with terms of the same order of magnitude would either both converge or both diverge. If $a_n = (-1)^{n-1}\frac{1}{n}$ and $b_n = \frac{1}{n}$, then

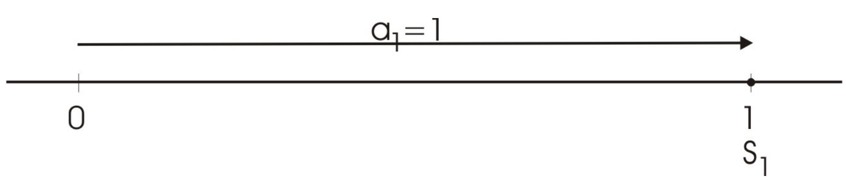
$$| a_n |= \frac{1}{n} =| b_n |$$

so the terms a_n and b_n are of the same order of magnitude. Our *rule of thumb* would suggest that since $\sum\limits_{n=1}^{\infty}\frac{1}{n}$ diverges, we might expect that $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ would also diverge. However, $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ **is not a positive series**.

Let $S_j = \sum\limits_{n=1}^{j}(-1)^{n-1}\frac{1}{n}$ be the j-th partial sum of the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$. Then

$$S_1 = 1.$$

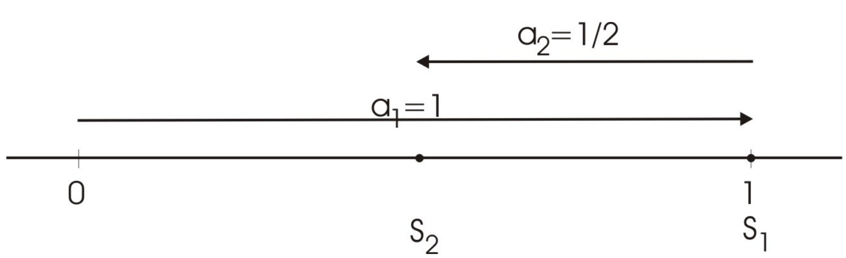
We can represent this graphically by beginning at 0 and then moving 1 unit to the *right* to reach $S_1 = 1$.



Next we have that

$$S_2 = 1 - \frac{1}{2} = S_1 - \frac{1}{2} = \frac{1}{2}.$$

Again, this can be represented graphically. This time we begin at $S_1 = 1$ and then move $\frac{1}{2}$ units to the *left* to reach $S_2 = \frac{1}{2}$.



Notice that since we have moved to the left, we have $S_2 < S_1$, but since $\frac{1}{2}$ was smaller than 1, we did not get back to 0. This means that

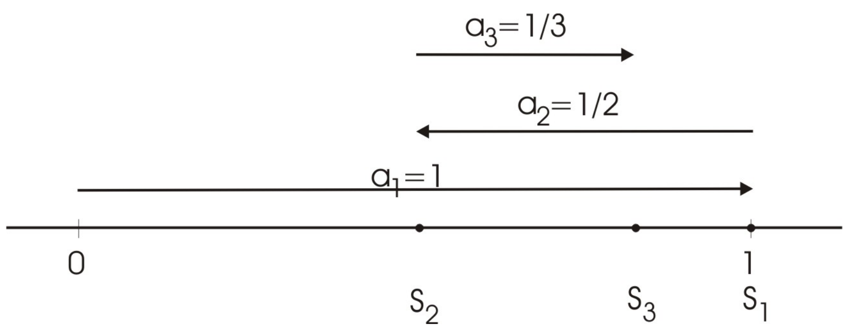
$$0 < S_2 < S_1.$$

In the third step, we get

$$S_3 = 1 - \frac{1}{2} + \frac{1}{3} = S_2 + \frac{1}{3} = \frac{5}{6}.$$

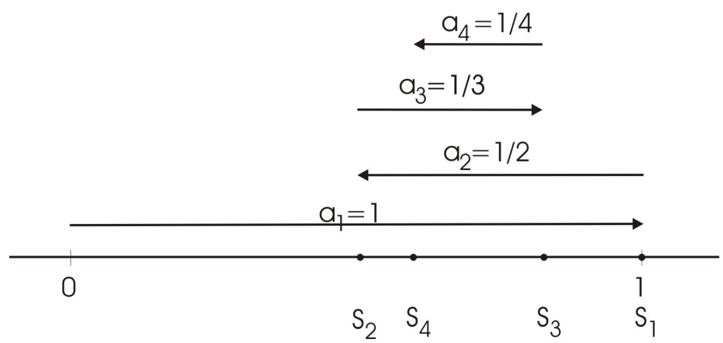
To reach S_3 we move to the *right* a total of $\frac{1}{3}$ units. It is also very important to note that because $\frac{1}{3} < \frac{1}{2}$, we do not get all the way back to S_1. That is,

$$0 < S_2 < S_3 < S_1.$$



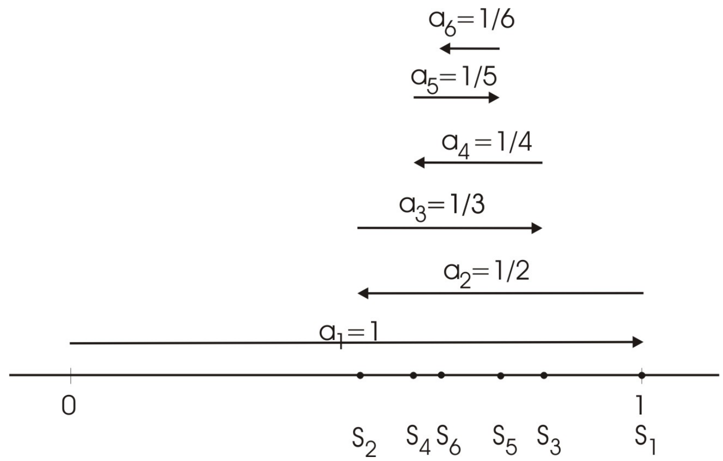
The fourth step will take us to the *left* a total of $\frac{1}{4}$ units. Since our previous move to the right was $\frac{1}{3}$ units and clearly $\frac{1}{4} < \frac{1}{3}$, we now have

$$0 < S_2 < S_4 < S_3 < S_1.$$

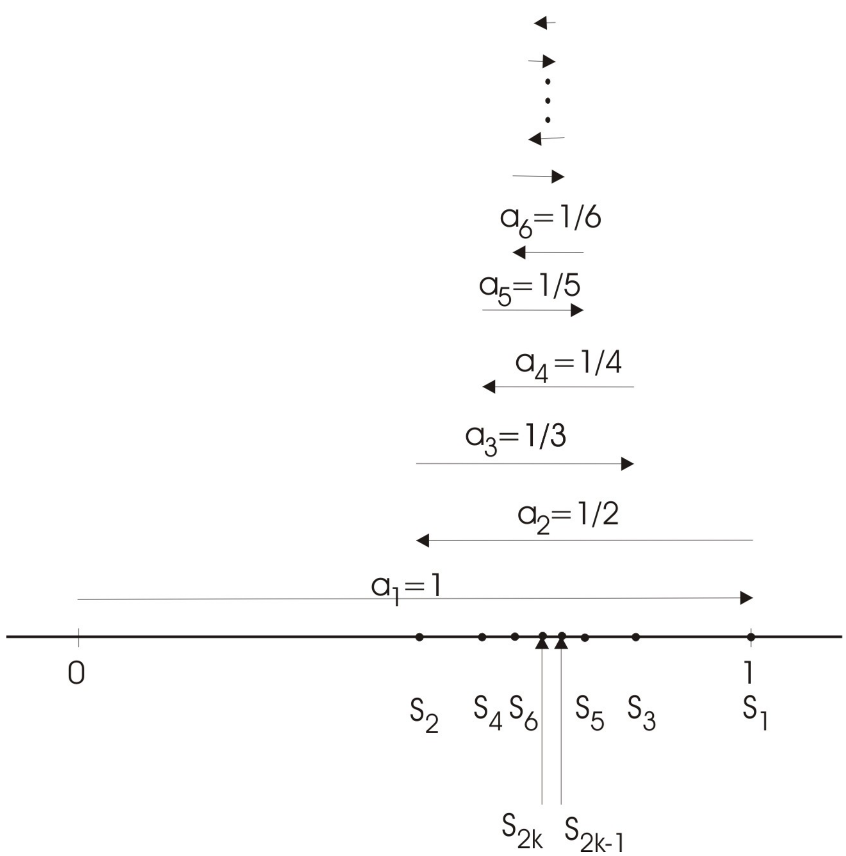


After two more steps, we see a clear pattern emerging with

$$0 < S_2 < S_4 < S_6 < S_5 < S_3 < S_1.$$



The terms with even indices are getting larger, while the terms with odd indices are decreasing. In fact, if we continue on we will see a picture that looks as follows:



If we denote the odd indexed terms by S_{2k-1} for $k = 1, 2, 3, \cdots$ and the even indexed terms by S_{2k} for $k = 1, 2, 3, \cdots$, then after $2k$ steps we have

$$0 < S_2 < S_4 < S_6 < \cdots < S_{2k} < S_{2k-1} < \cdots < S_5 < S_3 < S_1.$$

Eventually, we will have two sequences consisting of the odd partial sums $\{S_{2k-1}\}$ with

$$S_1 > S_3 > S_5 > \cdots > S_{2k-1} > S_{2k+1} > \cdots > 0$$

and the even partial sums $\{S_{2k}\}$ with

$$S_2 < S_4 < S_6 < \cdots < S_{2k} < S_{2k+2} < \cdots < 1.$$

Both of the sequences are monotonic and bounded. The Monotone Convergence Theorem shows that they both converge. Let

$$\lim_{k\to\infty} S_{2k-1} = M$$

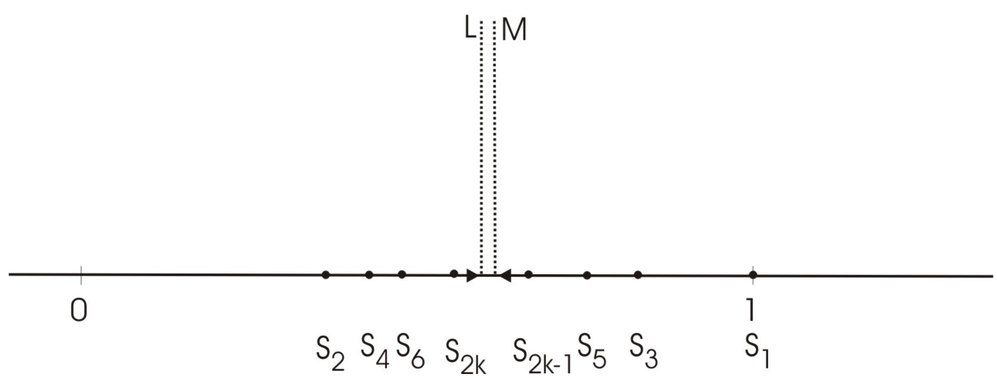
and

$$\lim_{k\to\infty} S_{2k} = L.$$

Since the elements of $\{S_{2k-1}\}$ decrease to M and the elements of $\{S_{2k}\}$ increase to L, we can show that

$$L \leq M$$

.



Moreover, since the odd terms and the even terms combine to give the entire sequence $\{S_j\}$ of partial sums, to show that $\{S_j\}$ converges we need only show that $M = L$. The key observation is that

$$S_{2k} < L \leq M < S_{2k-1}$$

for every k. But to get to S_{2k} from S_{2k-1}, we subtract $\frac{1}{2k}$. This is equivalent to stating that

$$S_{2k-1} - S_{2k} = \frac{1}{2k}.$$

Moreover, the distance between M and L is less than the distance from S_{k-1} to S_{2k}.

Putting this all together gives us

$$0 \leq M - L \leq \frac{1}{2k}$$

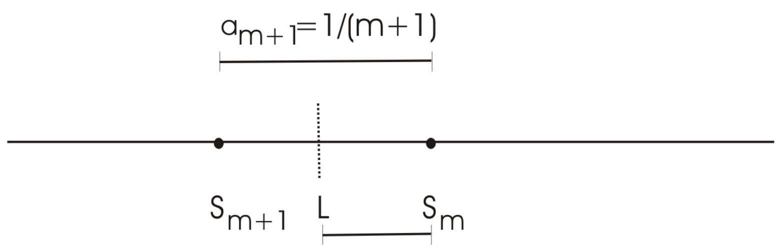
for every $k \in \mathbb{N}$. Since $\frac{1}{2k}$ can be made as small as we would like, the last statement can only be true if $L = M$.

We have just succeeded in showing that the sequence $\{S_j\}$ of partial sums of the series $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ converges. This means that $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ also converges.

There is one more observation that we can make. The process above shows that any two consecutive partial sums S_m and S_{m+1} will always sit on opposite sides of the final sum. If we denote $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ by L, then this means that the distance from S_m to L is less than the distance from S_m to S_{m+1}.

However, to get to S_{m+1} from S_m, we either add or subtract $\frac{1}{m+1}$ units depending on whether m is odd or even. Either way this tells us that the distance from S_m to S_{m+1} is exactly $\frac{1}{m+1}$. Therefore, we get that for any m,

$$| S_m - L | \leq | S_m - S_{m+1} | = \frac{1}{m+1}.$$



That is, the partial sum S_m approximates the sum of the series $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ with an error of less than $\frac{1}{m+1}$. ◀

Important Observation: A careful observation of the analysis of the series $\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ will show that we used the following properties of the sequence $a_n = \frac{1}{n}$ to show that the series converges:

1. $a_n > 0$ for all n.
2. $a_{n+1} \leq a_n$ for all n.
3. $\lim_{n\to\infty} a_n = 0$.

In fact our analysis is valid for any alternating series with these properties. We can summarize this in the following theorem:

THEOREM 10 **Alternating Series Test (AST)**

Assume that

1. $a_n > 0$ for all n.
2. $a_{n+1} \leq a_n$ for all n.
3. $\lim\limits_{n\to\infty} a_n = 0$.

Then the alternating series

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

converges.

If $S_k = \sum\limits_{n=1}^{k} (-1)^{n-1} a_n$, then S_k approximates the sum $S = \sum\limits_{n=1}^{\infty} (-1)^{n-1} a_n$ with an error that is at most a_{k+1}. That is

$$| S_k - S | \leq a_{k+1}.$$

PROOF

We will show that the two subsequences of partial sums $\{S_{2k-1}\}$ and $\{S_{2k}\}$ converge to the same limit L.

We first prove that both subsequences are monotonic. We have

$$\begin{aligned} S_{2(k+1)-1} - S_{2k-1} &= S_{2k+1} - S_{2k-1} \\ &= \sum_{n=1}^{2k+1} (-1)^{n-1} a_n - \sum_{n=1}^{2k-1} (-1)^{n-1} a_n \\ &= (-1)^{2k-1} a_{2k} + (-1)^{(2k+1)-1} a_{2k+1} \\ &= -a_{2k} + a_{2k+1} \\ &\leq 0 \end{aligned}$$

This shows that $\{S_{2k-1}\}$ is decreasing. Similarly,

$$\begin{aligned} S_{2(k+1)} - S_{2k} &= S_{2k+2} - S_{2k} \\ &= \sum_{n=1}^{2k+2} (-1)^{n-1} a_n - \sum_{n=1}^{2k} (-1)^{n-1} a_n \\ &= (-1)^{(2k+1)-1} a_{2k+1} + (-1)^{(2k+2)-1} a_{2k+2} \\ &= a_{2k+1} - a_{2k+2} \\ &\geq 0 \end{aligned}$$

This shows that $\{S_{2k}\}$ is increasing.

Now we will show that both subsequences are bounded. We have

$$\begin{aligned} S_{2k-1} &= (a_1 - a_2) + (a_3 - a_4) + \cdots + (a_{2k-3} - a_{2k-2}) + a_{2k-1} \\ &\geq 0 + 0 + \cdots + 0 + a_{2k-1} \\ &\geq 0 \end{aligned}$$

and

$$\begin{aligned} S_{2k} &= a_1 - (a_2 - a_3) - (a_4 - a_5) - \cdots - (a_{2k-2} - a_{2k-1}) - a_{2k} \\ &\leq a_1 - 0 - 0 - \cdots - 0 - a_{2k} \\ &\leq a_1 \end{aligned}$$

Hence $\{S_{2k-1}\}$ is bounded below by 0 and $\{S_{2k}\}$ is bounded above by a_1. By the Monotone Convergence Theorem, $\lim\limits_{k\to\infty} S_{2k-1} = L \in \mathbb{R}$ and $\lim\limits_{k\to\infty} S_{2k} = M \in \mathbb{R}$.

Next we show that $L = M$. To see why this is the case we note that

$$|S_{2k} - S_{2k-1}| = \left| \sum_{n=1}^{2k} (-1)^{n-1} a_n - \sum_{n=1}^{2k-1} (-1)^{n-1} a_n \right| = a_{2k}.$$

Then

$$\begin{aligned} |M - L| &= \lim_{k\to\infty} |S_{2k} - S_{2k-1}| \\ &= \lim_{k\to\infty} a_{2k} \\ &= 0 \end{aligned}$$

so $L = M$ and the series converges. Next we let

$$S = \sum_{n=1}^{\infty} (-1)^{n-1} a_n.$$

Finally, since $\{S_{2k}\}$ increases to S and $\{S_{2k-1}\}$ decreases to S, we have

$$S_{2k} \leq S \leq S_{2k-1}$$

for all $k \in \mathbb{N}$. This shows that S sits between S_k and S_{k+1} for each $k \in \mathbb{N}$.

Thus we get that

$$|S_k - S| \leq |S_k - S_{k+1}| = |(-1)^k a_{k+1}| = a_{k+1}$$

for all $k \in \mathbb{N}$. This proves the last part of the theorem. ■

Remark: There are a few important observations we should make concerning this theorem.

1. Historically, the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ is the most important of the alternating series. For this reason it is usually called *The Alternating Series*.

2. All three of the conditions in the statement of the theorem are important for the theorem to be valid. However, it is actually sufficient for the first two to hold for all $n \geq M$ where M is some fixed integer. In this case, the error estimate will only be valid when $k \geq M$.

3. For the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}a_n$, we have

$$\begin{aligned} 0 \quad &\leq \quad S_2 \leq S_4 \leq S_6 \leq \cdots \leq S_{2k} \leq \cdots \sum_{n=1}^{\infty}(-1)^{n-1}a_n \\ \cdots \quad &\leq S_{2k-1} \leq \cdots \leq S_5 \leq S_3 \leq S_1 = a_1. \end{aligned}$$

   Therefore, if j is even, S_j under estimates the sum and if j is odd, S_j over estimates the sum.

4. With the obvious changes, the theorem remains valid for series of the form

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

   provided that the three assumptions on the sequence $\{a_n\}$ hold.

EXAMPLE 23 Show that the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3}$ converges and determine how large k must be so that

$$\mid S_k - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3} \mid < 10^{-6}.$$

Let $a_n = \frac{1}{n^3}$. Then

$$0 < \frac{1}{(n+1)^3} < \frac{1}{n^3}$$

and $\lim\limits_{n\to\infty}\frac{1}{n^3} = 0$. Therefore, the series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3}$ converges by the Alternating Series Test.

The Alternating Series Test tells us that

$$\mid S_k - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3} \mid \leq a_{k+1} = \frac{1}{(k+1)^3}.$$

We must choose k large enough so that

$$\frac{1}{(k+1)^3} < 10^{-6}.$$

After cross-multiplying this inequality is equivalent to

$$10^6 < (k+1)^3.$$

Taking cube roots of both sides gives

$$10^2 < k+1$$

and hence that

$$99 < k.$$

Therefore, if $k \geq 100$, then

$$\mid S_k - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3} \mid < 10^{-6}.$$

It is not surprising that this series converges since the terms become quite small very quickly. In fact, the series $\sum\limits_{n=1}^{\infty} \frac{1}{n^3}$ also converges.

We can use the Integral Test to see how many terms are needed so that

$$\mid T_k - \sum_{n=1}^{\infty}\frac{1}{n^3} \mid < 10^{-6}$$

where $T_k = \sum\limits_{n=1}^{k} \frac{1}{n^3}$. The Integral Test tells us that

$$\mid T_k - \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3} \mid < \int_k^{\infty} \frac{1}{x^3}\,dx.$$

Now

$$\begin{aligned}
\int_k^{\infty} \frac{1}{x^3}\,dx &= \lim_{b\to\infty}\int_k^{b} \frac{1}{x^3}\,dx \\
&= \lim_{b\to\infty} \frac{-1}{2x^2}\bigg|_k^b \\
&= \lim_{b\to\infty}\left(\frac{-1}{2b^2} + \frac{1}{2k^2}\right) \\
&= \frac{1}{2k^2}
\end{aligned}$$

We would like to choose k large enough so that

$$\frac{1}{2k^2} < 10^{-6}$$

or equivalently so that

$$\frac{10^6}{2} < k^2.$$

Taking square roots of both sides of this inequality shows us that we require

$$\frac{1000}{\sqrt{2}} < k$$

and since $\frac{1000}{\sqrt{2}} = 707.1068$, we get that k must be at least 708 to ensure that the error in approximating $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^3}$ by T_k is no more than 10^{-6}. ◀

EXAMPLE 24 The previous example illustrates the fact that alternating series converge much more quickly than positive series with terms of equal magnitude. An even more extreme example of this phenomenon can be seen by comparing the number of terms it would take for the partial sums of the alternating series

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

versus the partial sums of the positive series

$$\sum_{n=2}^{\infty}\frac{1}{n(\ln(n))^2}$$

to be within 10^{-2} of the corresponding sum.

In the case of the series

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

we can use the Alternating Series Test to show not only that the series converges, but also that

$$\left|\sum_{n=2}^{k}(-1)^{n-1}\frac{1}{n(\ln(n))^2} - \sum_{n=2}^{\infty}(-1)^{n-1}\frac{1}{n(\ln(n))^2}\right| \leq \frac{1}{(k+1)(\ln(k+1))^2}.$$

If we sum up to $k = 14$, then since

$$\frac{1}{(k+1)(\ln(k+1))^2} = \frac{1}{(15)(\ln(15))^2} = 0.009091 < \frac{1}{100}$$

we will be within 10^{-2} of the final sum.

We can show that the series

$$\sum_{n=2}^{\infty}\frac{1}{n(\ln(n))^2}$$

converges using the Integral Test. Moreover, we have already seen that the Integral Test shows us that to approximate the final sum within a tolerance of 10^{-2} we would need to use approximately e^{100} terms. This number is larger than 10^{43} which as we mentioned before is unimaginably big! In particular, this example shows us that we could not find a reasonable approximation to this latter sum by simply asking a computer to add the terms one by one. ◀

5.8 Absolute versus Conditional Convergence

Recall that the Harmonic Series $\sum\limits_{n=1}^{\infty} \frac{1}{n}$ diverges, while the Alternating Series $\sum\limits_{n=1}^{\infty} (-1)^{n-1}\frac{1}{n}$ converges even though the terms have the same order of magnitude. In fact,

$$\left|\frac{1}{n}\right| = \frac{1}{n} = \left|(-1)^{n-1}\frac{1}{n}\right|$$

and the second series converges because of the cancellation that occurs as the terms of the series alternate in sign.

On the other hand, both $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ and $\sum\limits_{n=1}^{\infty} (-1)^{n-1}\frac{1}{n^2}$ converge because $\frac{1}{n^2}$ is *small enough*!

We will see that there are differences between series that converge because the magnitude of the terms is small and those that rely on cancellation. We begin with the following definition:

DEFINITION

Absolute vs Conditional Convergence

A series $\sum\limits_{n=1}^{\infty} a_n$ is said to *converge absolutely* if

$$\sum_{n=1}^{\infty} |\, a_n \,|$$

converges.

A series $\sum\limits_{n=1}^{\infty} a_n$ is said to *converge conditionally* if

$$\sum_{n=1}^{\infty} |\, a_n \,|$$

diverges while

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

converges.

EXAMPLE 25 The series $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ converges by the Alternating Series Test. However, $\sum\limits_{n=1}^{\infty}|(-1)^{n-1}\frac{1}{n}| = \sum\limits_{n=1}^{\infty}\frac{1}{n}$ diverges so that $\sum\limits_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ is conditionally convergent. ◀

EXAMPLE 26 If $a_n \geq 0$ for each n, then $|a_n| = a_n$, so the series $\sum\limits_{n=1}^{\infty} a_n$ either converges absolutely or it diverges. ◀

EXAMPLE 27 The series $\sum\limits_{n=0}^{\infty}(\frac{-1}{2})^n$ converges absolutely since $\sum\limits_{n=0}^{\infty}|(\frac{-1}{2})^n| = \sum\limits_{n=0}^{\infty}(\frac{1}{2})^n$ converges by the Geometric Series Test. ◀

The terminology for absolute versus conditional convergence seems to suggest that if a series converges absolutely it should also converge without the absolute values.

Question: Is it possible that $\sum\limits_{n=1}^{\infty}|a_n|$ converges while $\sum\limits_{n=1}^{\infty} a_n$ does not?

It turns out that such a scenario is not possible as the following theorem illustrates.

THEOREM 11 **Absolute Convergence Theorem**

If $\sum\limits_{n=1}^{\infty}|a_n|$ converges, then so does $\sum\limits_{n=1}^{\infty} a_n$.

Note: The sums $\sum\limits_{n=1}^{\infty}|a_n|$ and $\sum\limits_{n=1}^{\infty} a_n$ will converge to different values unless $a_n \geq 0$ for all n.

PROOF

The proof is an application of the Comparison Test.

Assume that $\sum\limits_{n=1}^{\infty}|a_n|$ converges. Then so does

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

Next observe that

$$0 \leq a_n + |a_n| \leq 2|a_n|.$$

(This is true because if $a_n \geq 0$, then $a_n = |a_n|$, so $a_n + |a_n| = 2|a_n|$ and if $a_n < 0$, then $a_n = -|a_n|$ so $a_n + |a_n| = 0$.)

Since $0 \leq a_n + \mid a_n \mid \leq 2 \mid a_n \mid$, we can apply the Comparison Test to show that

$$\sum_{n=1}^{\infty} (a_n + \mid a_n \mid)$$

converges. Finally, we have that

$$a_n = (a_n + \mid a_n \mid) - \mid a_n \mid$$

and hence

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (a_n + \mid a_n \mid) - \sum_{n=1}^{\infty} \mid a_n \mid .$$

The two series on the right-hand side both converge. Therefore, $\sum\limits_{n=1}^{\infty} a_n$ will also converge. ■

Remark: This theorem is very useful because there are many more tests for determining the convergence of positive series than there are for general series. In fact, the only test we have so far that will determine if a non-positive series converges is the Alternating Series Test. However, the conditions under which the Alternating Series Test applies are rather restrictive.

EXAMPLE 28 Show that the series $\sum\limits_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ converges.

It can be shown that as n goes from 1 to ∞ the values of $\cos(n)$ will be positive infinitely often and negative infinitely often. Moreover, since

$$\begin{array}{rcr} \cos(1) & = & 0.540 \\ \cos(2) & = & -0.416 \\ \cos(3) & = & -0.9899 \\ \cos(4) & = & -0.6536 \\ \cos(5) & = & 0.2836 \end{array}$$

$\sum\limits_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ is not an alternating series. Strictly speaking, none of the tests we have discussed up to this point apply to this series. However, if we can show that $\sum\limits_{n=1}^{\infty} \left|\frac{\cos(n)}{n^2}\right|$ converges, then $\sum\limits_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ also converges.

We know that $\mid \cos(x) \mid \leq 1$ for any value of x. Therefore, for every n

$$0 \leq \left|\frac{\cos(n)}{n^2}\right| \leq \frac{1}{n^2}.$$

The p-Series Test shows that $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ converges. Therefore, by the Comparison Test

$$\sum_{n=1}^{\infty} \left| \frac{\cos(n)}{n^2} \right|$$

converges. Hence, by the Absolute Convergence Theorem, the series $\sum\limits_{n=1}^{\infty} \frac{\cos(n)}{n^2}$ converges. ◀

EXAMPLE 29 Show that the series $\sum\limits_{n=0}^{\infty} \dfrac{x^n}{2^n(n+1)}$ converges absolutely if $|\, x \,| < 2$, converges conditionally at $x = -2$ and diverges if $x = 2$.

Choose x_0 with $|\, x_0 \,| < 2$. Let $a_n = \dfrac{x_0^n}{2^n(n+1)}$. Then

$$\begin{aligned} |\, a_n \,| &= \left| \frac{x_0^n}{2^n(n+1)} \right| \\ &= \frac{1}{n+1} \left| \frac{x_0^n}{2^n} \right| \\ &= \frac{1}{n+1} \left| \frac{x_0}{2} \right|^n \\ &\leq \left| \frac{x_0}{2} \right|^n \end{aligned}$$

If $|\, x_0 \,| < 2$, then $|\, \frac{x_0}{2} \,| < 1$ and so by the Geometric Series Test,

$$\sum_{n=0}^{\infty} \left| \frac{x_0}{2} \right|^n$$

converges. The Comparison Test shows that

$$\sum_{n=0}^{\infty} \left| \frac{x_0^n}{2^n(n+1)} \right|$$

converges.

If $x = 2$, the series becomes

$$\begin{aligned} \sum_{n=0}^{\infty} \frac{2^n}{2^n(n+1)} &= \sum_{n=0}^{\infty} \frac{1}{n+1} \\ &= \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots \end{aligned}$$

which is the Harmonic Series and as such diverges.

If $x = -2$, the series is

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{(-2)^n}{2^n(n+1)} &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \\ &= \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \cdots\end{aligned}$$

This is the Alternating Series and as such it converges conditionally.

◀

Remark: We have seen that absolutely convergent series also converge and that testing for absolute convergence allows us to use most of the tools we have developed. There is one more important reason why we would want to know if a series converges absolutely.

If we have a finite sum

$$a_1 + a_2 + a_3$$

we can add the terms in any order and we will get the same sum. For example,

$$a_1 + a_2 + a_3 = a_3 + a_2 + a_1$$

and

$$a_1 + a_2 + a_3 = a_2 + a_3 + a_1.$$

We would hope that this would also be true for infinite series. Unfortunately, this is **not** the case. However, if the series converges absolutely then it is true. That is, **no matter how we rearrange the terms, the result will be a new series that has the same sum as the original**. Hence,

$$a_{17} + a_{357} + a_{45} + a_{10437} + \cdots = a_1 + a_2 + a_3 + a_4 + \cdots .$$

Alternately, if a series $\sum\limits_{n=1}^{\infty} a_n$ converges conditionally, then given any $\alpha \in \mathbb{R}$ or $\alpha = \pm\infty$, there is a new series $\sum\limits_{n=1}^{\infty} b_n$ consisting of exactly the same terms as our original series except in a different order but with

$$\sum_{n=1}^{\infty} b_n = \alpha.$$

This means that absolutely convergent series are *very stable*, whereas conditionally convergent series are not.

We can make this remark more formal.

DEFINITION

Rearrangement of a Series

Given a series $\sum_{n=1}^{\infty} a_n$ and a 1-1 and onto function $\phi : \mathbb{N} \to \mathbb{N}$, if we let

$$b_n = a_{\phi(n)},$$

then the series

$$\sum_{n=1}^{\infty} b_n$$

is called a *rearrangement* of $\sum_{n=1}^{\infty} a_n$.

This definition leads to the following theorem.

THEOREM 12

Rearrangement Theorem

1) Let $\sum_{n=1}^{\infty} a_n$ be an absolutely convergent series. If $\sum_{n=1}^{\infty} b_n$ is any rearrangement of $\sum_{n=1}^{\infty} a_n$, then $\sum_{n=1}^{\infty} b_n$ also converges and

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

2) Let $\sum_{n=1}^{\infty} a_n$ be a conditionally convergent series. Let $\alpha \in \mathbb{R}$ or $\alpha = \pm\infty$. Then there exists a rearrangement $\sum_{n=1}^{\infty} b_n$ of $\sum_{n=1}^{\infty} a_n$ such that

$$\sum_{n=1}^{\infty} b_n = \alpha.$$

Remark: In summary, whenever you must test a series with terms of mixed signs for convergence it is always a good idea to first check if the series converges absolutely.

5.9 Ratio Test

Recall that the Geometric Series Test states that a geometric series

$$\sum_{n=0}^{\infty} r^n$$

will converge if and only if $|\, r \,| < 1$.

Suppose that we had a series where

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{2}.$$

Then, for a large number N, we would have

$$\begin{array}{rcl}
|\, a_{N+1} \,| & \cong & \frac{1}{2}\,|\, a_N \,| \\
|\, a_{N+2} \,| & \cong & \frac{1}{2}\,|\, a_{N+1} \,| \cong \left(\frac{1}{2}\right)^2 |\, a_N \,| \\
|\, a_{N+3} \,| & \cong & \frac{1}{2}\,|\, a_{N+2} \,| \cong \left(\frac{1}{2}\right)^3 |\, a_N \,| \\
|\, a_{N+4} \,| & \cong & \frac{1}{2}\,|\, a_{N+3} \,| \cong \left(\frac{1}{2}\right)^4 |\, a_N \,| \\
 & \vdots & \\
|\, a_{N+k} \,| & \cong & \frac{1}{2}\,|\, a_{N+k-1} \,| \cong \left(\frac{1}{2}\right)^k |\, a_N \,|
\end{array}$$

Since $|\, a_N \,| = 1\cdot |\, a_N \,| = \left(\frac{1}{2}\right)^0 |\, a_N \,|$ this would suggest that

$$\sum_{k=0}^{\infty} |\, a_{N+k} \,| \cong \sum_{k=0}^{\infty} |\, a_N \,| \left(\frac{1}{2}\right)^k$$

The series

$$\sum_{k=0}^{\infty} |\, a_N \,| \left(\frac{1}{2}\right)^k = |\, a_N \,| \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k$$

converges by the Geometric Series Test. Therefore, we might expect that

$$\sum_{k=0}^{\infty} |\, a_{N+k} \,|$$

also converges. Since

$$\sum_{k=0}^{\infty} |\, a_{N+k} \,| = \sum_{n=N}^{\infty} |\, a_n \,|,$$

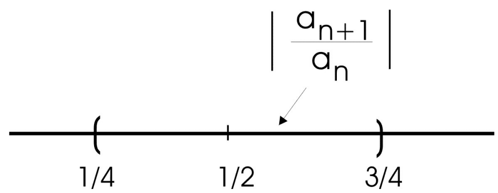
if $\sum\limits_{k=0}^{\infty} |\, a_{N+k} \,|$ converges, so does $\sum\limits_{n=N}^{\infty} |\, a_n \,|$. Finally, from the properties for series, we could conclude that $\sum\limits_{n=1}^{\infty} |\, a_n \,|$ would converge.

This argument seems plausible, but can we make the argument above more rigorous? In fact we can.

If we assume that

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{2},$$

then we can find an N large enough so that if $n \geq N$ then $\mid \frac{a_{n+1}}{a_n} \mid$ approximates $\frac{1}{2}$ with an error of less than $\frac{1}{4}$. This means that for $n \geq N$, $\mid \frac{a_{n+1}}{a_n} \mid$ must be in the interval $\left(\frac{1}{4}, \frac{3}{4}\right)$.



In particular, if $n \geq N$ we get that $\mid \frac{a_{n+1}}{a_n} \mid < \frac{3}{4}$ or equivalently, that

$$\mid a_{n+1} \mid < \frac{3}{4} \mid a_n \mid$$

for every $n \geq N$. This shows that

$$\begin{aligned}
\mid a_{N+1} \mid &< \frac{3}{4} \mid a_N \mid \\
\mid a_{N+2} \mid &< \frac{3}{4} \mid a_{N+1} \mid < \left(\frac{3}{4}\right)^2 \mid a_N \mid \\
\mid a_{N+3} \mid &< \frac{3}{4} \mid a_{N+2} \mid < \left(\frac{3}{4}\right)^3 \mid a_N \mid \\
\mid a_{N+4} \mid &< \frac{3}{4} \mid a_{N+3} \mid < \left(\frac{3}{4}\right)^4 \mid a_N \mid \\
&\vdots \\
\mid a_{N+k} \mid &< \frac{3}{4} \mid a_{N+k-1} \mid < \left(\frac{3}{4}\right)^k \mid a_N \mid
\end{aligned}$$

The Geometric Series Test shows that

$$\sum_{k=0}^{\infty}\left(\frac{3}{4}\right)^k \mid a_N \mid = \mid a_N \mid \sum_{k=0}^{\infty}\left(\frac{3}{4}\right)^k$$

converges. Since

$$\mid a_{N+k} \mid < (\frac{3}{4})^k \mid a_N \mid$$

the Comparison Test tells us that

$$\sum_{k=0}^{\infty} \mid a_{N+k} \mid$$

also converges. From this we can conclude that

$$\sum_{n=0}^{\infty} |\ a_n\ |$$

converges.

In fact, if $\lim\limits_{n\to\infty} |\ \frac{a_{n+1}}{a_n}\ | = L$ and $0 \leq L < 1$, then a similar method would show that $\sum\limits_{n=0}^{\infty} |\ a_n\ |$ converges.

On the other hand, if $\lim\limits_{n\to\infty} |\ \frac{a_{n+1}}{a_n}\ | = 2$, then for large n we would have that

$$|\ a_{n+1}\ | \cong 2\ |\ a_n\ |\ .$$

This means that rather than going to 0, the terms in the tail are getting larger. Since we would then have that $\lim\limits_{n\to\infty} a_n \neq 0$, the Divergence Test would tell us that $\sum\limits_{n=0}^{\infty} a_n$ diverges.

A similar statement would hold whenever $\lim\limits_{n\to\infty} |\ \frac{a_{n+1}}{a_n}\ | = L$ and $L > 1$.

This is summarized in the next theorem which gives us one of the most important tests for convergence of series.

THEOREM 13

Ratio Test

Given a series $\sum\limits_{n=0}^{\infty} a_n$, assume that

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L$$

where $L \in \mathbb{R}$ or $L = \infty$.

1. If $0 \leq L < 1$, then $\sum\limits_{n=0}^{\infty} a_n$ converges absolutely.
2. If $L > 1$, then $\sum\limits_{n=0}^{\infty} a_n$ diverges.
3. If $L = 1$, then no conclusion is possible.

Remarks:

1) If $0 \leq L < 1$, the Ratio Test shows that the given series converges absolutely and hence that the original series also converges.

2) If $\lim\limits_{n\to\infty} |\ \frac{a_{n+1}}{a_n}\ | = L$ exists with $L \neq 1$, then the series $\sum\limits_{n=0}^{\infty} a_n$ behaves like the geometric series $\sum\limits_{n=0}^{\infty} L^n$ as far as convergence is concerned.

3) While the Ratio Test is one of the most important tests for convergence, we will see that it cannot detect convergence or divergence for many of the series we have seen so far. In fact, it can only detect convergence if the terms a_n approach 0 very rapidly, and it can only detect divergence if $\lim\limits_{n\to\infty} |a_n| = \infty$. This means that the Ratio Test is appropriate for a very special class of series.

EXAMPLE 30 Show that $\sum\limits_{n=0}^{\infty} \frac{1}{n!}$ converges.

We have already seen how this could be done using the Comparison Test. However, the Ratio Test is perfectly suited to series involving factorials.

With $a_n = \frac{1}{n!}$, we see that

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\frac{1}{(n+1)!}}{\frac{1}{n!}} \\ &= \frac{n!}{(n+1)!} \\ &= \frac{1}{n+1}\end{aligned}$$

Therefore

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{n+1} = 0.$$

The Ratio Test shows that $\sum\limits_{n=0}^{\infty} \frac{1}{n!}$ converges. ◀

EXAMPLE 31 In the previous example, we saw how that Ratio Test could be used to show that the series $\sum\limits_{n=0}^{\infty} \frac{1}{n!}$ converges. This series actually converges very rapidly since $n!$ grows very quickly. However, if we let

$$a_n = \frac{1000000^n}{n!}$$

the situation is quite different. For example, $a_{10} > 10^{50}$.

Still,

$$\begin{aligned}\frac{a_{n+1}}{a_n} &= \frac{\frac{1000000^{n+1}}{(n+1)!}}{\frac{1000000^n}{n!}} \\ &= \frac{1000000^{n+1} n!}{1000000^n (n+1)!} \\ &= \frac{1000000}{n+1}\end{aligned}$$

and so we again have that

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1000000}{n+1} = 0.$$

This means that despite the enormous size of 1000000^n in the numerator, $n!$ eventually dominates. Consequently, we can use the Ratio Test to show that $\sum\limits_{n=0}^{\infty} \frac{1000000^n}{n!}$ converges. ◀

EXAMPLE 32 Based on the two previous examples, for which values of x would the series

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

converge?

To answer this question, we first note that if $x = 0$

$$\sum_{n=0}^{\infty} \frac{0^n}{n!} = 1 + 0 + 0 + 0 + \cdots = 1$$

so the series converges when $x = 0$. Next fix a value for $x \neq 0$.

If $a_n = \dfrac{x^n}{n!}$, then

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \left| \frac{\frac{x^{n+1}}{(n+1)!}}{\frac{x^n}{n!}} \right| \\ &= \frac{|\,x\,|^{n+1}\; n!}{|\,x\,|^n\; (n+1)!} \\ &= \frac{|\,x\,|}{n+1} \end{aligned}$$

Then

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n\to\infty} \frac{|\,x\,|}{n+1} = 0.$$

The Ratio Test shows that

$$\sum_{n=0}^{\infty} \left| \frac{x^n}{n!} \right|$$

converges and hence that

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

also converges. ◀

Remark: One final observation can be made from the previous example. Since $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges for any x and since the Divergence Test tells us that the terms of a convergent series must approach 0 we have the following theorem.

THEOREM 14 **Polynomial vs Factorial Growth**

For any $x \in \mathbb{R}$

$$\lim_{n\to\infty} \frac{x^n}{n!} = 0.$$

Remark: This important limit tells us that exponentials are of a lower order of magnitude compared to factorials. That is, for any fixed $x_0 \in \mathbb{R}$, $|\, x_0 \,|^n \ll n!$

Note: We know that the series $\sum_{n=0}^{\infty} r^n$ will diverge if $|\, r \,| = 1$. Therefore, since the conclusions of the Ratio Test are based on the Geometric Series Test, it might be surprising that if $L = 1$, the Ratio Test would not show that the series diverges. However, it is important to recognize that $\lim_{n\to\infty} |\, \frac{a_{n+1}}{a_n} \,| = 1$ does not actually mean that $|\, a_{N+k} \,| = |\, a_N \,|$ for large N as would be the case if the ratio was exactly 1.

The next two examples show that when $L = 1$ we could have either convergence or divergence.

EXAMPLE 33 Apply the Ratio Test to the divergent series $\sum_{n=1}^{\infty} \frac{1}{n}$.

In this example, $a_n = \frac{1}{n}$ so that

$$\begin{aligned} \left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{1}{n+1}}{\frac{1}{n}} \\ &= \frac{n}{n+1} \\ &= \frac{1}{1+\frac{1}{n}} \end{aligned}$$

Therefore, $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} = 1$ and the Ratio Test fails. ◀

EXAMPLE 34 Apply the Ratio Test to the convergent series $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

In this example, $a_n = \frac{1}{n^2}$ so that

$$\begin{aligned}\left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} \\ &= \frac{n^2}{n^2+2n+1} \\ &= \frac{1}{1+\frac{2}{n}+\frac{1}{n^2}}\end{aligned}$$

Therefore, $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \lim_{n\to\infty} \frac{1}{1+\frac{2}{n}+\frac{1}{n^2}} = 1$ and the Ratio Test fails again. ◀

We see that the Ratio Test cannot detect convergence or divergence of many other series similar to the previous two examples.

Fact: If $p(x) = a_0 + a_1x + \cdots + a_kx^k$ and $q(x) = b_0 + b_1x + \cdots + b_mx^m$ are two polynomials, then the Ratio Test will *always fail* for the series

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

EXAMPLE 35 Consider the series

$$\sum_{n=1}^{\infty} \frac{n^n}{n!}.$$

It is even more difficult to predict how the growth of n^n compares with that of $n!$ than it was for x^n versus $n!$. Since the base is increasing as well as the exponent, n^n will get very large, very quickly. However, this is also true of $n!$. For example, there is no easy way to determine the value of $\frac{(10^6)^{10^6}}{10^6!}$. Instead, observe that

$$\frac{n^n}{n!} = \frac{n\cdot n\cdot n\cdots n}{1\cdot 2\cdot 3\cdots n} = \frac{n}{1}\cdot\frac{n}{2}\cdot\frac{n}{3}\cdot\cdots\cdot\frac{n}{n}\geq 1.$$

Since the terms of the series are always larger than 1, they cannot converge to 0 and hence the series diverges.

The Ratio Test would also show this result because

$$\begin{aligned}
\left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{(n+1)^{n+1}}{(n+1)!}}{\frac{n^n}{n!}} \\
&= \frac{(n+1)^{n+1}n!}{n^n(n+1)!} \\
&= \frac{(n+1)^{n+1}}{n^n(n+1)} \\
&= \frac{(n+1)^n}{n^n} \\
&= \left(\frac{n+1}{n}\right)^n \\
&= \left(1+\frac{1}{n}\right)^n
\end{aligned}$$

Recall that

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e.$$

Therefore, $\lim\limits_{n\to\infty} \frac{a_{n+1}}{a_n} = e > 1$. The Ratio Test shows that the series diverges.

This also shows that

$$\lim_{n\to\infty}\frac{n^n}{n!} = \infty$$

since for large N we will have that

$$a_{N+k} \cong a_N e^k \to \infty.$$

Hence

$$n! \ll n^n.$$

Alternately, if we consider the series

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

then

$$\begin{aligned}
\left|\frac{a_{n+1}}{a_n}\right| &= \frac{\frac{n^n}{n!}}{\frac{(n+1)^{n+1}}{(n+1)!}} \\
&= \frac{n^n(n+1)!}{(n+1)^{n+1}n!} \\
&= \frac{n^n(n+1)}{(n+1)^{n+1}} \\
&= \frac{(n)^n}{(n+1)^n} \\
&= \left(\frac{n}{n+1}\right)^n \\
&= \frac{1}{\left(\frac{n+1}{n}\right)^n} \\
&= \frac{1}{\left(1+\frac{1}{n}\right)^n}
\end{aligned}$$

This time, we get

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n} = \lim_{n\to\infty}\frac{1}{(1+\frac{1}{n})^n} = \frac{1}{e} < 1$$

so the series

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

converges. ◀

The following is a summary of what we have learned about the order of magnitude of various functions:

$$\ln(n) \ll n^p \ll x^n \ll n! \ll n^n$$

for $|\, x \,| > 1$.

Therefore,

$$\frac{1}{n^n} \ll \frac{1}{n!} \ll \frac{1}{x^n} \ll \frac{1}{n^p} \ll \frac{1}{\ln(n)}.$$

5.10 Root Test

The next test is related to the Ratio Test. In fact, it can be derived in a manner similar to the Ratio Test by comparing the series with a suitably chosen geometric series.

THEOREM 15 **Root Test**

Given a series $\sum_{n=1}^{\infty} a_n$, assume that

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

where $L \in \mathbb{R}$ or $L = \infty$.

1. If $0 \le L < 1$, then $\sum_{n=1}^{\infty} a_n$ converges absolutely.
2. If $L > 1$, then $\sum_{n=1}^{\infty} a_n$ diverges.
3. If $L = 1$, then no conclusion is possible.

EXAMPLE 36 Does the series $\sum_{n=1}^{\infty} \left(\frac{3n^2+1}{4n^2+n-1}\right)^n$ converge or diverge?

Let $a_n = \left(\frac{3n^2+1}{4n^2+n-1}\right)^n$. Then $\sqrt[n]{a_n} = \left(\frac{3n^2+1}{4n^2+n-1}\right)$.

We know that

$$\lim_{n\to\infty} \sqrt[n]{a_n} = \lim_{n\to\infty} \left(\frac{3n^2+1}{4n^2+n-1}\right) = \frac{3}{4} < 1.$$

It follows from the Root Test that $\sum_{n=1}^{\infty} \left(\frac{3n^2+1}{4n^2+n-1}\right)^n$ converges. ◀

EXAMPLE 37 Does the series $\sum_{n=1}^{\infty} \left(1+\frac{1}{n}\right)^{n^2}$ converge or diverge?

Let $a_n = \left(1+\frac{1}{n}\right)^{n^2}$. Then $\sqrt[n]{a_n} = \sqrt[n]{\left(1+\frac{1}{n}\right)^{n^2}} = \left(1+\frac{1}{n}\right)^n$.

We know that

$$\lim_{n\to\infty} \left(1+\frac{1}{n}\right)^n = e > 1.$$

Hence, by the Root Test the series $\sum_{n=1}^{\infty} \left(1+\frac{1}{n}\right)^{n^2}$ diverges. ◀

Chapter 6

Power Series

So far all of the series we have considered have been numerical. That is, they have consisted of an infinite sum of real numbers which either converged or diverged. In this chapter, we will introduce a type of series called a power series which resembles a polynomial of infinite degree.

6.1 Introduction to Power Series

In this section, we will introduce an important class of series called *power series*.

DEFINITION

Power Series

A power series centered at $x = a$ is a series of the form

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

where x is considered a variable and the value a_n is called the *coefficient* of the term $(x-a)^n$.

Once we assign a value to the variable x, the series becomes a *numerical series*. In particular, the following is the fundamental problem that we must answer.

Problem: For which values of x does the power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ converge?

Note that if $x = a$, the series becomes

$$\sum_{n=0}^{\infty} a_n(0)^n.$$

Since our convention is that $0^0 = 1$, we have

$$\begin{aligned}\sum_{n=0}^{\infty} a_n(0)^n &= a_0(1) + a_1(0) + a_2(0) + \cdots \\ &= a_0 + 0 + 0 + 0 + \cdots \\ &= a_0\end{aligned}$$

This shows that every power series centered at $x = a$ will converge at $x = a$ to the value a_0. The problem now is to determine what happens for other values of x.

As you might expect, the answer depends on the coefficients a_n.

Note: Before we proceed to study the convergence properties for power series we note that if we let $u = x - a$, then

$$\sum_{n=0}^{\infty} a_n(x-a)^n = \sum_{n=0}^{\infty} a_n u^n.$$

This allows us to *transform* a power series centered at $x = a$ into a power series centered at $u = 0$. Moreover, the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ converges at a point x_0 if and only if the power series $\sum_{n=0}^{\infty} a_n u^n$ converges at $x_0 - a$. This means that if we know all of the values of u at which the series $\sum_{n=0}^{\infty} a_n u^n$ converges, then we know all the values of x for which the series $\sum_{n=0}^{\infty} a_n(x-a)^n$ converges. For this reason, in studying the basic properties of convergence of power series we can focus almost exclusively on the case where $a = 0$. Therefore, unless otherwise specified, we will be dealing with power series of the form

$$\sum_{n=0}^{\infty} a_n x^n.$$

◀

EXAMPLE 1 If we choose $a_n = 1$ for each n, we get the power series

$$\sum_{n=0}^{\infty} x^n.$$

We know from the Geometric Series Test that this series will converge if and only if $| x |< 1$. ◀

EXAMPLE 2 We saw as an application of the Ratio Test that the series

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

will converge no matter what value we assign to x. Moreover, it converges absolutely for each x. ◀

EXAMPLE 3 For which values of x does the series

$$\sum_{n=0}^{\infty} \frac{x^n}{n+1}$$

converge?

We can use the Ratio Test to answer this question.

Fix a value for x and let

$$b_n = \frac{x^n}{n+1}.$$

Then

$$\begin{aligned}
\lim_{n\to\infty} \left|\frac{b_{n+1}}{b_n}\right| &= \lim_{n\to\infty} \left|\frac{\frac{(x)^{n+1}}{n+2}}{\frac{(x)^n}{n+1}}\right| \\
&= \lim_{n\to\infty} \frac{(n+1)}{(n+2)} \frac{|\, x\,|^{n+1}}{|\, x\,|^n} \\
&= \lim_{n\to\infty} \frac{(n+1)}{(n+2)} \,|\, x\,| \\
&= |\, x\,| \lim_{n\to\infty} \frac{(n+1)}{(n+2)} \\
&= |\, x\,| \cdot 1 \\
&= |\, x\,|
\end{aligned}$$

The Ratio Test shows that the series $\sum\limits_{n=0}^{\infty} \frac{x^n}{n+1} = \sum\limits_{n=0}^{\infty} b_n$ converges (absolutely) if

$$\lim_{n\to\infty} |\, \frac{b_{n+1}}{b_n} \,| = |\, x\,| < 1$$

and it diverges if

$$\lim_{n\to\infty} |\, \frac{b_{n+1}}{b_n} \,| = |\, x\,| > 1.$$

The Ratio Test fails to tell us what happens if $|\, x\,| = 1$, so we must consider this case separately.

When $x = 1$, the series is

$$\sum_{n=0}^{\infty} \frac{1}{n+1}$$

which is just the Harmonic Series written in a different form. Therefore, when $x = 1$, the series diverges.

When $x = -1$ the series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}.$$

This is just the Alternating Series and as such when $x = -1$, the series converges.

In summary, the power series $\sum_{n=0}^{\infty} \frac{x^n}{n+1}$ converges absolutely if $| x | < 1$, diverges if $| x | > 1$ and if $x = 1$, and converges conditionally at $x = -1$. That is, the series converges if

$$x \in [-1, 1)$$

If we take a closer look at these three examples we will see that they have certain common features. First observe that the series $\sum_{n=0}^{\infty} x^n$ converges if $x \in (-1, 1)$. In fact it converges absolutely on this interval. The series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges absolutely on the interval $(-\infty, \infty)$. The series $\sum_{n=0}^{\infty} \frac{x^n}{n+1}$ converges if $x \in [-1, 1)$ and it converges absolutely if $x \in (-1, 1)$.

These three power series have the following two properties:

Property 1: The set of points on which the power series converge is an interval centered around $x = 0$.

Property 2: There exists an R with either $R \in [0, \infty)$ or $R = \infty$ such that the power series converges absolutely if $| x | < R$ and it diverges if $| x | > R$ when $R \in [0, \infty)$, and the series converges absolutely at each $x \in \mathbb{R}$ when $R = \infty$. In the first and third examples $R = 1$, while in the second example $R = \infty$.

We will now show that two properties are shared by all power series centered at $x = 0$. Moreover, if a power series is centered at $x = a$, then Property 1 will hold with an interval centered around $x = a$ and Property 2 will hold if we replace $| x |$ by $| x - a |$.

Key Observation:

Assume that the power series $\sum_{n=0}^{\infty} a_n x^n$ converges at $x = x_0$. If $0 \leq | x_1 | < | x_0 |$, then we claim that the series

$$\sum_{n=0}^{\infty} | a_n x_1^n |$$

also converges. To see why this is the case, we first note that since

$$\sum_{n=0}^{\infty} a_n x_0^n$$

converges, the Divergence Test shows that

$$\lim_{n\to\infty} |a_n x_0^n| = 0.$$

It follows that there exists an N_0 such that

$$|a_n x_0^n| < 1$$

for all $n \geq N_0$.

Next observe that

$$|a_n x_1^n| = |a_n x_0^n| \cdot \left|\frac{x_1}{x_0}\right|^n \leq \left|\frac{x_1}{x_0}\right|^n$$

for all $n \geq N_0$. But $\left|\frac{x_1}{x_0}\right| < 1$ so the series

$$\sum_{n=N_0}^{\infty} \left|\frac{x_1}{x_0}\right|^n$$

converges by the Geometric Series Test. Since we know that $|a_n x_1^n| \leq \left|\frac{x_1}{x_0}\right|^n$ the Comparison Test shows that

$$\sum_{n=N_0}^{\infty} |a_n x_1^n|$$

also converges. Finally we get that

$$\sum_{n=0}^{\infty} |a_n x_1^n|$$

converges as claimed.

Summary: We have actually shown that if we are given a series $\sum\limits_{n=0}^{\infty} a_n x^n$ then the set

$$I = \{x_0 \mid \sum_{n=0}^{\infty} \mid a_n x_0^n \text{ converges}\}$$

is an interval centered at $x = 0$. Moreover, a similar result is true for a series of the form $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$.

This leads us to the following definition:

DEFINITION

Interval and Radius of Convergence

Given a power series of the form $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$, the set

$$I = \{x_0 \in \mathbb{R} \mid \sum_{n=0}^{\infty} a_n(x_0-a)^n \text{ converges}\}$$

is an interval centered at $x = a$ which we call *the interval of convergence* for the power series.

Let

$$R := \begin{cases} lub(\{|x_0 - a| \mid x_0 \in I\}) & \text{if } I \text{ is bounded,} \\ \infty & \text{if } I \text{ is not bounded.} \end{cases}$$

Then R is called the *radius of convergence* of the power series.

The following theorem summarizes what we now know about convergence of a power series.

THEOREM 1

Fundamental Convergence Theorem for Power Series

Given a power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ centered at $x = a$, let R be the radius of convergence.

1. If $R = 0$, then $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ converges for $x = a$ but it diverges for all other values of x.
2. If $0 < R < \infty$, then the series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ converges absolutely for every $x \in (a-R, a+R)$ and diverges if $| x - a |> R$.
3. If $R = \infty$, then the series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ converges absolutely for every $x \in \mathbb{R}$.

In particular, $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ converges on an interval that is centered at $x = a$ which may or may not include one or both of the endpoints.

The following are some important observations concerning the interval of convergence.

1. As the previous theorem states, if $0 < R < \infty$, the power series $\sum\limits_{n=0}^{\infty} a_n x^n$ converges absolutely on the interval $(-R, R)$. It may or may not converge at $x = R$ or at $x = -R$. These points must be tested separately. As we will see later, the interval of convergence could be $(-R, R)$, $[-R, R)$, $(-R, R]$ or $[-R, R]$. The next few examples show that all four cases are possible.

2. If $R = \infty$, then the power series has interval of convergence $(-\infty, \infty)$ and it converges absolutely at each point. For example, the series $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ has $(-\infty, \infty)$ as its interval of convergence.
3. If $R = 0$, then the power series only converges at $x = 0$. That is, "the interval of convergence" is just the single point $\{0\}$. An example of such a power series is $\sum_{n=0}^{\infty} n!x^n$.

EXAMPLE 4 The following power series all have radius of convergence $R = 1$. The interval of convergence is specified.

1. $\sum_{n=0}^{\infty} x^n$ has interval of convergence $(-1, 1)$ since when $x = 1$, the series

$$\sum_{n=0}^{\infty} 1^n$$

diverges and when $x = -1$, the series

$$\sum_{n=0}^{\infty} (-1)^n$$

also diverges.

2. $\sum_{n=1}^{\infty} \frac{x^n}{n}$ has interval of convergence $[-1, 1)$ since when $x = 1$, the series

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

diverges but when $x = -1$, the series

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

converges.

3. $\sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n}$ has interval of convergence $(-1, 1]$ since when $x = 1$, the series

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

converges but when $x = -1$, the series

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

diverges.

4. $\sum\limits_{n=1}^{\infty} \frac{x^n}{n^2}$ has interval of convergence $[-1, 1]$ since when $x = 1$, the series

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges and when $x = -1$, the series

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

also converges.

6.1.1 Finding the Radius of Convergence

Given a power series $\sum\limits_{n=0}^{\infty} a_n x^n$, we have seen that we can often use the Ratio Test to find the radius of convergence. To see how this works in general, assume that

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = L$$

where $0 \leq L < \infty$ For $x \neq 0$, let

$$b_n = a_n x^n.$$

Then

$$\begin{aligned}
\lim_{n\to\infty} \left| \frac{b_{n+1}}{b_n} \right| &= \lim_{n\to\infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| \\
&= \lim_{n\to\infty} \mid x \mid \left| \frac{a_{n+1}}{a_n} \right| \\
&= \mid x \mid \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \\
&= L \mid x \mid
\end{aligned}$$

The Ratio Test shows that the series $\sum\limits_{n=0}^{\infty} b_n = \sum\limits_{n=0}^{\infty} a_n x^n$ converges absolutely if $L \mid x \mid < 1$ and diverges if $L \mid x \mid > 1$.

Assume that $0 < L < \infty$. Then $L \mid x \mid < 1$ if and only if $\mid x \mid < \frac{1}{L}$. Therefore, the radius of convergence is $R = \frac{1}{L}$.

If $L = 0$, then no matter what x is, we have $L \mid x \mid= 0 < 1$. Therefore, $R = \infty$. At this point we might be tempted to cheat with our notation and write $\infty \cong \frac{1}{0}$, since this is consistent with the first case. However, the symbol

$$\frac{1}{0}$$

actually **has no numerical meaning**.

If

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L = \infty,$$

then the same calculation would show

$$\lim_{n\to\infty} \mid \frac{b_{n+1}}{b_n} \mid= \infty$$

and so the series diverges for all non-zero x. However, we know that the series must converge at $x = 0$ so the radius of convergence is 0. Again, we might be tempted write $0 = \frac{1}{\infty}$ but once more

$$\frac{1}{\infty}$$

is not actually defined. Instead we can simply make use of the following theorem which summarizes what we have just discussed.

THEOREM 2 **Test for the Radius of Convergence**

Let $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ be a power series for which

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = L$$

where $0 \leq L < \infty$ or $L = \infty$. Let R be the radius of convergence of the power series.

1. If $0 < L < \infty$, then $R = \frac{1}{L}$.
2. If $L = 0$, then $R = \infty$.
3. If $L = \infty$, then $R = 0$.

EXAMPLE 5 Find the radius and interval of convergence for the power series

$$\sum_{n=0}^{\infty} \frac{x^n}{3^n(n^2+1)}.$$

We begin by trying to calculate $\lim\limits_{n\to\infty} \mid \frac{a_{n+1}}{a_n} \mid$. In this case, $a_n = \frac{1}{3^n(n^2+1)}$.

Therefore

$$
\begin{aligned}
\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| &= \lim_{n\to\infty}\left|\frac{\frac{1}{3^{n+1}((n+1)^2+1)}}{\frac{1}{3^n(n^2+1)}}\right| \\
&= \lim_{n\to\infty}\left|\frac{3^n(n^2+1)}{3^{n+1}((n+1)^2+1)}\right| \\
&= \lim_{n\to\infty}\frac{1}{3}\left(\frac{n^2+1}{n^2+2n+2}\right) \\
&= \frac{1}{3}\lim_{n\to\infty}\frac{1+\frac{1}{n^2}}{1+\frac{2}{n}+\frac{2}{n^2}} \\
&= \frac{1}{3}\cdot(1) \\
&= \frac{1}{3}
\end{aligned}
$$

It follows from the Test for Radius of Convergence that $R = \frac{1}{\frac{1}{3}} = 3$.

We know that the series converges absolutely on $(-3, 3)$. We must check for convergence at $x = 3$ and $x = -3$.

For $x = 3$, the series becomes

$$\sum_{n=0}^{\infty}\frac{3^n}{3^n(n^2+1)} = \sum_{n=0}^{\infty}\frac{1}{n^2+1}.$$

Since

$$\frac{1}{n^2+1} < \frac{1}{n^2}$$

and since $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}$ converges, the Comparison Test shows that $\sum\limits_{n=1}^{\infty}\frac{1}{n^2+1}$ converges and hence that

$$\sum_{n=0}^{\infty}\frac{1}{n^2+1}$$

converges.

Similarly, if $x = -3$, the series becomes

$$\sum_{n=0}^{\infty}\frac{(-3)^n}{3^n(n^2+1)} = \sum_{n=0}^{\infty}\frac{(-1)^n}{n^2+1}.$$

Notice that the Alternating Series Test applies and so this series also converges. Alternatively, we have

$$\left|\frac{(-1)^n}{n^2+1}\right| = \frac{1}{n^2+1}.$$

We have just shown that the series $\sum_{n=0}^{\infty} \frac{1}{n^2+1}$ converges, which shows that $\sum_{n=0}^{\infty} \frac{(-1)^n}{n^2+1}$ converges absolutely.

We have just shown that the interval of convergence includes both endpoints, therefore the interval of convergence is $[-3, 3]$.

If the previous calculation is repeated, it would show that the power series $\sum_{n=0}^{\infty} \frac{x^n}{3^n(n^2+1)}$ and the power series $\sum_{n=0}^{\infty} \frac{x^n}{3^n}$ have the same radius of convergence, namely 3, though they have a different interval of convergence. (A close look at the second series shows that it really looks like a geometric series with $r = \frac{x}{3}$ so it must converge when $\left| \frac{x}{3} \right| < 1$ and diverge when $\left| \frac{x}{3} \right| > 1$.) ◀

The following theorem will be very useful when we consider differentiation and integration for functions obtained from power series. It may also help to find the radius of convergence of many series more quickly. It essentially says that multiplying or dividing the terms of a power series by a fixed polynomial in n will not change the radius of convergence, though it is important to remember that it may change the interval of convergence.

THEOREM 3 **Equivalence of Radius of Convergence**

Let p and q be non-zero polynomials where $q(n) \neq 0$ for $n \geq k$. Then the following series have the same radius of convergence:

1. $\sum_{n=k}^{\infty} a_n(x-a)^n$
2. $\sum_{n=k}^{\infty} \frac{a_n p(n)(x-a)^n}{q(n)}$

However, they may have different intervals of convergence.

Note: The limit $\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$ may not always exist. In this case, it might require some clever thought to find the radius of convergence. ◀

EXAMPLE 6 Consider the power series

$$\sum_{n=0}^{\infty} a_n x^n = 1 + 2x + x^2 + 2x^3 + x^4 + 2x^5 + \cdots$$

with $a_n = 1$ if n is even and $a_n = 2$ if n is odd. When $x = 1$ the series is

$$1 + 2 + 1 + 2 + 1 + 2 + \cdots$$

which diverges. In fact the series also diverges for $x = -1$. This shows that $R \leq 1$.

Next pick $x_0 \in (-1, 1)$. Then $|x_0| < 1$ then

$$|a_n x_0^n| \leq 2|x_0^n|$$

and since the series

$$\sum_{n=0}^{\infty} 2\left|x_0^n\right|$$

converges, the Comparison Test shows that the original series converges absolutely at $x = x_0$. This shows that the interval of convergence is $(-1, 1)$ and that $R = 1$.

However, the sequence $\left\{\left|\frac{a_{n+1}}{a_n}\right|\right\}$ osculates between 2 and $\frac{1}{2}$ so $\lim\limits_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$ does not exist.

◀

6.2 Functions Represented by Power Series

In this section, we consider power series centered at $x = a$ with radius of convergence $R > 0$. We will see how these series define functions with particularly nice properties.

EXAMPLE 7 Recall that the geometric series

$$\sum_{n=0}^{\infty} x^n$$

converges for each x with $|\, x \,| < 1$. Moreover, if $|\, x \,| < 1$, then we also know that

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n.$$

This means that the series provides us with a means to *represent* the function $f(x) = \frac{1}{1-x}$ on the interval $(-1, 1)$.

◀

DEFINITION

Functions Represented by a Power Series

Let $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ be a power series with radius of convergence $R > 0$. Let I be the interval of convergence for $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$. Let f be the function defined on the interval I by the formula

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

for each $x \in I$.

We say that the function $f(x)$ *is represented by the power series* $\sum_{n=0}^{\infty} a_n(x-a)^n$ *on* I.

The next theorem tells us that a function represented by a power series must be continuous on its interval of convergence.

THEOREM 4 **Abel's Theorem: Continuity of Power Series**

Assume that the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ has interval of convergence I. Let

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

for each $x \in I$. Then $f(x)$ is continuous on I.

The proof of this theorem is beyond the scope of this course.

6.2.1 Building Power Series Representations

Suppose that f and g are two functions represented by power series centered at x=a with

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

and

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

and that the two series have intervals of convergence I_f and I_g, respectively.

Question: Can this information be used to build a power series for the sum $f+g$?

To see why this is possible, we first start by noting that if both series converge at a point $x_0 \in I_f \cap I_g$, then

$$\begin{aligned}
(f+g)(x_0) &= f(x_0)+g(x_0) \\
&= \lim_{k\to\infty} \sum_{n=0}^{k} a_n(x_0-a)^n + \lim_{k\to\infty} \sum_{n=0}^{k} b_n(x_0-a)^n \\
&= \lim_{k\to\infty} \sum_{n=0}^{k} (a_n+b_n)(x_0-a)^n \\
&= \sum_{n=0}^{\infty} (a_n+b_n)(x_0-a)^n
\end{aligned}$$

This tells us that the function $f + g$ can be represented by the power series

$$\sum_{n=0}^{\infty}(a_n + b_n)(x-a)^n$$

on $I_f \cap I_g$.

THEOREM 5

Addition of Power Series

Assume that f and g are represented by power series centered at $x = a$ with

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

and

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

respectively.

Assume also that the radii of convergence of these series are R_f and R_g with intervals of convergence I_f and I_g. Then

$$(f+g)(x) = \sum_{n=0}^{\infty}(a_n + b_n)(x-a)^n.$$

Moreover, if $R_f \neq R_g$, then the radius of convergence of the power series representing $f + g$ is $R = \min\{R_f, R_g\}$ and the interval of convergence is $I = I_f \cap I_g$.

If $R_f = R_g$, then $R \geq R_f$.

Next assume that $h(x) = (x-a)^m f(x)$ where $m \in \mathbb{N}$. We might guess that $h(x)$ would be represented by the following power series centered at $x = a$:

$$h(x) = (x-a)^m \sum_{n=0}^{\infty} a_n(x-a)^n = \sum_{n=0}^{\infty} a_n(x-a)^{n+m}.$$

To see why this is the case, we note that if $x_0 \in I_f$, then

$$\begin{aligned} h(x_0) &= (x_0-a)^m f(x_0) \\ &= (x_0-a)^m \lim_{k\to\infty} \sum_{n=0}^{k} a_n(x_0-a)^n \\ &= \lim_{k\to\infty} \sum_{n=0}^{k} a_n(x_0-a)^{n+m} \\ &= \sum_{n=0}^{\infty} a_n(x_0-a)^{n+m} \end{aligned}$$

We can summarize this in the following Theorem.

THEOREM 6 **Multiplication of a Power Series by $(x-a)^m$**

Assume that f is represented by a power series centered at $x = a$ as

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

with radius of convergence R_f and interval of convergence I_f.

Assume that $h(x) = (x-a)^m f(x)$ where $m \in \mathbb{N}$. Then $h(x)$ can also be represented by a power series centered at $x = a$ with

$$h(x) = \sum_{n=0}^{\infty} a_n(x-a)^{n+m}$$

Moreover, the series that represents h has the same radius of convergence and the same interval of convergence as the series that represents f.

Finally, assume that f has a power series representation

$$f(u) = \sum_{n=0}^{\infty} a_n u^n$$

centered at $u = 0$ with interval of convergence I_f.

Question: Can we find a power series representation for $h(x) = f(c \cdot x^m)$ centered at $x = 0$ where c is some non-zero constant?

In fact, if we choose x_0 so that $c \cdot x_0^m \in I_f$, then substituting $c \cdot x_0^m$ for u gives us

$$h(x) = f(c \cdot x_0^m) = \sum_{n=0}^{\infty} a_n(c \cdot x_0^m)^n = \sum_{n=0}^{\infty} (a_n \cdot c^n) x_0^{mn}.$$

This leads to the following Theorem.

THEOREM 7 **Power Series of Composite Functions**

Assume that f has a power series representation

$$f(u) = \sum_{n=0}^{\infty} a_n u^n$$

centered at $u = 0$ with radius of convergence R_f and interval of convergence I_f. Let $h(x) = f(c \cdot x^m)$ where c is a non-zero constant. Then h has a power series representation centered at $x = 0$ of the form

$$h(x) = f(c \cdot x^m) = \sum_{n=0}^{\infty} (a_n \cdot c^n) x^{mn}$$

The interval of convergence is

$$I_h = \{x \in \mathbb{R} \,|\, c \cdot x^m \in I_f\}$$

and the radius of convergence is $R_h = \sqrt[m]{\frac{R_f}{|c|}}$ if $R_f < \infty$ and $R_h = \infty$ otherwise.

EXAMPLE 8 Find a power series representation for $f(x) = \dfrac{x}{1 - 2x^2}$ centered at $x = 0$.

We know that

$$\frac{1}{1-u} = \sum_{n=0}^{\infty} u^n$$

for $u \in (-1, 1)$. Then

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

provided that $2x^2 \in (-1, 1)$. However, $2x^2 \in (-1, 1)$ if and only if $x^2 \in \left(-\frac{1}{2}, \frac{1}{2}\right)$. Therefore,

$$\frac{x}{1-2x^2} = x \cdot \sum_{n=0}^{\infty} 2^n x^{2n} = \sum_{n=0}^{\infty} 2^n x^{2n+1}$$

if and only if $x \in \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. ◀

6.3 Differentiation of Power Series

While the Continuity of Power Series Theorem has important theoretical applications, we will be more interested to see if we can differentiate or integrate functions that can be represented by power series.

Let $f(x) = \sum\limits_{n=0}^{\infty} a_n (x - a)^n$ be a function that is represented by a power series. We could naively try to *differentiate f by differentiating the series one term at a time.* We would hope that we can do this because we know that for the sum of two functions we have

$$\frac{d}{dx}(g(x) + h(x)) = g'(x) + f'(x).$$

Since $\frac{d}{dx}(a_n(x-a)^n) = na_n(x-a)^{n-1}$, we would also hope that $f'(x)$ would exist and that

$$f'(x) = \sum_{n=0}^{\infty} na_n(x-a)^{n-1} = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}$$

where the last equality holds since multiplying by 0 gives 0.

Unfortunately, since we are dealing with infinite sums rather than finite sums, we do not even know if the above series will converge. Even if it does, it is not at all obvious that it will sum to $f'(x)$. At the very least, the process of *term-by-term differentiation* does produce another power series centered at $x = a$.

DEFINITION

The Formal Derivative of a Power Series

Given a power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$, the *formal derivative* is the series

$$\sum_{n=0}^{\infty} na_n(x-a)^{n-1} = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}.$$

We are left to consider two fundamental problems.

Problem 1: For which values of x does the formal power series

$$\sum_{n=1}^{\infty} na_n(x-a)^{n-1}$$

converge? In particular, does this series converge for the same values as does the original series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$?

Problem 2: If both the series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ and $\sum\limits_{n=1}^{\infty} na_n(x-a)^{n-1}$ converge at the same at x, must it be the case that

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}?$$

In other words, why must the statement that *the derivative of a sum is the sum of the derivatives* carry over from finite sums to infinite sums?

Fortunately, Problem 1 is not too difficult.

In the previous section, we saw that multiplying the terms of a power series by a polynomial in n does not change the radius of convergence R. Therefore, the series $\sum\limits_{n=0}^{\infty} a_nx^n$ and the series $\sum\limits_{n=0}^{\infty} na_nx^n$ have the same radius of convergence. A minor modification of this result shows that the series $\sum_{n=0}^{\infty} a_n(x-a)^n$ and its formal derivative $\sum_{n=1}^{\infty} na_n(x-a)^{n-1}$ also have the same radius of convergence, though the interval of convergence may be different. Therefore, we have that

$$g(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}$$

is defined for all $x \in (a-R, a+R)$. The question that remains is:

$$\text{Does } g(x) = f'(x)?$$

A rigorous answer to this question is very difficult. It involves a number of very sophisticated results including the Fundamental Theorem of Calculus. However, the following theorem tells us that functions represented by power series do indeed have the remarkable property that we require.

THEOREM 8 **Term-by-Term Differentiation of Power Series**

Assume that the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$ has radius of convergence $R > 0$. Let

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

for all $x \in (a-R, a+R)$. Then f is differentiable on $(a-R, a+R)$ and for each $x \in (a-R, a+R)$,

$$f'(x) = \sum_{n=1}^{\infty} na_n(x-a)^{n-1}.$$

Remarks:

1) It may be tempting to dismiss the previous theorem as obvious since we would expect differentiation to be preserved by sums. However, this theorem illustrates a very special property of power series that is not shared by other types of series of functions. For example, it can be shown as an application of the Comparison Test and the Geometric Series Test that for each $x \in \mathbb{R}$, the series

$$\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \sin(9^n x)$$

converges. This allows us to define a function

$$f(x) = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \sin(9^n x)$$

that is defined for all $x \in \mathbb{R}$.

Notice that f is the sum of infinitely many differentiable functions, namely

$$f_n(x) = \left(\frac{3}{4}\right)^n \sin(9^n x)$$

for each n. Based on our experience with power series, we might be tempted to think that $f(x)$ should be differentiable and that its derivative would be the sum of all of the derivatives of the f_n's. Unfortunately, this is **false**. In fact, it turns out that the function $f(x)$ is a *continuous function that is not differentiable at any point on the real line*. This is about as strange as a continuous function can be!

This example, which is one of a class of everywhere continuous but nowhere differentiable functions discovered by the 19th century mathematician Karl Weierstrass. This is a historically important example because up until Weierstrass' discovery it was thought that a continuous function had to have some points of differentiability. The graph of such a function has "fractal-like" properties. This means that no matter how close we zoom in to any given point, the graph of the function does not start to resemble a straight line.

2) While it may not be immediately obvious, the Differentiation Theorem for Power Series actually says much more about a function f that can be represented by a power series than just simply that it is differentiable. To see why, observe that if $f(x) = \sum\limits_{n=0}^{\infty} a_n(x-a)^n$ is represented by a power series on the interval $(a-R, a+R)$, then $f'(x) = \sum\limits_{n=1}^{\infty} na_n(x-a)^{n-1}$ is also represented by a power series on $(a-R, a+R)$. Applying the previous theorem again shows that $f'(x)$ is also differentiable on $(a-R, a+R)$. Moreover, we can calculate $\frac{d}{dx}(f'(x)) = f''(x)$ by differentiating the series representing $f'(x)$ term-by-term. Therefore

$$f''(x) = \sum_{n=1}^{\infty} n(n-1)a_n(x-a)^{n-2} = \sum_{n=2}^{\infty} n(n-1)a_n(x-a)^{n-2}.$$

(**Note:** We begin the series at $n = 2$ since for $n = 1$, $n(n-1) = 0$.)

Furthermore, we need not stop here. We can term-by-term differentiate again to show that

$$f'''(x) = \sum_{n=3}^{\infty} n(n-1)(n-2)a_n(x-a)^{n-3}$$

and that

$$f^{(4)}(x) = \sum_{n=4}^{\infty} n(n-1)(n-2)(n-3)a_n(x-a)^{n-4}.$$

In fact, we can show that $f(x)$ has derivatives of all orders on $(a-R, a+R)$ with

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)(n-3)\cdots(n-k+1)a_n(x-a)^{n-k}$$

for each k. ◀

EXAMPLE 9 We know that if $|x| < 1$, then

$$\frac{1}{1-x} = \sum_{n=0}^{\infty} x^n.$$

Let $f(x) = \frac{1}{1-x} = (1-x)^{-1}$. Differentiating, we get

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

We also get a power series representation for $f'(x)$ by term-by-term differentiation of the series $\sum\limits_{n=0}^{\infty} x^n$. It follows that

$$f'(x) = \frac{1}{(1-x)^2} = \sum_{n=1}^{\infty} nx^{n-1}.$$

We can use this to evaluate the sum of the series

$$\sum_{n=1}^{\infty} \frac{n}{2^{n-1}}.$$

Observe that this series is obtained from $\sum\limits_{n=1}^{\infty} nx^{n-1}$ by letting $x = \frac{1}{2}$. Therefore,

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{n}{2^{n-1}} &= f'\left(\frac{1}{2}\right) \\ &= \frac{1}{\left(1-\frac{1}{2}\right)^2} \\ &= 4 \end{aligned}$$

◀

EXAMPLE 10 We have seen that if a function f satisfies the differential equation $y' = y$, then there is a constant C such that $f(x) = Ce^x$. We also know that the series $\sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$ converges for every value of x. Let

$$g(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.$$

We can find $g'(x)$ by term-by-term differentiation. Therefore

$$g'(x) = \sum_{n=1}^{\infty} \frac{nx^{n-1}}{n!}.$$

If we take a close look at this series we see immediately that it becomes

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!}.$$

Writing out the terms, we get

$$\frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

which is identical to the series that gave us $g(x)$. This shows that

$$g'(x) = g(x)$$

for every x. Therefore, there is a C such that $g(x) = Ce^x$. Moreover, $g(0) = Ce^0 = C$. But

$$\begin{aligned} g(0) &= \sum_{n=0}^{\infty} \frac{0^n}{n!} \\ &= \frac{0^0}{0!} + \frac{0^1}{1!} + \frac{0^2}{2!} + \frac{0^3}{3!} + \cdots \\ &= 1 + 0 + 0 + 0 + \cdots \\ &= 1 \end{aligned}$$

Hence $C = 1$ and

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

for every $x \in \mathbb{R}$.

This is an extremely important example that we will come back to again and again. one immediate application is that

$$e = \sum_{n=0}^{\infty} \frac{1^n}{n!} = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

This series converges quite rapidly so we can get a very accurate approximation for e by summing a relatively small number of terms. ◀

EXAMPLE 11 Find a power series representation for the function $f(x) = e^{-x^2}$.

We have just seen for any $u \in \mathbb{R}$ that

$$e^u = \sum_{n=0}^{\infty} \frac{u^n}{n!} \qquad (*)$$

Given $x \in \mathbb{R}$, let $u = -x^2$ and substitute for u in the expression $(*)$. This gives us

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

At first glance

$$e^{-x^2} = 1 - \frac{x^2}{1!} + \frac{x^4}{2!} - \frac{x^6}{3!} + \cdots + (-1)^n \frac{x^{2n}}{n!} + \cdots$$

may not look like a power series since there are no terms involving x^n when n is odd. But in fact, it is a power series where the coefficients are of the form $a_{2k-1} = 0$ and $a_{2k} = (-1)^k \frac{1}{k!}$ for each $k = 0, 1, 2, 3, 4, \ldots$.

Moreover, since the original series converges for all $u \in \mathbb{R}$, the power series for $f(x) = e^{-x^2}$ will also converge for all $x \in \mathbb{R}$. That is, its radius of convergence is $R = \infty$. ◀

Key Observation: We have seen that if

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

for all $x \in (-R, R)$, then $f(x)$ is infinitely differentiable on $(-R, R)$. We can now use this to show that once a function $f(x)$ has a power series representation at $x = 0$, the coefficients are uniquely determined by the various values of the derivatives of $f(x)$. In particular, once we fix the center $x = 0$, the function f can only be represented by one such power series at $x = 0$ (though there may well be other representations of f with different centers). To see why this is true, recall that for any function

$$g(x) = \sum_{n=0}^{\infty} b_n x^n$$

represented by a power series, if we substitute in 0 for x, we get that

$$g(0) = b_0.$$

That is, $g(0)$ is simply the coefficient of the term x^0 in the power series representation of $g(x)$. Therefore, if

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

is *any* power series representation for the function $f(x)$, then $a_0 = f(0)$. This shows that a_0 is in fact, uniquely determined by the value of $f(x)$ at $x = 0$.

We can show that something similar occurs for all of the other coefficients. For example, we note that since

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$$

then $f'(0)$ is the coefficient of x^0 in this series new representation. But we get x^0 when $n = 1$ and when $n = 1$, the coefficient is $(1)(a_1)$. It follows that

$$f'(0) = (1)a_1 = a_1.$$

Alternatively, we could get this directly by substituting $x = 0$ in the expression for $f'(x)$ to get

$$\begin{aligned} f'(0) &= (1)a_1(0)^0 + (2)a_2(0)^1 + (3)a_3(0)^2 + (4)a_4(0)^4 + \cdots \\ &= a_1 + 0 + 0 + 0 + 0 + \cdots \\ &= a_1 \end{aligned}$$

By either method, we see that $f'(0) = a_1$, so the coefficient a_1 is also uniquely determined by $f(x)$.

A similar calculation shows that $f''(0)$ is the coefficient of x^0 in

$$f''(x) = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}.$$

This time to get $x^{n-2} = x^0$, we let $n = 2$. From this we see that the coefficient of x^0 in the previous expression is given by $2(2-1)a_2$ or $2 \cdot 1a_2 = 2!a_2$. As a result we have

$$f''(0) = 2!a_2$$

or that

$$a_2 = \frac{f''(0)}{2!}.$$

Once again, a_2 is uniquely determined.

To find a_3, start with

$$f'''(x) = \sum_{n=3}^{\infty} n(n-1)(n-2)a_n x^{n-3}.$$

Then $f'''(0)$ is the coefficient of x^0 in this series and therefore

$$f'''(0) = (3)(3-1)(3-2)a_3 = 3 \cdot 2 \cdot 1a_3 = 3!a_3.$$

Solving for a_3 shows that

$$a_3 = \frac{f'''(0)}{3!}.$$

In fact, for any $k \geq 2$, we have

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)(n-3)\cdots(n-k+1)a_n x^{n-k}$$

so that the coefficient of x^0 is obtained when $n = k$ and hence

$$\begin{aligned} f^{(k)}(0) &= (k)(k-1)(k-2)\cdots(k-(k-2))(k-(k-1))a_k \\ &= k\cdot(k-1)\cdot(k-2)\cdots 2\cdot 1 a_k \\ &= k!a_k \end{aligned}$$

Solving this expression for a_k shows that for $k \geq 2$,

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

If we note that

$$a_0 = f(0) = \frac{f^{(0)}(0)}{0!}$$

and

$$a_1 = f'(0) = \frac{f^{(1)}(0)}{1!}$$

then we see that for every k,

$$a_k = \frac{f^{(k)}(0)}{k!}.$$

This tells us if a function can be represented by a power series $\sum_{n=0}^{\infty} a_n x^n$ centered at $x = 0$, then the function $f(x)$ and its various derivatives uniquely determine the coefficients.

A similar argument shows that the previous observation holds for a power series centered at $x = a$. This situation motivates the next theorem.

THEOREM 9 **Uniqueness of Power Series Representations**

Suppose that

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

for all $x \in (a-R, a+R)$ where $R > 0$. Then

$$a_n = \frac{f^{(n)}(a)}{n!}.$$

In particular, if

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

then

$$b_n = a_n$$

for each $n = 0, 1, 2, 3, \cdots$.

6.4 Integration of Power Series

Up until now we have seen that a function which is represented by a power series has the remarkable property that it is infinitely differentiable and that its derivatives can be obtained by repeated term-by-term differentiation of the power series. It would make sense to determine if a similar statement could be made with respect to anti-differentiation and integration.

We begin with the following definition:

DEFINITION

Formal Antiderivative of a Power Series

Given a power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$, we define *the formal antiderivative* to be the power series

$$\sum_{n=0}^{\infty} \int a_n(x-a)^n\,dx = C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}.$$

where C is an arbitrary constant.

In the definition, we called the series a *formal* antiderivative because at this point we do not know if it is an actual antiderivative of f. In fact we do not even know if the formal antiderivative converges at any point other than the center.

Problem: Suppose that the power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ has radius of convergence $R > 0$. Let

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

be the function that is represented by this power series on the interval $(a-R, a+R)$. Are the formal anitderivatives

$$C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

true antiderivatives of the function f?

To see why this is in fact the case we first note that the series

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^n$$

is obtained from $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ by dividing each coefficient a_n by $q(n) = n+1$. We know that this will not change the radius of convergence. Hence, $\sum\limits_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^n$ will also converge on $(a-R, a+R)$. Then the formal antiderivative

$$C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

will have radius of convergence R. This means that we can define a function

$$F(x) = C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

on $(a-R, a+R)$.

Next we show that F is an antiderivative of f.

Since F is represented by a power series, it is differentiable. Moreover, its derivative can be obtained by term-by-term differentiation. Therefore

$$\begin{aligned} F'(x) &= \frac{d}{dx}(C) + \sum_{n=0}^{\infty} \frac{d}{dx}\left(\frac{a_n}{n+1}(x-a)^{n+1}\right) \\ &= 0 + \sum_{n=0}^{\infty}(n+1)\left(\frac{a_n}{n+1}\right)(x-a)^n \\ &= \sum_{n=0}^{\infty} a_n(x-a)^n \\ &= f(x) \end{aligned}$$

so F is an antiderivative of f. Moreover, since the constant C is arbitrary, **all antiderivatives have this form**.

The next theorem summarizes what we have just discussed. It also tells us that we can evaluate definite integrals using term-by-term methods. (The proof of this last statement is beyond the scope of the course.)

THEOREM 10 **Term-by-Term Integration of Power Series**

Assume that the power series $\sum\limits_{n=0}^{\infty} a_n(x-a)^n$ has radius of convergence $R > 0$. Let

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

for every $x \in (a-R, a+R)$. Then the series

$$\sum_{n=0}^{\infty} \int a_n(x-a)^n\, dx = C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

also has radius of convergence R and if

$$F(x) = C + \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x-a)^{n+1}$$

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

Furthermore, if $[c, b] \subset (a - R, a + R)$, then

$$\begin{aligned}\int_c^b f(x)\,dx &= \int_c^b \sum_{n=0}^{\infty} a_n(x-a)^n\,dx \\ &= \sum_{n=0}^{\infty} \int_c^b a_n(x-a)^n\,dx \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n+1} \cdot ((b-a)^{n+1} - (c-a)^{n+1})\end{aligned}$$

Note: Similar to the case with the term-by-term differentiation rule for power series, it should seem natural that we are also able to integrate term-by-term the functions that are represented by a power series. But as was the case with differentiation, the reason we are able to do this is because of some very special properties impacting how power series converge. If

$$F(x) = \sum_{n=1}^{\infty} f_n(x)$$

for each $x \in [a, b]$, then we might hope that

$$\int_a^b F(x)\,dx = \sum_{n=0}^{\infty} \int_a^b f_n(x)\,dx.$$

Unfortunately, if we do not make any additional assumptions about the nature of the functions f_n or how the series converges, then this statement could be false. In fact, it is possible that the function F need not even be integrable on $[a, b]$. ◀

We can use the previous theorem to build power series representations for many functions.

EXAMPLE 12 Find a power series representation for $\ln(1 + x)$.

We begin by recognizing that

$$\frac{d}{dx}(\ln(1+x)) = \frac{1}{1+x}.$$

If we had a power series representation for $\frac{1}{1+x}$, we could use the previous theorem to build a representation for $\ln(1 + x)$.

We know that for any $u \in (-1, 1)$,

$$\frac{1}{1-u} = \sum_{n=0}^{\infty} u^n.$$

Let $x \in (-1,1)$. Then $(-x) \in (-1,1)$. If we let $u = -x$ and substitute this into the previous equation, then

$$\frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-x)^n$$

so that for any $(-x) \in (-1,1)$,

$$\begin{aligned} \frac{1}{1+x} &= \frac{1}{1-(-x)} \\ &= \sum_{n=0}^{\infty} (-x)^n \\ &= \sum_{n=0}^{\infty} (-1)^n x^n \end{aligned}$$

We now know $\frac{d}{dx}(\ln(1+x)) = \frac{1}{1+x}$ and that

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n.$$

Therefore, there is a constant C such that

$$\begin{aligned} \ln(1+x) &= C + \sum_{n=0}^{\infty} \int (-1)^n x^n \, dx \\ &= C + \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1} \end{aligned}$$

for all $x \in (-1,1)$.

To find C, let $x = 0$. Then

$$\begin{aligned} 0 &= \ln(1+0) \\ &= C + \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} 0^{n+1} \\ &= C \end{aligned}$$

This shows that for all $x \in (-1,1)$,

$$\ln(1+x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1}.$$

There is one more very useful observation that we can make regarding this example. We know that the series

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} x^{n+1}$$

has radius of convergence $R = 1$. However, if we let $x = 1$, the series becomes

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} 1^{n+1} = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

which is exactly the Alternating Series. Therefore, the series also converges at $x = 1$. However, the equation

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

is actually valid wherever this new series converges. This tells us that

$$\ln(2) = \ln(1+1) = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots$$

so we have found the sum of the alternating series.

◀

6.5 Review of Taylor Polynomials

Recall that if f is differentiable at $x = a$, then if $x \cong a$

$$f'(a) \cong \frac{f(x) - f(a)}{x - a}.$$

Cross-multiplying gives us

$$f'(a)(x-a) \cong f(x) - f(a)$$

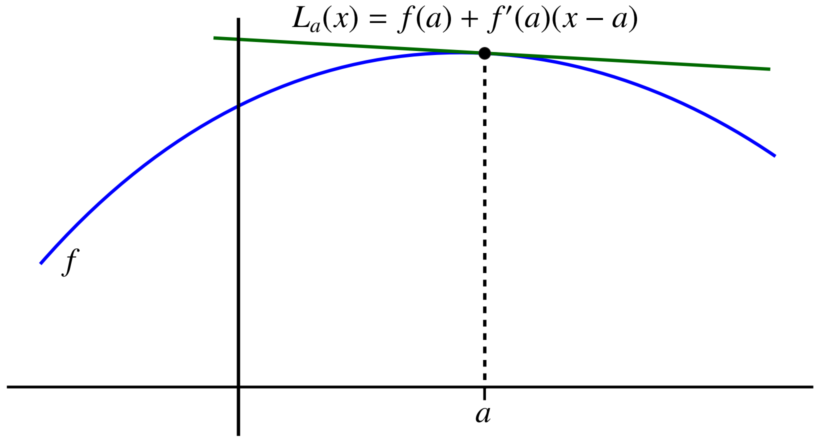
and finally that

$$f(x) \cong f(a) + f'(a)(x-a).$$

This led us to define the *linear approximation to f at $x = a$* to be the function

$$L_a(x) = f(a) + f'(a)(x-a).$$

We saw that the geometrical significance of the linear approximation is that its graph is the tangent line to the graph of f through the point $(a, f(a))$.



Recall also that the linear approximation has the following two important properties:

1. $L_a(a) = f(a)$.
2. $L'_a(a) = f'(a)$.

In fact, amongst all polynomials of degree at most 1, that is functions of the form

$$p(x) = c_0 + c_1(x - a),$$

the linear approximation is the only one with both properties (1) and (2) and as such, the only one that encodes both the value of the function at $x = a$ and its derivative.

We know that for x near a that

$$f(x) \cong L_a(x).$$

This means that we can use the simple linear function L_a to approximate what could be a rather complicated function f at points near $x = a$. However, any time we use a process to approximate a value, it is best that we understand as much as possible about the **error** in the procedure. In this case, the error in the linear approximation is

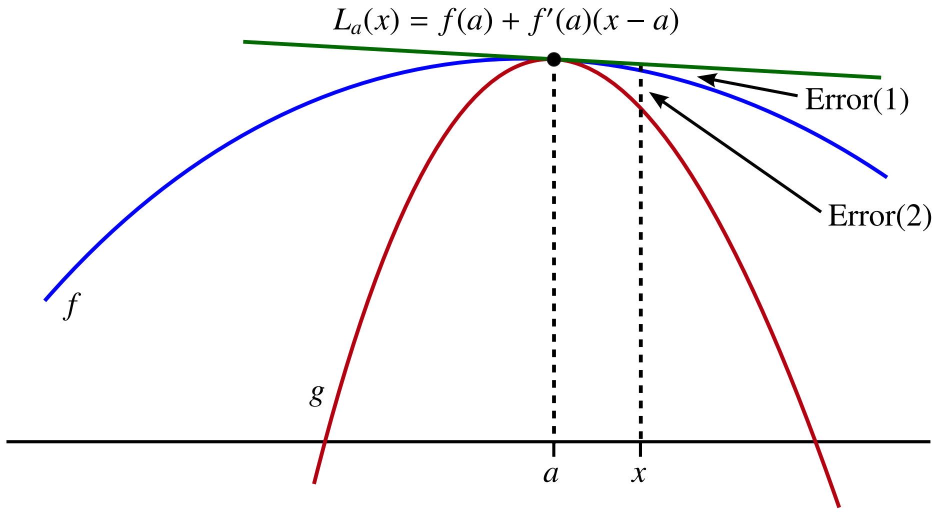
$$\text{Error}(x) = \mid f(x) - L_a(x) \mid$$

and at $x = a$ the estimate is exact since $L_a(a) = f(a)$.

There are two basic factors that affect the potential size of the error in using linear approximation. These are

1. The distance between x and a. That is, how large is $\mid x - a \mid$?
2. How *curved* the graph is near $x = a$?

Note that the larger $\mid f''(x) \mid$ is, the more rapidly the tangent lines turn, and hence the more curved the graph of f. For this reason the second factor affecting the size of the error can be expressed in terms of the size of $\mid f''(x) \mid$. Generally speaking, the further x is away from a and the more curved the graph of f, the larger the potential for error in using linear approximation. This is illustrated in the following diagram which shows two different functions, f and g, with the same tangent line at $x = a$. The error in using the linear approximation is the length of the vertical line joining the graph of the function and the graph of the linear approximation.



Notice that in the diagram, the graph of g is much more curved near $x = a$ than is the graph of f. You can also see that at the chosen point x the error

$$\text{Error}(1) = | f(x) - L_a(x) |$$

in using $L_a(x)$ to estimate the value of $f(x)$ is extremely small, whereas the error

$$\text{Error}(2) = | g(x) - L_a(x) |$$

in using $L_a(x)$ to estimate the value of $g(x)$ is noticeably larger. The diagram also shows that for both f and g, the further away x is from a, the larger the error is in the linear approximation process.

In the case of the function g, its graph looks more like a parabola (second degree polynomial) than it does a line. This suggests that it would make more sense to try and approximate g with a function of the form

$$p(x) = c_0 + c_1(x - a) + c_2(x - a)^2.$$

(Notice that the form for this polynomial looks somewhat unusual. You will see that we write it this way because this form makes it easier to properly encode the information about f at $x = a$).

In constructing the linear approximation, we encoded the value of the function and of its derivative at the point $x = a$. We want to again encode this local information, but we want to do more. If we can include the second derivative, we might be able to capture the curvature of the function that was missing in the linear approximation. In summary, we would like to find constants c_0, c_1, and c_2, so that

1. $p(a) = f(a)$,
2. $p'(a) = f'(a)$, and
3. $p''(a) = f''(a)$.

It may not seem immediately obvious that we can find such constants. However, this task is actually not too difficult. For example, if we want $p(a) = f(a)$, then by noting that

$$p(a) = c_0 + c_1(a-a) + c_2(a-a)^2 = c_0$$

we immediately know that we should let $c_0 = f(a)$.

We can use the standard rules of differentiation to show that

$$p'(x) = c_1 + 2c_2(x-a).$$

In order that $p'(a) = f'(a)$, we have

$$f'(a) = p'(a) = c_1 + 2c_2(a-a) = c_1.$$

Finally, since

$$p''(x) = 2c_2$$

for all x, if we let $c_2 = \frac{f''(a)}{2}$, we have

$$p''(a) = 2c_2 = 2(\frac{f''(a)}{2}) = f''(a)$$

exactly as required. This shows that if

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2,$$

then p is the unique polynomial of degree 2 or less such that

1. $p(a) = f(a)$,
2. $p'(a) = f'(a)$, and
3. $p''(a) = f''(a)$.

The polynomial p is called the *second degree Taylor polynomial for f centered at $x = a$*. We denote this Taylor polynomial by $T_{2,a}$.

EXAMPLE 13 Let $f(x) = \cos(x)$. Then,

$$f(0) = \cos(0) = 1, \quad \text{and} \quad f'(0) = -\sin(0) = 0,$$

and

$$f''(0) = -\cos(0) = -1.$$

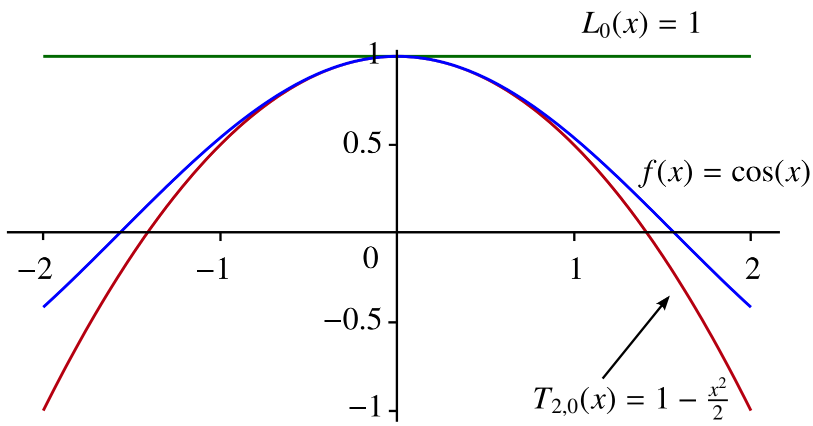
It follows that

$$L_0(x) = f(0) + f'(0)(x-0) = 1 + 0(x-0) = 1$$

for all x while

$$\begin{aligned} T_{2,0}(x) &= f(0) + f'(0)(x-0) + \frac{f''(0)}{2}(x-0)^2 \\ &= 1 + 0(x-0) + \frac{-1}{2}(x-0)^2 \\ &= 1 - \frac{x^2}{2}. \end{aligned}$$

The following diagram shows $\cos(x)$ with its linear approximation and its second degree Taylor polynomial centered at $x = 0$.



Notice that the second degree Taylor polynomial $T_{2,0}$ does a much better job approximating $\cos(x)$ over the interval $[-2, 2]$ than does the linear approximation L_0.

◀

We might guess that if f has a third derivative at $x = a$, then by encoding the value $f'''(a)$ along with $f(a)$, $f'(a)$ and $f''(a)$, we may do an even better job of approximating $f(x)$ near $x = a$ than we did with either L_a or with $T_{2,a}$. As such we would be looking for a polynomial of the form

$$p(x) = c_0 + c_1(x-a) + c_2(x-a)^2 + c_3(x-a)^3$$

such that

1. $p(a) = f(a)$,
2. $p'(a) = f'(a)$,
3. $p''(a) = f''(a)$, and
4. $p'''(a) = f'''(a)$.

To find such a p, we follow the same steps that we outlined before. We want $p(a) = f(a)$, but $p(a) = c_0 + c_1(a-a) + c_2(a-a)^2 + c_3(a-a)^3 = c_0$, so we can let $c_0 = f(a)$.

Differentiating p we get

$$p'(x) = c_1 + 2c_2(x-a) + 3c_3(x-a)^2$$

so that

$$p'(a) = c_1 + 2c_2(a-a) + 3c_3(a-a)^2 = c_1.$$

Therefore, if we let $c_1 = f'(a)$ as before, then we will get $p'(a) = f'(a)$.

Differentiating p' gives us

$$p''(x) = 2c_2 + 3(2)c_3(x-a).$$

Therefore,

$$p''(a) = 2c_2 + 3(2)c_3(a-a) = 2c_2.$$

Now if we let $c_2 = \frac{f''(a)}{2}$, we get

$$p''(a) = f''(a).$$

Finally, observe that

$$p'''(x) = 3(2)c_3 = 3(2)(1)c_3 = 3!c_3$$

for all x, so if we require

$$p'''(a) = 3!c_3 = f'''(a),$$

then we need only let $c_3 = \frac{f'''(a)}{3!}$.

It follows that if

$$p(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3,$$

then

1. $p(a) = f(a)$,
2. $p'(a) = f'(a)$,
3. $p''(a) = f''(a)$, and
4. $p'''(a) = f'''(a)$.

In this case, we call p the *third degree Taylor polynomial centered at $x = a$* and denote it by $T_{3,a}$.

Given a function f, we could also write

$$T_{0,a}(x) = f(a)$$

and

$$T_{1,a}(x) = L_a(x) = f(a) + f'(a)(x-a)$$

and call these polynomials the *zero-th degree* and the *first degree Taylor polynomials of f centered at $x = a$*, respectively.

Observe that using the convention where $0! = 1! = 1$ and $(x-a)^0 = 1$, we have the following:

$$\begin{aligned}
T_{0,a}(x) &= \frac{f(a)}{0!}(x-a)^0 \\
T_{1,a}(x) &= \frac{f(a)}{0!}(x-a)^0 + \frac{f'(a)}{1!}(x-a)^1 \\
T_{2,a}(x) &= \frac{f(a)}{0!}(x-a)^0 + \frac{f'(a)}{1!}(x-a)^1 + \frac{f''(a)}{2!}(x-a)^2 \\
T_{3,a}(x) &= \frac{f(a)}{0!}(x-a)^0 + \frac{f'(a)}{1!}(x-a)^1 + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3.
\end{aligned}$$

Recall that $f^{(k)}(a)$ denotes the k-th derivative of f at $x = a$. By convention, $f^{(0)}(x) = f(x)$. Then using summation notation, we have

$$T_{0,a}(x) = \sum_{k=0}^{0} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

$$T_{1,a}(x) = \sum_{k=0}^{1} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

$$T_{2,a}(x) = \sum_{k=0}^{2} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

and

$$T_{3,a}(x) = \sum_{k=0}^{3} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

This leads us to the following definition:

DEFINITION

Taylor Polynomials

Assume that f is n-times differentiable at $x = a$. The *n-th degree Taylor polynomial for f centered at $x = a$* is the polynomial

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

NOTE

A remarkable property about $T_{n,a}$ is that for any k between 0 and n,

$$T_{n,a}^{(k)}(a) = f^{(k)}(a).$$

That is, $T_{n,a}$ encodes not only the value of $f(x)$ at $x = a$ but all of its first n derivatives as well. Moreover, this is the *only* polynomial of degree n or less that does so! ◀

EXAMPLE 14 Find all of the Taylor polynomials up to degree 5 for the function $f(x) = \cos(x)$ with center $x = 0$.

We have already seen that $f(0) = \cos(0) = 1$, $f'(0) = -\sin(0) = 0$, and $f''(0) = -\cos(0) = -1$. It follows that

$$T_{0,0}(x) = 1,$$

and

$$T_{1,0}(x) = L_0(x) = 1 + 0(x-0) = 1$$

for all x, while

$$T_{2,0}(x) = 1 + 0(x-0) + \frac{-1}{2!}(x-0)^2 = 1 - \frac{x^2}{2}.$$

Since $f'''(x) = \sin(x)$, $f^{(4)}(x) = \cos(x)$, and $f^{(5)}(x) = -\sin(x)$, we get $f'''(0) = \sin(0) = 0$, $f^{(4)}(0) = \cos(0) = 1$ and $f^{(5)}(0) = -\sin(0) = 0$. Hence,

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

We also have that

$$\begin{aligned} T_{4,0}(x) &= 1 + 0(x-0) + \frac{-1}{2!}(x-0)^2 + \frac{0}{3!}(x-0)^3 + \frac{1}{4!}(x-0)^4 \\ &= 1 - \frac{x^2}{2} + \frac{x^4}{24} \end{aligned}$$

and

$$\begin{aligned} T_{5,0}(x) &= 1 + 0(x-0) + \frac{-1}{2!}(x-0)^2 + \frac{0}{3!}(x-0)^3 + \frac{1}{4!}(x-0)^4 + \frac{0}{5!}(x-0)^5 \\ &= 1 - \frac{x^2}{2} + \frac{x^4}{24} \\ &= T_{4,0}(x) \end{aligned}$$

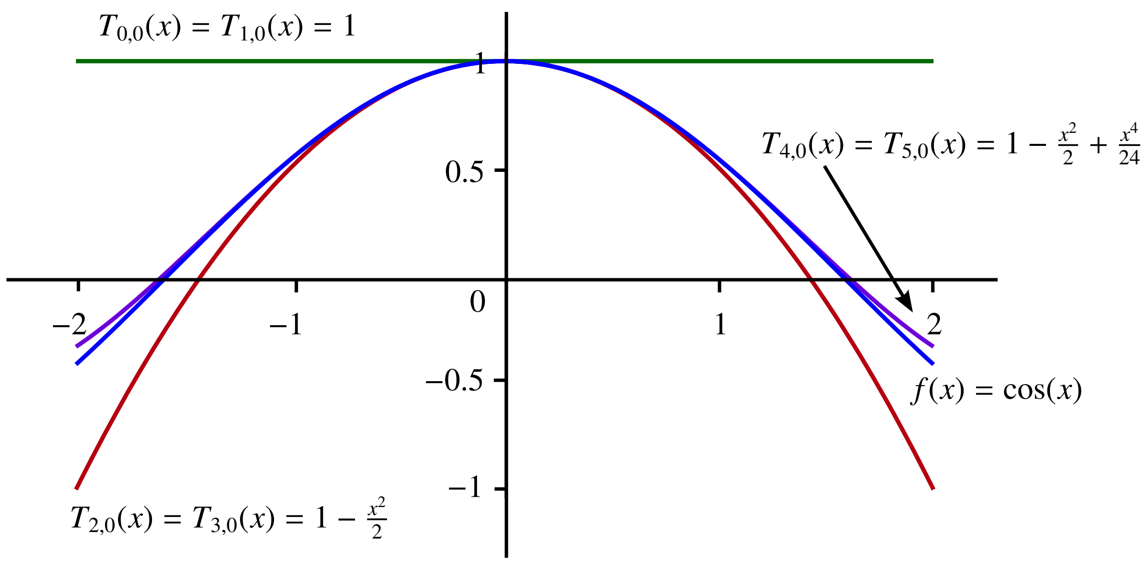
◀

An important observation to make is that not all of these polynomials are distinct. In fact, $T_{0,0}(x) = T_{1,0}(x)$, $T_{2,0}(x) = T_{3,0}(x)$, and $T_{4,0}(x) = T_{5,0}(x)$. In general, this equality of different order Taylor polynomials happens when one of the derivatives is 0 at $x = a$. (In this example at $x = 0$.) This can be seen by observing that for any n

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

so if $f^{(n+1)}(a) = 0$, we get $T_{n+1,a}(x) = T_{n,a}(x)$.

The following diagram shows $\cos(x)$ and its Taylor polynomials up to degree 5. You will notice that there are only four distinct graphs.



In the next example, we will calculate the Taylor Polynomials for $f(x) = \sin(x)$.

EXAMPLE 15 Find all of the Taylor polynomials up to degree 5 for the function $f(x) = \sin(x)$ with center $x = 0$.

We can see that $f(0) = \sin(0) = 0$, $f'(0) = \cos(0) = 1$, $f''(0) = -\sin(0) = 0$, $f'''(0) = -\cos(0) = -1$, $f^{(4)}(0) = \sin(0) = 0$, and $f^{(5)}(0) = \cos(0) = 1$. It follows that

$$T_{0,0}(x) = 0,$$

and

$$T_{1,0}(x) = L_0(x) = 0 + 1(x-0) = x$$

and

$$\begin{aligned} T_{2,0}(x) &= 0 + 1(x-0) + \frac{0}{2!}(x-0)^2 \\ &= x \\ &= T_{1,0}(x). \end{aligned}$$

Next we have

$$\begin{aligned} T_{3,0}(x) &= 0 + 1(x-0) + \frac{0}{2!}(x-0)^2 + \frac{-1}{3!}(x-0)^3 \\ &= x - \frac{x^3}{6} \end{aligned}$$

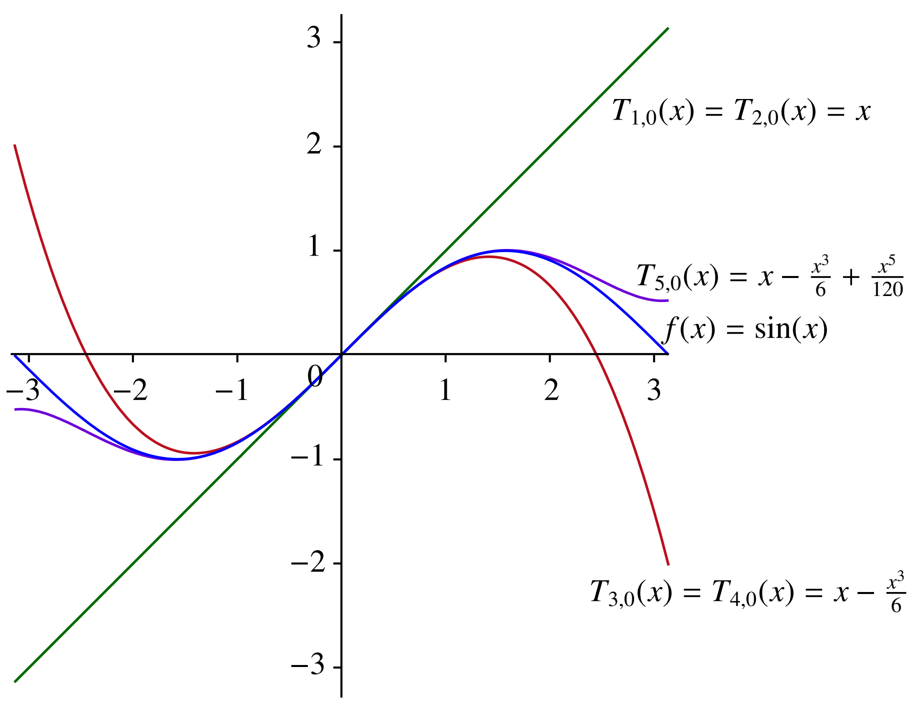
and that

$$\begin{aligned} T_{4,0}(x) &= 0 + 1(x-0) + \frac{0}{2!}(x-0)^2 + \frac{-1}{3!}(x-0)^3 + \frac{0}{4!}(x-0)^4 \\ &= x - \frac{x^3}{6} \\ &= T_{3,0}(x). \end{aligned}$$

Finally,

$$\begin{aligned} T_{5,0}(x) &= 0 + 1(x-0) + \frac{0}{2!}(x-0)^2 + \frac{-1}{3!}(x-0)^3 + \frac{0}{4!}(x-0)^4 + \frac{1}{5!}(x-0)^5 \\ &= x - \frac{x^3}{6} + \frac{x^5}{5!} \\ &= x - \frac{x^3}{6} + \frac{x^5}{120}\,. \end{aligned}$$

The following diagram includes the graph of $\sin(x)$ with its Taylor polynomials up to degree 5, excluding $T_{0,0}$ since its graph is the x-axis.

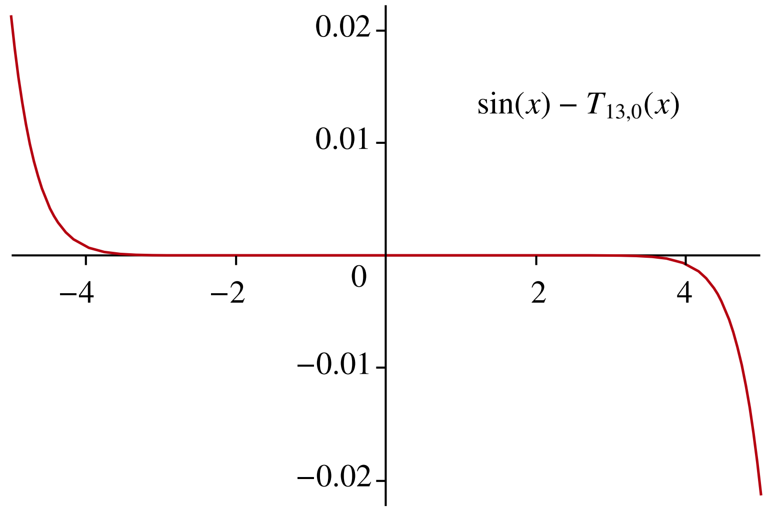


Notice again that the polynomials are not distinct though, in general, as the degree increases so does the accuracy of the estimate near $x = 0$.

To illustrate the power of using Taylor polynomials to approximate functions, we can use a computer to aid us in showing that for $f(x) = \sin(x)$ and $a = 0$, we have

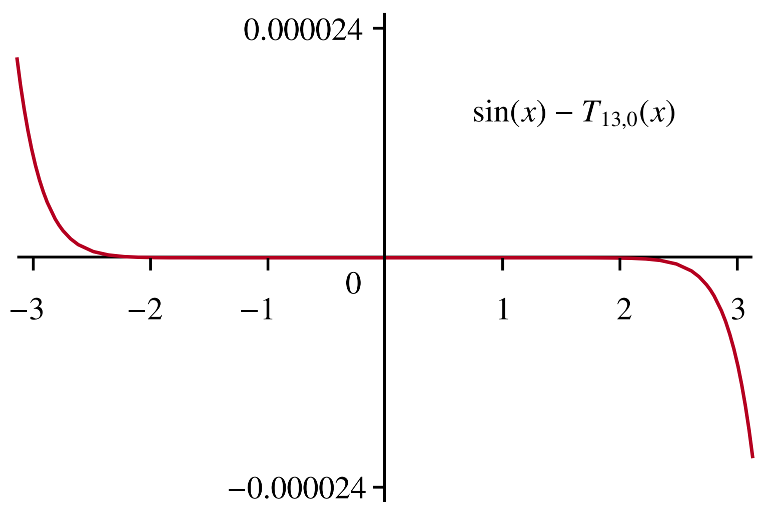
$$T_{13,0}(x) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5 - \frac{1}{5040}x^7 + \frac{1}{362880}x^9 - \frac{1}{39916800}x^{11} + \frac{1}{6227020800}x^{13}$$

The next diagram represents a plot of the function $\sin(x) - T_{13,0}(x)$. (This represents the error between the actual value of $\sin(x)$ and the approximated value of $T_{13,0}(x)$.)



Notice that the error is very small until x approaches 4 or −4. However, the y-scale is different from that of the x-axis, so even near $x = 4$ or $x = -4$ the actual error is still quite small. The diagram suggests that on the slightly more restrictive interval $[-\pi, \pi]$, $T_{13,0}(x)$ does an exceptionally good job of approximating $\sin(x)$.

To strengthen this point even further, we have provided the plot of the graph of $\sin(x) - T_{13,0}(x)$ on the interval $[-\pi, \pi]$.



Note again the scale for the y-axis. It is clear that near 0, $T_{13,0}(x)$ and $\sin(x)$ are essentially indistinguishable. In fact, we will soon have the tools to show that for $x \in [-1, 1]$,

$$| \sin(x) - T_{13,0}(x) | < 10^{-12}$$

while for $x \in [-0.01, 0.01]$,

$$| \sin(x) - T_{13,0}(x) | < 10^{-42}.$$

Indeed, in using $T_{13,0}(x)$ to estimate $\sin(x)$ for very small values of x, round-off errors and the limitations of the accuracy in floating-point arithmetic become much more significant than the true difference between the functions.

EXAMPLE 16 The function $f(x) = e^x$ is particularly well-suited to the process of creating estimates using Taylor polynomials. This is because for any k, the k-th derivative of e^x is again e^x. This means that for any n,

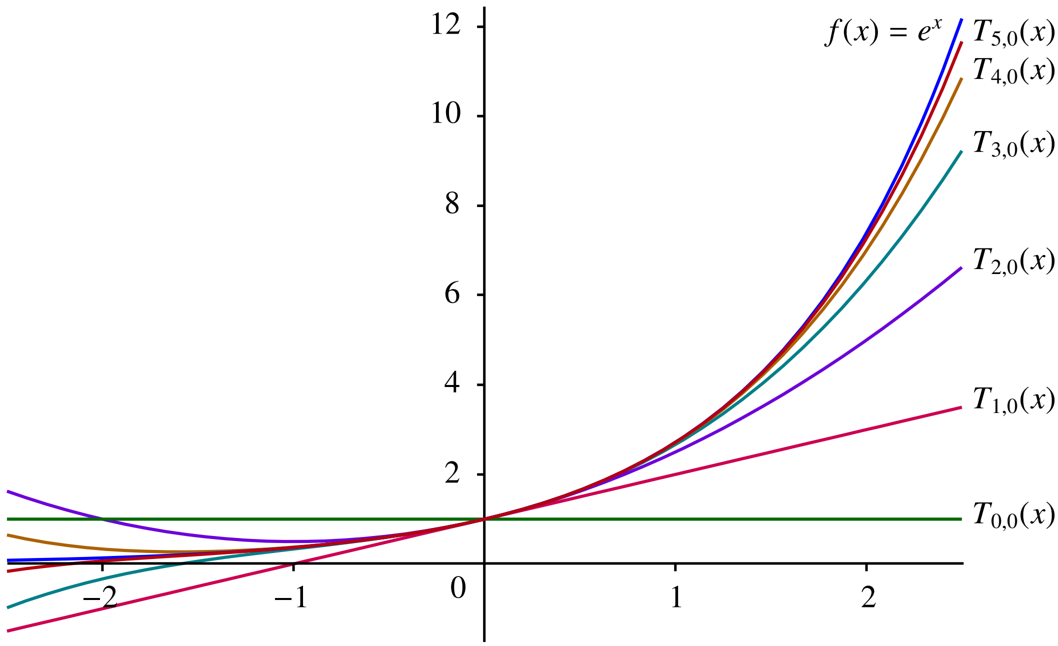
$$\begin{aligned} T_{n,0}(x) &= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \\ &= \sum_{k=0}^{n} \frac{e^0}{k!}(x-0)^k \\ &= \sum_{k=0}^{n} \frac{x^k}{k!}. \end{aligned}$$

In particular,

$$\begin{aligned} T_{0,0}(x) &= 1, \\ T_{1,0}(x) &= 1 + x, \\ T_{2,0}(x) &= 1 + x + \frac{x^2}{2}, \\ T_{3,0}(x) &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6}, \\ T_{4,0}(x) &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}, \text{ and} \\ T_{5,0}(x) &= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120}. \end{aligned}$$

Observe that in the case of e^x, the Taylor polynomials are distinct since e^x, and hence all of its derivatives, is never 0. ◀

The next diagram shows the graphs of e^x and its Taylor polynomials up to degree 5.



6.6 Taylor's Theorem and Errors in Approximations

We have seen that using linear approximation and higher order Taylor polynomials enable us to approximate potentially complicated functions with much simpler ones with surprising accuracy. However, up until now we have only had qualitative information about the behavior of the potential error. We saw that the error in using Taylor polynomials to approximate a function seems to depend on how close we are to the center point. We have also seen that the error in linear approximation seems to depend on the potential size of the second derivative and that the approximations seem to improve as we encode more local information. However, we do not have any precise mathematical statements to substantiate these claims. In this section, we will correct this deficiency by introducing an upgraded version of the Mean Value Theorem called *Taylor's Theorem*.

We begin by introducing some useful notation.

DEFINITION

Taylor Remainder

Assume that f is n times differentiable at $x = a$. Let

$$R_{n,a}(x) = f(x) - T_{n,a}(x).$$

$R_{n,a}(x)$ is called the *n-th degree Taylor remainder function centered at $x = a$*.

The error in using the Taylor polynomial to approximate f is given by

$$\textbf{Error} = \mid R_{n,a}(x) \mid \ .$$

The following is the central problem for this approximation process.

Problem: Given a function f and a point $x = a$, how do we estimate the size of $R_{n,a}(x)$?

The following theorem provides us with the answer to this question.

THEOREM 11 **Taylor's Theorem**

Assume that f is $n + 1$-times differentiable on an interval I containing $x = a$. Let $x \in I$. Then there exists a point c between x and a such that

$$f(x) - T_{n,a}(x) = R_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

We will make **three important observations about Taylor's theorem.**

1) First, since $T_{1,a}(x) = L_a(x)$, when $n = 1$ the absolute value of the remainder $R_{1,a}(x)$ represents the error in using the linear approximation. Taylor's Theorem shows that for some c,

$$| R_{1,a}(x) | = |\frac{f''(c)}{2}(x-a)^2|.$$

This shows explicitly how the error in linear approximation depends on the potential size of $f''(x)$ and on $| x - a |$, the distance from x to a.

2) The second observation involves the case when $n = 0$. In this case, the theorem requires that f be differentiable on I and its conclusion states that for any $x \in I$ there exists a point c between x and a such that

$$f(x) - T_{0,a}(x) = f'(c)(x-a).$$

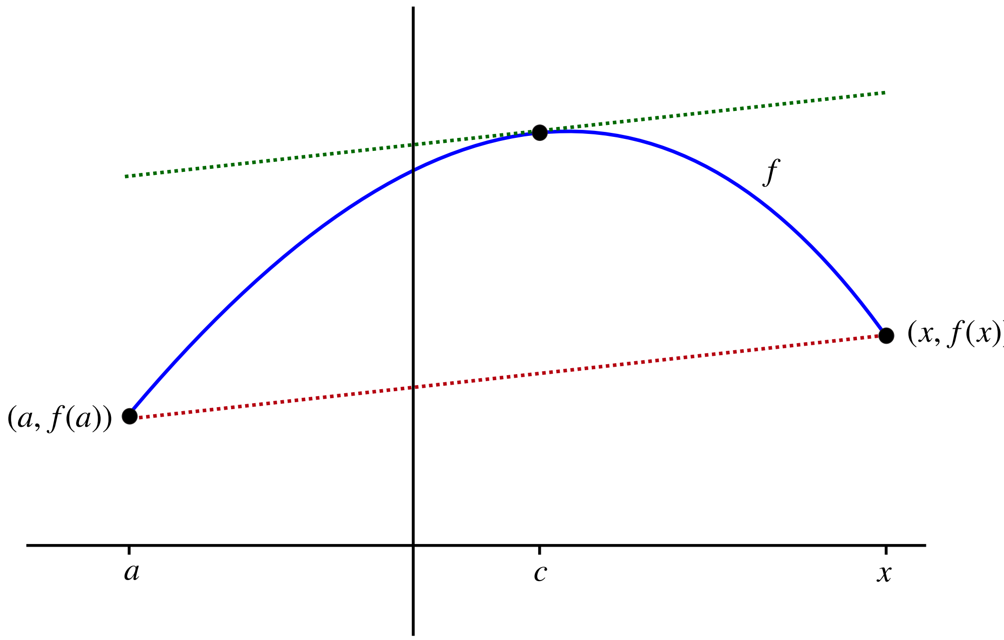
But $T_{0,a}(x) = f(a)$, so we have

$$f(x) - f(a) = f'(c)(x-a).$$

Dividing by $x - a$ shows that there is a point c between x and a such that

$$\frac{f(x) - f(a)}{x - a} = f'(c).$$

This is exactly the statement of the Mean Value Theorem. Therefore, Taylor's Theorem is really a higher-order version of the MVT.



3) Finally, Taylor's Theorem does not tell us how to find the point c, but rather that such a point exists. It turns out that for the theorem to be of any value, we really need to be able to say something intelligent about how large $|\, f^{(n+1)}(c)\, |$ might be without knowing c. For an arbitrary function, this might be a difficult task since higher order derivatives have a habit of being very complicated. However, the good news is that for some of the most important functions in mathematics, such as $\sin(x)$, $\cos(x)$, and e^x, we can determine roughly how large $|\, f^{(n+1)}(c)\, |$ might be and in so doing, show that the estimates obtained for these functions can be extremely accurate.

EXAMPLE 17 Use linear approximation to estimate $\sin(.01)$ and show that the error in using this approximation is less than 10^{-4}.

SOLUTION We know that $f(0) = \sin(0) = 0$ and that $f'(0) = \cos(0) = 1$, so

$$L_0(x) = T_{1,0}(x) = x.$$

Therefore, the estimate we obtain for $\sin(.01)$ using linear approximation is

$$\sin(.01) \cong L_0(.01) = .01$$

Taylor's Theorem applies since $\sin(x)$ is always differentiable. Moreover, if $f(x) = \sin(x)$, then $f'(x) = \cos(x)$ and $f''(x) = -\sin(x)$. It follows that there exists some c between 0 and .01 such that the error in the linear approximation is given by

$$\left|R_{1,0}(.01)\right| = \left|\frac{f''(c)}{2}(.01-0)^2\right| = \left|\frac{-\sin(c)}{2}(.01)^2\right|$$

Recall that the theorem does not tell us the value of c, but rather just that it exists. Not knowing the value of c may seem to make it impossible to say anything significant about the error, but this is actually not the case. The key observation in this example is that regardless the value of point c, $|-\sin(c)| \leq 1$. Therefore,

$$\begin{aligned} |R_{1,0}(.01)| &= \left|\frac{-\sin(c)}{2}(.01)^2\right| \\ &\leq \frac{1}{2}(.01)^2 \\ &< 10^{-4}\,. \end{aligned}$$

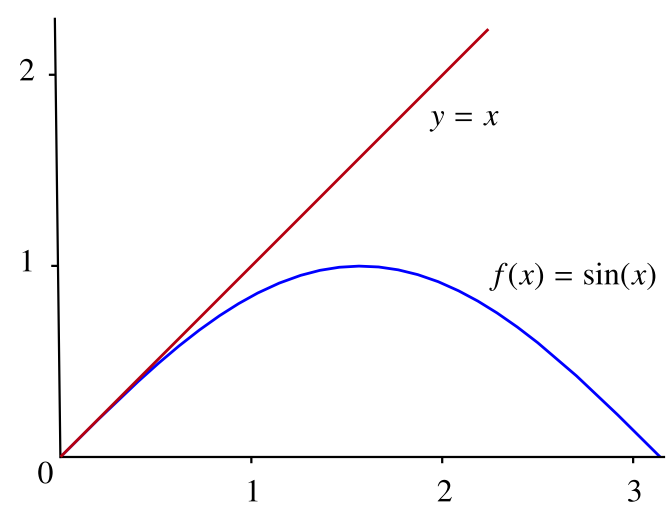
This simple process seems to be remarkably accurate. In fact, it turns out that this estimate is actually much better than the calculation suggests. This is true because not only does $T_{1,0}(x) = x$, but we also have that $T_{2,0}(x) = T_{1,0}(x) = x$. This means that there is a new number c between 0 and .01 such that

$$\begin{aligned} |\sin(.01) - .01| &= |R_{2,0}(.01)| \\ &= \left|\frac{f'''(c)}{6}(.01-0)^3\right| \\ &= \left|\frac{-\cos(c)}{6}(.01)^3\right| \\ &< 10^{-6} \end{aligned}$$

since $|-\cos(c)| \leq 1$ for all values of c.

This shows that the estimate $\sin(.01) \cong .01$ is accurate to six decimal places. In fact, the actual error is approximately $-1.666658333 \times 10^{-7}$.

Finally, we know that for $0 < x < \frac{\pi}{2}$, the tangent line to the graph of $f(x) = \sin(x)$ is above the graph of f since $\sin(x)$ is concave downward on this interval. (In fact, the Mean Value Theorem can be used to show that $\sin(x) \leq x$ for every $x \geq 0$.) Since the tangent line is the graph of the linear approximation, this means that our estimate is actually too large.



Taylor's Theorem can be used to confirm this because

$$\sin(x) - x = R_{1,0}(x) = \frac{-\sin(c)}{2}(x)^2 < 0$$

since $\sin(c) > 0$ for any $c \in (0, \frac{\pi}{2})$. ◀

In the next example we will see how Taylor's Theorem can help in calculating various limits. In order to simplify the notation, we will only consider limits as $x \to 0$.

EXAMPLE 18 Find $\lim\limits_{x\to 0} \dfrac{\sin(x) - x}{x^2}$.

SOLUTION First notice that this is an indeterminate limit of the type $\frac{0}{0}$.

We know that if $f(x) = \sin(x)$, then $T_{1,0}(x) = T_{2,0}(x) = x$. We will assume that we are working with $T_{2,0}$. Then Taylor's Theorem shows that for any $x \in [-1, 1]$, there exists a c between 0 and x such that

$$|\sin(x) - x| = \left|\frac{-\cos(c)}{3!}x^3\right| \leq \frac{1}{6}|x|^3$$

since $|-\cos(c)| \leq 1$ regardless where c is located. This inequality is equivalent to

$$\frac{-1}{6}|x|^3 \leq \sin(x) - x \leq \frac{1}{6}|x|^3 .$$

If $x \neq 0$, we can divide all of the terms by x^2 to get that for $x \in [-1, 1]$

$$\frac{-|x|^3}{6x^2} \leq \frac{\sin(x) - x}{x^2} \leq \frac{|x|^3}{6x^2}$$

or equivalently that

$$\frac{-|x|}{6} \leq \frac{\sin(x) - x}{x^2} \leq \frac{|x|}{6}.$$

We also know that

$$\lim_{x\to 0} \frac{-|x|}{6} = \lim_{x\to 0} \frac{|x|}{6} = 0$$

The Squeeze Theorem guarantees that

$$\lim_{x\to 0} \frac{\sin(x) - x}{x^2} = 0.$$

◀

The technique we outlined in the previous example can be used in much more generality. However, we require the following observation.

Suppose that $f^{(k+1)}$ is a continuous function on $[-1,1]$. Then so is the function

$$g(x) = \left|\frac{f^{(k+1)}(x)}{(k+1)!}\right|.$$

The Extreme Value Theorem tells us that g has a maximum on $[-1,1]$. Therefore, there is an M such that

$$\left|\frac{f^{(k+1)}(x)}{(k+1)!}\right| \leq M$$

for all $x \in [-1,1]$.

Let $x \in [-1,1]$. Taylor's Theorem assures us that there is a c between x and 0 such that

$$|\, R_{k,0}(x)\,| = \left|\frac{f^{(k+1)}(c)}{(k+1)!}x^{k+1}\right|.$$

Therefore,

$$\begin{aligned} |\, f(x) - T_{k,0}(x)\,| &= |\, R_{k,0}(x)\,| \\ &= \left|\frac{f^{(k+1)}(c)}{(k+1)!}x^{k+1}\right| \\ &\leq M\,|\,x\,|^{k+1} \end{aligned}$$

since c is also in $[-1,1]$.

It follows that

$$-M\,|\,x\,|^{k+1} \leq f(x) - T_{k,0}(x) \leq M\,|\,x\,|^{k+1}\ .$$

We summarize this technique as follows:

THEOREM 12 **Taylor's Approximation Theorem I**

Assume that $f^{(k+1)}$ is continuous on $[-1,1]$. Then there exists a constant $M > 0$ such that

$$|\, f(x) - T_{k,0}(x)\,| \leq M\,|\,x\,|^{k+1}$$

or equivalently that

$$-M\,|\,x\,|^{k+1} \leq f(x) - T_{k,0}(x) \leq M\,|\,x\,|^{k+1}$$

for each $x \in [-1,1]$.

This theorem is very helpful in calculating many limits.

EXAMPLE 19 Calculate $\lim\limits_{x\to 0} \dfrac{\cos(x)-1}{x^2}$.

SOLUTION We know that for $f(x) = \cos(x)$ we have $T_{2,0} = 1 - \frac{x^2}{2}$. Moreover, all of the derivatives of $\cos(x)$ are continuous everywhere. The Taylor Approximation Theorem tells us that there is a constant M such that

$$-M\mid x\mid^3 \le \cos(x) - (1 - \frac{x^2}{2}) \le M\mid x\mid^3$$

for all $x \in [-1, 1]$. Dividing by x^2 with $x \neq 0$ we have that

$$-M\mid x\mid \le \frac{\cos(x) - (1 - \frac{x^2}{2})}{x^2} \le M\mid x\mid$$

for all $x \in [-1, 1]$. Simplifying the previous expression produces

$$-M\mid x\mid \le \frac{\cos(x)-1}{x^2} + \frac{1}{2} \le M\mid x\mid$$

for all $x \in [-1, 1]$.

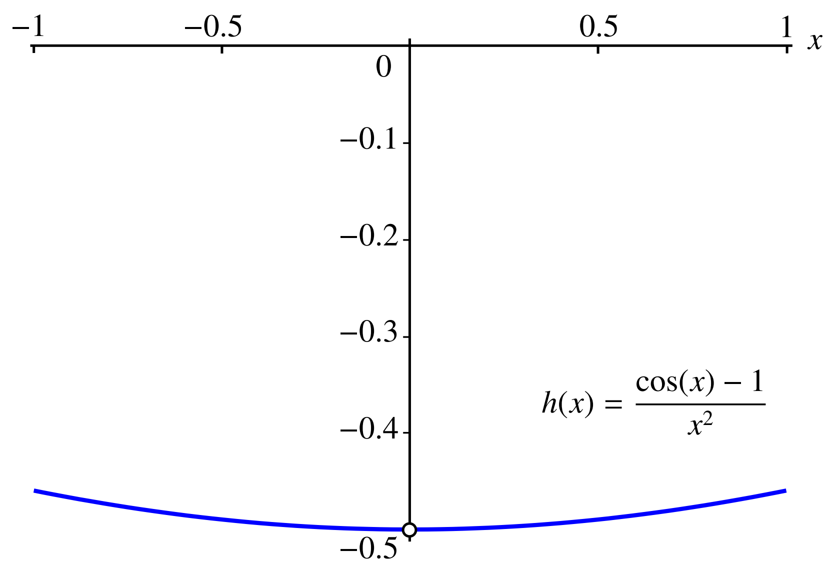
Applying the Squeeze Theorem we have that

$$\lim_{x\to 0} \frac{\cos(x)-1}{x^2} + \frac{1}{2} = 0$$

which is equivalent to

$$\lim_{x\to 0} \frac{\cos(x)-1}{x^2} = \frac{-1}{2}.$$

This limit is consistent with the behavior of the function $h(x) = \frac{\cos(x)-1}{x^2}$ near 0. This is illustrated in the following graph.



◀

The previous limit can actually be calculated quite easily using L'Hôpital's Rule. As an exercise, you should try to verify the answer using this rule. The next example would require much more work using L'Hôpital's Rule. It is provided to show you how powerful Taylor's Theorem can be for finding limits.

EXAMPLE 20 Find $\lim\limits_{x\to 0} \dfrac{e^{\frac{x^4}{2}} - \cos(x^2)}{x^4}$.

SOLUTION This is an indeterminate limit of type $\frac{0}{0}$. We know from Taylor's Approximation Theorem that we can find a constant M_1 such that for any $u \in [-1, 1]$

$$-M_1u^2 \leq e^u - (1 + u) \leq M_1u^2$$

since $1 + u$ is the first degree Taylor polynomial of e^u. Now if $x \in [-1, 1]$, then $u = \frac{x^4}{2} \in [-1, 1]$. In fact, $u \in [0, \frac{1}{2}]$. It follows that if $x \in [-1, 1]$ and we substitute $u = \frac{x^4}{2}$, then we get

$$\frac{-M_1x^8}{4} \leq e^{\frac{x^4}{2}} - (1 + \frac{x^4}{2}) \leq \frac{M_1x^8}{4}\,.$$

We also can show that there exists a constant M_2 such that for any $v \in [-1, 1]$

$$-M_2v^4 \leq \cos(v) - (1 - \frac{v^2}{2}) \leq M_2v^4$$

since $1 - \frac{v^2}{2}$ is the third degree Taylor polynomial for $\cos(v)$.

If $x \in [-1, 1]$ then so is x^2. If we let $v = x^2$, then we see that

$$-M_2x^8 \leq \cos(x^2) - (1 - \frac{x^4}{2}) \leq M_2x^8\,.$$

The next step is to multiply each term in the previous inequality by -1 to get

$$-M_2x^8 \leq (1 - \frac{x^4}{2}) - \cos(x^2) \leq M_2x^8.$$

(Remember, multiplying by a negative number reverses the inequality.)

Now add the two inequalities together:

$$-(\frac{M_1}{4} + M_2)x^8 \leq e^{\frac{x^4}{2}} - (1 + \frac{x^4}{2}) + (1 - \frac{x^4}{2}) - \cos(x^2) \leq (\frac{M_1}{4} + M_2)x^8\,.$$

If we let $M = \frac{M_1}{4} + M_2$ and simplify, this inequality becomes

$$-Mx^8 \leq e^{\frac{x^4}{2}} - \cos(x^2) - x^4 \leq Mx^8$$

for all $x \in [-1, 1]$. Dividing by x^4 gives us that

$$-Mx^4 \leq \frac{e^{\frac{x^4}{2}} - \cos(x^2)}{x^4} - 1 \leq Mx^4\,.$$

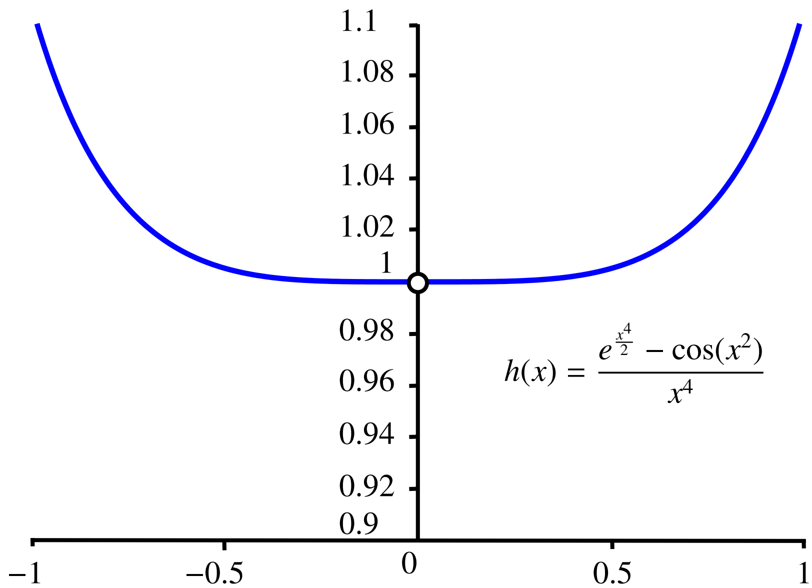
The final step is to apply the Squeeze Theorem to show that

$$\lim_{x\to 0}\frac{e^{\frac{x^4}{2}}-\cos(x^2)}{x^4}-1=0$$

or equivalently that

$$\lim_{x\to 0}\frac{e^{\frac{x^4}{2}}-\cos(x^2)}{x^4}=1.$$

This limit can be confirmed visually from the graph of the function $h(x)=\frac{e^{\frac{x^4}{2}}-\cos(x^2)}{x^4}$.



◀

The previous example involved a rather complicated argument. However, with a little practice using Taylor polynomials and the mastery of a few techniques, limits like this can actually be done by inspection!

6.7 Introduction to Taylor Series

Given a function $f(x)$ that can be represented by a power series

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

centered at $x=a$ with radius of convergence $R>0$, we have seen that $f(x)$ has derivatives of all orders at $x=a$ and that

$$a_n=\frac{f^{(n)}(a)}{n!}.$$

In fact,

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

If we assume that a function f has derivatives of all orders at $a\in\mathbb{R}$ then this series can certainly be constructed.

DEFINITION

Taylor Series

Assume that f has derivatives of all orders at $a\in\mathbb{R}$. The series $\sum\limits_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n$ is called the *Taylor series* for f centered at $x=a$.

We write

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

In the special case where $a=0$, the series is referred to as the *Maclaurin series* for f.

Remark:

Up until now, we have started with a function that was represented by a power series on its interval of convergence. In this case, the series that represents the function must be the Taylor Series.

However, suppose that f is any function for which $f^{(n)}(a)$ exists for each n. Then we can build the power series

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

However, we do not know the following:

1) For which values of x does the series

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

converge?

2) If the series converges at x_0, is it true that

$$f(x_0)=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x_0-a)^n\ ?$$

◀

These two questions essentially ask whether a function f can be fully reconstructed from the data set consisting of the values the derivatives of all order at a point $a\in\mathbb{R}$.

The answer to the first question can be answered by using the method developed for finding the interval of convergence of a power series.

The second problem seems intuitively like it should be true at any point where the series converges. However, a closer look reveals why this may not be true. Essentially we are trying to rebuild a function over an interval that could very well be the entire Real line by using only the information provided by the function at one single point. In this respect, it seems that using only information about e^x at $x = 0$ to get

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

and as such to completely reproduce the function for all values of x seems quite remarkable and indeed it is! To further illustrate why e^x is such a remarkable function in this regard, consider the following example.

EXAMPLE 21 Consider the function g which is obtained by modifying $f(x) = e^x$ outside the interval $[-1, 1]$:

$$g(x) = \begin{cases} \frac{1}{e} & \text{if } x < -1 \\ e^x & \text{if } -1 \leq x \leq 1 \\ e & \text{if } x > 1 \end{cases}$$

On the interval $[-1, 1]$, $g(x)$ behaves exactly like e^x. In particular, $g(0) = e^0 = 1$ and $g^{(n)}(0) = e^0 = 1$ for every n. This means that the Taylor series centered at $x = 0$ for $g(x)$ is

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

which is exactly the same Taylor Series for e^x. We already know that this series converges for all $x \in \mathbb{R}$ and that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This means that the Taylor series for g centered at $x = 0$ also converges for all $x \in \mathbb{R}$ and in particular at $x = 2$. However at $x = 2$, $g(2) = e$ while

$$\sum_{n=0}^{\infty} \frac{2^n}{n!} = e^2 \neq g(2).$$

Hence, this is an example of a function g with the property that its Taylor Series converges at a point x_0 but

$$g(x_0) \neq \sum_{n=0}^{\infty} \frac{g^{(n)}(a)}{n!}(x_0 - a)^n.$$

◀

EXAMPLE 22 Find the Taylor series centered at $x = 0$ for $f(x) = \cos(x)$ and $g(x) = \sin(x)$.

We have that

$$\begin{array}{rclcrclcr}
f'(x) & = & -\sin(x) & \Longrightarrow & f'(0) & = & -\sin(0) & = & 0 \\
f''(x) & = & -\cos(x) & \Longrightarrow & f''(0) & = & -\cos(0) & = & -1 \\
f'''(x) & = & \sin(x) & \Longrightarrow & f'''(0) & = & \sin(0) & = & 0 \\
f^{(4)}(x) & = & \cos(x) & \Longrightarrow & f^{(4)}(0) & = & \cos(0) & = & 1 \\
f^{(5)}(x) & = & -\sin(x) & \Longrightarrow & f^{(5)}(0) & = & -\sin(0) & = & 0 \\
f^{(6)}(x) & = & -\cos(x) & \Longrightarrow & f^{(6)}(0) & = & -\cos(0) & = & -1 \\
f^{(7)}(x) & = & \sin(x) & \Longrightarrow & f^{(7)}(0) & = & \sin(0) & = & 0 \\
f^{(8)}(x) & = & \cos(x) & \Longrightarrow & f^{(8)}(0) & = & \cos(0) & = & 1 \\
 & & & \vdots & & & & &
\end{array}$$

with the cycle repeating itself every four derivatives. Therefore

$$\begin{array}{rclcrclcr}
f^{(4k)}(x) & = & \cos(x) & \Longrightarrow & f^{(4k)}(0) & = & \cos(0) & = & 1 \\
f^{(4k+1)}(x) & = & -\sin(x) & \Longrightarrow & f^{(4k+1)}(0) & = & -\sin(0) & = & 0 \\
f^{(4k+2)}(x) & = & -\cos(x) & \Longrightarrow & f^{(4k+2)}(0) & = & -\cos(0) & = & -1 \\
f^{(4k+3)}(x) & = & \sin(x) & \Longrightarrow & f^{(4k+3)}(0) & = & \sin(0) & = & 0
\end{array}$$

Hence,

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

A similar calculation shows that

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

The problems that remain is to determine if

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$$

and if

$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

◀

To answer these questions we will need to use Taylor's Theorem.

6.8 Convergence of Taylor Series

In this section we return to a question that we asked earlier.

Question: Given a function f that is infinitely differentiable at $x = a$, is f equal to its Taylor Series? That is, does

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

for any x at which the Taylor Series converges?

Unfortunately, we saw that this need not be true even if the Taylor Series converges everywhere. Taylor's Theorem gives us a means to show that for many important functions this equality does hold. To see why this is the case, note that if we fix an x_0, then

$$T_{k,a}(x_0) = \sum_{n=0}^{k} \frac{f^{(n)}(a)}{n!}(x_0-a)^n$$

is not only the k-th partial sum of the Taylor Series for f centered at $x = a$, but it is also the k-th degree Taylor polynomial. As such, Taylor's Theorem shows that

$$\mid f(x_0) - T_{k,a}(x_0) \mid = \mid R_{k,a}(x_0) \mid .$$

If we can show that

$$\lim_{k\to\infty} R_{k,a}(x_0) = 0$$

then

$$f(x) = \lim_{k\to\infty} T_{k,a}(x_0) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Therefore, $f(x)$ agrees with its Taylor series precisely when the Taylor remainders

$$R_{k,a}(x) \to 0$$

as k goes to ∞.

Remark: Before we present the next example we need to recall the following limit which we previously established as a consequence of the Ratio Test.

Let $x_0 \in \mathbb{R}$. Then

$$\lim_{k\to\infty} \frac{M \mid x_0 \mid^k}{k!} = 0.$$

◀

EXAMPLE 23 Let $f(x) = \cos(x)$ and $a = 0$. Let x_0 be any point in $\mathbb{R}$. Taylor's Theorem shows that for each k there exists a point c_k between 0 and x_0 such that

$$| R_{k,a}(x_0) | = \left| \frac{f^{(k+1)}(c_k)}{(k+1)!} x_0^{k+1} \right|$$

We have seen that if $f(x) = \cos(x)$, then $f'(x) = -\sin(x)$, $f''(x) = -\cos(x)$, $f'''(x) = \sin(x)$ and $f^{(4)}(x) = \cos(x)$. Since the fourth derivative is again $\cos(x)$, the 5-th, 6-th, 7-th and 8-th derivative will be, respectively, $f^{(5)}(x) = -\sin(x)$, $f^{(6)}(x) = -\cos(x)$, $f^{(7)}(x) = \sin(x)$ and $f^{(8)}(x) = \cos(x)$. This pattern will be repeated for the 9-th, 10-th, 11-th and 12-th derivatives, and then for every group of four derivatives thereafter. In fact, what we have just shown is that if $f(x) = \cos(x)$, then for any k

$$f^{(k)}(x) = \begin{cases} \cos(x) & \text{if } k = 4j \\ -\sin(x) & \text{if } k = 4j+1 \\ -\cos(x) & \text{if } k = 4j+2 \\ \sin(x) & \text{if } k = 4j+3 \end{cases}$$

where $j = 0, 1, 2, \cdots$. However, this means that regardless the value of k or where c_k is, we will have

$$| f^{(k+1)}(c_k) | \leq 1.$$

It follows immediately that

$$\left|R_{k,a}(x_0)\right| = \left| \frac{f^{(k+1)}(c_k)}{(k+1)!} x_0^{k+1} \right| \leq \frac{| x_0 |^{k+1}}{(k+1)!}.$$

However, we know that $\lim\limits_{k\to\infty} \frac{|x_0|^k}{k!} = 0$, so the Squeeze Theorem shows that

$$\lim_{k\to\infty} R_{k,a}(x_0) = 0.$$

Therefore, since x_0 was chosen at random, for $f(x) = \cos(x)$ and any $x \in \mathbb{R}$, we have

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

In particular, for any $x \in \mathbb{R}$

$$\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{2k!}.$$

A similar argument applies to $\sin(x)$ as it did for $\cos(x)$ to show that for any $x \in \mathbb{R}$, $\sin(x)$ agrees with the value of its Taylor series. That is,

$$\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

◀

Remark: Notice that in each of the previous examples that if either $f(x) = \cos(x)$ or $f(x) = \sin(x)$, then the function f had the property that for any $k = 0, 1, 2, 3, \ldots$ and for each $x \in \mathbb{R}$, then

$$\left| f^{(k)}(x) \right| \leq 1.$$

◀

The fact that we can find a simultaneous *uniform bound* for the size of all of the derivatives of f over all of $\mathbb{R}$ was the key to showing that both $\cos(x)$ and $\sin(x)$ agree with their Taylor series. In fact, these two examples suggest the following very useful theorem.

THEOREM 13 **Convergence Theorem for Taylor Series**

Assume that $f(x)$ has derivatives of all orders on an interval I containing $x = a$. Assume also that there exists an M such that

$$| f^{(k)}(x) | \leq M$$

for all k and for all $x \in I$. Then

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

for all $x \in I$.

PROOF

We know that $T_{k,a}(x)$ is the k-th partial sum of the Taylor series centered at $x = a$. We also know that the Taylor series converges at $x = a$ and that

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

so we only need to show the theorem holds for $x_0 \in I$ with $x_0 \neq a$.

Choose $x_0 \in I$ with $x_0 \neq a$. Let $k \in \mathbb{N} \cup \{0\}$. Then Taylor's Theorem tells us that there exists a c between a and x_0 so that

$$|f(x_0) - T_{k,a}(x_0)| = \frac{|f^{(k+1)}(c)|}{(k+1)!} |x_0 - a|^{k+1}.$$

But since

$$| f^{(k+1)}(c) | \leq M$$

we have that

$$0 \leq |f(x_0) - T_{k,a}(x_0)| \leq M \cdot \frac{|x_0 - a|^{k+1}}{(k+1)!}.$$

Since

$$\lim_{k\to\infty} M \cdot \frac{|x_0 - a|^{k+1}}{(k+1)!} = M \cdot \lim_{k\to\infty} \frac{|x_0 - a|^{k+1}}{(k+1)!} = 0$$

the Squeeze Theorem shows that

$$\lim_{k\to\infty} |f(x_0) - T_{k,a}(x_0)| = 0$$

and hence that

$$f(x_0) = \lim_{k\to\infty} T_{k,a}(x_0) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x_0 - a)^n$$

for all $x_0 \in I$. ■

We showed using term-by-term differentiation of power series that

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

We can now provide a different proof of this important result.

EXAMPLE 24 Let $f(x) = e^x$ and let $a = 0$. Let $I = [-B, B]$. We know that for each k, $f^{(k)}(x) = e^x$. Moreover, since e^x is increasing,

$$0 < e^{-B} \leq e^x \leq e^B$$

for all $x \in [-B, B]$. This means that if $M = e^B$, then for all $x \in [-B, B]$ and all k, we have

$$| f^{(k)}(x) |= e^x \leq e^B = M.$$

All of the conditions of the Convergence Theorem for Taylor Series are satisfied. It follows that for any $x \in [-B, B]$,

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

Finally, we see that this would work regardless of what B we choose. However, given any $x \in \mathbb{R}$, if we pick a B such that $| x |< B$, then $x \in [-B, B]$. This means that for this x

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Hence for every $x \in \mathbb{R}$, the equality

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

holds. ◀

6.9 Binomial Series

Consider the following version of the Binomial Theorem:

THEOREM 14 **Binomial Theorem**

Let $a \in \mathbb{R}$ and $n \in \mathbb{N}$. Then for each $x \in \mathbb{R}$ we have that

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

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

In particular, when $a = 1$ we have

$$(1+x)^n = 1 + \sum_{k=1}^{n} \frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} x^k.$$

Remark: Consider the expression

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}.$$

Typically we are only concerned with the case where $k \in \{0, 1, 2, \ldots, n\}$. But the expression actually makes sense for any $k \in \mathbb{N} \cup \{0\}$. If $k > n$, then one of the terms in the expression

$$n(n-1)(n-2)\cdots(n-k+1)$$

will be 0 and so

$$\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = 0.$$

Consequently,

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

This leaves us to make the rather strange observation that the polynomial function $(1+x)^n$ is actually represented by the power series

$$1+\sum_{k=1}^{\infty}\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}x^k$$

In other words, $1+\sum\limits_{k=1}^{\infty}\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}x^k$ is the Taylor Series centered at $x=0$ for the function $(1+x)^n$.

By itself the observation above does not tell us anything new about the function $(1+x)^n$. However it does give us an important clue towards answering the following question.

Question: Suppose that $\alpha \in \mathbb{R}$. Is there an analog of the Binomial Theorem for the function

$$(1+x)^\alpha?$$

To answer this question, one strategy would be to mimic what happens with the classical Binomial Theorem. We begin by defining the generalized binomial coefficients and the generalized binomial series.

DEFINITION

Generalized Binomial Coefficients and Binomial Series

Let $\alpha \in \mathbb{R}$ and let $k \in \{0,1,2,3,\ldots\}$. Then we define the generalized binomial coefficient

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

if $k \neq 0$ and

$$\binom{\alpha}{0}=1.$$

We also define the generalized binomial series for α to be the power series

$$1+\sum_{k=1}^{\infty}\frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!}x^k=\sum_{k=0}^{\infty}\binom{\alpha}{k}x^k.$$

Remark: The first problem is to find the radius of convergence for the generalized binomial series. To do this let $b_k = \mid \binom{\alpha}{k} \mid$. Then a straight-forward calculation shows that if $k \geq 1$

$$\frac{b_{k+1}}{b_k}=\frac{\mid \alpha-k \mid}{k+1}.$$

It follows that

$$\lim_{k\to\infty} \frac{b_{k+1}}{b_k} = \lim_{k\to\infty} \frac{\mid \alpha - k \mid}{k+1} = 1.$$

This tells us that the radius of convergence for the binomial series is 1. In particular, the series converges absolutely on $(-1, 1)$.

Next we must determine if

$$(1+x)^\alpha = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!} x^k = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k\ ?$$

To see why this is true we start with the following calculation which shows that if $k \geq 1$, then

$$\begin{aligned}
\binom{\alpha}{k+1}(k+1) + \binom{\alpha}{k}k &= \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}(\alpha-k) \\
&+ \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}(k) \\
&= \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}(\alpha) \\
&= \alpha\binom{\alpha}{k}
\end{aligned}$$

Next let

$$f(x) = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!} x^k = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$$

for each $x \in (-1, 1)$. We claim that

$$f'(x) + xf'(x) = \alpha f(x)$$

for each $x \in (-1, 1)$. To see why this is true we use term-by-term differentiation to get that

$$\begin{aligned}
f'(x) + xf'(x) &= \sum_{k=1}^{\infty} \binom{\alpha}{k} kx^{k-1} + \sum_{k=1}^{\infty} \binom{\alpha}{k} kx^k \\
&= \binom{\alpha}{1} + \sum_{k=2}^{\infty} \binom{\alpha}{k} kx^{k-1} + \sum_{k=1}^{\infty} \binom{\alpha}{k} kx^k \\
&= \alpha + \sum_{k=1}^{\infty} \binom{\alpha}{k+1}(k+1)x^k + \sum_{k=1}^{\infty} \binom{\alpha}{k} kx^k \\
&= \alpha + \sum_{k=1}^{\infty} (\binom{\alpha}{k+1}(k+1) + \binom{\alpha}{k}k)x^k
\end{aligned}$$

But if $k \geq 1$ we have

$$\binom{\alpha}{k+1}(k+1) + \binom{\alpha}{k}k = \alpha\binom{\alpha}{k}.$$

It follows that

$$\begin{aligned} f'(x) + xf'(x) &= \alpha + \alpha \sum_{k=1}^{\infty} \binom{\alpha}{k} x^k \\ &= \alpha \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \\ &= \alpha f(x) \end{aligned}$$

as claimed.

Next let

$$g(x) = \frac{f(x)}{(1+x)^{\alpha}}.$$

Then g is differentiable on $(-1, 1)$ with

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

since $\alpha f(x) = (1+x)f'(x)$.

Since $g'(x) = 0$ for all $x \in (-1, 1)$, $g(x)$ is constant on this interval. However,

$$g(0) = f(0) = 1.$$

Therefore, $g(x) = 1$ for all $x \in (-1, 1)$. It follows that

$$f(x) = (1+x)^{\alpha}$$

for all $x \in (-1, 1)$.

THEOREM 15 **Generalized Binomial Theorem**

Let $\alpha \in \mathbb{R}$. Then for each $x \in (-1, 1)$ we have that

$$(1+x)^{\alpha} = 1 + \sum_{k=1}^{\infty} \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-k+1)}{k!} x^k = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

EXAMPLE 25 Use the Generalized Binomial Theorem to find a power series representation for $(1+x)^{-2}$.

The Generalized Binomial Theorem shows that

$$(1+x)^{-2} = \sum_{k=0}^{\infty} \binom{-2}{k} x^k.$$

For $k \geq 1$,

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

It is also true that

$$\binom{-2}{0} = 1 = (-1)^0(0+1).$$

Therefore,

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

We can also use term-by-term differentiation to verify the previous calculation. First begin with

$$\frac{1}{1-u} = \sum_{k=0}^{\infty} u^k$$

for all $u \in (-1, 1)$. Differentiating both sides gives us

$$\frac{1}{(1-u)^2} = \sum_{k=1}^{\infty} ku^{k-1}$$

for all $u \in (-1, 1)$.

Finally, if we let $u = -x$, we have

$$\frac{1}{(1+x)^2} = \sum_{k=1}^{\infty} k(-x)^{k-1} = \sum_{k=1}^{\infty}(-1)^{k-1}kx^{k-1}$$

for all $x \in (-1, 1)$.

◀

6.10 Additional Examples and Applications of Taylor Series

In this section we will present some further examples of functions that are representable by their Taylor series and see what this tells us about these functions.

EXAMPLE 26 Find a power series representation for $f(x) = \arctan(x)$ and determine the interval on which the representation is valid.

We begin with the observation that $\frac{d}{dx}(\arctan(x)) = \frac{1}{1+x^2}$. Therefore, if we can find a power series representation for $\frac{1}{1+x^2}$, we can use the integration techniques to find a representation for $\arctan(x)$.

We know that for any $u \in (-1, 1)$,

$$\frac{1}{1-u} = \sum_{n=0}^{\infty} u^n.$$

Let $x \in (-1, 1)$. If we let $u = -x^2$, then $u \in (-1, 1)$. It follows that

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

Since $\arctan(x)$ is an antiderivative of $\frac{1}{1+x^2}$, the Integration of Power Series Theorem shows that there exists a constant C such that

$$\begin{aligned}\arctan(x) &= C + \sum_{n=0}^{\infty} \int (-1)^n x^{2n}\, dx \\ &= C + \sum_{n=0}^{\infty}(-1)^n \frac{x^{2n+1}}{2n+1}\end{aligned}$$

To find C, note that $\arctan(0) = 0$, so

$$\begin{aligned}0 &= \arctan(0) \\ &= C + \sum_{n=0}^{\infty}(-1)^n \frac{0^{2n+1}}{2n+1} \\ &= C\end{aligned}$$

Therefore, we have shown that if $x \in (-1, 1)$

$$\arctan(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

The radius of convergence of this series is 1. At $x = 1$, the series becomes

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

which converges by the Alternating Series Test.

When $x = -1$, the series is

$$\sum_{n=0}^{\infty} (-1)^n \frac{(-1)^{2n+1}}{2n+1} = \sum_{n=0}^{\infty} (-1)^{3n+1} \frac{1}{2n+1}.$$

Since

$$(-1)^{3n+1} = \begin{cases} 1 & \text{if } n \text{ is odd} \\ -1 & \text{if } n \text{ is even} \end{cases}$$

the series is the same as

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

which also converges by the Alternating Series Test.

Therefore, the Continuity Theorem for Power Series shows that

$$\arctan(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$$

for all $x \in [-1, 1]$.

This last statement has an interesting application. We get that

$$\begin{aligned} \frac{\pi}{4} &= \arctan(1) \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} \end{aligned}$$

Multiplying both sides of this equation by 4 gives

$$\begin{aligned} \pi &= 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots\right) \\ &= 4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} + \frac{4}{9} - \cdots \end{aligned}$$

Note: This series representation for $\arctan(x)$ is called the *Gregory's series* after the Scottish mathematician of the same name. The famous series expansion for π which we derived from Gregory's series is called *Leibniz's formula for* π. ◀

EXAMPLE 27 **(i)** Find the Taylor series centered at $x = 0$ for the integral function

$$F(x) = \int_0^x \cos(t^2)\,dt.$$

(ii) Find $F^{(9)}(0)$ and $F^{(16)}(0)$.

(iii) Estimate $\int_0^{0.1} \cos(t^2)\,dt$ with an error of less than $\frac{1}{10^6}$.

SOLUTIONS

(i) For any $u \in \mathbb{R}$,

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

Let $u = t^2$ to get that for any $t \in \mathbb{R}$,

$$\cos(t^2) = \sum_{n=0}^{\infty} (-1)^n \frac{(t^2)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n}}{(2n)!}.$$

The Integration Theorem for Power Series gives us that

$$\begin{aligned}
F(x) &= \int_0^x \cos(t^2)\,dt \\
&= \int_0^x \sum_{n=0}^{\infty} (-1)^n \frac{t^{4n}}{(2n)!}\,dt \\
&= \sum_{n=0}^{\infty} \int_0^x (-1)^n \frac{t^{4n}}{(2n)!}\,dt \\
&= \sum_{n=0}^{\infty} \left[(-1)^n \frac{t^{4n+1}}{(4n+1)(2n)!} \Bigg|_0^x \right] \\
&= \sum_{n=0}^{\infty} (-1)^n \frac{x^{4n+1}}{(4n+1)(2n)!}
\end{aligned}$$

This is valid for any $x \in \mathbb{R}$. Moreover, by the Uniqueness Theorem for Power Series Representations, this must be the Taylor series centered at $x = 0$ for F.

(ii) To find $F^{(9)}(0)$, we recall that if

$$F(x) = \sum_{k=0}^{\infty} a_k x^k$$

then
$$a_9 = \frac{F^{(9)}(0)}{9!}.$$

This tells us that to find $F^{(9)}(0)$ we must first identify the coefficient of x^9 in
$$\sum_{n=0}^{\infty}(-1)^n\frac{x^{4n+1}}{(4n+1)(2n)!}.$$

Notice for x^9, we let $n = 2$. The coefficient is then
$$(-1)^2\left(\frac{1}{(4(2)+1)(2(2))!}\right) = \frac{1}{9(4!)}$$

Therefore,
$$a_9 = \frac{1}{9(4!)}.$$

Finally, this means
$$\begin{aligned} F^{(9)}(0) &= 9!a_9 \\ &= \frac{9!}{9(4!)} \\ &= 5\cdot 6\cdot 7\cdot 8 \\ &= 1680 \end{aligned}$$

Next, to find $F^{(16)}(0)$ we look for the coefficient of x^{16} in the Taylor series for $F(x)$. However, this time there is no n such that $x^{4n+1} = x^{16}$. This means that $a_{16} = 0$ and hence that
$$F^{(16)}(0) = 0.$$

(iii) Since
$$F(x) = \sum_{n=0}^{\infty}(-1)^n\frac{x^{4n+1}}{(4n+1)(2n)!}$$
we have
$$\int_0^{0.1}\cos(t^2)\,dt = F(0.1) = \sum_{n=0}^{\infty}(-1)^n\frac{(0.1)^{4n+1}}{(4n+1)(2n)!}.$$
This is an alternating series with
$$a_n = \frac{(0.1)^{4n+1}}{(4n+1)(2n)!}.$$

Moreover, we see that
$$a_1 = \frac{(0.1)^5}{(5)(2)!} = \frac{1}{10^6}$$

and

$$\sum_{n=0}^{0}(-1)^n\frac{(0.1)^{4n+1}}{(4n+1)(2n)!} = \frac{(0.1)}{1}.$$

Using the error estimate in the Alternating Series Test we get that

$$\left|\int_0^{0.1}\cos(t^2)\,dt - 0.1\right| < a_1 = \frac{1}{10^6}.$$

◀

Suppose that we wanted to know the value of

$$\int_0^{\frac{1}{2}}\frac{1}{1+x^9}\,dx.$$

Since $\frac{1}{1+x^9}$ looks like a rather simple rational function, we might be tempted to use partial fractions to try and calculate the integral exactly. At least theoretically, this should work. However, in practice, this would require us to factor the polynomial $1+x^9$ which would certainly require the aid of a sophisticated computer algebra program such as Maple. Even then, the answer that we would get would not be very useful. (Try it!)

Fortunately, we can use what we know about series to get an extremely accurate approximation to this integral with surprisingly little effort.

EXAMPLE 28 Estimate $\int_0^{\frac{1}{2}}\frac{1}{1+x^9}\,dx$ with an error of less than 10^{-12}.

We know that for any $-1 < u < 1$,

$$\frac{1}{1-u} = \sum_{n=0}^{\infty}u^n.$$

If $0 \le x \le \frac{1}{2}$, then $-1 < -x^9 < 1$, so we can let $u = -x^9$ to get

$$\frac{1}{1+x^9} = \frac{1}{1-(-x^9)} = \sum_{n=0}^{\infty}(-1)^n x^{9n}$$

for all $x \in [0, \frac{1}{2}]$.

The Integration Theorem for Power Series shows that

$$\begin{aligned}
\int_0^{\frac{1}{2}} \frac{1}{1+x^9}\,dx &= \int_0^{\frac{1}{2}} \sum_{n=0}^{\infty}(-1)^n x^{9n}\,dx \\
&= \sum_{n=0}^{\infty}(-1)^n \int_0^{\frac{1}{2}} x^{9n}\,dx \\
&= \sum_{n=0}^{\infty}(-1)^n \left.\frac{x^{9n+1}}{9n+1}\right|_0^{\frac{1}{2}} \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{(\frac{1}{2})^{9n+1}}{9n+1}
\end{aligned}$$

Notice that the numerical series we have just obtained satisfies the conditions of the Alternating Series Test. In particular, we can use the error estimation in the Alternating Series Test to conclude that

$$\begin{aligned}
\left| \int_0^{\frac{1}{2}} \frac{1}{1+x^9}\,dx - \sum_{n=0}^{k}(-1)^n \frac{(\frac{1}{2})^{9n+1}}{9n+1} \right| &\leq \frac{(\frac{1}{2})^{9(k+1)+1}}{9(k+1)+1} \\
&= \frac{(\frac{1}{2})^{9k+10}}{9k+10}
\end{aligned}$$

If we let $k = 3$, we get

$$\int_0^{\frac{1}{2}} \frac{1}{1+x^9}\,dx \cong \sum_{n=0}^{3}(-1)^n \frac{(\frac{1}{2})^{9n+1}}{9n+1} = \frac{1}{2} - \frac{1}{10(2^{10})} + \frac{1}{19(2^{19})} - \frac{1}{28(2^{28})}$$

with an error that is less than

$$\begin{aligned}
\frac{1}{2^{9(3)+10}[9(3)+10]} &= \frac{1}{37(2^{37})} \\
&= \frac{1}{5.08524 \times 10^{12}}
\end{aligned}$$

◀